COLORED OPERADS, SERIES ON COLORED OPERADS, AND COMBINATORIAL GENERATING SYSTEMS

SAMUELE GIRAUDO

Abstract. Bud generating systems, sorts of combinatorial generating systems, are introduced. They are devices for specifying sets of various kinds of combinatorial objects, called languages. They can emulate context-free grammars, regular tree grammars, and synchronous grammars, allowing us to work with all these generating systems in a unified way. The theory of bud generating systems presented here heavily uses the one of colored operads. Indeed, an object is generated by a bud generating system if it satisfies a certain equation in a colored operad. With the aim to compute the generating series of the languages of bud generating systems, we introduce formal power series on colored operads and several operations on these. Series on colored operads intervene to express the languages specified by bud generating systems and allow us to enumerate combinatorial objects with respect to some statistics. Some examples of bud generating systems are constructed, in particular to specify some sorts of balanced trees and to obtain recursive formulas enumerating these.

Contents

Date: September 6, 2018.

2010 *Mathematics Subject Classification.* 05C05, 18D50, 68Q42, 32A05.

Key words and phrases. Tree; Formal power series; Combinatorial generation; Grammar; Colored operad.

Introduction

Coming from theoretical computer science and formal language theory, formal grammars [Har78, HMU06] are powerful tools having many applications in several fields of mathematics. A formal grammar is a device which describes—more or less concisely and with more or less restrictions—a set of words, called a language. There are several variations in the definitions of formal grammars and some sorts of these are classified by the Chomsky-Schützenberger hierarchy [Cho59, CS63] according to four different categories, taking into account their expressive power. In an increasing order of power, there are the classes of Type-3 to Type-0 grammars, known respectively as regular grammars, context-free grammars, context-sensitive grammars, and unrestricted grammars. One of the most striking similarities between all these variations of formal grammars is that they work by constructing words by applying rewrite rules [BN98]. Indeed, a word of the language described by a formal grammar is obtained by considering a starting word and by iteratively altering some of its factors in accordance with the production rules of the grammar.

Similar mechanisms and ideas are translatable into the world of trees, instead only those of words. Grammars of trees [CDG+07] are thus the natural counterpart of formal grammars to describe sets of trees, and here also, there exist several very different types of grammars. One can cite for instance tree grammars, regular tree grammars [GS84], and synchronous grammars [Gir12], which are devices providing a way to describe sets of various kinds of treelike structures. Here also, one of the common points between these grammars is that they work by applying rewrite rules on trees. In this framework, trees are constructed by growing from the root to the leaves by replacing some subtrees by other ones.

Furthermore, the theory of operads seems to have virtually no link with the one of formal grammars. Operads are algebraic structures introduced in the context of algebraic topology [May72, BV73] (see also [Mar08, LV12, Mén15] for a modern conspectus of the theory). This theory has somewhat been almost neglected during the first two decades after its discovery. In the 1990s, the theory of operads enjoyed a renaissance raised by Loday [Lod96] and, from the 2000s, many links between the theory of operads and combinatorics have been developed (see, for instance [CL01, Cha08, CG14]). Therefore, in the last years, a lot of operads involving various sets of combinatorial objects have been defined, so that almost every classical object can be seen as an element of at least one operad (see the previous references and for instance [Zin12, Gir15, Gir16a, FFM18]). From an intuitive point of view, an operad is a set of abstract operators with several inputs and one output that can be composed in many ways. More precisely, if x is an operator with n inputs and y is an operator with m inputs, $x \circ_i y$ denotes the operator with $n + m - 1$ inputs obtained by gluing the output of y to the i-th input of x. Operads are algebraic structures related to trees in the same ways as monoids are algebraic structures related to words. For this reason, the study of operads has many connections with the one of combinatorial properties of trees.

The initial spark of this work has been caused by the following simple observation. The partial composition $x \circ_i y$ of two elements x and y of an operad $\mathcal{O}$ can be regarded as the application of a rewrite rule on x to obtain a new element of $\mathcal{O}$—the rewrite rule being encoded essentially by y. This leads to the idea consisting in considering an operad $\mathcal{O}$ to define grammars generating some subsets of $\mathcal{O}$. In this way, according to the nature of the elements of $\mathcal{O}$, this provides a way to define grammars which generate objects different than words (as in the case of formal grammars) and than trees (as in the case of grammars of trees). We rely in this work on colored operads [BV73, Yau16], a generalization of operads. In a colored operad $\mathcal{C}$, every input and every output for the elements of $\mathcal{C}$ has a color, taken from a fixed set. These colors lead to the creation of constraints for the partial compositions of two elements. Indeed, $x \circ_i y$ is defined only if the color of the output of y is the same as the color of the i-th input of x. Colored operads are the suitable devices to our aim of defining a new kind of grammars since the restrictions provided by the colors allow a precise control on how the rewrite rules can be applied.

Thus, we introduce in this work a new kind of grammars, the bud generating systems. They are defined mainly from a ground operad $\mathcal{O}$, a set $\mathscr{C}$ of colors, and a set $\mathfrak{R}$ of production rules. A bud generating system describes a subset of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$—the colored operad obtained by augmenting the elements of $\mathcal{O}$ with input and output colors taken from $\mathscr{C}$. The generation of an element works by iteratively altering an element x of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ by composing it, if possible, with an element y of $\mathfrak{R}$. In this context, the colors play the role analogous of the one of nonterminal symbols in the formal grammars and in the grammars of trees. Any bud generating system $\mathcal{B}$ specifies two sets of objects: its language $\mathrm{L}(\mathcal{B})$ and its synchronous language $\mathrm{L}_{\mathrm{S}}(\mathcal{B})$. Thereby, bud generating systems can be used to describe several sets of combinatorial objects. For instance, they can be used to describe sets of Motkzin paths with some constraints, the set of $\{2, 3\}$-perfect trees [MPRS79, CLRS09] and some of its generalizations, and the set of balanced binary trees [AVL62]. One remarkable fact is that bud generating systems can emulate both context-free grammars and regular tree grammars, and allow us to see both of these in a unified manner. In the first case, context-free grammars are emulated by bud generating systems with the associative operad $\mathbf{As}$ as ground operad and in the second case, regular tree grammars are emulated by bud generating systems with a free operad $\mathbf{Free}(C)$ as ground operad, where C is a suitable set of generators.

A very normal combinatorial question consists, given a bud generating system $\mathcal{B}$, in computing the generating series $\mathbf{s}_{\mathrm{L}(\mathcal{B})}(t)$ and $\mathbf{s}_{\mathrm{L}_{\mathrm{S}}(\mathcal{B})}(t)$, respectively counting the elements of the language and of the synchronous language of $\mathcal{B}$ with respect to the arity of the elements. To achieve this objective, we develop a new generalization of formal power series, namely series on colored operads. Any bud generating system $\mathcal{B}$ leads to the definition of three series on colored operads: its hook generating series $\mathrm{hook}(\mathcal{B})$, its syntactic generating series $\mathrm{synt}(\mathcal{B})$, and its synchronous generating series $\mathrm{sync}(\mathcal{B})$. The hook generating series allows us to define analogues of the hook-length statistic of binary trees [Knu98] for objects belonging to the language of $\mathcal{B}$, possibly different than trees. The syntactic (resp. synchronous) generating series leads to obtain functional equations and recurrence formulas to compute the coefficients of $\mathbf{s}_{\mathrm{L}(\mathcal{B})}(t)$ and $\mathbf{s}_{\mathrm{L}_{\mathrm{S}}(\mathcal{B})}(t)$.

One has to observe that since the introduction of formal power series, a lot of generalizations were proposed in order to extend the range of problems they can help to solve. The most obvious ones are multivariate series allowing us to count objects not only with respect to their sizes but also with respect to various other statistics. Another one consists in considering noncommutative series on words [Eil74, SS78, BR10], or even, pushing the generalization one step further, on elements of a monoid [Sak09]. Besides, as another generalization, series on trees have been considered [BR82, Boz01]. Series on (noncolored) operads increase the list of these generalizations. Chapoton is the first to have considered such series on operads [Cha02,Cha08,Cha09]. Several authors have contributed to this field by considering slight variations in the definitions of these series. Among these, one can cite van der Laan [vdL04], Frabetti [Fra08], and Loday and Nikolov [LN13]. Our notion of series on colored operads developed in this work is a natural generalization of series on operads.

This paper is organized as follows. Section 1 is devoted to set our notations and definitions about operads and colored operads, and to introduce the construction $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ producing a colored operad from a noncolored one $\mathcal{O}$ and a set $\mathscr{C}$ of colors. Section 2 contains the main definition on which this work this work is based on: bud generating systems. We establish some of properties of these. Next, we introduce formal power series on colored operads in Section 3, define several products on these, and explain how these series can be used to obtained enumerative results from bud generating systems. This article ends by Section 4 which contains a collection of examples for most of the notions introduced by this work. We have taken the freedom to put all the examples in this section. For this reason, the reader is encouraged to consult this section whilst he reads the first ones, by following the references we shall give.

Acknowledgements. The author would like to thank the anonymous referees for their very valuable suggestions, improving the presentation of this paper.

General notations and conventions. We denote by $\delta_{x,y}$ the Kronecker delta function (that is, for any elements x and y of a same set, $\delta_{x,y} = 1$ if $x = y$ and $\delta_{x,y} = 0$ otherwise). 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 $\#S$. For any finite multiset $S := \lbag s_1, \dots, s_n \rbag$ of nonnegative integers, we denote by $\sum S$ the sum

$$\sum S := s_1 + \cdots + s_n \tag{0.0.1}$$

of its elements and by $S!$ the multinomial coefficient

$$S! := \binom{\sum S}{s_1, \dots, s_n}. \tag{0.0.2}$$

For any set A, A^* denotes the set of all finite sequences, called words, of elements of A. We denote by A^+ the subset of A^* consisting in nonempty words. For any $n \geqslant 0$, A^n is the set of all words on A of length n. If u is a word, its letters are indexed from left to right from 1 to its length $|u|$. For any $i \in [|u|]$, u_i is the letter of u at position i. If a is a letter and n is a nonnegative integer, a^n denotes the word consisting in n occurrences of a. Notice that a^0 is the empty word ϵ.

In our graphical representations of trees, the uppermost nodes are always roots. Moreover, internal nodes are represented by circles ○, leaves by squares ▫, and edges by segments |. To distinguish trees and syntax trees, we shall draw the latter without circles for internal nodes and without squares for leaves (only the labels of the nodes are depicted).

In graphical representations of multigraphs, labels of edges denote their multiplicities. All unlabeled edges have 1 as multiplicity.

1. Colored operads and bud operads

The aim of this section is to set our notations about operads, colored operads, and colored syntax trees. We also establish some properties of treelike expressions in colored operads and present a construction producing colored operads from operads.

1.1. Colored operads.

Let us recall here the definitions of colored graded collections and colored operads.

1.1.1. *Colored graded collections.* Let $\mathscr{C}$ be a finite set, called *set of colors*. A *$\mathscr{C}$-colored graded collection* is a graded set

$$C := \bigsqcup_{n \geqslant 1} C(n) \tag{1.1.1}$$

together with two maps $\mathrm{out} : C \to \mathscr{C}$ and $\mathrm{in} : C(n) \to \mathscr{C}^n, n \geqslant 1$, respectively sending any $x \in C(n)$ to its *output color* $\mathrm{out}(x)$ and to its *word of input colors* $\mathrm{in}(x)$. The *i-th input color* of x is the i-th letter of $\mathrm{in}(x)$, denoted by $\mathrm{in}_i(x)$. For any $n \geqslant 1$ and $x \in C(n)$, the *arity* $|x|$ of x is n. We say that C is *locally finite* if for all $n \geqslant 1$, the $C(n)$ are finite sets. A *monochrome graded collection* is a $\mathscr{C}$-colored graded collection where $\mathscr{C}$ is a singleton. If C_1 and C_2 are two $\mathscr{C}$-colored graded collections, a map $\phi : C_1 \to C_2$ is a *$\mathscr{C}$-colored graded collection morphism* if it preserves arities. Besides, C_2 is a *$\mathscr{C}$-colored graded subcollection* of C_1 if for all $n \geqslant 1$, $C_2(n) \subseteq C_1(n)$, and C_1 and C_2 have the same maps out and in.

1.1.2. *Hilbert series.* In all this work, we consider that $\mathscr{C}$ has cardinal k and that the colors of $\mathscr{C}$ are arbitrarily indexed so that $\mathscr{C} = \{c_1, \dots, c_k\}$. Let $\mathbb{X}_{\mathscr{C}} := \{\mathrm{x}_{c_1}, \dots, \mathrm{x}_{c_k}\}$ and $\mathbb{Y}_{\mathscr{C}} := \{\mathrm{y}_{c_1}, \dots, \mathrm{y}_{c_k}\}$ be two alphabets of mutually commutative parameters and $\mathbb{N}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$ be the set of commutative multivariate series on $\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}$ with nonnegative integer coefficients. As usual, if $\mathbf{s}$ is a series of $\mathbb{N}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$, $\langle m, \mathbf{s} \rangle$ denotes the coefficient of the monomial m in $\mathbf{s}$.

For any $\mathscr{C}$-colored graded collection C, the *Hilbert series* $\mathbf{h}_C$ of C is the series of $\mathbb{N}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$ defined by

$$\mathbf{h}_C := \sum_{x \in C} \left(\mathrm{x}_{\mathrm{out}(x)} \prod_{i \in [|x|]} \mathrm{y}_{\mathrm{in}_i(x)} \right). \tag{1.1.2}$$

The coefficient of $\mathrm{x}_a \mathrm{y}_{c_1}^{\alpha_1} \dots \mathrm{y}_{c_k}^{\alpha_k}$ in $\mathbf{h}_C$ thus counts the elements of C having a as output color and α_j inputs of color c_j for any $j \in [k]$. Note that (1.1.2) is defined only if there are only finitely many such elements for any $a \in \mathscr{C}$ and any $\alpha_j \geqslant 0$, $j \in [k]$. This is the case when C is locally finite.

Besides, the *generating series* of C is the series $\mathbf{s}_C$ of $\mathbb{N}[[t]]$ defined as the specialization of $\mathbf{h}_C$ at $\mathrm{x}_a := 1$ and $\mathrm{y}_a := t$ for all $a \in \mathscr{C}$. Therefore, for any $n \geqslant 1$ the coefficient $\langle t^n, \mathbf{s}_C \rangle$ counts the elements of arity n in C.

1.1.3. *Colored operads.* A *nonsymmetric colored set-operad* on $\mathscr{C}$, or a *$\mathscr{C}$-colored operad* for short, is a colored graded collection $\mathcal{C}$ together with partially defined maps

$$\circ_i : \mathcal{C}(n) \times \mathcal{C}(m) \to \mathcal{C}(n + m - 1), \qquad 1 \leqslant i \leqslant n, \;\; 1 \leqslant m, \tag{1.1.3}$$

called *partial compositions*, and a subset $\{\mathbb{1}_a : a \in \mathscr{C}\}$ of $\mathcal{C}(1)$ such that any $\mathbb{1}_a$, $a \in \mathscr{C}$, is called *unit of color a* and satisfies $\mathrm{out}(\mathbb{1}_a) = \mathrm{in}(\mathbb{1}_a) = a$. This data has to satisfy, for any $x, y, z \in \mathcal{C}$, the following constraints. First, for any $i \in [|x|]$, $x \circ_i y$ is defined if and only if $\mathrm{out}(y) = \mathrm{in}_i(x)$. Moreover, the relations

$$(x \circ_i y) \circ_{i+j-1} z = x \circ_i \left(y \circ_j z\right), \qquad 1 \leqslant i \leqslant |x|, 1 \leqslant j \leqslant |y|, \tag{1.1.4a}$$

$$(x \circ_i y) \circ_{j+|y|-1} z = \left(x \circ_j z\right) \circ_i y, \qquad 1 \leqslant i < j \leqslant |x|, \tag{1.1.4b}$$

$$\mathbb{1}_a \circ_1 x = x = x \circ_i \mathbb{1}_b, \qquad 1 \leqslant i \leqslant |x|, \;\; a, b \in \mathscr{C}, \tag{1.1.4c}$$

have to hold when they are well-defined.

The *complete composition map* of $\mathcal{C}$ is the partially defined map

$$\circ : \mathcal{C}(n) \times \mathcal{C}(m_1) \times \cdots \times \mathcal{C}(m_n) \to \mathcal{C}(m_1 + \cdots + m_n), \tag{1.1.5}$$

defined from the partial composition maps in the following way. For any $x \in \mathcal{C}(n)$ and $y_1, \dots, y_n \in \mathcal{C}$ such that $\mathrm{out}(y_i) = \mathrm{in}_i(y)$ for all $i \in [n]$, we set

$$x \circ [y_1, \dots, y_n] := (\dots((x \circ_n y_n) \circ_{n-1} y_{n-1}) \dots) \circ_1 y_1. \tag{1.1.6}$$

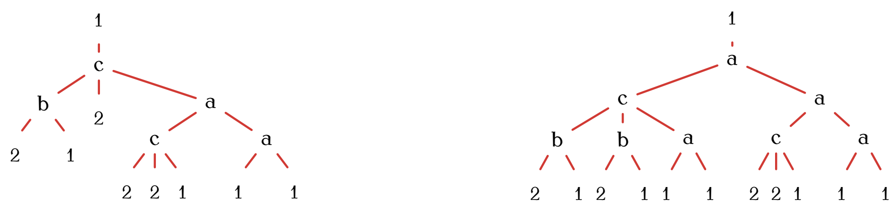
Let $\mathcal{C}_1$ and $\mathcal{C}_2$ are two $\mathscr{C}$-colored operads. A $\mathscr{C}$-colored graded collection morphism $\phi : \mathcal{C}_1 \to \mathcal{C}_2$ is a *$\mathscr{C}$-colored operad morphism* if it sends any unit of color $a \in \mathscr{C}$ of $\mathcal{C}_1$ to the unit of color a of $\mathcal{C}_2$, if it commutes with partial composition maps and, if for any $x, y \in \mathcal{C}_1$ and $i \in [|x|]$, if $x \circ_i y$ is defined in $\mathcal{C}_1$, then $\phi(x) \circ_i \phi(y)$ is defined in $\mathcal{C}_2$. Besides, $\mathcal{C}_2$ is a *colored suboperad* of $\mathcal{C}_1$ if $\mathcal{C}_2$ is a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}_1$ and $\mathcal{C}_1$ and $\mathcal{C}_2$ have the same colored units and the same partial composition maps. If G is a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}$, we denote by $\mathcal{C}^G$ the *$\mathscr{C}$-colored operad generated* by G, that is the smallest $\mathscr{C}$-colored suboperad of $\mathcal{C}$ containing G. When the $\mathscr{C}$-colored operad generated by G is $\mathcal{C}$ itself, G is a *generating $\mathscr{C}$-colored graded collection* of $\mathcal{C}$. Moreover, when G is minimal with respect to inclusion among the $\mathscr{C}$-colored graded subcollections of $\mathcal{C}$ satisfying this property, G is a *minimal generating $\mathscr{C}$-colored graded collection of $\mathcal{C}$*. We say that $\mathcal{C}$ is *locally finite* if, as a colored graded collection, $\mathcal{C}$ is locally finite.

A *monochrome operad* (or an *operad* for short) $\mathcal{O}$ is a $\mathscr{C}$-colored operad with a monochrome graded collection as underlying set. In this case, $\mathscr{C}$ is a singleton $\{c_1\}$ and, since for all $x \in \mathcal{O}(n)$, we necessarily have $\mathrm{out}(x) = c_1$ and $\mathrm{in}(x) = c_1^n$, for all $x, y \in \mathcal{O}$ and $i \in [|x|]$, all partial compositions $x \circ_i y$ are defined. In this case, $\mathscr{C}$ and its single element c_1 do not play any role. For this reason, in the future definitions of monochrome operads, we shall not define their set of colors $\mathscr{C}$.

1.2. **Free colored operads.**

Free colored operads and more particularly colored syntax trees play an important role in this work. We recall here the definitions of these two notions and establish some of their properties.

1.2.1. *Colored syntax trees.* Unless otherwise specified, we use in the sequel the standard terminology (*i.e.*, *node*, *edge*, *root*, *parent*, *child*, *path*, *etc.*) about planar rooted trees [Knu97]. Let $\mathscr{C}$ be a set of colors and C be a $\mathscr{C}$-colored graded collection. A *$\mathscr{C}$-colored C-syntax tree* is a planar rooted tree $\mathfrak{t}$ such that, for any $n \geqslant 1$, any internal node of $\mathfrak{t}$ having n children is labeled by an element of arity n of C and, for any internal nodes u and v of $\mathfrak{t}$ such that v is the i-th child of u, $\mathrm{out}(y) = \mathrm{in}_i(x)$ where x (resp. y) is the label of u (resp. v). In our graphical representations of a $\mathscr{C}$-colored C-syntax tree $\mathfrak{t}$, we write the colors of the leaves of $\mathfrak{t}$ below them and the color of the edge exiting the root of $\mathfrak{t}$ above it (see Figure 1).



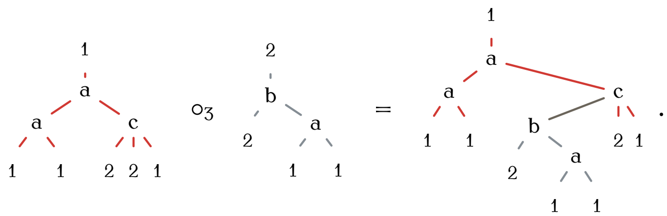
(A) The degree of this $\mathscr{C}$-colored C-syntax tree is 5, its arity is 8, and its height is 3

(B) A perfect $\mathscr{C}$-colored C-syntax tree. The degree of this colored syntax tree is 8, its arity is 11, and its height is 3.

FIGURE 1. Two $\mathscr{C}$-colored C-syntax trees, where $\mathscr{C}$ is the set of colors $\{1, 2\}$ and C is the $\mathscr{C}$-graded colored collection defined by $C := C(2) \sqcup C(3)$ with $C(2) := \{\mathrm{a}, \mathrm{b}\}$, $C(3) := \{\mathrm{c}\}$, $\mathrm{out}(\mathrm{a}) := 1$, $\mathrm{out}(\mathrm{b}) := 2$, $\mathrm{out}(\mathrm{c}) := 1$, $\mathrm{in}(\mathrm{a}) := 11$, $\mathrm{in}(\mathrm{b}) := 21$, and $\mathrm{in}(\mathrm{c}) := 221$.

Let $\mathfrak{t}$ be a $\mathscr{C}$-colored C-syntax tree. The *arity* of an internal node v of $\mathfrak{t}$ is its number $|v|$ of children and its *label* is the element of C labeling it and denoted by $\mathfrak{t}(v)$. The *degree* $\deg(\mathfrak{t})$ (resp. *arity* $|\mathfrak{t}|$) of $\mathfrak{t}$ is its number of internal nodes (resp. leaves). We say that $\mathfrak{t}$ is a *corolla* if $\deg(\mathfrak{t}) = 1$. The *height* of $\mathfrak{t}$ is the length $\mathrm{ht}(\mathfrak{t})$ of a longest path connecting the root of $\mathfrak{t}$ to one of its leaves. For instance, the height of a colored syntax tree of degree 0 is 0 and the one of a corolla is 1. The set of all internal nodes of $\mathfrak{t}$ is denoted by $\mathcal{N}(\mathfrak{t})$. For any $v \in \mathcal{N}(\mathfrak{t})$, $\mathfrak{t}_v$ is the subtree of $\mathfrak{t}$ rooted at the node v. We say that $\mathfrak{t}$ is *perfect* if all paths connecting the root of $\mathfrak{t}$ to its leaves have the same length. Finally, $\mathfrak{t}$ is a *monochrome C-syntax tree* if C is a monochrome graded collection.

1.2.2. *Free colored operads.* The *free $\mathscr{C}$-colored operad over C* is the operad $\mathbf{Free}(C)$ wherein for any $n \geqslant 1$, $\mathbf{Free}(C)(n)$ is the set of all $\mathscr{C}$-colored C-syntax trees of arity n. For any $\mathfrak{t} \in \mathbf{Free}(C)$, $\mathrm{out}(\mathfrak{t})$ is the output color of the label of the root of $\mathfrak{t}$ and $\mathrm{in}(\mathfrak{t})$ is the word obtained by reading, from left to right, the input colors of the leaves of $\mathfrak{t}$. For any $\mathfrak{s}, \mathfrak{t} \in \mathbf{Free}(C)$, the partial composition $\mathfrak{s} \circ_i \mathfrak{t}$, defined if and only if the output color of $\mathfrak{t}$ is the input color of the i-th leaf of $\mathfrak{s}$, is the tree obtained by grafting the root of $\mathfrak{t}$ to the i-th leaf of $\mathfrak{s}$. For instance, with the $\mathscr{C}$-colored graded collection C defined in Figure 1, one has in $\mathbf{Free}(C)$,



(1.2.1)

1.2.3. *Treelike expressions and finitely factorizing sets.* For any $\mathscr{C}$-colored operad $\mathcal{C}$, the *evaluation map* of $\mathcal{C}$ is the map $\mathrm{ev}_{\mathcal{C}} : \mathbf{Free}(\mathcal{C}) \to \mathcal{C}$, defined as the unique surjective morphism of colored operads satisfying $\mathrm{ev}_{\mathcal{C}}(\mathfrak{t}) = x$ where $\mathfrak{t}$ is a tree of degree 1 having its root labeled by x. If S is a colored graded subcollection of $\mathcal{C}$, an *S-treelike expression* of $x \in \mathcal{C}$ is a tree $\mathfrak{t}$ of $\mathbf{Free}(\mathcal{C})$ such that $\mathrm{ev}_{\mathcal{C}}(\mathfrak{t}) = x$ and all internal nodes of $\mathfrak{t}$ are labeled on S.

Besides, when S is such that any $x \in \mathcal{C}$ admits finitely many S-treelike expressions, we say that S *finitely factorizes* $\mathcal{C}$. This notion is important in the sequel and is used as sufficient condition for the well-definition of some formal power series on colored operads.

