

Some Combinatorial Operators in Language Theory

J.-G. Luque, L. Mignot and F. Nicart

May 1, 2014

Contents

1	Introduction	1
2	What are operads?	2
2.1	Operadic structures	3
2.2	Free operad, morphisms, suboperads, quotients <i>etc.</i>	5
3	Multitildes operad	6
3.1	Extended to multityle regular expression	6
3.2	The operators \ggg and $\circ\rightarrow$	8
3.3	The Operad of multitildes	10
4	Actions on languages	11
4.1	An operad of boolean vectors and its action on languages	11
4.2	Action of \mathcal{T} on languages	12
5	\mathcal{T}, RAS and POSet	14
5.1	From \mathcal{T} to RAS^{\leq}	14
5.2	From RAS^{\leq} to POSet^{\leq}	15
5.3	Pseudotransitive tildes	16
6	Consequences and perspectives	17
6.1	Enumeration	17
6.2	Finite languages	18
6.3	Regular languages	20
7	Conclusion	20

Abstract

Multitildes are regular operators that were introduced by Caron *et al.* in order to increase the number of Glushkov automata. In this paper, we study the family of the multityle operators from an algebraic point of view using the notion of operad. This leads to a combinatorial description of already known results as well as new results on compositions, actions and enumerations.

1 Introduction

Regular expressions have been studied from several years since they have numerous applications: pattern matching, compilation, verification, bio-informatics *etc.*. From a well known result (Kleene [10]), regular languages can be represented by both automata and regular expressions. From an

expression, it is possible to compute an automaton whose number of states is a linear function of the alphabetical width (*i.e.* the number of occurrences of alphabet symbols) [16, 9, 1, 5]. In the opposite direction, there exists no construction providing the linearity. For instance, Ehrenfeucht and Zeiger [7] showed a one parameter family of automata whose shortest equivalent regular expressions have a width exponentially growing with the numbers of states. However this property occurs when the automaton is the Gluskov automaton of an expression. Note the characterization of such an automaton is due to Caron and Ziadi [4].

Multitilde operators have been introduced by Caron *et al.* [3] in the aim to increase expressiveness of expressions for a bounded length. The study of the equivalence of two multitilde expressions leads naturally to a notion of composition. On the other hand, there is an algebraic way to represent the compositions of operators, namely operads. The aim of this paper is to describe some properties of multitildes in terms of operads.

Usually, operads are used to encode types of algebras by describing the universal operations which act on the elements of any algebras of a given type together with the ways of composing them. To be more concise, an operad is given by a set of symbols (the operations) and a composition law which satisfies some rules such as associativity. Such a study of the compositions of operations appeared earlier in the work of M. Lazard [11] and was referred as *analyseurs*. The word *operad* is the contraction of *operations* and *monad*. This terminology first appeared in the field of algebraic topology in a paper of May about the geometry of iterated loop spaces [15]. The notion arose simultaneously in Kelly's categorical works on coherence and the definitions used in [15] have been constructed after conversations between the two mathematicians.

In the 1990's, the interest in the topic was renewed with the works of Ginzburg and Kapranov on the Koszul duality [8]. One of the recent developments of this theory deals with Hopf algebras and combinatorics [6].

Readers interested by the history of this notion and its recent development can refer to [14, 12, 13].

Several materials about operads are recalled in Section 2. In his PhD thesis [17], one of the authors introduced a composition on multitildes. We show (Section 3) that this confers an operadic structure to the set of multitildes. Throughout the paper, we describe derived structures (isomorphic operads, suboperads and quotient operads) which allow us to model several behaviors and properties of the multitilde operators. For instance, we define in Section 4 an operadic structure on boolean vectors and use it to describe the action of multitildes on languages. In Section 5, we introduce an operadic structure on POSets in the aim to formalize the equivalence of the actions of two multitildes. In Section 6, we show that the representation by POSets is optimal in the sense that two different operators have different actions on k -tuples of languages. As a consequence, we enumerate the different operations which can be encoded by a multitilde. Finally, we give applications to the representation of finite languages and also regular languages.

2 What are operads?

Operad theory has been developed in the aim to study prototypical algebras that model classical properties such as commutativity, associativity *etc.*. In most cases, operads are considered to study algebras with an underlying vector space but the structure can be more generally used to study \mathbb{K} -modules on monoid (\mathbb{K} semiring). In the most general context, operads are described in terms of category theory. In our case, since we apply the operad theory to the theory of languages