Lemma 1.2.1. *Let $\mathcal{C}$ be a locally finite $\mathscr{C}$-colored operad and S be a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}$ such that $S(1)$ finitely factorizes $\mathcal{C}$. Then, S finitely factorizes $\mathcal{C}$.*

Proof. Since $\mathcal{C}$ is locally finite and $S(1)$ finitely factorizes $\mathcal{C}$, there is a nonnegative integer k such that k is the degree of a $\mathscr{C}$-colored $S(1)$-syntax tree with a maximal number of internal nodes. Let x be an element of $\mathcal{C}$ of arity n admitting an S-treelike expression $\mathfrak{t}$. Observe first that $\mathfrak{t}$ has at most $n-1$ non-unary internal nodes and at most $2n-1$ edges. Moreover, by the pigeonhole principle, if $\mathfrak{t}$ would have more than $(2n-1)k$ unary internal nodes, there would be a chain made of more than k unary internal nodes in $\mathfrak{t}$. This cannot happen since, by hypothesis, it is not possible to form any $\mathscr{C}$-colored $S(1)$-syntax tree with more than k nodes. Therefore, we have shown that all S-treelike expressions of x are of degrees at most $n-1+(2n-1)k$. Moreover, since $\mathcal{C}$ is locally finite and S is a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}$, all $S(m)$ are finite for all $m \geqslant 1$. Therefore, there are finitely many S-treelike expressions of x. $\square$

Assume now that $\mathcal{C}$ is locally finite and that S is a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}$ such that $S(1)$ finitely factorizes $\mathcal{C}$. For any element x of $\mathcal{C}^S$, the colored suboperad of $\mathcal{C}$ generated by S, the *S-degree* of x is defined by

$$\deg_S(x) := \max\left\{\deg(\mathfrak{t}) : \mathfrak{t} \in \mathbf{Free}(S) \text{ and } \mathrm{ev}_{\mathcal{C}}(\mathfrak{t}) = x\right\}. \tag{1.2.2}$$

Thanks to the fact that, by hypothesis, x admits at least one S-treelike expression and, by Lemma 1.2.1, the fact that x admits finitely many S-treelike expressions, $\deg_S(x)$ is well-defined.

1.2.4. *Left expressions and hook-length formula.* Let S be a $\mathscr{C}$-colored graded subcollection of $\mathcal{O}$ and $x \in \mathcal{O}$. An *S-left expression* of x is an expression for x in $\mathcal{O}$ of the form

$$x = \left(\dots\left(\left(\mathbb{1}_{\mathrm{out}(x)} \circ_1 s_1\right) \circ_{i_1} s_2\right) \circ_{i_2} \dots\right) \circ_{i_{\ell-1}} s_\ell \tag{1.2.3}$$

where $s_1, \dots, s_\ell \in S$ and $i_1, \dots, i_{\ell-1} \in \mathbb{N}$. Besides, if $\mathfrak{t}$ is an S-treelike expression of x such that $\mathcal{N}(\mathfrak{t}) = \{e_1, \dots, e_\ell\}$, a sequence $(e_1, \dots, e_\ell)$ is a *linear extension* of $\mathfrak{t}$ if the sequence is a linear extension of the poset induced by $\mathfrak{t}$ seen as an Hasse diagram where the root of $\mathfrak{t}$ is the smallest element.

Lemma 1.2.2. *Let $\mathcal{O}$ be a locally finite $\mathscr{C}$-colored operad and S be a $\mathscr{C}$-colored graded subcollection of $\mathcal{O}$. Then, for any $x \in \mathcal{O}$, the set of all S-left expressions of x is in one-to-one correspondence with the set of all pairs $(\mathfrak{t}, e)$ where $\mathfrak{t}$ is an S-treelike expression of x and e is a linear extension of $\mathfrak{t}$.*

Proof. Let ϕ_x be the map sending any S-left expression of the form (1.2.3) of $x \in \mathcal{O}$ to the pair $(\mathfrak{t}, e)$ where $\mathfrak{t}$ is the colored syntax tree of $\mathbf{Free}(S)$ obtained by interpreting (1.2.3) in $\mathbf{Free}(S)$, *i.e.*, by replacing any s_j, $j \in [\ell]$, in (1.2.3) by a corolla $\mathfrak{s}_j$ of $\mathbf{Free}(S)$ labeled by s_j, and where e is the sequence $(e_1, \dots, e_\ell)$ of the internal nodes of $\mathfrak{t}$, where any e_j, $j \in [\ell]$, is the node of $\mathfrak{t}$ coming from $\mathfrak{s}_j$. We then have

$$\mathfrak{t} = \left(\dots\left(\left(\mathbb{1}_{\mathrm{out}(x)} \circ_1 \mathfrak{s}_1\right) \circ_{i_1} \mathfrak{s}_2\right) \circ_{i_2} \dots\right) \circ_{i_{\ell-1}} \mathfrak{s}_\ell \tag{1.2.4}$$

and by construction, $\mathfrak{t}$ is an S-treelike expression of x. Moreover, immediately from the definition of the partial composition in free $\mathscr{C}$-colored operads, $(e_1, \dots, e_\ell)$ is a linear extension of $\mathfrak{t}$. Therefore, we have shown that ϕ_x sends any S-left expression of x to a pair $(\mathfrak{t}, e)$ where $\mathfrak{t}$ is an S-treelike expression of x and e is a linear extension of $\mathfrak{t}$.

Let $\mathfrak{t}$ be an S-treelike expression of $x \in \mathcal{O}$ and e be a linear extension $(e_1, \dots, e_\ell)$ of $\mathfrak{t}$. It follows by induction on the degree ℓ of $\mathfrak{t}$ that $\mathfrak{t}$ can be expressed by an expression of the form (1.2.4) where any e_j, $j \in [\ell]$, is the node of $\mathfrak{t}$ coming from $\mathfrak{s}_j$. Now, the interpretation of (1.2.4) in $\mathcal{O}$, *i.e.*, by replacing any corolla $\mathfrak{s}_j$, $j \in [\ell]$, in (1.2.4) by its label s_j, is an S-left expression of the form (1.2.3) for x. Since (1.2.3) is the only antecedent of $(\mathfrak{t}, e)$ by ϕ_x, it follows that ϕ_x, with domain the set of all S-left expressions of x and with codomain the set of all pairs $(\mathfrak{t}, e)$ where $\mathfrak{t}$ is an S-treelike expression of x and e is a linear extension of $\mathfrak{t}$, is a bijection. □

A famous result of Knuth [Knu98], known as the *hook-length formula for trees*, stated here in our setting, says that given a $\mathscr{C}$-colored syntax tree $\mathfrak{t}$, the number of linear extensions of $\mathfrak{t}$ is

$$\frac{\deg(\mathfrak{t})!}{\prod\limits_{v \in \mathcal{N}(\mathfrak{t})} \deg\left(\mathfrak{t}_v\right)}. \tag{1.2.5}$$

When $S(1)$ finitely factorizes $\mathcal{O}$, by Lemma 1.2.1, the number of S-treelike expressions for any $x \in \mathcal{O}$ is finite. Hence, in this case, we deduce from Lemma 1.2.2 and (1.2.5) that the number of S-left expressions of x is

$$\sum_{\substack{\mathfrak{t} \in \mathbf{Free}(S) \\ \mathrm{ev}_{\mathcal{O}}(\mathfrak{t}) = x}} \frac{\deg(\mathfrak{t})!}{\prod\limits_{v \in \mathcal{N}(\mathfrak{t})} \deg\left(\mathfrak{t}_v\right)}. \tag{1.2.6}$$

1.3. **Bud operads.** Let us now present a simple construction producing colored operads from operads.

1.3.1. *From monochrome operads to colored operads.* If $\mathcal{O}$ is a monochrome operad and $\mathscr{C}$ is a finite set of colors, we denote by $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ the $\mathscr{C}$-colored graded collection defined by

$$\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})(n) := \mathscr{C} \times \mathcal{O}(n) \times \mathscr{C}^n, \qquad n \geqslant 1, \tag{1.3.1}$$

and for all $(a, x, u) \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, $\mathrm{out}((a, x, u)) := a$ and $\mathrm{in}((a, x, u)) := u$. We endow $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ with the partially defined partial composition $\circ_i$ satisfying, for all triples (a, x, u) and (b, y, v) of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, and $i \in [|x|]$ such that $\mathrm{out}((b, y, v)) = \mathrm{in}_i((a, x, u))$,

$$(a, x, u) \circ_i (b, y, v) := (a, x \circ_i y, u \hookleftarrow_i v)\,, \tag{1.3.2}$$

where $u \hookleftarrow_i v$ is the word obtained by replacing the i-th letter of u by v. Besides, if $\mathcal{O}_1$ and $\mathcal{O}_2$ are two operads and $\phi : \mathcal{O}_1 \to \mathcal{O}_2$ is an operad morphism, we denote by $\mathbf{Bud}_{\mathscr{C}}(\phi)$ the map

$$\mathbf{Bud}_{\mathscr{C}}(\phi) : \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}_1) \to \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}_2) \tag{1.3.3}$$

defined by

$$\mathbf{Bud}_{\mathscr{C}}(\phi)((a, x, u)) := (a, \phi(x), u)\,. \tag{1.3.4}$$

Proposition 1.3.1. *For any set of colors $\mathscr{C}$, the construction $(\mathcal{O}, \phi) \mapsto (\mathbf{Bud}_{\mathscr{C}}(\mathcal{O}), \mathbf{Bud}_{\mathscr{C}}(\phi))$ is a functor from the category of monochrome operads to the category of $\mathscr{C}$-colored operads.*

We omit the proof of Proposition 1.3.1 since it is very straightforward. This result shows that $\mathbf{Bud}_{\mathscr{C}}$ is a functorial construction producing colored operads from monochrome ones. We call $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ the *$\mathscr{C}$-bud operad* of $\mathcal{O}$[a]. When $\mathscr{C}$ is a singleton, $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ is by definition a monochrome operad isomorphic to $\mathcal{O}$. For this reason, in this case, we identify $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ with $\mathcal{O}$.

As a side observation, remark that in general, the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ of a free operad $\mathcal{O}$ is not a free $\mathscr{C}$-colored operad. Indeed, consider for instance the bud operad $\mathbf{Bud}_{\{1,2\}}(\mathcal{O})$, where $\mathcal{O} := \mathbf{Free}(C)$ and C is the monochrome graded collection defined by $C := C(1) := \{\mathtt{a}\}$. Then, a minimal generating set of $\mathbf{Bud}_{\{1,2\}}(\mathcal{O})$ is

$$\left\{ \left(1, \mathtt{a}, 1\right), \left(1, \mathtt{a}, 2\right), \left(2, \mathtt{a}, 1\right), \left(2, \mathtt{a}, 2\right) \right\}. \tag{1.3.5}$$

These elements are subjected to the nontrivial relations

$$\left(d, \mathtt{a}, 1\right) \circ_1 \left(1, \mathtt{a}, e\right) = \left(d, \begin{smallmatrix}\mathtt{a}\\ \mathtt{a}\end{smallmatrix}, e\right) = \left(d, \mathtt{a}, 2\right) \circ_1 \left(2, \mathtt{a}, e\right), \tag{1.3.6}$$

where $d, e \in \{1, 2\}$, implying that $\mathbf{Bud}_{\{1,2\}}(\mathcal{O})$ is not free.

[a]See examples of monochrome operads and their bud operads in Section 4.1.

1.3.2. *The associative operad.* The *associative operad* **As** is the monochrome operad defined by $\mathbf{As}(n) := \{\star_n\}$, $n \geqslant 1$, and wherein partial composition maps are defined by

$$\star_n \circ_i \star_m := \star_{n+m-1}, \qquad 1 \leqslant i \leqslant n, \ 1 \leqslant m. \tag{1.3.7}$$

For any set of colors $\mathscr{C}$, the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathbf{As})$ is the set of all triples

$$(a, \star_n, u_1 \dots u_n) \tag{1.3.8}$$

where $a \in \mathscr{C}$ and $u_1, \dots, u_n \in \mathscr{C}$. For $\mathscr{C} := \{1, 2, 3\}$, one has for instance the partial composition

$$(2, \star_4, 3112) \circ_2 (1, \star_3, 233) = (2, \star_6, 323312)\,. \tag{1.3.9}$$

The associative operad and its bud operads will play an important role in the sequel. For this reason, to gain readability, we shall simply denote by (a, u) any element $\left(a, \star_{|u|}, u\right)$ of $\mathbf{Bud}_{\mathscr{C}}(\mathbf{As})$ without any loss of information.

1.3.3. *Pruning map.* Here, we use the fact that any monochrome operad $\mathcal{O}$ can be seen as a $\mathscr{C}$-colored operad where all output and input colors of its elements are equal to c_1, where c_1 is the first color of $\mathscr{C}$ (see Section 1.1.3). Let

$$\mathfrak{pru} : \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \to \mathcal{O} \tag{1.3.10}$$

be the morphism of $\mathscr{C}$-colored operads defined, for any $(a, x, u) \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by

$$\mathfrak{pru}((a, x, u)) := x. \tag{1.3.11}$$

We call $\mathfrak{pru}$ the *pruning map* on $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$.

2. Bud generating systems and combinatorial generation

In this section, we introduce bud generating systems. A bud generating system relies on an operad $\mathcal{O}$, a set of colors $\mathscr{C}$, and the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. The principal interest of these objects is that they allow us to specify sets of objects of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. We shall also establish some first properties of bud generating systems by showing that they can emulate context-free grammars, regular tree grammars, and synchronous grammars.

2.1. **Bud generating systems.** We introduce here the main definitions and the main tools about bud generating systems.

2.1.1. *Bud generating systems.* A *bud generating system* is a tuple $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ where $\mathcal{O}$ is an operad called *ground operad*, $\mathscr{C}$ is a finite set of colors, $\mathfrak{R}$ is a finite $\mathscr{C}$-colored graded subcollection of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ called *set of rules*, I is a subset of $\mathscr{C}$ called *set of initial colors*, and T is a subset of $\mathscr{C}$ called *set of terminal colors*.

A *monochrome bud generating system* is a bud generating system whose set $\mathscr{C}$ of colors contains a single color, and whose sets of initial and terminal colors are equal to $\mathscr{C}$. In this case, as explained in Section 1.3.1, $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ and $\mathcal{O}$ are identified. These particular generating systems are thus simply denoted by pairs $(\mathcal{O}, \mathfrak{R})$.

Let us explain how bud generating systems specify, in two different ways, two $\mathscr{C}$-colored graded subcollections of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. In what follows, $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ is a bud generating system.

2.1.2. *Generation.* We say that $x_2 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ is *derivable in one step* from $x_1 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ if there is a rule $r \in \mathfrak{R}$ and an integer i such that such that $x_2 = x_1 \circ_i r$. We denote this property by $x_1 \to x_2$. When $x_1, x_2 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ are such that $x_1 = x_2$ or there are $y_1, \dots, y_{\ell-1} \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, $\ell \geqslant 1$, satisfying

$$x_1 \to y_1 \to \cdots \to y_{\ell-1} \to x_2, \tag{2.1.1}$$

we say that x_2 is *derivable* from x_1. Moreover, $\mathcal{B}$ *generates* $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ if there is a color a of I such that x is derivable from $\mathbb{1}_a$ and all colors of $\mathrm{in}(x)$ are in T. The *language* $\mathrm{L}(\mathcal{B})$ of $\mathcal{B}$ is the set of all the elements of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ generated by $\mathcal{B}$.

The *derivation graph* of $\mathcal{B}$ is the oriented multigraph $\mathrm{G}(\mathcal{B})$ with the set of elements derivable from $\mathbb{1}_a$, $a \in I$, as set of vertices. In $\mathrm{G}(\mathcal{B})$, for any $x_1, x_2 \in \mathrm{L}(\mathcal{B})$ such that $x_1 \to x_2$, there are ℓ edges from x_1 to x_2, where ℓ is the number of pairs $(i, r) \in \mathbb{N} \times \mathfrak{R}$ such that $x_2 = x_1 \circ_i r$[b].

2.1.3. *Synchronous generation.* We say that $x_2 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ is *synchronously derivable in one step* from $x_1 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ if there are rules $r_1, \dots, r_{|x_1|}$ of $\mathfrak{R}$ such that $x_2 = x_1 \circ \left[r_1, \dots, r_{|x_1|}\right]$. We denote this property by $x_1 \rightsquigarrow x_2$. When $x_1, x_2 \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ are such that $x_1 = x_2$ or there are $y_1, \dots, y_{\ell-1} \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, $\ell \geqslant 1$, satisfying

$$x_1 \rightsquigarrow y_1 \rightsquigarrow \cdots \rightsquigarrow y_{\ell-1} \rightsquigarrow x_2, \tag{2.1.2}$$

we say that x_2 is *synchronously derivable* from x_1. Moreover, $\mathcal{B}$ *synchronously generates* $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ if there is a color a of I such that x is synchronously derivable from $\mathbb{1}_a$ and all colors of $\mathrm{in}(x)$ are in T. The *synchronous language* $\mathrm{L_S}(\mathcal{B})$ of $\mathcal{B}$ is the set of all the elements of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ synchronously generated by $\mathcal{B}$.

The *synchronous derivation graph* of $\mathcal{B}$ is the oriented multigraph $\mathrm{G_S}(\mathcal{B})$ with the set of elements synchronously derivable from $\mathbb{1}_a$, $a \in I$, as set of vertices. In $\mathrm{G_S}(\mathcal{B})$, for any $x_1, x_2 \in \mathrm{L_S}(\mathcal{B})$ such that $x_1 \rightsquigarrow x_2$, there are ℓ edges from x_1 to x_2, where ℓ is the number of tuples $\left(r_1, \dots, r_{|x_1|}\right) \in \mathfrak{R}^{|x_1|}$ such that $x_2 = x_1 \circ \left[r_1, \dots, r_{|x_1|}\right]$[c].

2.2. **First properties.** We state now two properties about the languages and the synchronous languages of bud generating systems.

Lemma 2.2.1. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, x belongs to $\mathrm{L}(\mathcal{B})$ if and only if x admits an $\mathfrak{R}$-treelike expression with output color in I and all input colors in T.*

[b]See examples of bud generating systems and derivation graphs in Sections 4.3.1, 4.3.2, and 4.3.3.

[c]See examples of bud generating systems and synchronous derivation graphs in Sections 4.3.4 and 4.3.5.

Proof. Assume that x belongs to $\mathrm{L}(\mathcal{B})$. Then, by definition of the derivation relation $\to$, x admits an $\mathfrak{R}$-left expression. Lemma 1.2.2 implies in particular that x admits an $\mathfrak{R}$-treelike expression $\mathfrak{t}$. Moreover, since $\mathfrak{t}$ is a treelike expression for x, $\mathfrak{t}$ has the same output and input colors as those of x. Hence, because x belongs to $\mathrm{L}(\mathcal{B})$, its output color is in I and all its input colors are in T. Thus, $\mathfrak{t}$ satisfies the required properties.

Conversely, assume that x is an element of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ admitting an $\mathfrak{R}$-treelike expression $\mathfrak{t}$ with output color in I and all input colors in T. Lemma 1.2.2 implies in particular that x admits an $\mathfrak{R}$-left expression. Hence, by definition of the derivation relation $\to$, x is derivable from $\mathbb{1}_{\mathrm{out}(x)}$ and all its input colors are in T. Therefore, x belongs to $\mathrm{L}(\mathcal{B})$. □

Lemma 2.2.2. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, x belongs to $\mathrm{L}_{\mathrm{S}}(\mathcal{B})$ if and only if x admits an $\mathfrak{R}$-treelike expression with output color in I and all input colors in T and which is a perfect tree.*

Proof. The proof of the statement of the lemma is very similar to the one of Lemma 2.2.1. The only difference lies on the fact that the definition of synchronous languages uses the complete composition map $\circ$ instead of partial composition maps $\circ_i$, intervening in the definition of languages. Hence, in this context, $\mathfrak{R}$-treelike expressions are perfect trees. □

Proposition 2.2.3. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system. Then, the language of $\mathcal{B}$ satisfies*

$$\mathrm{L}(\mathcal{B}) = \left\{x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}} : \mathrm{out}(x) \in I \text{ and } \mathrm{in}(x) \in T^+\right\}, \tag{2.2.1}$$

where $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$ is the colored suboperad of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ generated by $\mathfrak{R}$.

Proof. By definition of suboperads generated by a set, as a $\mathscr{C}$-colored graded collection, $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$ consists in all the elements obtained by evaluating in $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ all $\mathscr{C}$-colored $\mathfrak{R}$-syntax trees. Therefore, the statement of the proposition is a consequence of Lemma 2.2.1. □

Proposition 2.2.4. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system. Then, the synchronous language of $\mathcal{B}$ is a subset of the language of $\mathcal{B}$. Moreover, when $\mathfrak{R}$ contains all the colored units of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, these two languages are equal.*

Proof. By Lemma 2.2.1 the language of $\mathcal{B}$ is the set of the elements obtained by evaluating in $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ all $\mathscr{C}$-colored $\mathfrak{R}$-syntax trees satisfying some conditions for their output and input colors. Lemma 2.2.2 says that the synchronous language of $\mathcal{B}$ is the set of the elements obtained by evaluating in $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ some $\mathscr{C}$-colored $\mathfrak{R}$-syntax trees satisfying at least the previous conditions. Hence, this implies the statement of the proposition.

The second part of the proposition follows from the fact that, if $x_1 \to x_2$ for two elements x_1 and x_2 of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, there is by definition $r \in \mathfrak{R}$ and an integer i such that $x_2 = x_1 \circ_i r$. Then, one has

$$x_2 = x_1 \circ \left[\mathbb{1}_{\mathrm{in}_1(x_1)}, \dots, \mathbb{1}_{\mathrm{in}_{i-1}(x_1)}, r, \mathbb{1}_{\mathrm{in}_{i+1}(x_1)}, \dots, \mathbb{1}_{\mathrm{in}_{|x_1|}(x_1)}\right]. \tag{2.2.2}$$

Since by hypothesis, all the colored units of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ are in $\mathfrak{R}$, this implies $x_1 \rightsquigarrow x_2$. Hence, as binary relations, $\to$ and $\rightsquigarrow$ are equal, establishing the second part of the statement of the proposition. □

2.3. **Links with other generating systems.** Context-free grammars, regular tree grammars, and synchronous grammars are already existing generating systems describing sets of words for the first, and sets of trees for the last two. We show here that any of these grammars can be emulated by bud generating systems.

2.3.1. *Context-free grammars.* A *context-free grammar* [Har78, HMU06] is a tuple $\mathcal{G} := (V, T, P, s)$ where V is a finite alphabet of *variables*, T is a finite alphabet of *terminal symbols*, P is a finite subset of $V \times (V \sqcup T)^*$ called *set of productions*, and s is a variable of V called *start symbol*. If x_1 and x_2 are two words of $(V \sqcup T)^*$, x_2 is *derivable in one step* from x_1 if x_1 is of the form $x_1 = uav$ and x_2 is of the form $x_2 = uwv$ where $u, v \in (V \sqcup T)^*$ and (a, w) is a production of P. This property is denoted by $x_1 \to x_2$, so that $\to$ is a binary relation on $(V \sqcup T)^*$. The reflexive and transitive closure of $\to$ is the *derivation relation*. A word $x \in T^*$ is *generated* by $\mathcal{G}$ if x is derivable from the word s. The *language* of $\mathcal{G}$ is the set of all words generated by $\mathcal{G}$. We say that $\mathcal{G}$ is *proper* if, for any $(a, w) \in P$, w is not the empty word.

If $\mathcal{G} := (V, T, P, s)$ is a proper context-free grammar, we denote by $\mathrm{CFG}(\mathcal{G})$ the bud generating system

$$\mathrm{CFG}(\mathcal{G}) := (\mathbf{As}, V \sqcup T, \mathfrak{R}, \{s\}, T) \tag{2.3.1}$$

wherein $\mathfrak{R}$ is the set of rules

$$\mathfrak{R} := \{(a, u) \in \mathbf{Bud}_{V \sqcup T}(\mathbf{As}) : (a, u) \in P\}. \tag{2.3.2}$$

Proposition 2.3.1. *Let $\mathcal{G}$ be a proper context-free grammar. Then, the restriction of the map* in*, sending any $(a, u) \in \mathbf{Bud}_{V \sqcup T}(\mathbf{As})$ to u, on the domain $\mathrm{L}(\mathrm{CFG}(\mathcal{G}))$ is a bijection between $\mathrm{L}(\mathrm{CFG}(\mathcal{G}))$ and the language of $\mathcal{G}$.*

Proof. Let us denote by V the set of variables, by T the set of terminal symbols, by P the set of productions, and by s the start symbol of $\mathcal{G}$.

Let $(a, x) \in \mathbf{Bud}_{V \sqcup T}(\mathbf{As})$, $\ell \geqslant 1$, and $y_1, \dots, y_{\ell-1} \in (V \sqcup T)^*$. Then, by definition of CFG, there are in $\mathrm{CFG}(\mathcal{G})$ the derivations

$$\mathbb{1}_s \to (s, y_1) \to \dots \to (s, y_{\ell-1}) \to (a, x) \tag{2.3.3}$$

if and only if $a = s$ and there are in $\mathcal{G}$ the derivations

$$s \to y_1 \to \dots \to y_{\ell-1} \to x. \tag{2.3.4}$$

Then, (a, x) belongs to $\mathrm{L}(\mathrm{CFG}(\mathcal{G}))$ if and only if $a = s$ and x belongs to the language of $\mathcal{G}$. The fact that $\mathrm{in}((s, x)) = x$ completes the proof. □

2.3.2. *Regular tree grammars.* Let $V := V(0)$ be a finite graded alphabet of *variables* and $T := \bigsqcup_{n \geqslant 0} T(n)$ be a finite graded alphabet of *terminal symbols*. For any $n \geqslant 0$ and $a \in T(n)$, the arity $|a|$ of a is n. The tuple (V, T) is called a *signature*.

A *(V, T)-tree* is an element of $\mathbf{Bud}_{V \sqcup T(0)}(\mathbf{Free}(T \setminus T(0)))$, where $T \setminus T(0)$ is seen as a monochrome graded collection. In other words, a (V, T)-tree is a planar rooted tree $\mathfrak{t}$ such that, for any $n \geqslant 1$, any internal node of $\mathfrak{t}$ having n children is labeled by an element of arity n of T, and the output and all leaves of $\mathfrak{t}$ are labeled on $V \sqcup T(0)$.

A *regular tree grammar* [GS84, CDG+07] is a tuple $\mathcal{G} := (V, T, P, s)$ where (V, T) is a signature, P is a set of pairs of the form $(v, \mathfrak{s})$ called *productions* where $v \in V$ and $\mathfrak{s}$ is a (V, T)-tree, and s is a variable of V called *start symbol*. If $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are two (V, T)-trees, $\mathfrak{t}_2$ is *derivable in one step* from $\mathfrak{t}_1$ if $\mathfrak{t}_1$ has a leaf y labeled by a and the tree obtained by replacing y by the root of $\mathfrak{s}$ in $\mathfrak{t}_1$ is $\mathfrak{t}_2$, provided that $(a, \mathfrak{s})$ is a production of P. This property is denoted by $\mathfrak{t}_1 \to \mathfrak{t}_2$, so that $\to$ is a binary relation on the set of all (V, T)-trees. The reflexive and transitive closure of $\to$ is the derivation relation. A (V, T)-tree $\mathfrak{t}$ is *generated* by $\mathcal{G}$ if $\mathfrak{t}$ is derivable from the tree $\mathbb{1}_s$ consisting in one leaf labeled by s and all leaves of $\mathfrak{t}$ are labeled on $T(0)$. The *language* of $\mathcal{G}$ is the set of all (V, T)-trees generated by $\mathcal{G}$.

If $\mathcal{G} := (V, T, P, s)$ is a regular tree grammar, we denote by $\mathrm{RTG}(\mathcal{G})$ the bud generating system

$$\mathrm{RTG}(\mathcal{G}) := (\mathbf{Free}(T \setminus T(0)), V \sqcup T(0), \mathfrak{R}, \{s\}, T(0)) \tag{2.3.5}$$

wherein $\mathfrak{R}$ is the set of rules

$$\mathfrak{R} := \left\{(a, \mathfrak{t}, u) \in \mathbf{Bud}_{V \sqcup T(0)}(\mathbf{Free}(T \setminus T(0))) : (a, \mathfrak{t}_{a,u}) \in P\right\}, \tag{2.3.6}$$

where, for any $\mathfrak{t} \in \mathbf{Free}(T \setminus T(0))$, $a \in V \sqcup T(0)$, and $u \in (V \sqcup T(0))^{|\mathfrak{t}|}$, $\mathfrak{t}_{a,u}$ is the (V, T)-tree obtained by labeling the output of $\mathfrak{t}$ by a and by labeling from left to right the leaves of $\mathfrak{t}$ by the letters of u.

Proposition 2.3.2. *Let $\mathcal{G}$ be a regular tree grammar. Then, the map $\phi : \mathrm{L}(\mathrm{RTG}(\mathcal{G})) \to L$ defined by $\phi((a, \mathfrak{t}, u)) := \mathfrak{t}_{a,u}$ is a bijection between the language of $\mathrm{RTG}(\mathcal{G})$ and the language L of $\mathcal{G}$.*

Proof. Let us denote by (V, T) the underlying signature and by s the start symbol of $\mathcal{G}$.

Let $(a, \mathfrak{t}, u) \in \mathbf{Bud}_{V \sqcup T(0)}(\mathbf{Free}(T \setminus T(0)))$, $\ell \geqslant 1$, and $\mathfrak{s}^{(1)}, \dots, \mathfrak{s}^{(\ell-1)} \in \mathbf{Free}(T \setminus T(0))$, and $v_1, \dots, v_{\ell-1} \in (V \sqcup T(0))^+$. Then, by definition of RTG, there are in $\mathrm{RTG}(\mathcal{G})$ the derivations

$$\mathbb{1}_s \to \left(s, \mathfrak{s}^{(1)}, v_1\right) \to \cdots \to \left(s, \mathfrak{s}^{(\ell-1)}, v_{\ell-1}\right) \to (a, \mathfrak{t}, u) \tag{2.3.7}$$

if and only if $a = s$ and there are in $\mathcal{G}$ the derivations

$$\mathbb{1}_s \to {\mathfrak{s}^{(1)}}_{s,v_1} \to \cdots \to {\mathfrak{s}^{(\ell-1)}}_{s,v_{\ell-1}} \to \mathfrak{t}_{a,u}. \tag{2.3.8}$$

Then, $(a, \mathfrak{t}, u)$ belongs to $\mathrm{L}(\mathrm{RTG}(\mathcal{G}))$ if and only if $a = s$ and $\mathfrak{t}_{a,u}$ belongs to the language of $\mathcal{G}$. The fact that $\phi((a, \mathfrak{t}, u)) = \mathfrak{t}_{a,u}$ completes the proof. □

2.3.3. *Synchronous grammars.* In this section, we shall denote by **Tree** the monochrome operad defined as the free operad generated by one operation a_n of arity n for all $n \geqslant 1$. The elements of this operad are planar rooted trees where internal nodes have an arbitrary arity. Observe by the way that **Tree** is not locally finite.

Let B be a finite alphabet. A *B-bud tree* is an element of $\mathbf{Bud}_B(\mathbf{Tree})$. In other words, a B-bud tree is a planar rooted tree $\mathfrak{t}$ such that the output and all leaves of $\mathfrak{t}$ are labeled on B. The leaves of a B-bud tree are indexed from 1 from left to right.

A *synchronous grammar* [Gir12] is a tuple $\mathcal{G} := (B, a, R)$ where B is a finite alphabet of *bud labels*, a is an element of B called *axiom*, and R is a finite set of pairs of the form $(b, \mathfrak{s})$ called *substitution rules* where $b \in B$ and $\mathfrak{s}$ is a B-bud tree. If $\mathfrak{t}_1$ and $\mathfrak{t}_2$ are two B-bud trees such that $\mathfrak{t}_1$ is of arity n, $\mathfrak{t}_2$ is *derivable in one step* from $\mathfrak{t}_1$ if there are substitution rules $(b_1, \mathfrak{s}_1), \dots, (b_n, \mathfrak{s}_n)$ of R such that for all $i \in [n]$, the i-th leaf of $\mathfrak{t}_1$ is labeled by b_i and $\mathfrak{t}_2$ is obtained by replacing the i-th leaf of $\mathfrak{t}_1$ by $\mathfrak{s}_i$ for all $i \in [n]$. This property is denoted by $\mathfrak{t}_1 \rightsquigarrow \mathfrak{t}_2$, so that $\rightsquigarrow$ is a binary relation on the set of all B-bud trees. The reflexive and transitive closure of $\rightsquigarrow$ is the derivation relation. A B-bud tree $\mathfrak{t}$ is *generated* by $\mathcal{G}$ if $\mathfrak{t}$ is derivable from the tree $\mathbb{1}_a$ consisting is one leaf labeled by a. The *language* of $\mathcal{G}$ is the set of all B-bud trees generated by $\mathcal{G}$.

If $\mathcal{G} := (B, a, R)$ is a synchronous grammar, we denote by $\mathrm{SG}(\mathcal{G})$ the bud generating system

$$\mathrm{SG}(\mathcal{G}) := (\mathbf{Tree}, B, \mathfrak{R}, \{a\}, B) \tag{2.3.9}$$

wherein $\mathfrak{R}$ is the set of rules

$$\mathfrak{R} := \{(b, \mathfrak{t}, u) \in \mathbf{Bud}_B(\mathbf{Tree}) : (b, \mathfrak{t}_{b,u}) \in R\}, \tag{2.3.10}$$

where, for any $\mathfrak{t} \in \mathbf{Bud}_B(\mathbf{Tree})$, $b \in B$, and $u \in B^+$, $\mathfrak{t}_{b,u}$ is the B-bud tree obtained by labeling the output of $\mathfrak{t}$ by b and by labeling from left to right the leaves of $\mathfrak{t}$ by the letters of u.

Proposition 2.3.3. *Let $\mathcal{G}$ be a synchronous grammar. Then, the map $\phi : \mathrm{L}_{\mathrm{S}}(\mathrm{SG}(\mathcal{G})) \to L$ defined by $\phi((b, \mathfrak{t}, u)) := \mathfrak{t}_{b,u}$ is a bijection between the synchronous language of $\mathrm{SG}(\mathcal{G})$ and the language L of $\mathcal{G}$.*

Proof. Let us denote by B the set of bud labels and by a the axiom of $\mathcal{G}$.

Let $(b, \mathfrak{t}, u) \in \mathbf{Bud}_B(\mathbf{Tree})$, $\ell \geqslant 1$, and $\mathfrak{s}^{(1)}, \dots, \mathfrak{s}^{(\ell-1)} \in \mathbf{Tree}$, and $v_1, \dots, v_{\ell-1} \in B^+$. Then, by definition of SG, there are in $\mathrm{SG}(\mathcal{G})$ the synchronous derivations

$$\mathbb{1}_a \rightsquigarrow \left(a, \mathfrak{s}^{(1)}, v_1\right) \rightsquigarrow \cdots \rightsquigarrow \left(a, \mathfrak{s}^{(\ell-1)}, v_{\ell-1}\right) \rightsquigarrow (b, \mathfrak{t}, u) \tag{2.3.11}$$

if and only if $b = a$ and there are in $\mathcal{G}$ the derivations

$$\mathbb{1}_a \rightsquigarrow \mathfrak{s}^{(1)}{}_{a,v_1} \rightsquigarrow \cdots \rightsquigarrow \mathfrak{s}^{(\ell-1)}{}_{a,v_{\ell-1}} \rightsquigarrow \mathfrak{t}_{b,u}. \tag{2.3.12}$$

Then, $(a, \mathfrak{t}, u)$ belongs to $\mathrm{L}_{\mathrm{S}}(\mathrm{SG}(\mathcal{G}))$ if and only if $b = a$ and $\mathfrak{t}_{b,u}$ belongs to the language of $\mathcal{G}$. The fact that $\phi((b, \mathfrak{t}, u)) = \mathfrak{t}_{b,u}$ completes the proof. $\square$

3. Series on colored operads and bud generating systems

We introduce in this section the concept of series on colored operads and define two binary products $\curvearrowleft$ and $\odot$ on these. We then explain how to use bud generating systems as tools to enumerate families of combinatorial objects. For this purpose, we will define and consider three series on colored operads extracted from bud generating systems. Each of these series brings information about their languages or the synchronous languages. One of a key issues is, given a bud generating system $\mathcal{B}$, to count arity by arity the elements of the language or the synchronous language of $\mathcal{B}$. In other terms, this amounts to compute the generating series $\mathbf{s}_{\mathrm{L}(\mathcal{B})}$ or $\mathbf{s}_{\mathrm{L}_{\mathrm{S}}(\mathcal{B})}$. As we shall see, these generating series can be computed from the series of colored operads extracted from $\mathcal{B}$.

From now, $\mathbb{K}$ is a field of characteristic zero. Moreover, in all this section, $\mathcal{C}$ is a $\mathscr{C}$-colored operad. Recall that the set $\mathscr{C}$ of colors is always considered on the form $\mathscr{C} = \{c_1, \dots, c_k\}$. Besides, $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ is a bud generating system.

3.1. **Series on colored operads.** We introduce here the main definitions about series on colored operads. We also explain how to encode usual noncommutative multivariate series and series on monoids by series on colored operads.

3.1.1. *Series on colored operads.* The linear span of the underlying set of $\mathcal{C}$ is denoted by $\mathbb{K}\langle\mathcal{C}\rangle$. Let $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ be the dual space of $\mathbb{K}\langle\mathcal{C}\rangle$. By definition, the elements of $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ are maps $\mathbf{f} : \mathcal{C} \to \mathbb{K}$, called *$\mathcal{C}$-formal power series* (or *$\mathcal{C}$-series* for short). Let $\mathbf{f} \in \mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$. The coefficient of any $x \in \mathcal{C}$ in $\mathbf{f}$ is denoted by $\langle x, \mathbf{f}\rangle$. The *support* of $\mathbf{f}$ is the set $\mathrm{Supp}(\mathbf{f}) := \{x \in \mathcal{C} : \langle x, \mathbf{f}\rangle \neq 0\}$. For any $\mathscr{C}$-colored graded subcollection S of $\mathcal{C}$, the *characteristic series* of S is the $\mathcal{C}$-series $\bar{S}$ defined for any $x \in \mathcal{C}$ by $\left\langle x, \bar{S}\right\rangle := 1$ if $x \in S$, and by $\left\langle x, \bar{S}\right\rangle := 0$ otherwise. The *series of colored units* of $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ is the series $\mathbf{u}$ defined as the characteristic of $\{\mathbb{1}_a : a \in \mathscr{C}\}$. This series will play a special role in the sequel. Since $\mathscr{C}$ is finite, $\mathbf{u}$ is a polynomial. By exploiting the vector space structure of $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$, any $\mathcal{C}$-series $\mathbf{f}$ expresses as

$$\mathbf{f} = \sum_{x \in \mathcal{C}} \langle x, \mathbf{f}\rangle\, x. \tag{3.1.1}$$

This notation using potentially infinite sums of elements of $\mathcal{C}$ accompanied with coefficients of $\mathbb{K}$ is common in the context of formal power series. In the sequel, we shall define and handle some $\mathcal{C}$-series using the notation (3.1.1).

Observe that $\mathcal{C}$-series are defined here on fields $\mathbb{K}$ instead on the much more general structures of semirings, as it is the case for series on monoids [Sak09]. We choose to tolerate this loss of generality because this considerably simplifies the theory. Furthermore, we shall use in the sequel $\mathcal{C}$-series as devices for combinatorial enumeration, so that it is sufficient to pick $\mathbb{K}$ as the field $\mathbb{Q}(q_0, q_1, q_2, \dots)$ of rational functions in an infinite number of commuting parameters with rational coefficients. The parameters $q_0, q_1, q_2, \dots$ intervene in the enumeration of colored graded subcollections of $\mathcal{C}$ with respect to several statistics[d].

3.1.2. *Colored operad morphisms and series.* If $\mathcal{C}_1$ and $\mathcal{C}_2$ are two $\mathscr{C}$-colored operads and $\phi : \mathcal{C}_1 \to \mathcal{C}_2$ is a morphism of colored operads, $\bar{\phi}$ is the map

$$\bar{\phi} : \mathbb{K}\langle\langle\mathcal{C}_1\rangle\rangle \to \mathbb{K}\langle\langle\mathcal{C}_2\rangle\rangle \tag{3.1.2}$$

defined, for any $\mathbf{f} \in \mathbb{K}\langle\langle\mathcal{C}_1\rangle\rangle$ and $y \in \mathcal{C}_2$, by

$$\left\langle y, \bar{\phi}(\mathbf{f})\right\rangle := \sum_{\substack{x \in \mathcal{C}_1 \\ \phi(x) = y}} \langle x, \mathbf{f}\rangle. \tag{3.1.3}$$

Observe first that $\bar{\phi}$ is a linear map. Moreover, notice that (3.1.3) could be undefined for arbitrary colored operads $\mathcal{C}_1$ and $\mathcal{C}_2$, and an arbitrary morphism of colored operads ϕ. However, when all fibers of ϕ are finite, for any $y \in \mathcal{C}_2$, the right member of (3.1.3) is well-defined since the sum has a finite number of terms.

[d]See examples of series on the bud operad of the operad **Motz** of Motzkin paths in Section 4.2.2.

3.1.3. *Pruned series and faithfulness.* Let $\mathbf{f}$ be a $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series. By a slight abuse of notation, we denote by

$$\mathfrak{pru} : \mathbb{K}\langle\langle \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})\rangle\rangle \to \mathbb{K}\langle\langle \mathcal{O}\rangle\rangle \tag{3.1.4}$$

the map $\bar{\mathfrak{pru}}$. Since $\mathscr{C}$ is finite, the series $\mathfrak{pru}(\mathbf{f})$ is well-defined and is called *pruned series* of $\mathbf{f}$. Intuitively, the series $\mathfrak{pru}(\mathbf{f})$ can be seen as a version of $\mathbf{f}$ wherein the colors of the elements of its support are forgotten[e]. Besides, $\mathbf{f}$ is said *faithful* if all coefficients of $\mathfrak{pru}(\mathbf{f})$ are equal to 0 or to 1.

We say that $\mathcal{B}$ is *faithful* (resp. *synchronously faithful*) if the characteristic series of $\mathrm{L}(\mathcal{B})$ (resp. $\mathrm{L}_{\mathrm{S}}(\mathcal{B})$) is faithful. Observe that all monochrome bud generating systems are faithful (resp. synchronously faithful). One of the reasons for requiring faithfulness (resp. synchronous faithfulness) for bud generating systems appears when $\mathcal{B}$ is utilized for specifying objects of $\mathcal{O}$ by pruning the objects of $\mathrm{L}(\mathcal{B})$ (resp. $\mathrm{L}_{\mathrm{S}}(\mathcal{B})$). In this case, if $\mathcal{B}$ is not faithful (resp. synchronously faithful), there would be several distinct elements (a, x, u) of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ generated (resp. synchronously generated) by $\mathcal{B}$ whose image by $\mathfrak{pru}$ is x. This could make very hard the enumeration of the pruned version of the language (resp. synchronous language) of $\mathcal{B}$.

3.1.4. *Series of colors.* Let

$$\mathfrak{col} : \mathcal{G} \to \mathbf{Bud}_{\mathscr{C}}(\mathbf{As}) \tag{3.1.5}$$

be the morphism of colored operads defined for any $x \in \mathcal{G}$ by

$$\mathfrak{col}(x) := (\mathrm{out}(x), \mathrm{in}(x)) . \tag{3.1.6}$$

By a slight abuse of notation, we denote by

$$\mathfrak{col} : \mathbb{K}\langle\langle \mathcal{G}\rangle\rangle \to \mathbb{K}\langle\langle \mathbf{Bud}_{\mathscr{C}}(\mathbf{As})\rangle\rangle \tag{3.1.7}$$

the map $\bar{\mathfrak{col}}$. If $\mathbf{f}$ is a $\mathcal{G}$-series, we call $\mathfrak{col}(\mathbf{f})$ the *series of colors* of $\mathbf{f}$. Intuitively, the series $\mathfrak{col}(\mathbf{f})$ can be seen as a version of $\mathbf{f}$ wherein only the colors of the elements of its support are taken into account[f].

3.1.5. *Series of color types.* The *$\mathscr{C}$-type* of a word $u \in \mathscr{C}^+$ is the word $\mathrm{type}(u)$ of $\mathbb{N}^k$ defined by

$$\mathrm{type}(u) := |u|_{c_1} \dots |u|_{c_k}, \tag{3.1.8}$$

where for any $a \in \mathscr{C}$, $|u|_a$ is the number of occurrences of a in u. By extension, we shall call *$\mathscr{C}$-type* any word of $\mathbb{N}^k$ with at least a nonzero letter and we denote by $\mathcal{T}_{\mathscr{C}}$ the set of all $\mathscr{C}$-types. The *degree* $\deg(\alpha)$ of $\alpha \in \mathcal{T}_{\mathscr{C}}$ is the sum of the letters of α. We denote by $\mathscr{C}^{\alpha}$ the word $c_1^{\alpha_1} \dots c_k^{\alpha_k}$.

Assume that $\mathbb{Z}_{\mathscr{C}} := \{\mathrm{z}_{c_1}, \dots, \mathrm{z}_{c_k}\}$ is any alphabet of commutative letters. For any type α, we denote by $\mathbb{Z}_{\mathscr{C}}^{\alpha}$ the monomial $\mathrm{z}_{c_1}^{\alpha_1} \dots \mathrm{z}_{c_k}^{\alpha_k}$ of $\mathbb{K}[\mathbb{Z}_{\mathscr{C}}]$. Moreover, for any two types α and β, the *sum* $\alpha \dot{+} \beta$ of α and β is the type satisfying $(\alpha \dot{+} \beta)_i := \alpha_i + \beta_i$ for all $i \in [k]$. Observe that with this notation, $\mathbb{Z}_{\mathscr{C}}^{\alpha} \mathbb{Z}_{\mathscr{C}}^{\beta} = \mathbb{Z}_{\mathscr{C}}^{\alpha \dot{+} \beta}$.

---

[e]See an example of pruned series in Section 4.2.2.

[f]See examples of series of colors in Section 4.2.1.

Consider now the map

$$\mathfrak{colt} : \mathbb{K} \langle\langle \mathcal{G} \rangle\rangle \to \mathbb{K} [[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]] , \tag{3.1.9}$$

defined for all $\alpha, \beta \in \mathcal{T}_{\mathscr{C}}$ by

$$\left\langle \mathbb{X}_{\mathscr{C}}^{\alpha} \mathbb{Y}_{\mathscr{C}}^{\beta}, \mathfrak{colt}(\mathbf{f}) \right\rangle := \sum_{\substack{(a,u) \in \mathbf{Bud}_{\mathscr{C}}(\mathbf{As}) \\ \mathrm{type}(a) = \alpha \\ \mathrm{type}(u) = \beta}} \langle (a, u), \mathfrak{col}(\mathbf{f}) \rangle . \tag{3.1.10}$$

By the definition of the map $\mathfrak{col}$,

$$\mathfrak{colt}(\mathbf{f}) = \sum_{x \in \mathcal{G}} \langle x, \mathbf{f} \rangle \ \mathbb{X}_{\mathscr{C}}^{\mathrm{type}(\mathrm{out}(x))} \ \mathbb{Y}_{\mathscr{C}}^{\mathrm{type}(\mathrm{in}(x))}. \tag{3.1.11}$$

Observe that for all $\alpha, \beta \in \mathcal{T}_{\mathscr{C}}$ such that $\deg(\alpha) \neq 1$, the coefficients of $\mathbb{X}_{\mathscr{C}}^{\alpha} \mathbb{Y}_{\mathscr{C}}^{\beta}$ in $\mathfrak{colt}(\mathbf{f})$ are zero. In intuitive terms, the series $\mathfrak{colt}(\mathbf{f})$, called *series of color types* of $\mathbf{f}$, can be seen as a version of $\mathfrak{col}(\mathbf{f})$ wherein only the output colors and the types of the input colors of the elements of its support are taken into account, the variables of $\mathbb{X}_{\mathscr{C}}$ encoding output colors and the variables of $\mathbb{Y}_{\mathscr{C}}$ encoding input colors[g]. In the sequel, we shall be concerned by the computation of the coefficients of $\mathfrak{colt}(\mathbf{f})$ for some $\mathcal{G}$-series $\mathbf{f}$.

3.1.6. *Elementary series of bud generating systems.* We assume here that $\mathcal{O}$ is a locally finite monochrome operad. We shall denote by $\mathbf{r}$ the characteristic series of $\mathfrak{R}$, by $\mathbf{i}$ the characteristic series of $\{\mathbb{1}_a : a \in I\}$, and by $\mathbf{t}$ the characteristic series of $\{\mathbb{1}_a : a \in T\}$. For all colors $a \in \mathscr{C}$ and types $\alpha \in \mathcal{T}_{\mathscr{C}}$, let

$$\chi_{a,\alpha} := \# \{ r \in \mathfrak{R} : (\mathrm{out}(r), \mathrm{type}(\mathrm{in}(r))) = (a, \alpha) \} . \tag{3.1.12}$$

For any $a \in \mathscr{C}$, let $\mathbf{g}_a (\mathsf{y}_{c_1}, \dots, \mathsf{y}_{c_k})$ be the series of $\mathbb{K} [[\mathbb{Y}_{\mathscr{C}}]]$ defined by

$$\mathbf{g}_a (\mathsf{y}_{c_1}, \dots, \mathsf{y}_{c_k}) := \sum_{\gamma \in \mathcal{T}_{\mathscr{C}}} \chi_{a,\gamma} \ \mathbb{Y}_{\mathscr{C}}^{\gamma} = \sum_{\substack{r \in \mathfrak{R} \\ \mathrm{out}(r) = a}} \mathbb{Y}_{\mathscr{C}}^{\mathrm{type}(\mathrm{in}(r))}. \tag{3.1.13}$$

Since $\mathfrak{R}$ is finite, this series is a polynomial[h].

In the sequel, we shall use maps $\phi : \mathscr{C} \times \mathcal{T}_{\mathscr{C}} \to \mathbb{N}$ such that $\phi(a, \gamma) \neq 0$ for a finite number of pairs $(a, \gamma) \in \mathscr{C} \times \mathcal{T}_{\mathscr{C}}$, to express in a concise manner some recurrence relations for the coefficients of series on colored operads. We shall consider the two following notations. If ϕ is such a map and $a \in \mathscr{C}$, we define $\phi^{(a)}$ as the natural number

$$\phi^{(a)} := \sum_{\substack{b \in \mathscr{C} \\ \gamma \in \mathcal{T}_{\mathscr{C}}}} \phi(b, \gamma) \gamma_a \tag{3.1.14}$$

and ϕ_a as the finite multiset

$$\phi_a := \lbag \phi(a, \gamma) : \gamma \in \mathcal{T}_{\mathscr{C}} \rbag . \tag{3.1.15}$$

[g]See an example of a series of color types in Section 4.2.1.

[h]See examples of these definitions in Sections 4.4.4, 4.4.5, and 4.4.6.

3.2. **Products on series.** Two binary products $\curvearrowleft$ and $\odot$ on the space of $\mathcal{C}$-series are presented. The product $\curvearrowleft$ is a generalizations to series and to colored operads of a known product on monochrome operads, and $\odot$ is a generalization to colored operads of a known product on series on monochrome operads.

3.2.1. *Pre-Lie product.* Given two $\mathcal{C}$-series $\mathbf{f}, \mathbf{g} \in \mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$, the *pre-Lie product* of $\mathbf{f}$ and $\mathbf{g}$ is the $\mathcal{C}$-series $\mathbf{f} \curvearrowleft \mathbf{g}$ defined, for any $x \in \mathcal{C}$, by

$$\langle x, \mathbf{f} \curvearrowleft \mathbf{g}\rangle := \sum_{\substack{y,z \in \mathcal{C} \\ i \in [|y|] \\ x = y \circ_i z}} \langle y, \mathbf{f}\rangle \langle z, \mathbf{g}\rangle . \tag{3.2.1}$$

Observe that $\mathbf{f} \curvearrowleft \mathbf{g}$ could be undefined for arbitrary $\mathcal{C}$-series $\mathbf{f}$ and $\mathbf{g}$ on an arbitrary colored operad $\mathcal{C}$. Besides, notice from (3.2.1) that $\curvearrowleft$ is bilinear and that $\mathbf{u}$ is a left unit of $\curvearrowleft$. However, since

$$\mathbf{f} \curvearrowleft \mathbf{u} = \sum_{x \in \mathcal{C}} |x| \langle x, \mathbf{f}\rangle x, \tag{3.2.2}$$

the $\mathcal{C}$-series $\mathbf{u}$ is not a right unit of $\curvearrowleft$. This product is also nonassociative in the general case since we have, for instance in $\mathbb{K}\langle\langle\mathbf{As}\rangle\rangle$,

$$(\star_2 \curvearrowleft \star_2) \curvearrowleft \star_2 = 6\star_4 \neq 4\star_4 = \star_2 \curvearrowleft (\star_2 \curvearrowleft \star_2) . \tag{3.2.3}$$

Recall that a *$\mathbb{K}$-pre-Lie algebra* [Vin63, Ger63] (see also [CL01, Man11]) is a $\mathbb{K}$-vector space V endowed with a bilinear product $\curvearrowleft$ satisfying, for all $x, y, z \in V$, the relation

$$(x \curvearrowleft y) \curvearrowleft z - x \curvearrowleft (y \curvearrowleft z) = (x \curvearrowleft z) \curvearrowleft y - x \curvearrowleft (z \curvearrowleft y). \tag{3.2.4}$$

In this case, we say that $\curvearrowleft$ is a *pre-Lie product*. Observe that any associative product satisfies (3.2.4), so that associative algebras are pre-Lie algebras.

Proposition 3.2.1. *For any locally finite colored operad $\mathcal{C}$, the space $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ endowed with the binary product $\curvearrowleft$ is a pre-Lie algebra.*

This product $\curvearrowleft$ is a generalization of a pre-Lie product defined in [Ger63] (see also [vdL04, Cha08]), endowing the $\mathbb{K}$-linear span of the underlying monochrome graded collection of a monochrome operad with a pre-Lie algebra structure. Proposition 3.2.1 is based on similar arguments as the ones contained in the previous references.

3.2.2. *Composition product.* Given two $\mathcal{C}$-series $\mathbf{f}, \mathbf{g} \in \mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$, the *composition product* of $\mathbf{f}$ and $\mathbf{g}$ is the $\mathcal{C}$-series $\mathbf{f} \odot \mathbf{g}$ defined, for any $x \in \mathcal{C}$, by

$$\langle x, \mathbf{f} \odot \mathbf{g}\rangle := \sum_{\substack{y, z_1, \dots, z_{|y|} \in \mathcal{C} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \langle y, \mathbf{f}\rangle \prod_{i \in [|y|]} \langle z_i, \mathbf{g}\rangle . \tag{3.2.5}$$

Observe that $\mathbf{f} \odot \mathbf{g}$ could be undefined for arbitrary $\mathcal{C}$-series $\mathbf{f}$ and $\mathbf{g}$ on an arbitrary colored operad $\mathcal{C}$. Besides, notice from (3.2.5) that $\odot$ is linear on the left and that the series $\mathbf{u}$ is the left and right unit of $\odot$. However, this product is not linear on the right since we have, for instance in $\mathbb{K}\langle\langle\mathbf{As}\rangle\rangle$,

$$\star_2 \odot (\star_2 + \star_3) = \star_4 + 2\,\star_5 + \star_6 \neq \star_4 + \star_6 = \star_2 \odot \star_2 + \star_2 \odot \star_3. \tag{3.2.6}$$

Proposition 3.2.2. *For any locally finite colored operad $\mathcal{C}$, the space $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ endowed with the binary product $\odot$ and the unit $\mathbf{u}$ is a monoid.*

This product $\odot$ is a generalization of the composition product of series on operads of [Cha02, Cha09] (see also [vdL04, Fra08, Cha08, LV12, LN13]). In the case where $\mathcal{C}$ is a monochrome operad concentrated in arity 1, $\odot$ coincides with the Cauchy product on series of monoids considered in [Sak09]. Proposition 3.2.2 is based on similar arguments as the ones contained in the previous references.

Lemma 3.2.3. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system and $\mathbf{f}$ be a $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series. Then, $\mathbf{i} \odot \mathbf{f} \odot \mathbf{t}$ is the $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series satisfying, for all $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$,*

$$\langle x, \mathbf{i} \odot \mathbf{f} \odot \mathbf{t} \rangle = \begin{cases} \langle x, \mathbf{f} \rangle & \text{if } \mathrm{out}(x) \in I \text{ and } \mathrm{in}(x) \in T^{+}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2.7}$$

Proof. By definition of the operation $\odot$, composing $\mathbf{f}$ with $\mathbf{i}$ to the left and with $\mathbf{t}$ to the right with respect to $\odot$ amounts to annihilate the coefficients of the terms of $\mathbf{f}$ that have an output color which is not in I or an input color which is not in T. This implies the statement of the lemma. □

3.3. **Series and languages.** We introduce the Kleene star operation of the pre-Lie product $\curvearrowleft$ in order to define the hook generating series of a bud generating system $\mathcal{B}$. We also study the inverse of the composition product $\odot$ in order the define the syntactic generating series of $\mathcal{B}$. We relate both of these series with the language of $\mathcal{B}$ and provide ways to compute its coefficients.

3.3.1. *Pre-Lie star product.* For any $\mathcal{C}$-series $\mathbf{f} \in \mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$ and any $\ell \geqslant 0$, let $\mathbf{f}^{\curvearrowleft \ell}$ be the $\mathcal{C}$-series recursively defined by

$$\mathbf{f}^{\curvearrowleft \ell} := \begin{cases} \mathbf{u} & \text{if } \ell = 0, \\ \mathbf{f}^{\curvearrowleft \ell - 1} \curvearrowleft \mathbf{f} & \text{otherwise.} \end{cases} \tag{3.3.1}$$

Immediately from this definition and the definition of the pre-Lie product $\curvearrowleft$, the coefficients of $\mathbf{f}^{\curvearrowleft \ell}$, $\ell \geqslant 0$, satisfy for any $x \in \mathcal{C}$,

$$\langle x, \mathbf{f}^{\curvearrowleft \ell} \rangle = \begin{cases} \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} & \text{if } \ell = 0, \\ \displaystyle\sum_{\substack{y, z \in \mathcal{C} \\ i \in [|y|] \\ x = y \circ_i z}} \langle y, \mathbf{f}^{\curvearrowleft \ell - 1} \rangle \langle z, \mathbf{f} \rangle & \text{otherwise.} \end{cases} \tag{3.3.2}$$

Lemma 3.3.1. *Let $\mathcal{C}$ be a locally finite $\mathscr{C}$-colored operad and $\mathbf{f}$ be a series of $\mathbb{K}\langle\langle\mathcal{C}\rangle\rangle$. Then, the coefficients of $\mathbf{f}^{\curvearrowleft \ell+1}$, $\ell \geqslant 0$, satisfy for any $x \in \mathcal{C}$,*

$$\langle x, \mathbf{f}^{\curvearrowleft \ell+1} \rangle = \sum_{\substack{y_1, \dots, y_{\ell+1} \in \mathcal{C} \\ i_1, \dots, i_\ell \in \mathbb{N} \\ x = (\dots(y_1 \circ_{i_1} y_2) \circ_{i_2} \dots) \circ_{i_\ell} y_{\ell+1}}} \prod_{j \in [\ell+1]} \langle y_j, \mathbf{f} \rangle. \tag{3.3.3}$$

Proof. By Proposition 3.2.1, since $\mathcal{C}$ is locally finite, $\mathbf{f}^{\curvearrowleft \ell+1}$ is a well-defined $\mathcal{C}$-series. The statement of the lemma follows by induction on ℓ and by using (3.3.2). □

The *$\curvearrowleft$-star* of $\mathbf{f}$ is the series

$$\mathbf{f}^{\curvearrowleft *} := \sum_{\ell \geqslant 0} \mathbf{f}^{\curvearrowleft \ell} = \mathbf{u} + \mathbf{f} + \mathbf{f} \curvearrowleft \mathbf{f} + (\mathbf{f} \curvearrowleft \mathbf{f}) \curvearrowleft \mathbf{f} + ((\mathbf{f} \curvearrowleft \mathbf{f}) \curvearrowleft \mathbf{f}) \curvearrowleft \mathbf{f} + \cdots . \tag{3.3.4}$$

Observe that $\mathbf{f}^{\curvearrowleft *}$ could be undefined for an arbitrary $\mathcal{C}$-series $\mathbf{f}$.

Proposition 3.3.2. *Let $\mathcal{C}$ be a locally finite $\mathscr{C}$-colored operad and $\mathbf{f}$ be a series of $\mathbb{K}\langle\langle \mathcal{C} \rangle\rangle$ such that* $\mathrm{Supp}(\mathbf{f})(1)$ *finitely factorizes $\mathcal{C}$. Then,*

(i) the series $\mathbf{f}^{\curvearrowleft *}$ is well-defined;

(ii) for any $x \in \mathcal{C}$, the coefficient of x in $\mathbf{f}^{\curvearrowleft *}$ satisfies

$$\langle x, \mathbf{f}^{\curvearrowleft *} \rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y, z \in \mathcal{C} \\ i \in [|y|] \\ x = y \circ_i z}} \langle y, \mathbf{f}^{\curvearrowleft *} \rangle \langle z, \mathbf{f} \rangle; \tag{3.3.5}$$

(iii) the equation

$$\mathbf{x} - \mathbf{x} \curvearrowleft \mathbf{f} = \mathbf{u} \tag{3.3.6}$$

admits $\mathbf{x} = \mathbf{f}^{\curvearrowleft *}$ as unique solution.

Proof. Let $x \in \mathcal{C}$ and let us show that the coefficient $\langle x, \mathbf{f}^{\curvearrowleft *} \rangle$ is well-defined. Since $\mathcal{C}$ is locally finite and $\mathrm{Supp}(\mathbf{f})(1)$ is finitely factorizes $\mathcal{C}$, by Lemma 1.2.1, there are finitely many $\mathrm{Supp}(\mathbf{f})$-treelike expressions for x. Thus, for all $\ell \geqslant \deg_{\mathrm{Supp}(\mathbf{f})}(x) + 1$ there is in particular no expression for x of the form $x = (\dots(y_1 \circ_{i_1} y_2) \circ_{i_2} \dots) \circ_{i_{\ell-1}} y_\ell$ where $y_1, \dots, y_\ell \in \mathrm{Supp}(\mathbf{f})$ and $i_1, \dots, i_{\ell-1} \in \mathbb{N}$. This implies, together with Lemma 3.3.1, that $\langle x, \mathbf{f}^{\curvearrowleft \ell} \rangle = 0$. Therefore, by virtue of this observation and by definition of the $\curvearrowleft$-star operation, the coefficient of x in $\mathbf{f}^{\curvearrowleft *}$ is

$$\langle x, \mathbf{f}^{\curvearrowleft *} \rangle = \sum_{\ell \geqslant 0} \langle x, \mathbf{f}^{\curvearrowleft \ell} \rangle = \sum_{0 \leqslant \ell \leqslant \deg_{\mathrm{Supp}(\mathbf{f})}(x)} \langle x, \mathbf{f}^{\curvearrowleft \ell} \rangle, \tag{3.3.7}$$

showing that $\langle x, \mathbf{f}^{\curvearrowleft *} \rangle$ is a sum of a finite number of terms, all well-defined by Lemma 3.3.1. Thus, $\mathbf{f}^{\curvearrowleft *}$ is well-defined, so that *(i)* holds.

Point *(ii)* follows straightforwardly from the definition of the $\curvearrowleft$-star operation and (3.3.2).

By (3.3.6), we have $\mathbf{x} = \mathbf{u} + \mathbf{x} \curvearrowleft \mathbf{f}$ so that the coefficients of $\mathbf{x}$ satisfy, for any $x \in \mathcal{C}$,

$$\langle x, \mathbf{x} \rangle = \langle x, \mathbf{u} \rangle + \langle x, \mathbf{x} \curvearrowleft \mathbf{f} \rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y, z \in \mathcal{C} \\ i \in [|y|] \\ x = y \circ_i z}} \langle y, \mathbf{x} \rangle \langle z, \mathbf{f} \rangle. \tag{3.3.8}$$

By *(ii)*, this implies $\mathbf{x} = \mathbf{f}^{\curvearrowleft *}$ and the uniqueness of this solution, so that *(iii)* is established. □

In particular, Point *(ii)* of Proposition 3.3.2 gives a way, given a $\mathcal{C}$-series $\mathbf{f}$ satisfying the stated constraints, to compute recursively the coefficients of its $\curvearrowleft$-star $\mathbf{f}^{\curvearrowleft *}$.

3.3.2. *Hook generating series.* The *hook generating series* of $\mathcal{B}$ the $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series $\mathrm{hook}(\mathcal{B})$ defined by

$$\mathrm{hook}(\mathcal{B}) := \mathbf{i} \odot \mathbf{r}^{\curvearrowleft *} \odot \mathbf{t}. \tag{3.3.9}$$

Observe that (3.3.9) could be undefined for an arbitrary set of rules $\mathfrak{R}$ of $\mathcal{B}$. Nevertheless, when $\mathbf{r}$ satisfies the conditions of Proposition 3.3.2, that is, when $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, $\mathrm{hook}(\mathcal{B})$ is well-defined.

The aim of the following is to provide an expression to compute the coefficients of $\mathrm{hook}(\mathcal{B})$.

Lemma 3.3.3. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$,*

$$\langle x, \mathbf{r}^{\curvearrowleft *} \rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ z \in \mathfrak{R} \\ i \in [|y|] \\ x = y \circ_i z}} \langle y, \mathbf{r}^{\curvearrowleft *} \rangle . \tag{3.3.10}$$

Proof. Since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.2, $\mathbf{r}^{\curvearrowleft *}$ is a well-defined series. Now, (3.3.10) is a consequence of Point *(ii)* of Proposition 3.3.2 together with the fact that all coefficients of $\mathbf{r}$ are equal to 0 or to 1. □

Proposition 3.3.4. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ such that $\mathrm{out}(x) \in I$, the coefficient $\langle x, \mathbf{r}^{\curvearrowleft *} \rangle$ is the number of multipaths from $\mathbb{1}_{\mathrm{out}(x)}$ to x in the derivation graph of $\mathcal{B}$.*

Proof. First, since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.2, $\mathbf{r}^{\curvearrowleft *}$ is a well-defined series. If $x = \mathbb{1}_a$ for an $a \in I$, since $\langle \mathbb{1}_a, \mathbf{r}^{\curvearrowleft *} \rangle = 1$, the statement of the proposition holds. Let us now assume that x is different from a colored unit and let us denote by λ_x the number of multipaths from $\mathbb{1}_{\mathrm{out}(x)}$ to x in the derivation graph $\mathrm{G}(\mathcal{B})$ of $\mathcal{B}$. By definition of $\mathrm{G}(\mathcal{B})$, by denoting by $\mu_{y,x}$ the number of edges from $y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ to x in $\mathrm{G}(\mathcal{B})$, we have

$$\lambda_x = \sum_{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})} \mu_{y,x} \lambda_y = \sum_{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})} \# \{(i, r) \in \mathbb{N} \times \mathfrak{R} : x = y \circ_i r\} \lambda_y = \sum_{\substack{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ i \in [|y|] \\ r \in \mathfrak{R} \\ x = y \circ_i r}} \lambda_y . \tag{3.3.11}$$

We observe that Relation (3.3.11) satisfied by the λ_x is the same as Relation (3.3.10) of Lemma 3.3.3 satisfied by the $\langle x, \mathbf{r}^{\curvearrowleft *} \rangle$. This implies the statement of the proposition. □

Theorem 3.3.5. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the hook generating series of $\mathcal{B}$ satisfies*

$$\mathrm{hook}(\mathcal{B}) = \sum_{\substack{\mathfrak{t} \in \mathbf{Free}(\mathfrak{R}) \\ \mathrm{out}(\mathfrak{t}) \in I \\ \mathrm{in}(\mathfrak{t}) \in T^+}} \frac{\deg(\mathfrak{t})!}{\prod\limits_{v \in \mathcal{N}(\mathfrak{t})} \deg(\mathfrak{t}_v)} \, \mathrm{ev}_{\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})}(\mathfrak{t}). \tag{3.3.12}$$

Proof. By definition of $\mathrm{L}(\mathcal{B})$ and $\mathrm{G}(\mathcal{B})$, any $x \in \mathrm{L}(\mathcal{B})$ can be reached from $\mathbb{1}_{\mathrm{out}(x)}$ by a multipath

$$\mathbb{1}_{\mathrm{out}(x)} \to y_1 \to y_2 \to \cdots \to y_{\ell-1} \to x \tag{3.3.13}$$

in $\mathrm{G}(\mathcal{B})$, where $y_1, \dots, y_{\ell-1}$ are elements of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ and $\mathbb{1}_{\mathrm{out}(x)} \in I$. Hence, by definition of $\to$, x admits an $\mathfrak{R}$-left expression

$$x = \left(\dots\left(\left(\mathbb{1}_{\mathrm{out}(x)} \circ_1 r_1\right) \circ_{i_1} r_2\right) \circ_{i_2} \dots\right) \circ_{i_{\ell-1}} r_\ell \tag{3.3.14}$$

where for any $j \in [\ell]$, $r_j \in \mathfrak{R}$, and for any $j \in [\ell - 1]$,

$$y_j = \left(\dots\left(\left(\mathbb{1}_{\mathrm{out}(x)} \circ_1 r_1\right) \circ_{i_1} r_2\right) \circ_{i_2} \dots\right) \circ_{i_{j-1}} r_j \tag{3.3.15}$$

and $i_j \in [|y_j|]$. This shows that the set of all multipaths from $\mathbb{1}_{\mathrm{out}(x)}$ to x in $\mathrm{G}(\mathcal{B})$ is in one-to-one correspondence with the set of all $\mathfrak{R}$-left expressions for x. Now, observe that since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.2, $\mathbf{r}^{\curvearrowleft *}$ is a well-defined series. If $x = \mathbb{1}_a$ for an $a \in I$, since $\langle \mathbb{1}_a, \mathbf{r}^{\curvearrowleft *} \rangle = 1$, the statement of the proposition holds. Let us now assume that x is different from a colored unit and let us denote by λ_x the number of multipaths from $\mathbb{1}_{\mathrm{out}(x)}$ to x in the derivation graph $\mathrm{G}(\mathcal{B})$ of $\mathcal{B}$. By d, $\mathbf{r}^{\curvearrowleft *}$ is a well-defined series. By Proposition 3.3.4, Lemmas 1.2.1 and 1.2.2, and (1.2.6), we obtain that

$$\langle x, \mathbf{r}^{\curvearrowleft *} \rangle = \sum_{\substack{\mathfrak{t} \in \mathbf{Free}(\mathfrak{R}) \\ \mathrm{ev}_{\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})}(\mathfrak{t}) = x}} \frac{\deg(\mathfrak{t})!}{\prod\limits_{v \in \mathcal{N}(\mathfrak{t})} \deg(\mathfrak{t}_v)}. \tag{3.3.16}$$

Finally, by Lemma 3.2.3, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ such that $\mathrm{out}(x) \in I$ and $\mathrm{in}(x) \in T^+$, we have $\langle x, \mathrm{hook}(\mathcal{B}) \rangle = \langle x, \mathbf{r}^{\curvearrowleft *} \rangle$. This shows that the right member of (3.3.12) is equal to $\mathrm{hook}(\mathcal{B})$. □

An alternative way to understand $\mathrm{hook}(\mathcal{B})$ thus offered by Theorem 3.3.5 consists is seeing the coefficient $\langle x, \mathrm{hook}(\mathcal{B}) \rangle$, $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, as the number of $\mathfrak{R}$-left expressions of x.

The following result establishes a link between the hook generating series of $\mathcal{B}$ and its language.

Proposition 3.3.6. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the support of the hook generating series of $\mathcal{B}$ is the language of $\mathcal{B}$.*

Proof. This is an immediate consequence of Theorem 3.3.5 and Lemma 2.2.1. □

Bud generating systems lead to the definition of analogues of the *hook-length statistic* [Knu98] for combinatorial objects possibly different than trees in the following way. Let $\mathcal{O}$ be a monochrome operad, G be a generating set of $\mathcal{O}$, and $\mathrm{HS}_{\mathcal{O},G} := (\mathcal{O}, G)$ be a monochrome bud generating system depending on $\mathcal{O}$ and G, called *hook bud generating system*. Since G is a generating set of $\mathcal{O}$, by Propositions 2.2.3 and 3.3.6, the support of $\mathrm{hook}\left(\mathrm{HS}_{\mathcal{O},G}\right)$ is equal to $\mathrm{L}\left(\mathrm{HS}_{\mathcal{O},G}\right)$. We define the *hook-length coefficient* of any element x of $\mathcal{O}$ as the coefficient $\langle x, \mathrm{hook}\left(\mathrm{HS}_{\mathcal{O},G}\right) \rangle$[i].

[i]See examples of definitions of a hook-length statistics for binary trees, words of $\mathbf{Dias}_\gamma$, and for Motzkin paths in Sections 4.4.1, 4.4.2, and 4.4.3.

3.3.3. *Invertible elements for the composition product.* Since by Proposition 3.2.2, $\odot$ is an associative product and $\mathbf{u}$ is its unit, the *$\odot$-inverse* of a $\mathcal{C}$-series $\mathbf{f}$ is defined as the unique $\mathcal{C}$-series $\mathbf{x}$ satisfying

$$\mathbf{f} \odot \mathbf{x} = \mathbf{u} = \mathbf{x} \odot \mathbf{f}. \tag{3.3.17}$$

This series $\mathbf{x}$ could be undefined for an arbitrary $\mathcal{C}$-series $\mathbf{f}$. The $\odot$-inverse of $\mathbf{f}$ is denoted by $\mathbf{f}^{\odot_{-1}}$ when it is well-defined.

Immediately from this definition and the definition of the composition product $\odot$, the coefficients of $\mathbf{f}^{\odot_{-1}}$ satisfy for any $x \in \mathcal{C}$,

$$\left\langle x, \mathbf{f}^{\odot_{-1}} \right\rangle = \frac{\delta_{x, \mathbb{1}_{\mathrm{out}(x)}}}{\left\langle \mathbb{1}_{\mathrm{out}(x)}, \mathbf{f} \right\rangle} - \frac{1}{\left\langle \mathbb{1}_{\mathrm{out}(x)}, \mathbf{f} \right\rangle} \sum_{\substack{y, z_1, \dots, z_{|y|} \in \mathcal{C} \\ y \neq \mathbb{1}_{\mathrm{out}(x)} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \langle y, \mathbf{f} \rangle \prod_{i \in [|y|]} \left\langle z_i, \mathbf{f}^{\odot_{-1}} \right\rangle . \tag{3.3.18}$$

Proposition 3.3.7. *Let $\mathcal{C}$ be a locally finite colored $\mathscr{C}$-operad and $\mathbf{f}$ be a series of $\mathbb{K}\langle\langle \mathcal{C} \rangle\rangle$ such that* $\mathrm{Supp}(\mathbf{f}) = \{\mathbb{1}_a : a \in \mathscr{C}\} \sqcup S$ *where S is a $\mathscr{C}$-colored graded subcollection of $\mathcal{C}$ such that $S(1)$ is finitely factorizes $\mathcal{C}$. Then,*

- *(i) the series $\mathbf{f}^{\odot_{-1}}$ is well-defined;*
- *(ii) for any $x \in \mathcal{C}$, the coefficient of x in $\mathbf{f}^{\odot_{-1}}$ satisfies*

$$\left\langle x, \mathbf{f}^{\odot_{-1}} \right\rangle = \frac{1}{\left\langle \mathbb{1}_{\mathrm{out}(x)}, \mathbf{f} \right\rangle} \sum_{\substack{\mathfrak{t} \in \mathbf{Free}(S) \\ \mathrm{ev}_{\mathcal{C}}(\mathfrak{t}) = x}} (-1)^{\deg(\mathfrak{t})} \prod_{v \in \mathcal{N}(\mathfrak{t})} \frac{\langle \mathfrak{t}(v), \mathbf{f} \rangle}{\prod_{j \in [|v|]} \left\langle \mathbb{1}_{\mathrm{in}_j(v)}, \mathbf{f} \right\rangle}. \tag{3.3.19}$$

Proof. Let us first assume that x does not belong to $\mathcal{C}^S$, the colored suboperad of $\mathcal{C}$ generated by S. Hence, since there is no $\mathfrak{t} \in \mathbf{Free}(S)$ such that $\mathrm{ev}_{\mathcal{C}}(\mathfrak{t}) = x$, the right member of (3.3.19) is equal to zero. Moreover, since x does not belong to $\mathcal{C}^S$, for any $y \in \mathcal{C}$ and $z_1, \dots, z_{|y|} \in \mathcal{C}$ such that $y \neq \mathbb{1}_{\mathrm{out}(x)}$ and $x = y \circ \left[z_1, \dots, z_{|y|}\right]$, we have necessarily $y \notin S$ or $z_i \notin \mathcal{C}^S$ for at least one $i \in [|y|]$. By (3.3.18), this implies that $\langle x, \mathbf{f}^{\odot_{-1}} \rangle = 0$. This hence shows that (3.3.19) holds when $x \notin \mathcal{C}^S$.

Let us assume that x belongs to $\mathcal{C}^S$. By Lemma 1.2.1, since $\mathcal{C}$ is locally finite and $S(1)$ finitely factorizes $\mathcal{C}$, the S-degree $\deg_S(x)$ of x is well-defined. To prove (3.3.19), we proceed by induction on $\deg_S(x)$ and by using (3.3.18). A straightforward computation shows that (3.3.19) holds so that *(ii)* checks out.

Finally, by Lemma 1.2.1, there is a finite number of S-treelike expressions of x. This shows that (3.3.19) is well-defined and then $\mathbf{f}^{\odot_{-1}}$ also is. Hence, *(i)* holds. □

Given a $\mathcal{C}$-series $\mathbf{f}$ such that $\mathbf{f}^{\odot_{-1}}$ is well-defined, Equation (3.3.18) (resp. Proposition 3.3.7) provides a recursive (resp. direct) way to compute the coefficients of $\mathbf{f}^{\odot_{-1}}$.

Besides, the set of all the $\mathcal{C}$-series satisfying the conditions of Proposition 3.3.7 forms a submonoid of $\mathbb{K}\langle\langle C \rangle\rangle$ for the composition product which is also a group. This group is a generalization of the groups constructed from operads of [Cha02, Cha09] (see also [vdL04, Fra08, Cha08, LV12, LN13]).

3.3.4. *Syntactic generating series.* The *syntactic generating series* of $\mathcal{B}$ the $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series $\mathrm{synt}(\mathcal{B})$ defined by

$$\mathrm{synt}(\mathcal{B}) := \mathbf{i} \odot (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \odot \mathbf{t}. \tag{3.3.20}$$

Observe that (3.3.20) could be undefined for an arbitrary set of rules $\mathfrak{R}$ of $\mathcal{B}$. Nevertheless, when $\mathbf{u} - \mathbf{r}$ satisfies the conditions of Proposition 3.3.7, $\mathrm{synt}(\mathcal{B})$ is well-defined. Remark that this condition is satisfied whenever $\mathcal{O}$ is locally finite and $\mathfrak{R}(1)$ factorizes finitely $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$.

The aim of this section is to provide an expression to compute the coefficients of $\mathrm{synt}(\mathcal{B})$.

Lemma 3.3.8. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$,*

$$\left\langle x, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \right\rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y \in \mathfrak{R} \\ z_1, \dots, z_{|y|} \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \prod_{i \in [|y|]} \left\langle z_i, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \right\rangle. \tag{3.3.21}$$

Proof. Since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.7, $(\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ is a well-defined series. Now, (3.3.21) is a consequence of Point *(ii)* of Proposition 3.3.7 and Equation (3.3.18) for the $\odot$-inverse, together with the fact that all coefficients of $\mathbf{r}$ are equal to 0 or to 1. □

Theorem 3.3.9. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the syntactic generating series of $\mathcal{B}$ satisfies*

$$\mathrm{synt}(\mathcal{B}) = \sum_{\substack{\mathfrak{t} \in \mathbf{Free}(\mathfrak{R}) \\ \mathrm{out}(\mathfrak{t}) \in I \\ \mathrm{in}(\mathfrak{t}) \in T^+}} \mathrm{ev}_{\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})}(\mathfrak{t}). \tag{3.3.22}$$

Proof. Let, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, λ_x be the number of $\mathfrak{R}$-treelike expressions for x. Since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Lemma 1.2.1, all λ_x are well-defined integers. Moreover, since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.7, $(\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ is a well-defined series. Let us show that $\langle x, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \rangle = \lambda_x$. First, when x does not belong to $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$, by Point *(ii)* of Proposition 3.3.7, $\langle x, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \rangle = 0$. Since, in this case $\lambda_x = 0$, the property holds. Let us now assume that x belongs to $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$. Again by Lemma 1.2.1, the $\mathfrak{R}$-degree of x is well-defined. Therefore, we proceed by induction on $\deg_{\mathfrak{R}}(x)$. By Lemma 3.3.8, when x is a colored unit $\mathbb{1}_a$, $a \in \mathscr{C}$, one has $\langle x, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \rangle = 1$. Since there is exactly one $\mathfrak{R}$-treelike expression for $\mathbb{1}_a$, namely the syntax tree consisting in one leaf of output and input color a, $\lambda_{\mathbb{1}_a} = 1$ so that the base case holds. Otherwise, again by Lemma 3.3.8, we have, by using induction hypothesis,

$$\left\langle x, (\mathbf{u} - \mathbf{r})^{\odot_{-1}} \right\rangle = \sum_{\substack{y \in \mathfrak{R} \\ z_1, \dots, z_{|y|} \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \prod_{i \in [|y|]} \lambda_{z_i} = \lambda_x. \tag{3.3.23}$$

Notice that one can apply the induction hypothesis to state (3.3.23) since one has $\deg_{\mathfrak{R}}(x) \geqslant 1 + \deg_{\mathfrak{R}}(z_i)$ for all $i \in [|y|]$.

Now, from (3.3.23) and by using Lemma 3.2.3, we obtain that for all $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ such that $\mathrm{out}(x) \in I$ and $\mathrm{in}(x) \in T^+$, $\langle x, \mathrm{synt}(\mathcal{B})\rangle = \lambda_x$. Denoting by $\mathbf{f}$ the series of the right member of (3.3.22), we have $\langle x, \mathbf{f}\rangle = \lambda_x$ if $\mathrm{out}(x) \in I$ and $\mathrm{in}(x) \in T^+$, and $\langle x, \mathbf{f}\rangle = 0$ otherwise. This shows that this expression is equal to $\mathrm{synt}(\mathcal{B})$. □

Theorem 3.3.9 explains the name of syntactic generating series for $\mathrm{synt}(\mathcal{B})$ because this series can be expressed following (3.3.22) as a sum of evaluations of syntax trees. An alternative way to see $\mathrm{synt}(\mathcal{B})$ is that for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, the coefficient $\langle x, \mathrm{synt}(\mathcal{B})\rangle$ is the number of $\mathfrak{R}$-treelike expressions for x.

The following result establishes a link between the syntactic generating series of $\mathcal{B}$ and its language.

Proposition 3.3.10. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the support of the syntactic generating series of $\mathcal{B}$ is the language of $\mathcal{B}$.*

Proof. This is an immediate consequence of Theorem 3.3.9 and Lemma 2.2.1. □

By Propositions 3.3.6 and 3.3.10, the series $\mathrm{hook}(\mathcal{B})$ and $\mathrm{synt}(\mathcal{B})$ have the same support. The main difference between these two series is that the coefficient of an $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ in $\mathrm{synt}(\mathcal{B})$ is the number of $\mathfrak{R}$-treelike expressions for x, while in $\mathrm{hook}(\mathcal{B})$ this coefficient is the number of ways to generate x in $\mathcal{B}$.

We say that $\mathcal{B}$ is *unambiguous* if all coefficients of $\mathrm{synt}(\mathcal{B})$ are equal to 0 or to 1. This property is important from a combinatorial and enumerative point of view. Indeed, when $\mathcal{B}$ is unambiguous, its syntactic generating series is the characteristic series of its language. As a consequence, by definition of the series of colors $\mathfrak{col}$ (see Section 3.1.4) and Proposition 3.3.10, the coefficient of $(a, u) \in \mathbf{Bud}_{\mathscr{C}}(\mathbf{As})$ in the series $\mathfrak{col}(\mathrm{synt}(\mathcal{B}))$ is the number of elements x of $\mathrm{L}(\mathcal{B})$ such that $(\mathrm{out}(x), \mathrm{in}(x)) = (a, u)$.

As a side remark, observe that Theorem 3.3.9 implies in particular that for any bud generating system of the form $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, \mathscr{C}, \mathscr{C})$, if $\mathrm{synt}(\mathcal{B})$ is unambiguous, then the colored suboperad of $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ generated by $\mathfrak{R}$ is free. The converse property does not hold.

Let us now describe the coefficients of $\mathfrak{colt}(\mathrm{synt}(\mathcal{B}))$, the series of color types of the syntactic series of $\mathcal{B}$, in the particular case when $\mathcal{B}$ is unambiguous. We shall give two descriptions: a first one involving a system of equations of series of $\mathbb{K}[[\mathbb{Y}_{\mathscr{C}}]]$, and a second one involving a recurrence relation on the coefficients of a series of $\mathbb{K}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$.

Lemma 3.3.11. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be an unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for all colors $a \in I$ and all types $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^{\alpha} \in T^+$, the coefficients $\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathrm{synt}(\mathcal{B}))\rangle$ count the number of elements x of $\mathrm{L}(\mathcal{B})$ such that $(\mathrm{out}(x), \mathrm{type}(\mathrm{in}(x))) = (a, \alpha)$.*

Proof. By Proposition 3.3.10 and since $\mathcal{B}$ is unambiguous, $\mathrm{synt}(\mathcal{B})$ is the characteristic series of $\mathrm{L}(\mathcal{B})$. The statement of the lemma follows immediately from the definition (3.1.10) of $\mathfrak{colt}$. □

Proposition 3.3.12. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be an unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. For all $a \in \mathscr{C}$, let $\mathbf{f}_a(y_{c_1}, \dots, y_{c_k})$ be the series of $\mathbb{K}[[\mathbb{Y}_{\mathscr{C}}]]$ satisfying*

$$\mathbf{f}_a(y_{c_1}, \dots, y_{c_k}) = y_a + \mathbf{g}_a(\mathbf{f}_{c_1}(y_{c_1}, \dots, y_{c_k}), \dots, \mathbf{f}_{c_k}(y_{c_1}, \dots, y_{c_k})). \tag{3.3.24}$$

Then, for any color $a \in I$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^\alpha \in T^+$, the coefficients $\langle x_a \mathbb{Y}^\alpha_{\mathscr{C}}, \mathfrak{colt}(\mathrm{synt}(\mathcal{B}))\rangle$ and $\langle \mathbb{Y}^\alpha_{\mathscr{C}}, \mathbf{f}_a\rangle$ are equal.

Proof. Let us set $\mathbf{h} := (\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ and, for all $a \in \mathscr{C}$, $\mathbf{h}_a := \mathbb{1}_a \odot \mathbf{h}$. Since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.7, $\mathbf{h}$ and $\mathbf{h}_a$ are well-defined series. Equation (3.3.17) implies that any $\mathbf{h}_a$, $a \in \mathscr{C}$, satisfies the relation

$$\mathbf{h}_a = \mathbb{1}_a + \mathbf{r}_a \odot \mathbf{h} \tag{3.3.25}$$

where $\mathbf{r}_a := \mathbb{1}_a \odot \mathbf{r}$. Observe that for any $a \in \mathscr{C}$, $\mathfrak{colt}(\mathbf{r}_a) = \mathbf{g}_a(y_{c_1}, \dots, y_{c_k})$. Moreover, from the definitions of $\mathfrak{colt}$ and the operation $\odot$, we obtain that $\mathfrak{colt}(\mathbf{r}_a \odot \mathbf{h})$ can be computed by a functional composition of the series $\mathbf{g}_a(y_{c_1}, \dots, y_{c_k})$ with $\mathbf{f}_{c_1}(y_{c_1}, \dots, y_{c_k})$, $\dots$, $\mathbf{f}_{c_k}(y_{c_1}, \dots, y_{c_k})$. Hence, Relation (3.3.25) leads to

$$\begin{aligned} \mathfrak{colt}(\mathbf{h}_a) &= \mathfrak{colt}(\mathbb{1}_a) + \mathfrak{colt}(\mathbf{r}_a \odot \mathbf{h}) \\ &= y_a + \mathbf{g}_a(\mathbf{f}_{c_1}(y_{c_1}, \dots, y_{c_k}), \dots, \mathbf{f}_{c_k}(y_{c_1}, \dots, y_{c_k})) \\ &= \mathbf{f}_a(y_{c_1}, \dots, y_{c_k}). \end{aligned} \tag{3.3.26}$$

Finally, Lemma 3.2.3 implies that, when $a \in I$ and $\mathscr{C}^\alpha \in T^+$, $\langle x_a \mathbb{Y}^\alpha_{\mathscr{C}}, \mathfrak{colt}(\mathrm{synt}(\mathcal{B}))\rangle$ and $\langle \mathbb{Y}^\alpha_{\mathscr{C}}, \mathbf{f}_a\rangle$ are equal. □

When $\mathcal{B}$ is a bud generating system satisfying the conditions of Proposition 3.3.12, the generating series of the language of $\mathcal{B}$ satisfies

$$\mathbf{s}_{\mathrm{L}(\mathcal{B})} = \sum_{a \in I} \mathbf{f}^T_a, \tag{3.3.27}$$

where $\mathbf{f}^T_a$ is the specialization of the series $\mathbf{f}_a(y_{c_1}, \dots, y_{c_k})$ at $y_b := t$ for all $b \in T$ and at $y_c := 0$ for all $c \in \mathscr{C} \setminus T$. Therefore, the resolution of the system of equations given by Proposition 3.3.12 provides a way to compute the coefficients of $\mathbf{s}_{\mathrm{L}(\mathcal{B})}$.

Theorem 3.3.13. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be an unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the generating series $\mathbf{s}_{\mathrm{L}(\mathcal{B})}$ of the language of $\mathcal{B}$ is algebraic.*

Proof. Proposition 3.3.12 shows that each series $\mathbf{f}_a$ satisfies an algebraic equation involving variables of $\mathbb{Y}_{\mathscr{C}}$ and series $\mathbf{f}_b$, $b \in \mathscr{C}$. Hence, $\mathbf{f}_a$ is algebraic. Moreover, the fact that, by (3.3.27), $\mathbf{s}_{\mathrm{L}(\mathcal{B})}$ is a specialized sum of some $\mathbf{f}_a$ implies the statement of the theorem. □

Theorem 3.3.14. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be an unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ Let $\mathbf{f}$ be the series of $\mathbb{K}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$ satisfying, for any $a \in \mathscr{C}$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$,*

$$\langle x_a \mathbb{Y}^\alpha_{\mathscr{C}}, \mathbf{f}\rangle = \delta_{\alpha, \mathrm{type}(a)} + \sum_{\substack{\phi : \mathscr{C} \times \mathcal{T}_{\mathscr{C}} \to \mathbb{N} \\ \alpha = \phi^{(c_1)} \dots \phi^{(c_k)}}} \chi_{a, \sum \phi_{c_1} \cdots \sum \phi_{c_k}} \left(\prod_{b \in \mathscr{C}} \phi_b !\right) \left(\prod_{\substack{b \in \mathscr{C} \\ \gamma \in \mathcal{T}_{\mathscr{C}}}} \langle x_b \mathbb{Y}^\gamma_{\mathscr{C}}, \mathbf{f}\rangle^{\phi(b,\gamma)}\right). \tag{3.3.28}$$

Then, for any color $a \in I$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^{\alpha} \in T^{+}$, the coefficients $\langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathsf{synt}(\mathscr{B}))\rangle$ and $\langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f}\rangle$ are equal.

Proof. First, since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.3.7, $(\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ is a well-defined series. Moreover, by (3.3.17), $(\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ satisfies the identity of series

$$(\mathbf{u} - \mathbf{r}) \odot (\mathbf{u} - \mathbf{r})^{\odot_{-1}} = \mathbf{u}. \tag{3.3.29}$$

Since the map $\mathfrak{col}$ commutes with the addition of series, with the composition product $\odot$, and with the inverse with respect to $\odot$, (3.3.29) leads to the equation

$$\mathfrak{col}(\mathbf{u} - \mathbf{r}) \odot \mathfrak{col}(\mathbf{u} - \mathbf{r})^{\odot_{-1}} = \mathfrak{col}(\mathbf{u}). \tag{3.3.30}$$

By Point *(ii)* of Proposition 3.3.7, by (3.3.18), and by definition of the composition map of $\mathbf{Bud}_{\mathscr{C}}(\mathbf{As})$, the coefficients of $\mathfrak{col}(\mathbf{u} - \mathbf{r})^{\odot_{-1}}$ satisfy, for all $(a, u) \in \mathbf{Bud}_{\mathscr{C}}(\mathbf{As})$, the recurrence relation

$$\left\langle (a, u), \mathfrak{col}(\mathbf{u} - \mathbf{r})^{\odot_{-1}} \right\rangle = \delta_{u,a} + \sum_{\substack{w \in \mathscr{C}^{+} \\ w \neq a}} \lambda_{a,w} \sum_{\substack{v^{(1)}, \dots, v^{(|w|)} \in \mathscr{C}^{+} \\ u = v^{(1)} \dots v^{(|w|)}}} \prod_{i \in [|w|]} \left\langle \left(w_i, v^{(i)}\right), \mathfrak{col}(\mathbf{u} - \mathbf{r})^{\odot_{-1}} \right\rangle, \tag{3.3.31}$$

where $\lambda_{a,w}$ denotes the number of rules $r \in \mathfrak{R}$ such that $\mathrm{out}(r) = a$ and $\mathrm{in}(r) = w$. By definition of $\mathfrak{colt}$ and by (3.3.31), a straightforward computation shows that the coefficients of $\mathfrak{colt}\left((\mathbf{u} - \mathbf{r})^{\odot_{-1}}\right)$ express for any $\alpha \in \mathcal{T}_{\mathscr{C}}$, as

$$\begin{aligned} &\left\langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}\left((\mathbf{u} - \mathbf{r})^{\odot_{-1}}\right) \right\rangle \\ &\quad = \delta_{\alpha,\mathrm{type}(a)} + \sum_{\substack{\gamma \in \mathcal{T}_{\mathscr{C}} \\ \gamma \neq \mathrm{type}(a)}} \chi_{a,\gamma} \sum_{\substack{\beta^{(1)}, \dots, \beta^{(\deg(\gamma))} \in \mathcal{T}_{\mathscr{C}} \\ \alpha = \beta^{(1)} \dotplus \dots \dotplus \beta^{(\deg(\gamma))}}} \prod_{i \in [\deg(\gamma)]} \left\langle \mathsf{x}_{\mathscr{C}_i^{\gamma}} \mathbb{Y}_{\mathscr{C}}^{\beta^{(i)}}, \mathfrak{colt}\left((\mathbf{u} - \mathbf{r})^{\odot_{-1}}\right) \right\rangle. \end{aligned} \tag{3.3.32}$$

Therefore, (3.3.32) provides a recurrence relation for the coefficients of $\mathfrak{colt}\left((\mathbf{u} - \mathbf{r})^{\odot_{-1}}\right)$. By using the notations introduced in Section 3.1.6 about mappings $\phi : \mathscr{C} \times \mathcal{T}_{\mathscr{C}} \to \mathbb{N}$, we obtain that the coefficients of $\mathfrak{colt}\left((\mathbf{u} - \mathbf{r})^{\odot_{-1}}\right)$ satisfy the same recurrence relation (3.3.28) as the ones of $\mathbf{f}$. Finally, Lemma 3.2.3 implies that, when $a \in I$ and $\mathscr{C}^{\alpha} \in T^{+}$, $\langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathsf{synt}(\mathscr{B}))\rangle$ and $\langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f}\rangle$ are equal. □

When $\mathscr{B}$ is a bud generating system satisfying the conditions of Theorem 3.3.14 (which are the same as the ones required by Proposition 3.3.12), one has for any $n \geqslant 1$,

$$\left\langle t^{n}, \mathbf{s}_{\mathrm{L}(\mathscr{B})} \right\rangle = \sum_{a \in I} \sum_{\substack{\alpha \in \mathcal{T}_{\mathscr{C}} \\ \alpha_i = 0, c_i \in \mathscr{C} \setminus T}} \langle \mathsf{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f} \rangle. \tag{3.3.33}$$

Therefore, this provides an alternative and recursive way to compute the coefficients of $\mathbf{s}_{\mathrm{L}(\mathscr{B})}$, different from the one of Proposition 3.3.12[j] .

[j]See an example of computation of a series $\mathbf{s}_{\mathrm{L}(\mathscr{B})}$ in Section 4.4.4.

3.4. **Series and synchronous languages.** We introduce the Kleene star operation of the composition product $\odot$ in order to define the synchronous generating series of a bud generating system $\mathcal{B}$. We relate this series with the synchronous language of $\mathcal{B}$ and provide ways to compute its coefficients.

Proofs of some results of this section are very similar to ones of Section 3.3. For this reason, some proofs are sketched here.

3.4.1. *Composition star product.* For any $\mathcal{O}$-series $\mathbf{f} \in \mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$ and any $\ell \geqslant 0$, let $\mathbf{f}^{\odot_\ell}$ be the series defined by

$$\mathbf{f}^{\odot_\ell} := \prod_{1 \leqslant i \leqslant \ell} \mathbf{f}, \tag{3.4.1}$$

where the product of (3.4.1) denotes the iterated version of the composition product $\odot$. Observe that since $\odot$ is associative (see Proposition 3.2.2), this definition is consistent. Immediately from this definition and the definition of the composition product $\odot$, the coefficient of $\mathbf{f}^{\odot_\ell}$, $\ell \geqslant 0$, satisfies for any $x \in \mathcal{O}$,

$$\left\langle x, \mathbf{f}^{\odot_\ell} \right\rangle = \begin{cases} \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} & \text{if } \ell = 0, \\ \displaystyle\sum_{\substack{y, z_1, \dots, z_{|y|} \in \mathcal{O} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \left\langle y, \mathbf{f}^{\odot_{\ell-1}} \right\rangle \prod_{i \in [|y|]} \langle z_i, \mathbf{f} \rangle & \text{otherwise.} \end{cases} \tag{3.4.2}$$

Lemma 3.4.1. *Let $\mathcal{O}$ be a locally finite $\mathscr{C}$-colored operad and $\mathbf{f}$ be a series of $\mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$. Then, the coefficients of $\mathbf{f}^{\odot_{\ell+1}}$, $\ell \geqslant 0$, satisfy for any $x \in \mathcal{O}$,*

$$\left\langle x, \mathbf{f}^{\odot_{\ell+1}} \right\rangle = \sum_{\substack{\mathfrak{t} \in \mathbf{Free}_{\mathrm{perf}}(\mathcal{O}) \\ \mathrm{ht}(\mathfrak{t}) = \ell + 1 \\ \mathrm{ev}_{\mathcal{O}}(\mathfrak{t}) = x}} \prod_{v \in \mathcal{N}(\mathfrak{t})} \langle \mathfrak{t}(v), \mathbf{f} \rangle . \tag{3.4.3}$$

Proof. By Proposition 3.2.2, since $\mathcal{O}$ is locally finite, $\mathbf{f}^{\odot_{\ell+1}}$ is a well-defined $\mathcal{O}$-series. The statement of the lemma follows by induction on ℓ and by using (3.4.2). □

The *$\odot$-star* of $\mathbf{f}$ is the series

$$\mathbf{f}^{\odot_*} := \sum_{\ell \geqslant 0} \mathbf{f}^{\odot_\ell} = \mathbf{u} + \mathbf{f} + \mathbf{f} \odot \mathbf{f} + \mathbf{f} \odot \mathbf{f} \odot \mathbf{f} + \mathbf{f} \odot \mathbf{f} \odot \mathbf{f} \odot \mathbf{f} + \cdots . \tag{3.4.4}$$

Observe that $\mathbf{f}^{\odot_*}$ could be undefined for an arbitrary $\mathcal{O}$-series $\mathbf{f}$.

Proposition 3.4.2. *Let $\mathcal{O}$ be a locally finite $\mathscr{C}$-colored operad and $\mathbf{f}$ be a series of $\mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$ such that* $\mathrm{Supp}(\mathbf{f})(1)$ *finitely factorizes $\mathcal{O}$. Then,*

(i) the series $\mathbf{f}^{\odot_*}$ is well-defined;

(ii) for any $x \in \mathcal{O}$, the coefficient of x in $\mathbf{f}^{\odot_*}$ satisfies

$$\left\langle x, \mathbf{f}^{\odot_*} \right\rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y, z_1, \dots, z_{|y|} \in \mathcal{O} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \left\langle y, \mathbf{f}^{\odot_*} \right\rangle \prod_{i \in [|y|]} \langle z_i, \mathbf{f} \rangle ; \tag{3.4.5}$$

(iii) the equation

$$\mathbf{x} - \mathbf{x} \odot \mathbf{f} = \mathbf{u} \tag{3.4.6}$$

admits $\mathbf{x} = \mathbf{f}^{\odot_*}$ as unique solution.

Proof. The proof is similar to the one of Proposition 3.3.2 and uses (3.4.2) and Lemmas 1.2.1 and 3.4.1. □

In particular, Point *(ii)* of Proposition 3.4.2 gives a way, given a $\mathcal{C}$-series $\mathbf{f}$ satisfying the stated constraints, to compute recursively the coefficients of its $\odot$-star $\mathbf{f}^{\odot_*}$.

3.4.2. *Synchronous generating series.* The *synchronous generating series* of $\mathcal{B}$ the $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$-series $\mathrm{sync}(\mathcal{B})$ defined by

$$\mathrm{sync}(\mathcal{B}) := \mathbf{i} \odot \mathbf{r}^{\odot_*} \odot \mathbf{t}. \tag{3.4.7}$$

Observe that (3.4.7) could be undefined for an arbitrary set of rules $\mathfrak{R}$ of $\mathcal{B}$. Nevertheless, when $\mathbf{r}$ satisfies the conditions of Proposition 3.4.2, that is, when $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, $\mathrm{sync}(\mathcal{B})$ is well-defined.

The aim of the following is to provide an expression to compute the coefficients of $\mathrm{sync}(\mathcal{B})$.

Lemma 3.4.3. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$,*

$$\left\langle x, \mathbf{r}^{\odot_*} \right\rangle = \delta_{x, \mathbb{1}_{\mathrm{out}(x)}} + \sum_{\substack{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ z_1, \dots, z_{|y|} \in \mathfrak{R} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \left\langle y, \mathbf{r}^{\odot_*} \right\rangle. \tag{3.4.8}$$

Proof. The proof is similar to the one of Lemma 3.3.3 and uses Proposition 3.4.2. □

Theorem 3.4.4. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the synchronous generating series of $\mathcal{B}$ satisfies*

$$\mathrm{sync}(\mathcal{B}) = \sum_{\substack{\mathfrak{t} \in \mathbf{Free}_{\mathrm{perf}}(\mathfrak{R}) \\ \mathrm{out}(\mathfrak{t}) \in I \\ \mathrm{in}(\mathfrak{t}) \in T^+}} \mathrm{ev}_{\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})}(\mathfrak{t}). \tag{3.4.9}$$

Proof. Let, for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, λ_x be the number of perfect $\mathfrak{R}$-treelike expressions for x. Since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Lemma 1.2.1, all λ_x are well-defined integers. Moreover, since $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, by Point *(i)* of Proposition 3.4.2, $\mathbf{r}^{\odot_*}$ is a well-defined series. Let us show that $\langle x, \mathbf{r}^{\odot_*} \rangle = \lambda_x$. First, when x does not belong to $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$, by Lemma 3.4.3, $\langle x, \mathbf{r}^{\odot_*} \rangle = 0$. Since, in this case $\lambda_x = 0$, the property holds. Let us now assume that x belongs to $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})^{\mathfrak{R}}$. Again by Lemma 1.2.1, the $\mathfrak{R}$-degree of x is well-defined. Therefore, we proceed by induction on $\deg_{\mathfrak{R}}(x)$. By Lemma 3.4.3, when x is a colored unit $\mathbb{1}_a$, $a \in \mathscr{C}$, one has $\langle x, \mathbf{r}^{\odot_*} \rangle = 1$. Since there is exactly one treelike expression which is a perfect tree for $\mathbb{1}_a$, namely the syntax tree consisting in one leaf of output and input color a, $\lambda_{\mathbb{1}_a} = 1$ so that the base case holds. Otherwise, again by Lemma 3.4.3, we have, by using induction hypothesis,

$$\left\langle x, \mathbf{r}^{\odot_*} \right\rangle = \sum_{\substack{y \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O}) \\ z_1, \dots, z_{|y|} \in \mathfrak{R} \\ x = y \circ \left[z_1, \dots, z_{|y|}\right]}} \lambda_y = \lambda_x. \tag{3.4.10}$$

Notice that one can apply the induction hypothesis to state (3.4.10) since one has $\deg_{\mathfrak{R}}(x) \geqslant 1 + \deg_{\mathfrak{R}}(y)$.

Now, from (3.4.10) and by using Lemma 3.2.3, we obtain that for all $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$ such that $\mathrm{out}(x) \in I$ and $\mathrm{in}(x) \in T^{+}$, $\langle x, \mathrm{sync}(\mathcal{B})\rangle = \lambda_x$. By denoting by $\mathbf{f}$ the series of the right member of (3.4.9), we have $\langle x, \mathbf{f}\rangle = \lambda_x$ if $\mathrm{out}(x) \in I$ and $\mathrm{in}(x) \in T^{*}$, and $\langle x, \mathbf{f}\rangle = 0$ otherwise. This shows that this expression is equal to $\mathrm{sync}(\mathcal{B})$. □

Theorem 3.4.4 implies that for any $x \in \mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$, the coefficient of $\langle x, \mathrm{sync}(\mathcal{B})\rangle$ is the number of $\mathfrak{R}$-treelike expressions for x which are perfect trees.

The following result establishes a link between the synchronous generating series of $\mathcal{B}$ and its synchronous language.

Proposition 3.4.5. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, the support of the synchronous generating series of $\mathcal{B}$ is the synchronous language of $\mathcal{B}$.*

Proof. This is an immediate consequence of Theorem 3.4.4 and Lemma 2.2.2. □

We say that $\mathcal{B}$ is *synchronously unambiguous* if all coefficients of $\mathrm{sync}(\mathcal{B})$ are equal to 0 or to 1. This property is important to describe the coefficients of $\mathfrak{col}(\mathrm{sync}(\mathcal{B}))$ for the same reasons as the ones concerning the unambiguity property exposed in Section 3.3.4.

Let us now describe the coefficients of $\mathfrak{colt}(\mathrm{sync}(\mathcal{B}))$, the series of colors types of the synchronous series of $\mathcal{B}$, in the particular case when $\mathcal{B}$ is unambiguous. We shall give two descriptions: a first one involving a system of functional equations of series of $\mathbb{K}\,[[\mathbb{Y}_{\mathscr{C}}]]$, and a second one involving a recurrence relation on the coefficients of a series of $\mathbb{K}\,[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$.

Lemma 3.4.6. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a synchronously unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Then, for all colors $a \in I$ and all types $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^{\alpha} \in T^{+}$, the coefficients $\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathrm{sync}(\mathcal{B}))\rangle$ count the number of elements x of $\mathrm{L}_{\mathrm{S}}(\mathcal{B})$ such that $(\mathrm{out}(x), \mathrm{type}(\mathrm{in}(x))) = (a, \alpha)$.*

Proof. The proof is similar to the one of Lemma 3.3.11 and uses (3.1.10) and Proposition 3.4.5. □

Proposition 3.4.7. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a synchronously unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. For all $a \in \mathscr{C}$, let $\mathbf{f}_a(\mathrm{y}_{c_1}, \dots, \mathrm{y}_{c_k})$ be the series of $\mathbb{K}\,[[\mathbb{Y}_{\mathscr{C}}]]$ satisfying*

$$\mathbf{f}_a(\mathrm{y}_{c_1}, \dots, \mathrm{y}_{c_k}) = \mathrm{y}_a + \mathbf{f}_a(\mathbf{g}_{c_1}(\mathrm{y}_{c_1}, \dots, \mathrm{y}_{c_k}), \dots, \mathbf{g}_{c_k}(\mathrm{y}_{c_1}, \dots, \mathrm{y}_{c_k})). \tag{3.4.11}$$

Then, for any color $a \in I$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^{\alpha} \in T^{+}$, the coefficients $\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathrm{sync}(\mathcal{B}))\rangle$ and $\langle \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f}_a\rangle$ are equal.

Proof. The proof is similar to the one of Proposition 3.3.12 and uses Lemma 3.2.3 and Proposition 3.4.2. □

When $\mathcal{B}$ is a bud generating system satisfying the conditions of Proposition 3.4.7, the generating series of the synchronous language of $\mathcal{B}$ satisfies

$$\mathbf{s}_{\mathrm{L}_{\mathrm{S}}(\mathcal{B})} = \sum_{a \in I} \mathbf{f}_a^{T}, \tag{3.4.12}$$

where $\mathbf{f}_a^T$ is the specialization of the series $\mathbf{f}_a(y_{c_1}, \dots, y_{c_k})$ at $y_b := t$ for all $b \in T$ and at $y_c := 0$ for all $c \in \mathscr{C} \setminus T$. Therefore, the resolution of the system of equations given by Proposition 3.4.7 provides a way to compute the coefficients of $\mathbf{s}_{\mathrm{L}_S(\mathcal{B})}$. This resolution can be made in most cases by iteration [BLL97,FS09][k] . Moreover, when $\mathcal{G}$ is a synchronous grammar [Gir12] (see also Section 2.3.3 for a description of these grammars) and when $\mathrm{SG}(\mathcal{G}) = \mathcal{B}$, the system of functional equations provided by Proposition 3.4.7 and (3.4.12) for $\mathbf{s}_{\mathrm{L}_S(\mathcal{B})}$ is the same as the one which can be extracted from $\mathcal{G}$.

Theorem 3.4.8. *Let $\mathcal{B} := (\mathcal{O}, \mathscr{C}, \mathfrak{R}, I, T)$ be a synchronously unambiguous bud generating system such that $\mathcal{O}$ is a locally finite operad and $\mathfrak{R}(1)$ finitely factorizes $\mathbf{Bud}_{\mathscr{C}}(\mathcal{O})$. Let $\mathbf{f}$ be the series of $\mathbb{K}[[\mathbb{X}_{\mathscr{C}} \sqcup \mathbb{Y}_{\mathscr{C}}]]$ satisfying, for any $a \in \mathscr{C}$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$,*

$$\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f} \rangle = \delta_{\alpha, \mathrm{type}(a)} + \sum_{\substack{\phi : \mathscr{C} \times \mathcal{T}_{\mathscr{C}} \to \mathbb{N} \\ \alpha = \phi^{(c_1)} \dots \phi^{(c_k)}}} \left(\prod_{b \in \mathscr{C}} \phi_b ! \right) \left(\prod_{\substack{b \in \mathscr{C} \\ \gamma \in \mathcal{T}_{\mathscr{C}}}} \chi_{b,\gamma}^{\phi(b,\gamma)} \right) \left\langle \mathrm{x}_a \prod_{b \in \mathscr{C}} y_b^{\sum \phi_b}, \mathbf{f} \right\rangle . \tag{3.4.13}$$

Then, for any color $a \in I$ and any type $\alpha \in \mathcal{T}_{\mathscr{C}}$ such that $\mathscr{C}^{\alpha} \in T^+$, the coefficients $\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathfrak{colt}(\mathrm{sync}(\mathcal{B})) \rangle$ and $\langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f} \rangle$ are equal.

Proof. The proof is similar to the one of Theorem 3.3.14 and uses Lemma 3.2.3 and Proposition 3.4.2. □

When $\mathcal{B}$ is a bud generating system satisfying the conditions of Theorem 3.4.8 (which are the same as the ones required by Proposition 3.4.7), one has for any $n \geqslant 1$,

$$\left\langle t^n, \mathbf{s}_{\mathrm{L}_S(\mathcal{B})} \right\rangle = \sum_{a \in I} \sum_{\substack{\alpha \in \mathcal{T}_{\mathscr{C}} \\ \alpha_i = 0, c_i \in \mathscr{C} \setminus T}} \langle \mathrm{x}_a \mathbb{Y}_{\mathscr{C}}^{\alpha}, \mathbf{f} \rangle . \tag{3.4.14}$$

Therefore, this provides an alternative and recursive way to compute the coefficients of $\mathbf{s}_{\mathrm{L}_S(\mathcal{B})}$, different from the one of Proposition 3.4.7[l] .

4. Examples

This final section is devoted to illustrate the notions and the results contained in the previous ones. We first define here some monochrome operads, then give examples of series on colored operads, and construct some bud generating systems. We end this section by explaining how bud generating systems can be used as tools for enumeration. For this purpose, we use the syntactic and synchronous generating series of several bud generating systems to compute the generating series of various combinatorial objects.

4.1. **Monochrome operads and bud operads.**

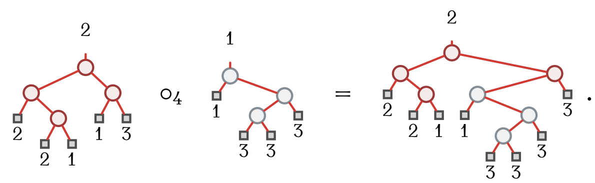
Let us start by defining three monochrome operads involving some classical combinatorial objects: binary trees, some words of integers, and Motzkin paths.

[k] See example of a computation of a series $\mathbf{s}_{\mathrm{L}_S(\mathcal{B})}$ by iteration in Section 4.4.6.

[l] See examples of computations of series $\mathbf{s}_{\mathrm{L}_S(\mathcal{B})}$ in in Sections 4.4.5 and 4.4.6.

4.1.1. *The magmatic operad.* A *binary tree* is a planar rooted tree $\mathfrak{t}$ such that any internal node of $\mathfrak{t}$ has two children. The *magmatic operad* **Mag** is the monochrome operad wherein $\mathbf{Mag}(n)$ is the set of all binary trees with n leaves. The partial composition $\mathfrak{s} \circ_i \mathfrak{t}$ of two binary trees $\mathfrak{s}$ and $\mathfrak{t}$ is the binary tree obtained by grafting the root of $\mathfrak{t}$ on the i-th leaf of $\mathfrak{s}$. The only tree of **Mag** consisting in exactly one leaf is denoted by $\perp$ and is the unit of **Mag**. Notice that **Mag** is isomorphic to the operad $\mathbf{Free}(C)$ where C is the monochrome graded collection defined by $C := C(2) := \{\mathsf{a}\}$.

For any set $\mathscr{C}$ of colors, the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathbf{Mag})$ is the $\mathscr{C}$-graded colored collection of all binary trees $\mathfrak{t}$ where all leaves (inputs) and the root (output) of $\mathfrak{t}$ are labeled on $\mathscr{C}$. For instance, in $\mathbf{Bud}_{\{1,2,3\}}(\mathbf{Mag})$, one has


(4.1.1)

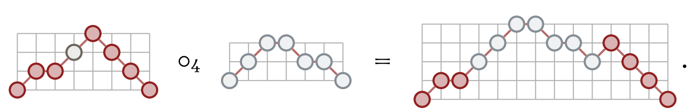
4.1.2. *The pluriassociative operad.* Let γ be a nonnegative integer. The *γ-pluriassociative operad* $\mathbf{Dias}_\gamma$ [Gir16b] is the monochrome operad wherein $\mathbf{Dias}_\gamma(n)$ is the set of all words of length n on the alphabet $\{0\} \cup [\gamma]$ with exactly one occurrence of 0. The partial composition $u \circ_i v$ of two such words u and v consists in replacing the i-th letter of u by v', where v' is the word obtained from v by replacing all its letters a by the greatest integer in $\{a, u_i\}$. For instance, in $\mathbf{Dias}_4$, one has

$$313321 \circ_4 4112 = 313433321. \tag{4.1.2}$$

Observe that $\mathbf{Dias}_0$ is the operad **As** and that $\mathbf{Dias}_1$ is the *diassociative operad* introduced by Loday [Lod01].

For any set $\mathscr{C}$ of colors, the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathbf{Dias}_\gamma)$ is the $\mathscr{C}$-colored graded collection of all words u of $\mathbf{Dias}_\gamma$ where all letters of u (inputs) and the whole word u (output) are labeled on $\mathscr{C}$.

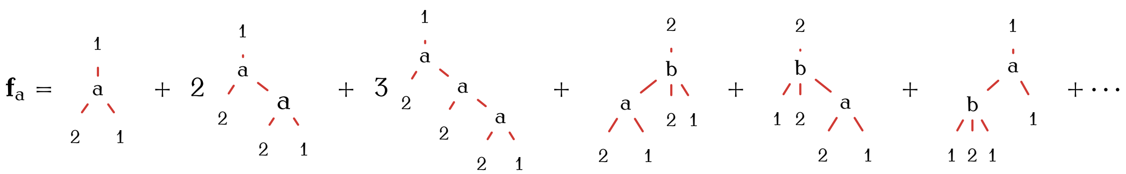
4.1.3. *The operad of Motzkin paths.* The *operad of Motzkin paths* **Motz** [Gir15] is a monochrome operad where $\mathbf{Motz}(n)$ is the set of all Motzkin paths consisting in $n-1$ steps. A *Motzkin path* of arity n is a path in $\mathbb{N}^2$ connecting the points $(0,0)$ and $(n-1,0)$, and made of steps $(1,0)$, $(1,1)$, and $(1,-1)$. If $\mathfrak{a}$ is a Motzkin path, the *i-th point* of $\mathfrak{a}$ is the point of $\mathfrak{a}$ of abscissa $i-1$. The partial composition $\mathfrak{a} \circ_i \mathfrak{b}$ of two Motzkin paths $\mathfrak{a}$ and $\mathfrak{b}$ consists in replacing the i-th point of $\mathfrak{a}$ by $\mathfrak{b}$. For instance, in **Motz**, one has


(4.1.3)

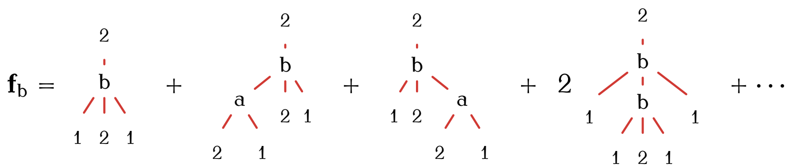
For any set $\mathscr{C}$ of colors, the bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathbf{Motz})$ is the $\mathscr{C}$-colored graded collection of all Motzkin paths $\mathfrak{a}$ where all points of $\mathfrak{a}$ (inputs) and the whole path $\mathfrak{a}$ (output) are labeled on $\mathscr{C}$.

4.2. **Series on colored operads.** Here, some examples of series on colored operads are constructed, as well as examples of series of colors, series of color types, and pruned series.

4.2.1. *Series of trees.* Let $\mathcal{O}$ be the free $\mathscr{C}$-colored operad over C where $\mathscr{C} := \{1, 2\}$ and C is the $\mathscr{C}$-graded collection defined by $C := C(2) \sqcup C(3)$ with $C(2) := \{\mathsf{a}\}$, $C(3) := \{\mathsf{b}\}$, $\mathrm{out}(\mathsf{a}) := 1$, $\mathrm{out}(\mathsf{b}) := 2$, $\mathrm{in}(\mathsf{a}) := 21$, and $\mathrm{in}(\mathsf{b}) := 121$. Let $\mathbf{f}_{\mathsf{a}}$ (resp. $\mathbf{f}_{\mathsf{b}}$) be the series of $\mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$ where for any syntax tree $\mathfrak{t}$ of $\mathcal{O}$, $\langle \mathfrak{t}, \mathbf{f}_{\mathsf{a}}\rangle$ (resp. $\langle \mathfrak{t}, \mathbf{f}_{\mathsf{b}}\rangle$) is the number of internal nodes of $\mathfrak{t}$ labeled by a (resp. b). The series $\mathbf{f}_{\mathsf{a}}$ and $\mathbf{f}_{\mathsf{b}}$ are of the form



(4.2.1a)



(4.2.1b)

The sum $\mathbf{f}_{\mathsf{a}} + \mathbf{f}_{\mathsf{b}}$ is the series wherein the coefficient of any syntax tree $\mathfrak{t}$ of $\mathcal{O}$ is its degree. Let also $\mathbf{f}_{|_1}$ (resp. $\mathbf{f}_{|_2}$) be the series of $\mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$ where for any syntax tree $\mathfrak{t}$ of $\mathcal{O}$, $\langle \mathfrak{t}, \mathbf{f}_{|_1}\rangle$ (resp. $\langle \mathfrak{t}, \mathbf{f}_{|_2}\rangle$) is the number of inputs colors 1 (resp. 2) of $\mathfrak{t}$. The sum $\mathbf{f}_{|_1} + \mathbf{f}_{|_2}$ is the series wherein the coefficient of any syntax tree $\mathfrak{t}$ of $\mathcal{O}$ is its arity. Moreover, the series $\mathbf{f}_{\mathsf{a}} + \mathbf{f}_{\mathsf{b}} + \mathbf{f}_{|_1} + \mathbf{f}_{|_2}$ is the series wherein the coefficient of any syntax tree $\mathfrak{t}$ of $\mathcal{O}$ is its total number of nodes.

The series of colors of $\mathbf{f}_{\mathsf{a}}$ is of the form

$$\mathfrak{col}(\mathbf{f}_{\mathsf{a}}) = (1, 21) + 2\,(1, 221) + 3\,(1, 2221) + (2, 2121) + (2, 1221) + (1, 1211) + \cdots, \tag{4.2.2}$$

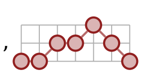
and the series of color types of $\mathbf{f}$ is of the form

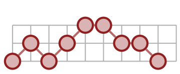
$$\mathfrak{colt}(\mathbf{f}_{\mathsf{a}}) = \mathsf{x}_1\mathsf{y}_1\mathsf{y}_2 + 2\mathsf{x}_1\mathsf{y}_1\mathsf{y}_2^2 + \mathsf{x}_1\mathsf{y}_1^3\mathsf{y}_2 + 3\mathsf{x}_1\mathsf{y}_1\mathsf{y}_2^3 + 2\mathsf{x}_2\mathsf{y}_1^2\mathsf{y}_2^2 + \cdots. \tag{4.2.3}$$

4.2.2. *Series of Motzkin paths.* Let $\mathcal{O}$ be the $\mathscr{C}$-bud operad $\mathbf{Bud}_{\mathscr{C}}(\mathbf{Motz})$, where $\mathscr{C} := \{-1, 1\}$. Let $\mathbf{f}$ be the series of $\mathbb{K}\langle\langle\mathcal{O}\rangle\rangle$ defined for any Motzkin path $\mathfrak{b}$, input color $a \in \mathscr{C}$, and word of input colors $u \in \mathscr{C}^{|\mathfrak{b}|}$ by

$$\langle (a, \mathfrak{b}, u), \mathbf{f}\rangle := \frac{1}{2^{|\mathfrak{b}|+1}} \left(\prod_{i \in [|\mathfrak{b}|]} q_{\mathrm{ht}_{\mathfrak{b}}(i)}^{u_i}\right)^{a}, \tag{4.2.4}$$

where $\mathrm{ht}_{\mathfrak{b}}(i)$ is the ordinate of the i-th point of $\mathfrak{b}$. One has for instance, where the notation $\bar{1}$ stands for -1,

$\left\langle\left(1,\right.\right.$  $\left.\left., 1\bar{1}111\bar{1}1\right), \mathbf{f}\right\rangle = \frac{1}{2^8}\left(q_0 q_0^{\bar{1}} q_1 q_1 q_2 q_1^{\bar{1}} q_0\right)^1 = \frac{q_0 q_1 q_2}{2^8},$ (4.2.5a)

$\left\langle\left(\bar{1},\right.\right.$  $\left.\left., \bar{1}1\bar{1}11\bar{1}1\bar{1}1\right), \mathbf{f}\right\rangle = \frac{1}{2^{10}}\left(q_0^{\bar{1}} q_1 q_0^{\bar{1}} q_1 q_2 q_2^{\bar{1}} q_1 q_1^{\bar{1}} q_0\right)^{\bar{1}} = \frac{q_0}{2^{10} q_1^2}.$ (4.2.5b)

Moreover, the coefficients of the pruned series of $\mathbf{f}$ satisfy, by definition of $\mathfrak{pru}$ and $\mathbf{f}$,

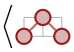
$$\langle \mathfrak{b}, \mathfrak{pru}(\mathbf{f}) \rangle = \sum_{\substack{a \in \{\bar{1}, 1\} \\ u \in \{\bar{1}, 1\}^{|\mathfrak{b}|}}} \langle (a, \mathfrak{b}, u), \mathbf{f} \rangle = \frac{1}{2^{|\mathfrak{b}|+1}} \sum_{u \in \{\bar{1}, 1\}^{|\mathfrak{b}|}} \left(\left(\prod_{i \in [|\mathfrak{b}|]} q_{\mathrm{ht}_{\mathfrak{b}}(i)}^{u_i} \right) + \left(\prod_{i \in [|\mathfrak{b}|]} q_{\mathrm{ht}_{\mathfrak{b}}(i)}^{-u_i} \right) \right). \tag{4.2.6}$$

These coefficients seem to factorize nicely. For instance,

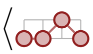
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{1 + q_0^2}{2q_0},$ (4.2.7a)

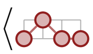
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^2}{4q_0^2},$ (4.2.7b)

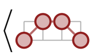
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3}{8q_0^3},$ (4.2.7c)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^2 \left(1 + q_1^2\right)}{8q_0^2 q_1},$ (4.2.7d)

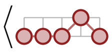
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^4}{16q_0^4},$ (4.2.7e)

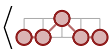
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3 \left(1 + q_1^2\right)}{16q_0^3 q_1},$ (4.2.7f)

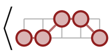
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3 \left(1 + q_1^2\right)}{16q_0^3 q_1},$ (4.2.7g)

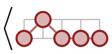
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^2 \left(1 + q_1^2\right)^2}{16q_0^2 q_1^2},$ (4.2.7h)

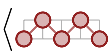
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^5}{32q_0^5},$ (4.2.7i)

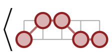
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^4 \left(1 + q_1^2\right)}{32q_0^4 q_1},$ (4.2.7j)

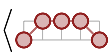
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^4 \left(1 + q_1^2\right)}{32q_0^4 q_1},$ (4.2.7k)

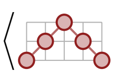
$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3 \left(1 + q_1^2\right)^2}{32q_0^3 q_1^2},$ (4.2.7l)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^4 \left(1 + q_1^2\right)}{32q_0^4 q_1},$ (4.2.7m)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3 \left(1 + q_1^2\right)^2}{32q_0^3 q_1^2},$ (4.2.7n)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^3 \left(1 + q_1^2\right)^2}{32q_0^3 q_1^2},$ (4.2.7o)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^2 \left(1 + q_1^2\right)^3}{32q_0^2 q_1^3},$ (4.2.7p)

$\left\langle$  $, \mathfrak{pru}(\mathbf{f}) \right\rangle = \dfrac{\left(1 + q_0^2\right)^2 \left(1 + q_1^2\right)^2 \left(1 + q_2^2\right)}{32q_0^2 q_1^2 q_2}.$ (4.2.7q)

Observe that the specializations at $q_0 := 1$, $q_1 := 1$, and $q_2 := 1$ of all these coefficients are equal to 1.

4.3. **Bud generating systems.** We rely on the monochrome operads defined in Section 4.1 to construct several bud generating systems. We review some properties of these, leaving the proofs to the reader.

4.3.1. *Monochrome bud generating systems from* $\mathbf{Dias}_\gamma$. Let γ be a nonnegative integer and consider the monochrome bud generating system $\mathcal{B}_{\mathrm{w},\gamma} := \left(\mathbf{Dias}_\gamma, \mathfrak{R}_\gamma\right)$ where

$$\mathfrak{R}_\gamma := \{0a, a0 : a \in [\gamma]\}. \tag{4.3.1}$$

The derivation graph of $\mathcal{B}_{\mathrm{w},1}$ is depicted by Figure 2 and the one of $\mathcal{B}_{\mathrm{w},2}$, by Figure 3.

Proposition 4.3.1. *For any $\gamma \geqslant 0$, the monochrome bud generating system $\mathcal{B}_{\mathrm{w},\gamma}$ satisfies the following properties.*

(i) It is faithful.
(ii) The set $\mathrm{L}\left(\mathcal{B}_{\mathrm{w},\gamma}\right)$ is equal to the underlying monochrome graded collection of $\mathbf{Dias}_\gamma$.
(iii) The set of rules $\mathfrak{R}_\gamma(1)$ factorizes finitely $\mathbf{Dias}_\gamma$.

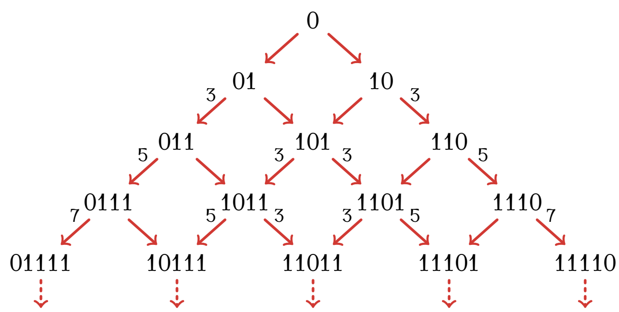


FIGURE 2. The derivation graph of $\mathcal{B}_{\mathrm{w},1}$.

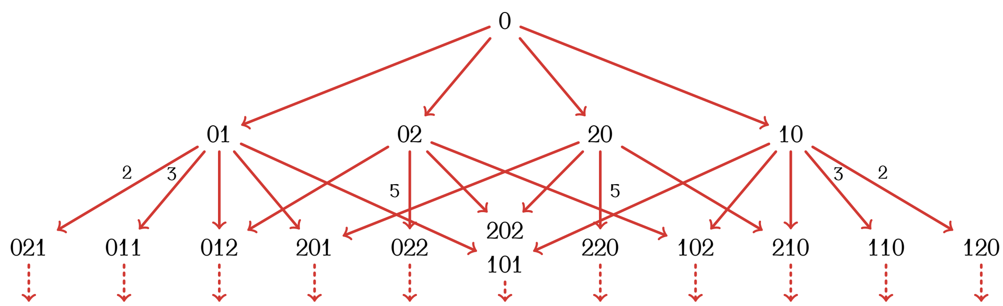


FIGURE 3. The derivation graph of $\mathcal{B}_{\mathrm{w},2}$.

Property *(i)* of Proposition 4.3.1 is a consequence of the fact that $\mathcal{B}_{\mathrm{w},\gamma}$ is monochrome and Property *(ii)* is implied by the fact that $\mathfrak{R}_\gamma$ is a generating set of $\mathbf{Dias}_\gamma$ [Gir15]. Moreover, observe that since the word $\gamma 0 \gamma$ of $\mathbf{Dias}_\gamma(3)$ admits exactly the two $\mathfrak{R}_\gamma$-treelike expressions $0\gamma \circ_1 \gamma 0$ and $\gamma 0 \circ_2 0 \gamma$, by Theorem 3.3.9, $\langle \gamma 0 \gamma, \mathrm{synt}(\mathcal{B}_{\mathrm{w},\gamma}) \rangle = 2$. Hence, $\mathcal{B}_{\mathrm{w},\gamma}$ is not unambiguous.

4.3.2. *A bud generating system for Motzkin paths.* Consider the bud generating system $\mathcal{B}_{\mathrm{p}} := (\mathbf{Motz}, \{1, 2\}, \mathfrak{R}, \{1\}, \{1, 2\})$ where

$$\mathfrak{R} := \left\{ (1, \quad, 22), \left(1, \quad, 111\right) \right\}. \tag{4.3.2}$$

Figure 4 shows a sequence of derivations in $\mathcal{B}_{\mathrm{p}}$ and Figure 5 shows the derivation graph of $\mathcal{B}_{\mathrm{p}}$.

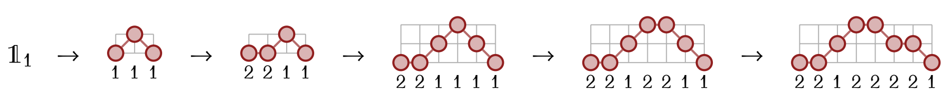


FIGURE 4. A sequence of derivations in $\mathcal{B}_{\mathrm{p}}$. The input colors of the elements of $\mathbf{Bud}_{\{1,2\}}(\mathbf{Motz})$ are depicted below the paths. The output color of all these elements is 1.

Let $L_{\mathcal{B}_{\mathrm{p}}}$ be the set of Motzkin paths with no consecutive horizontal steps.

Proposition 4.3.2. *The bud generating system $\mathcal{B}_{\mathrm{p}}$ satisfies the following properties.*

(i) It is faithful.

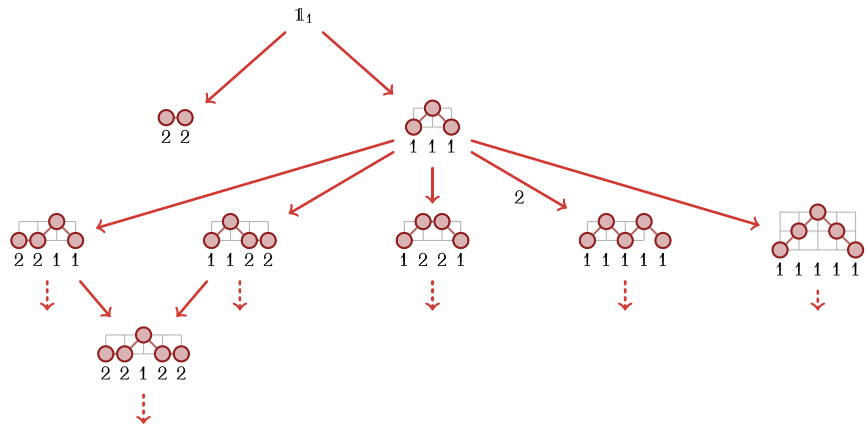


FIGURE 5. The derivation graph of $\mathcal{B}_{\mathrm{p}}$. The input colors of the elements of $\mathbf{Bud}_{\{1,2\}}(\mathbf{Motz})$ are depicted below the paths. The output color of all these elements is 1.

(ii) *The restriction of the pruning map* $\mathfrak{pru}$ *on the domain* $\mathrm{L}\left(\mathcal{B}_{\mathrm{p}}\right)$ *is a bijection between* $\mathrm{L}\left(\mathcal{B}_{\mathrm{p}}\right)$ *and* $L_{\mathcal{B}_{\mathrm{p}}}$.

(iii) *The set of rules* $\mathfrak{R}(1)$ *finitely factorizes* $\mathbf{Bud}_{\{1,2\}}(\mathbf{Motz})$.

Properties *(i)* and *(ii)* of Proposition 4.3.2 together say that the sequence enumerating the elements of $\mathrm{L}(\mathcal{B}_{\mathrm{p}})$ with respect to their arity is the one enumerating the Motzkin paths with no consecutive horizontal steps. This sequence is Sequence **A104545** of [Slo], starting by

$$1, 1, 1, 3, 5, 11, 25, 55, 129, 303, 721, 1743, 4241, 10415, 25761, 64095. \tag{4.3.3}$$

Moreover, since the Motzkin path of $\mathbf{Motz}(5)$ admits exactly the two $\mathfrak{R}$-treelike expressions $\circ_1$ and $\circ_3$, by Theorem 3.3.9, $\left\langle \left(1, \;, 11111\right), \mathrm{synt}(\mathcal{B}_{\mathrm{p}}) \right\rangle = 2$. Hence $\mathcal{B}_{\mathrm{p}}$ is not unambiguous.

4.3.3. *A bud generating system for unary-binary trees.* Let C be the monochrome graded collection defined by $C := C(1) \sqcup C(2)$ where $C(1) := \{\mathtt{a}, \mathtt{b}\}$ and $C(2) := \{\mathtt{c}\}$. Let $\mathcal{B}_{\mathrm{bu}} := (\mathbf{Free}(C), \{1, 2\}, \mathfrak{R}, \{1\}, \{2\})$ be the bud generating system where

$$\mathfrak{R} := \left\{ \left(1, \mathtt{a}, 2\right), \left(1, \mathtt{b}, 2\right), \left(2, \mathtt{c}, 11\right) \right\}. \tag{4.3.4}$$

Figure 6 shows a sequence of derivations in $\mathcal{B}_{\mathrm{bu}}$.

A *unary-binary tree* is a planar rooted tree $\mathfrak{t}$ such that all internal nodes of $\mathfrak{t}$ are of arities 1 or 2, all nodes of $\mathfrak{t}$ of arity 1 have a child which is an internal node of arity 2 or is a leaf, and all nodes of $\mathfrak{t}$ of arity 2 have two children which are internal nodes of arity 1 or are leaves.

Let $L_{\mathcal{B}_{\mathrm{bu}}}$ be the set of unary-binary trees with a root of arity 1, all parents of the leaves are of arity 1, and unary nodes are labeled by $\mathtt{a}$ or $\mathtt{b}$.

Proposition 4.3.3. *The bud generating system* $\mathcal{B}_{\mathrm{bu}}$ *satisfies the following properties.*

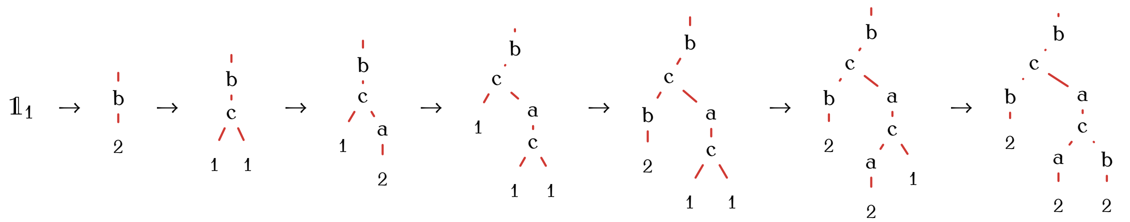


FIGURE 6. A sequence of derivations in $\mathscr{B}_{\mathrm{bu}}$. The input colors of the elements of $\mathbf{Bud}_{\{1,2\}}(\mathbf{Free}(C))$ are depicted below the leaves. The output color of all these elements is 1. Since all input colors of the last tree are 2, this tree is in $\mathrm{L}(\mathscr{B}_{\mathrm{bu}})$.

(i) It is faithful.
(ii) It is unambiguous.
(iii) The restriction of the pruning map $\mathfrak{pru}$ *on the domain* $\mathrm{L}(\mathscr{B}_{\mathrm{bu}})$ *is a bijection between* $\mathrm{L}(\mathscr{B}_{\mathrm{bu}})$ *and* $L_{\mathscr{B}_{\mathrm{bu}}}$.
(iv) The set of rules $\mathfrak{R}(1)$ *finitely factorizes* $\mathbf{Bud}_{\{1,2\}}(\mathbf{Free}(C))$.

4.3.4. *A bud generating system for B-perfect trees.* Let B be a finite set of positive integers and C_B be the monochrome graded collection defined by $C_B := \bigsqcup_{n \in B} C_B(n) := \bigsqcup_{n \in B} \{\mathsf{a}_n\}$. We consider the monochrome bud generating system $\mathscr{B}_{\mathrm{bt},B} := (\mathbf{Free}(C_B), \mathfrak{R}_B)$ where $\mathfrak{R}_B$ is the set of all corollas of $\mathbf{Free}(C_B)(1)$. Figure 7 shows the synchronous derivation graph of $\mathscr{B}_{\mathrm{bt},\{2,3\}}$.

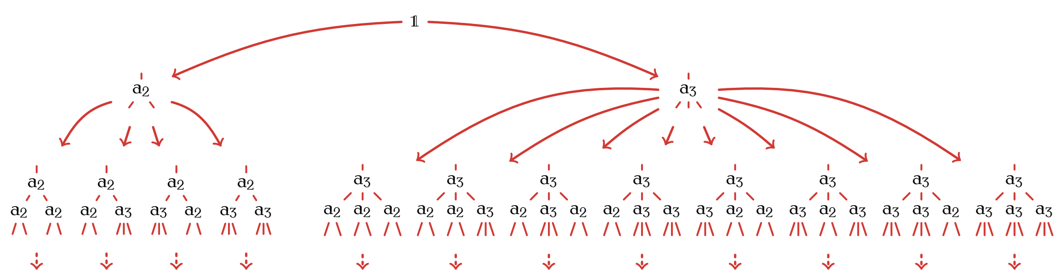


FIGURE 7. The synchronous derivation graph of $\mathscr{B}_{\mathrm{bt},\{2,3\}}$.

A *B-perfect tree* is a planar rooted tree $\mathfrak{t}$ such that all internal nodes of $\mathfrak{t}$ have an arity in B and all paths connecting the root of $\mathfrak{t}$ to its leaves have the same length. These trees and their generating series have been studied for the particular case $B := \{2, 3\}$ [MPRS79, CLRS09] and appear as data structures in computer science (see [Odl82, Knu98, FS09]).

Proposition 4.3.4. *For any finite set* B *of positive integers, the bud generating system* $\mathscr{B}_{\mathrm{bt},B}$ *satisfies the following properties.*

(i) It is synchronously faithful.
(ii) It is synchronously unambiguous.
(iii) The synchronous language $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bt},B})$ *of* $\mathscr{B}_{\mathrm{bt},B}$ *is the set of all B-perfect trees.*

(iv) *The set of rules* $\mathfrak{R}_B(1)$ *finitely factorizes* $\mathbf{Free}(C_B)$*.*

(v) *When* $1 \notin B$*, the generating series* $\mathbf{s}_{\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bt},B})}$ *of the synchronous language of* $\mathscr{B}_{\mathrm{bt},B}$ *is well-defined.*

Property (v) of Proposition 4.3.4 is a consequence of the fact that when $1 \notin B$, $\mathbf{Free}(C_B)$ is locally finite and therefore there are only finitely many elements in $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bt},B})$ of a given arity. By Property (iii) of Proposition 4.3.4, the sequences enumerating the elements of $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bt},B})$ with respect to their arity are, for instance, Sequence **A014535** of [Slo] for $B = \{2, 3\}$ which starts by

$$1, 1, 1, 1, 2, 2, 3, 4, 5, 8, 14, 23, 32, 43, 63, 97, 149, 224, 332, 489, \tag{4.3.5}$$

and Sequence **A037026** of [Slo] for $B = \{2, 3, 4\}$ which starts by

$$1, 1, 1, 2, 2, 4, 5, 9, 15, 28, 45, 73, 116, 199, 345601, 1021, 1738, 2987, 5244. \tag{4.3.6}$$

4.3.5. *A bud generating system for balanced binary trees.* Consider the bud generating system $\mathscr{B}_{\mathrm{bbt}} := (\mathbf{Mag}, \{1, 2\}, \mathfrak{R}, \{1\}, \{1\})$ where

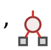
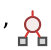
$\mathfrak{R} := \Big\{ \Big(1,$  $, 11\Big), \Big(1,$  $, 12\Big), \Big(1,$  $, 21\Big), (2,$  $, 1) \Big\}.$ (4.3.7)

Figure 8 shows a sequence of synchronous derivations in $\mathscr{B}_{\mathrm{bbt}}$.

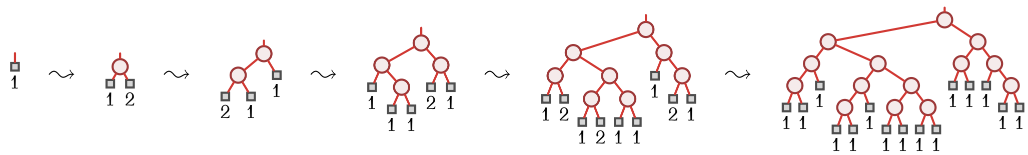


FIGURE 8. A sequence of synchronous derivations in $\mathscr{B}_{\mathrm{bbt}}$. The input colors of the elements of $\mathbf{Bud}_{\{1,2\}}(\mathbf{Mag})$ are depicted below the leaves. The output color of all these elements is 1. Since all input colors of the last tree are 1, this tree is in $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bbt}})$.

The *height* of a binary tree $\mathfrak{t}$ is the height of $\mathfrak{t}$ seen as a monochrome syntax tree. A *balanced binary tree* [AVL62] is a binary tree $\mathfrak{t}$ wherein, for any internal node x of $\mathfrak{t}$, the difference between the height of the left subtree and the height of the right subtree of x is -1, 0, or 1.

Proposition 4.3.5. *The bud generating system* $\mathscr{B}_{\mathrm{bbt}}$ *satisfies the following properties.*

(i) *It is synchronously faithful.*

(ii) *It is synchronously unambiguous.*

(iii) *The restriction of the pruning map* $\mathfrak{pru}$ *on the domain* $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bbt}})$ *is a bijection between* $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bbt}})$ *and the set of balanced binary trees.*

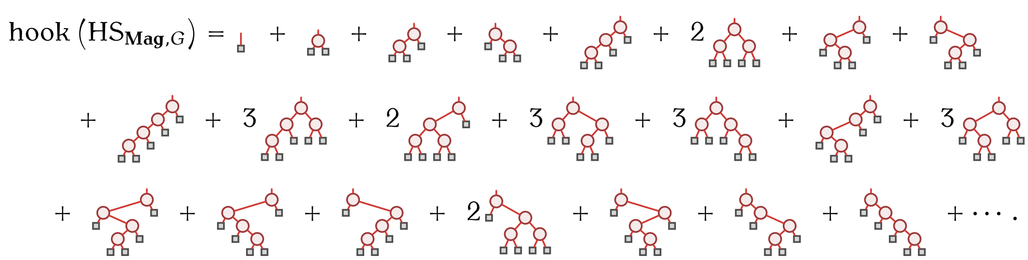
(iv) *The set of rules* $\mathfrak{R}(1)$ *finitely factorizes* $\mathbf{Bud}_{\{1,2\}}(\mathbf{Mag})$*.*

Properties (ii) and (iii) of Proposition 4.3.5 are based upon combinatorial properties of a synchronous grammar $\mathscr{G}$ of balanced binary trees defined in [Gir12] and satisfying $\mathrm{SG}(\mathscr{G}) = \mathscr{B}_{\mathrm{bbt}}$ (see Section 2.3.3 and Proposition 2.3.3). Besides, Properties (i) and (iii) of Proposition 4.3.5 together imply that the sequence enumerating the elements of $\mathrm{L}_{\mathrm{S}}(\mathscr{B}_{\mathrm{bbt}})$ with respect to their arity is the one enumerating the balanced binary trees. This sequence in Sequence A006265 of [Slo], starting by

$$1, 1, 2, 1, 4, 6, 4, 17, 32, 44, 60, 70, 184, 476, 872, 1553, 2720, 4288, 6312, 9004. \tag{4.3.8}$$

4.4. **Series of bud generating systems.** We now consider the bud generating systems constructed in Section 4.3 to give some examples of hook generating series. We also put into practice what we have exposed in Sections 3.3.4 and 3.4.2 to compute the generating series of languages or synchronous languages of bud generating systems by using syntactic generating series and synchronous generating series.

4.4.1. *Hook coefficients for binary trees.* Let us consider the hook bud generating system $\mathrm{HS}_{\mathbf{Mag},G}$ where $G := \{$  $\}$. This bud generating system leads to the definition of a statistic on binary trees, provided by the coefficients of the hook generating series $\mathrm{hook}\left(\mathrm{HS}_{\mathbf{Mag},G}\right)$ which begins by



(4.4.1)

Theorem 3.3.5 implies that for any binary tree $\mathfrak{t}$, the coefficient $\left\langle \mathfrak{t}, \mathrm{hook}\left(\mathrm{HS}_{\mathbf{Mag},G}\right)\right\rangle$ can be obtained by the usual hook-length formula of binary trees. This explains the name of hook generating series for $\mathrm{hook}(\mathcal{B})$, when $\mathcal{B}$ is a bud generating system. Alternatively, the coefficient $\left\langle \mathfrak{t}, \mathrm{hook}\left(\mathrm{HS}_{\mathbf{Mag},G}\right)\right\rangle$ is the cardinal of the sylvester class [HNT05] of permutations encoded by $\mathfrak{t}$.

4.4.2. *Hook coefficients for words of* $\mathbf{Dias}_\gamma$. Let us consider the monochrome bud generating system $\mathcal{B}_{\mathrm{w},\gamma}$ and its set of rules $\mathfrak{R}_\gamma$ introduced in Section 4.3.1. Since, by Proposition 4.3.1, $\mathfrak{R}_\gamma$ generates $\mathbf{Dias}_\gamma$, $\mathcal{B}_{\mathrm{w},\gamma}$ is a hook bud generating system $\mathrm{HS}_{\mathbf{Dias}_\gamma,\mathfrak{R}_\gamma}$ (see Section 3.3.2). This leads to the definition of a statistic on the words of $\mathbf{Dias}_\gamma$, provided by the coefficients of the hook generating series $\mathrm{hook}\left(\mathrm{HS}_{\mathbf{Dias}_\gamma,\mathfrak{R}_\gamma}\right)$ of $\mathrm{HS}_{\mathbf{Dias}_\gamma,\mathfrak{R}_\gamma}$ which begins, when $\gamma = 1$, by

$$\begin{aligned}\mathrm{hook}\left(\mathrm{HS}_{\mathbf{Dias}_1,\mathfrak{R}_1}\right) &= (0) + (01) + (10) + 3\,(011) + 2\,(101) + 3\,(110) + 15\,(0111) + 9\,(1011) \\ &+ 9\,(1101) + 15\,(1110) + 105\,(01111) + 60\,(10111) + 54\,(11011) + 60\,(11101) + 105\,(11110) \\ &+ 945\,(011111) + 525\,(101111) + 45\,(110111) + 450\,(111011) + 525\,(111101) + 945\,(111110) + \cdots .\end{aligned} \tag{4.4.2}$$

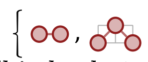
Let us set, for all $0 \leqslant a \leqslant n-1$, $h_{n,a} := \left\langle 1^a 0 1^{n-a-1}, \mathrm{hook}\left(\mathrm{HS}_{\mathbf{Dias}_1, \mathfrak{R}_1}\right)\right\rangle$. By Lemmas 3.2.3 and 3.3.3, the $h_{n,a}$ satisfy the recurrence

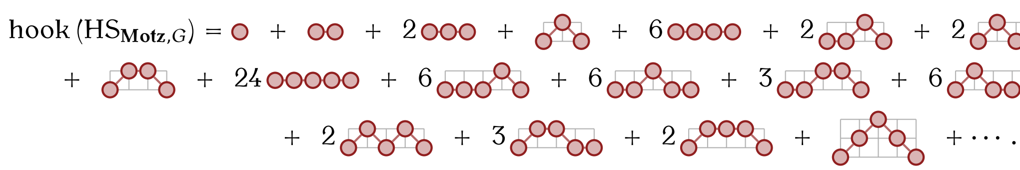
$$h_{n,a} = \begin{cases} 1 & \text{if } n = 1, \\ (2a-1)h_{n-1,a-1} & \text{if } a = n-1, \\ (2n-2a-3)h_{n-1,a} & \text{if } a = 0, \\ (2a-1)h_{n-1,a-1} + (2n-2a-3)h_{n-1,a} & \text{otherwise.} \end{cases} \tag{4.4.3}$$

The numbers $h_{n,a}$ form a triangle beginning by

$$\begin{array}{llllllll} 1 \\ 1 & 1 \\ 3 & 2 & 3 \\ 15 & 9 & 9 & 15 \\ 105 & 60 & 54 & 60 & 105 \\ 945 & 525 & 450 & 450 & 525 & 945 \\ 10395 & 5670 & 4725 & 4500 & 4725 & 5670 & 10395 \\ 135135 & 72765 & 59535 & 55125 & 55125 & 59535 & 72765 & 135135 \end{array} . \tag{4.4.4}$$

These numbers form Sequence A059366 of [Slo].

4.4.3. *Hook coefficients for Motkzin paths.* It is proven in [Gir15] that $G :=$  is a generating set of $\mathbf{Motz}$. Hence, $\mathrm{HS}_{\mathbf{Motz},G}$ is a hook generating system. This leads to the definition of a statistic on Motzkin paths, provided by the coefficients of the hook generating series $\mathrm{hook}(\mathrm{HS}_{\mathbf{Motz},G})$ of $\mathrm{HS}_{\mathbf{Motz},G}$ which begins by



(4.4.5)

4.4.4. *Generating series of some unary-binary trees.* Let us consider the bud generating system $\mathcal{B}_{\mathrm{bu}}$ introduced in Section 4.3.3. We have, for all $a \in \{1, 2\}$ and $\alpha \in \mathcal{T}_{\{1,2\}}$,

$$\chi_{a,\alpha} = \begin{cases} 2 & \text{if } (a, \alpha) = (1, 01), \\ 1 & \text{if } (a, \alpha) = (2, 20), \\ 0 & \text{otherwise,} \end{cases} \tag{4.4.6}$$

and

$$\mathbf{g}_1(y_1, y_2) = 2y_2, \tag{4.4.7a}$$

$$\mathbf{g}_2(y_1, y_2) = y_1^2. \tag{4.4.7b}$$

Since, by Proposition 4.3.3, $\mathcal{B}_{\mathrm{bu}}$ satisfies the conditions of Proposition 3.3.12, by this last proposition and (3.3.27), the generating series $\mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})}$ of $\mathrm{L}(\mathcal{B}_{\mathrm{bu}})$ satisfies $\mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})} = \mathbf{f}_1(0, t)$ where

$$\mathbf{f}_1\left(\mathsf{y}_1,\mathsf{y}_2\right) = \mathsf{y}_1 + 2\mathbf{f}_2\left(\mathsf{y}_1,\mathsf{y}_2\right), \qquad (4.4.8a) \qquad\qquad \mathbf{f}_2\left(\mathsf{y}_1,\mathsf{y}_2\right) = \mathsf{y}_2 + \mathbf{f}_1\left(\mathsf{y}_1,\mathsf{y}_2\right)^2. \qquad (4.4.8b)$$

We obtain that $\mathbf{f}_1(\mathsf{y}_1,\mathsf{y}_2)$ satisfies the functional equation

$$\mathsf{y}_1 + 2\mathsf{y}_2 - \mathbf{f}_1\left(\mathsf{y}_1,\mathsf{y}_2\right) + 2\mathbf{f}_1\left(\mathsf{y}_1,\mathsf{y}_2\right)^2 = 0. \tag{4.4.9}$$

Hence, $\mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})}$ satisfies

$$2t - \mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})} + 2\mathbf{s}^2_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})} = 0, \tag{4.4.10}$$

showing that the elements of $\mathrm{L}\left(\mathcal{B}_{\mathrm{bu}}\right)$ are enumerated, arity by arity, by Sequence **A052707** of [Slo] starting by

$$2, 8, 64, 640, 7168, 86016, 1081344, 14057472, 187432960, 2549088256. \tag{4.4.11}$$

Besides, since by Proposition 4.3.3, $\mathcal{B}_{\mathrm{bu}}$ satisfies the conditions of Theorem 3.3.14, by this last theorem and (3.3.33), $\mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})}$ satisfies, for any $n \geqslant 1$,

$$\left\langle t^n, \mathbf{s}_{\mathrm{L}(\mathcal{B}_{\mathrm{bu}})}\right\rangle = \left\langle \mathsf{x}_1\mathsf{y}_2^n, \mathbf{f}\right\rangle, \tag{4.4.12}$$

where $\mathbf{f}$ is the series satisfying, for any $a \in \mathcal{C}$ and any type $\alpha \in \mathcal{T}_{\{1,2\}}$, the recursive formula

$$\begin{aligned}\left\langle \mathsf{x}_a\mathsf{y}^{\alpha_1}\mathsf{y}^{\alpha_2}, \mathbf{f}\right\rangle = \delta_{\alpha,\mathrm{type}(a)} + \delta_{a,1}2\left\langle \mathsf{x}_2\mathsf{y}_1^{\alpha_1}\mathsf{y}_2^{\alpha_2}, \mathbf{f}\right\rangle \\ + \delta_{a,2}\sum_{\substack{d_1,d_2,d_3,d_4\in\mathbb{N}\\ \alpha_1=d_1+d_2\\ \alpha_2=d_3+d_4\\ (d_1,d_3)\neq(0,0)\neq(d_2,d_4)}}\left\langle \mathsf{x}_1\mathsf{y}_1^{d_1}\mathsf{y}_2^{d_3}, \mathbf{f}\right\rangle\left\langle \mathsf{x}_2\mathsf{y}_1^{d_2}\mathsf{y}_2^{d_4}, \mathbf{f}\right\rangle.\end{aligned} \tag{4.4.13}$$

4.4.5. *Generating series of B-perfect trees.* Let us consider the monochrome bud generating system $\mathcal{B}_{\mathrm{bt},B}$ and its set of rules $\mathfrak{R}_B$ introduced in Section 4.3.4. By Proposition 4.3.4, the generating series $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})}$ is well-defined when $1 \notin B$. For this reason, in all this section we restrict ourselves to the case where all elements of B are greater than or equal to 2. To maintain here homogeneous notations with the rest of the text, we consider that the set of colors $\mathcal{C}$ of $\mathcal{B}_{\mathrm{bt},B}$ is the singleton $\{1\}$. We have, for all $\alpha \in \mathcal{T}_{\{1\}}$,

$$\chi_{1,\alpha} = \begin{cases} 1 & \text{if } (a,\alpha) = (1,b) \text{ with } b \in B, \\ 0 & \text{otherwise,}\end{cases} \tag{4.4.14}$$

and

$$\mathbf{g}_1(\mathsf{y}_1) = \sum_{b\in B}\mathsf{y}_1^b. \tag{4.4.15}$$

Since by Proposition 4.3.4, $\mathcal{B}_{\mathrm{bt},B}$ satisfies the conditions of Proposition 3.4.7, by this last proposition and (3.4.12), the generating series $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})}$ of $\mathrm{L_S}\left(\mathcal{B}_{\mathrm{bt},B}\right)$ satisfies $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})} = \mathbf{f}_1(t)$ where

$$\mathbf{f}_1\left(\mathsf{y}_1\right) = \mathsf{y}_1 + \mathbf{f}_1\left(\sum_{b\in B}\mathsf{y}_1^b\right). \tag{4.4.16}$$

This functional equation for the generating series of B-perfect trees, in the case where $B = \{2,3\}$, is the one obtained in [Odl82, FS09, Gir12] by different methods.

Besides, since by Proposition 4.3.4, $\mathcal{B}_{\mathrm{bt},B}$ satisfies the conditions of Theorem 3.4.8, by this last theorem and (3.4.14), $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})}$ satisfies, for any $n \geqslant 1$, the recursive formula

$$\left\langle t^n, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})} \right\rangle = \delta_{n,1} + \sum_{\substack{d_b \in \mathbb{N}, b \in B \\ n = \sum_{b \in B} b\, d_b}} \wr d_b : b \in B \int ! \left\langle t^{\sum_{b \in B} d_b}, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})} \right\rangle \tag{4.4.17}$$

For instance, for $B := \{2, 3\}$, one has

$$\left\langle t^n, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},\{2,3\}})} \right\rangle = \delta_{n,1} + \sum_{\substack{d_2, d_3 \geqslant 0 \\ n = 2d_2 + 3d_3}} \binom{d_2 + d_3}{d_2} \left\langle t^{d_2 + d_3}, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},\{2,3\}})} \right\rangle, \tag{4.4.18}$$

which is a recursive formula to enumerate the $\{2, 3\}$-perfect trees known from [MPRS79], and for $B := \{2, 3, 4\}$,

$$\left\langle t^n, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},\{2,3,4\}})} \right\rangle = \delta_{n,1} + \sum_{\substack{d_2, d_3, d_4 \geqslant 0 \\ n = 2d_2 + 3d_3 + 4d_4}} \wr d_2, d_3, d_4 \int ! \left\langle t^{d_2 + d_3 + d_4}, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},\{2,3,4\}})} \right\rangle. \tag{4.4.19}$$

Moreover, it is possible to refine the enumeration of B-perfect trees to take into account of the number of internal nodes with a given arity in the trees. For this, we consider the series $\mathbf{s}_q$ satisfying the recurrence

$$\left\langle t^n, \mathbf{s}_q \right\rangle = \delta_{n,1} + \sum_{\substack{d_b \in \mathbb{N}, b \geqslant 2 \\ n = \sum_{b \geqslant 2} b\, d_b}} \wr d_b : b \geqslant 2 \int ! \left(\prod_{b \geqslant 2} q_b^{d_b} \right) \left\langle t^{\sum_{b \geqslant 2} d_b}, \mathbf{s}_q \right\rangle. \tag{4.4.20}$$

The coefficient of $\left(\prod_{b \geqslant 2} q_b^{d_b}\right) t^n$ in $\mathbf{s}_q$ is the number of $\mathbb{N} \setminus \{0, 1\}$-perfect trees with n leaves and with d_b internal nodes of arity b for all $b \geqslant 2$. The specialization of $\mathbf{s}_q$ at $q_b := 0$ for all $b \notin B$ and $q_b := t$ for all $b \in B$ is equal to the series $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bt},B})}$.

First coefficients of $\mathbf{s}_q$ are

$$\left\langle t, \mathbf{s}_q \right\rangle = 1, \tag{4.4.21a}$$
$$\left\langle t^2, \mathbf{s}_q \right\rangle = q_2, \tag{4.4.21b}$$
$$\left\langle t^3, \mathbf{s}_q \right\rangle = q_3, \tag{4.4.21c}$$
$$\left\langle t^4, \mathbf{s}_q \right\rangle = q_2^3 + q_4, \tag{4.4.21d}$$
$$\left\langle t^5, \mathbf{s}_q \right\rangle = 2q_2^2 q_3 + q_5, \tag{4.4.21e}$$
$$\left\langle t^6, \mathbf{s}_q \right\rangle = q_2^3 q_3 + q_2 q_3^2 + 2q_2^2 q_4 + q_6, \tag{4.4.21f}$$

$$\left\langle t^7, \mathbf{s}_q \right\rangle = 3q_2^2 q_3^2 + 2q_2 q_3 q_4 + 2q_2^2 q_5 + q_7, \tag{4.4.21g}$$
$$\left\langle t^8, \mathbf{s}_q \right\rangle = q_2^7 + q_2^4 q_4 + 3q_2 q_3^3 + 3q_2^2 q_3 q_4 + q_2 q_4^2 + 2q_2 q_3 q_5 + 2q_2^2 q_6 + q_8, \tag{4.4.21h}$$
$$\left\langle t^9, \mathbf{s}_q \right\rangle = 4q_2^6 q_3 + 4q_2^3 q_3 q_4 + q_3^4 + 6q_2 q_3^2 q_4 + 3q_2^2 q_3 q_5 + 2q_2 q_4 q_5 + 2q_2 q_3 q_6 + 2q_2^2 q_7 + q_9. \tag{4.4.21i}$$

4.4.6. *Generating series of balanced binary trees.* Let us consider the bud generating system $\mathcal{B}_{\mathrm{bbt}}$ introduced in Section 4.3.5. We have

$$\chi_{a,\alpha} = \begin{cases} 1 & \text{if } (a, \alpha) = (1, 20), \\ 2 & \text{if } (a, \alpha) = (1, 11), \\ 1 & \text{if } (a, \alpha) = (2, 10), \\ 0 & \text{otherwise}, \end{cases} \tag{4.4.22}$$

and

$$\mathbf{g}_1(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2, \qquad (4.4.23a) \qquad\qquad \mathbf{g}_2(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1. \qquad (4.4.23b)$$

Since by Proposition 4.3.5, $\mathcal{B}_{\mathrm{bbt}}$ satisfies the conditions of Proposition 3.4.7, by this last proposition and (3.4.12), the generating series $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})}$ of $\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})$ satisfies $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})} = \mathbf{f}_1(t, 0)$ where

$$\mathbf{f}_1(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1 + \mathbf{f}_1\left(\mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2, \mathsf{y}_1\right). \tag{4.4.24}$$

This functional equation for the generating series of balanced binary trees is the one obtained in [BLL88, BLL97, Knu98, Gir12] by different methods. As announced in Section 3.4.2, the coefficients of $\mathbf{f}_1$ (and hence, those of $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})}$) can be computed by iteration. This consists in defining, for any $\ell \geqslant 0$, the polynomials $\mathbf{f}_1^{(\ell)}(\mathsf{y}_1, \mathsf{y}_2)$ as

$$\mathbf{f}_1^{(\ell)}(\mathsf{y}_1, \mathsf{y}_2) := \begin{cases} \mathsf{y}_1 & \text{if } \ell = 0, \\ \mathsf{y}_1 + \mathbf{f}_1^{(\ell-1)}\left(\mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2, \mathsf{y}_1\right) & \text{otherwise.} \end{cases} \tag{4.4.25}$$

Since

$$\mathbf{f}_1(\mathsf{y}_1, \mathsf{y}_2) = \lim_{\ell \to \infty} \mathbf{f}_1^{(\ell)}(\mathsf{y}_1, \mathsf{y}_2), \tag{4.4.26}$$

Equation (4.4.25) provides a way to compute the coefficients of $\mathbf{f}_1(\mathsf{y}_1, \mathsf{y}_2)$. First polynomials $\mathbf{f}_1^{(\ell)}(\mathsf{y}_1, \mathsf{y}_2)$ are

$$\mathbf{f}_1^{(0)}(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1, \qquad (4.4.27a) \qquad\qquad \mathbf{f}_1^{(1)}(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1 + \mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2, \qquad (4.4.27b)$$

$$\mathbf{f}_1^{(2)}(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1 + \mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2 + 2\mathsf{y}_1^3 + 4\mathsf{y}_1^2\mathsf{y}_2 + \mathsf{y}_1^4 + 4\mathsf{y}_1^3\mathsf{y}_2 + 4\mathsf{y}_1^2\mathsf{y}_2^2, \tag{4.4.27c}$$

$$\begin{aligned} \mathbf{f}_1^{(3)}(\mathsf{y}_1, \mathsf{y}_2) = \mathsf{y}_1 &+ \mathsf{y}_1^2 + 2\mathsf{y}_1\mathsf{y}_2 + 2\mathsf{y}_1^3 + 4\mathsf{y}_1^2\mathsf{y}_2 + \mathsf{y}_1^4 + 4\mathsf{y}_1^3\mathsf{y}_2 + 4\mathsf{y}_1^2\mathsf{y}_2^2 + 4\mathsf{y}_1^5 \\ &+ 16\mathsf{y}_1^4\mathsf{y}_2 + 16\mathsf{y}_1^3\mathsf{y}_2^2 + 6\mathsf{y}_1^6 + 28\mathsf{y}_1^5\mathsf{y}_2 + 40\mathsf{y}_1^4\mathsf{y}_2^2 + 16\mathsf{y}_1^3\mathsf{y}_2^3 + 4\mathsf{y}_1^7 + 24\mathsf{y}_1^6\mathsf{y}_2 \\ &+ 48\mathsf{y}_1^5\mathsf{y}_2^2 + 32\mathsf{y}_1^4\mathsf{y}_2^3 + \mathsf{y}_1^8 + 8\mathsf{y}_1^7\mathsf{y}_2 + 24\mathsf{y}_1^6\mathsf{y}_2^2 + 32\mathsf{y}_1^5\mathsf{y}_2^3 + 16\mathsf{y}_1^4\mathsf{y}_2^4. \end{aligned} \tag{4.4.27d}$$

Besides, since by Proposition 4.3.5, $\mathcal{B}_{\mathrm{bbt}}$ satisfies the conditions of Theorem 3.4.8, by this last theorem and (3.4.14), $\mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})}$ satisfies, for any $n \geqslant 1$,

$$\left\langle t^n, \mathbf{s}_{\mathrm{L_S}(\mathcal{B}_{\mathrm{bbt}})} \right\rangle = \left\langle \mathsf{y}_1^n \mathsf{y}_2^0, \mathbf{f} \right\rangle, \tag{4.4.28}$$

where $\mathbf{f}$ is the series satisfying, for any type $\alpha \in \mathcal{T}_{\{1,2\}}$, the recursive formula

$$\langle \mathsf{y}_1^{\alpha_1} \mathsf{y}_2^{\alpha_2}, \mathbf{f} \rangle = \delta_{\alpha,(1,0)} + \sum_{\substack{d_1, d_2, d_3 \in \mathbb{N} \\ \alpha_1 = 2d_1 + d_2 + d_3 \\ \alpha_2 = d_2}} \binom{d_1 + \alpha_2}{d_1} 2^{d_2} \left\langle \mathsf{y}_1^{d_1 + d_2} \mathsf{y}_2^{d_3}, \mathbf{f} \right\rangle. \tag{4.4.29}$$

Conclusion and perspectives

In this paper, we have presented a framework for the generation of combinatorial objects by using colored operads. The described devices for combinatorial generation, called bud generating systems, are generalizations of context-free grammars [Har78, HMU06] generating words, of regular tree grammars [GS84, CDG+07] generating planar rooted trees, and of synchronous grammars [Gir12] generating some treelike structures. We have provided tools to enumerate the objects of the languages of bud generating systems or to define new statistics on these by using formal power series on colored operads and several products on these. There are many ways to extend this work. Here follow some few further research directions.

First, the notion of rationality an recognizability in usual formal power series [Sch61, Sch63, Eil74, BR88], in series on monoids [Sak09], and in series of trees [BR82] are fundamental. For instance, a series $\mathbf{s} \in \mathbb{K}\langle\langle \mathcal{M} \rangle\rangle$ on a monoid $\mathcal{M}$ is rational if it belongs to the closure of the set $\mathbb{K}\langle \mathcal{M} \rangle$ of polynomials on $\mathcal{M}$ with respect to the addition, the multiplication, and the Kleene star operations. Equivalently, $\mathbf{s}$ is rational if there exists a $\mathbb{K}$-weighted automaton accepting it. The equivalence between these two properties for the rationality property is remarkable. We ask here for the definition of an analogous and consistent notion of rationality for series on a colored operad $\mathcal{C}$. By consistent, we mean a property of rationality for $\mathcal{C}$-series which can be defined both by a closure property of the set $\mathbb{K}\langle \mathcal{C} \rangle$ of the polynomials on $\mathcal{C}$ with respect to some operations, and, at the same time, by a acceptance property involving a notion of a $\mathbb{K}$-weighted automaton on $\mathcal{C}$. The analogous question about the definition of a notion of recognizable series on colored operads also seems worth investigating.

A second research direction fits mostly in the contexts of computer science and compression theory. A *straight-line grammar* (see for instance [ZL78, SS82, Ryt04]) is a context-free grammar with a singleton as language. There exists also the analogous natural counterpart for regular tree grammars [LM06]. One of the main interests of straight-line grammars is that they offer a way to compress a word (resp. a tree) by encoding it by a context-free grammar (resp. a regular tree grammar). A word u can potentially be represented by a context-free grammar (as the unique element of its language) with less memory than the direct representation of u, provided that u is made of several repeating factors. The analogous definition for bud generating systems could potentially be used to compress a large variety of combinatorial objects. Indeed, given a suitable monochrome operad $\mathcal{O}$ defined on the objects we want to compress, we can encode an object x of $\mathcal{O}$ by a bud generating system $\mathcal{B}$ with $\mathcal{O}$ as ground operad and such that the language (or the synchronous language) of $\mathcal{B}$ is a singleton $\{y\}$ and $\mathfrak{pru}(y) = x$. Hence, we can hope to obtain a new and efficient method to compress arbitrary combinatorial objects.

Let us finally describe a third extension of this work. *Pros* are algebraic structures which naturally generalize operads. Indeed, a pro is a set of operators with several inputs and several outputs, unlike in operads where operators have only one output (see for instance [Mar08]). Surprisingly, pros appeared earlier than operads in the literature [ML65]. It seems fruitful to translate the main definitions and constructions of this work (as *e.g.*, bud operads, bud generating systems, series on colored operads, pre-Lie and composition products of series, star operations, *etc.*) with pros instead of operads. We can expect to obtain an even more general class of grammars and obtain a more general framework for combinatorial generation.

References

[AVL62] G.M. Adelson-Velsky and E. M. Landis. An algorithm for the organization of information. *Soviet Mathematics Doklady*, 3:1259–1263, 1962. 3, 40

[BLL88] F. Bergeron, G. Labelle, and P. Leroux. Functional Equations for Data Structures. *Lect. Notes Comput. Sc.*, 294:73–80, 1988. 45

[BLL97] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial Species and Tree-like Structures.* Cambridge University Press, 1997. 33, 45

[BN98] F. Baader and T. Nipkow. *Term rewriting and all that*. Cambridge University Press, Cambridge, New York, NY, USA, 1998. 2

[Boz01] S. Bozapalidis. Context-free series on trees. *Inform. Comput.*, 169(2):186–229, 2001. 4

[BR82] J. Berstel and C. Reutenauer. Recognizable formal power series on trees. *Theor. Comput. Sci.*, 18(2):115–148, 1982. 4, 46

[BR88] J. Berstel and C. Reutenauer. *Rational series and their languages*, volume 12 of *EATCS Monographs on Theoretical Computer Science*. Springer-Verlag, Berlin, 1988. 46

[BR10] J. Berstel and C. Reutenauer. *Noncommutative rational series with applications*, volume 137. Cambridge University Press, 2010. 3

[BV73] J. M. Boardman and R. M. Vogt. Homotopy invariant algebraic structures on topological spaces. *Lect. Notes Math.*, 347, 1973. 2, 3

[CDG+07] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, C. Löding, D. Lugiez, S. Tison, and M. Tommasi. *Tree Automata Techniques and Applications*. `http://www.grappa.univ-lille3.fr/tata`, 2007. 2, 15, 45

[CG14] F. Chapoton and S. Giraudo. Enveloping operads and bicoloured noncrossing configurations. *Exp. Math.*, 23:332–349, 2014. 2

[Cha02] F. Chapoton. Rooted trees and an exponential-like series. `arXiv:math/0209104 [math.QA]`, 2002. 4, 21, 25

[Cha08] F. Chapoton. Operads and algebraic combinatorics of trees. *Sém. Lothar. Combin.*, 58, 2008. 2, 4, 20, 21, 25

[Cha09] F. Chapoton. A rooted-trees q-series lifting a one-parameter family of Lie idempotents. *Algebra & Number Theory*, 3(6):611–636, 2009. 4, 21, 25

[Cho59] N. Chomsky. On certain formal properties of grammars. *Inform. Control*, 2:137–167, 1959. 2

[CL01] F. Chapoton and M. Livernet. Pre-Lie algebras and the rooted trees operad. *Int. Math. Res. Notices*, 8:395–408, 2001. 2, 20

[CLRS09] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms*. MIT Press, Cambridge, MA, third edition, 2009. 3, 39

[CS63] N. Chomsky and M. P. Schützenberger. The algebraic theory of context-free languages. In *Computer programming and formal systems*, pages 118–161. North-Holland, Amsterdam, 1963. 2

[Eil74] S. Eilenberg. *Automata, languages, and machines. Vol. A*. Academic Press, New York, 1974. Pure and Applied Mathematics, Vol. 58. 3, 46

[FFM18] F. Fauvet, L. Foissy, and D. Manchon. Operads of finite posets. *Electron. J. Combin.*, 25(1):Paper 1.44, 29, 2018. 2

[Fra08] A. Frabetti. Groups of tree-expanded series. *J. Algebra*, 319(1):377–413, 2008. 4, 21, 25

[FS09] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009. 33, 39, 43

[Ger63] M. Gerstenhaber. The cohomology structure of an associative ring. *Ann. Math.*, 78:267–288, 1963. 20

[Gir12] S. Giraudo. Intervals of balanced binary trees in the Tamari lattice. *Theor. Comput. Sci.*, 420:1–27, 2012. 2, 16, 33, 40, 43, 45

[Gir15] S. Giraudo. Combinatorial operads from monoids. *J. Algebr. Comb.*, 41(2):493–538, 2015. 2, 34, 37, 42

[Gir16a] S. Giraudo. Operads from posets and Koszul duality. *Eur. J. Combin.*, 56C:1–32, 2016. 2

[Gir16b] S. Giraudo. Pluriassociative algebras I: The pluriassociative operad. *Adv. Appl. Math.*, 77:1–42, 2016. 34

[GS84] F. Gécseg and M. Steinby. *Tree Automata*. Akadémia Kiadó, 1984. 2nd edition (2015) available at arXiv:1509.06233 [cs.FL]. 2, 15, 45

[Har78] M. A. Harrison. *Introduction to formal language theory*. Addison-Wesley Publishing Co., Reading, Mass., 1978. 2, 14, 45

[HMU06] J. E. Hopcroft, R. Matwani, and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Pearson, 3rd edition, 2006. 2, 14, 45

[HNT05] F. Hivert, J.-C. Novelli, and J.-Y. Thibon. The algebra of binary search trees. *Theor. Comput. Sci.*, 339(1):129–165, 2005. 41

[Knu97] D. Knuth. *The Art of Computer Programming, volume 1: Fundamental Algorithms*. Addison Wesley Longman, Redwood City, CA, USA, 3rd edition, 1997. 7

[Knu98] D. Knuth. *The Art of Computer Programming, Sorting and Searching*, volume 3. Addison Wesley, 2nd edition, 1998. 3, 9, 24, 39, 45

[LM06] M. Lohrey and S. Maneth. The complexity of tree automata and XPath on grammar-compressed trees. *Theor. Comput. Sci.*, 363(2):196–210, 2006. 46

[LN13] J.-L. Loday and N. M. Nikolov. Operadic construction of the renormalization group. In *Lie theory and its applications in physics*, volume 36 of *Springer Proc. Math. Stat.*, pages 191–211. Springer, Tokyo, 2013. 4, 21, 25

[Lod96] J.-L. Loday. La renaissance des opérades. *Astérisque*, 237:Exp. No. 792, 3, 47–74, 1996. Séminaire Bourbaki, Vol. 1994/95. 2

[Lod01] J.-L. Loday. Dialgebras. In *Dialgebras and related operads*, volume 1763 of *Lecture Notes in Math.*, pages 7–66. Springer, Berlin, 2001. 34

[LV12] J.-L. Loday and B. Vallette. *Algebraic Operads*, volume 346 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2012. 2, 21, 25

[Man11] D. Manchon. A short survey on pre-Lie algebras. In *Noncommutative geometry and physics: renormalisation, motives, index theory*, ESI Lect. Math. Phys., pages 89–102. Eur. Math. Soc., Zürich, 2011. 20

[Mar08] M. Markl. Operads and PROPs. In *Handbook of Algebra*, volume 5, pages 87–140. Elsevier/North-Holland, Amsterdam, 2008. 2, 46

[May72] J. P. May. *The geometry of iterated loop spaces*. Springer-Verlag, Berlin-New York, 1972. Lectures Notes in Mathematics, Vol. 271. 2

[Mén15] M. A. Méndez. *Set operads in combinatorics and computer science*. SpringerBriefs in Mathematics. Springer, Cham, 2015. 2

[ML65] S. Mac Lane. Categorical algebra. *Bull. Amer. Math. Soc.*, 71:40–106, 1965. 46

[MPRS79] R. E. Miller, N. Pippenger, A. L. Rosenberg, and L. Snyder. Optimal 2,3-Trees. *SIAM J. Comput.*, 8:42–59, 1979. 3, 39, 44

[Odl82] A. M. Odlyzko. Periodic Oscillations of Coefficients of Power Series That Satisfy Functional Equations. *Adv. in Math.*, 44:180–205, 1982. 39, 43

[Ryt04] W. Rytter. Grammar compression, Z-encodings, and string algorithms with implicit input. In *Automata, languages and programming*, volume 3142 of *Lect. Notes Comput. Sci.*, pages 15–27. Springer, Berlin, 2004. 46

[Sak09] J. Sakarovitch. *Elements of Automata Theory*. Cambridge University Press, 2009. 3, 17, 21, 46

[Sch61] M. P. Schützenberger. On the definition of a family of automata. *Inform. Control*, 4:245–270, 1961. 46

[Sch63] M. P. Schützenberger. Certain elementary families of automata. In *Proc. Sympos. Math. Theory of Automata (New York, 1962)*, pages 139–153. Polytechnic Press of Polytechnic Inst. of Brooklyn, Brooklyn, New York, 1963. 46

[Slo] N. J. A. Sloane. The On-Line Encyclopedia of Integer Sequences. `http://www.research.att.com/~njas/sequences/`. 38, 40, 41, 42, 43

[SS78] A. Salomaa and M. Soittola. *Automata-theoretic aspects of formal power series*. Springer-Verlag, New York-Heidelberg, 1978. Texts and Monographs in Computer Science. 3

[SS82] J. A. Storer and T. G. Szymanski. Data compression via textual substitution. *J. Assoc. Comput. Mach.*, 29(4):928–951, 1982. 46

[vdL04] P. van der Laan. *Operads. Hopf algebras and coloured Koszul duality*. PhD thesis, Universiteit Utrecht, 2004. 4, 20, 21, 25

[Vin63] È. B. Vinberg. The theory of homogeneous convex cones. *Trudy Moskovskogo Matematicheskogo Obshchestva*, 12:303–358, 1963. 20

[Yau16] D. Yau. *Colored Operads*. Graduate Studies in Mathematics. American Mathematical Society, 2016. 3

[Zin12] G. W. Zinbiel. Encyclopedia of types of algebras 2010. In *Operads and universal algebra*, volume 9 of *Nankai Ser. Pure Appl. Math. Theoret. Phys.*, pages 217–297. World Sci. Publ., Hackensack, NJ, 2012. 2

[ZL78] J. Ziv and A. Lempel. Compression of individual sequences via variable-rate coding. *IEEE T. Inform. Theory*, 24(5):530–536, 1978. 46

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

Email address: `samuele.giraudo@u-pem.fr`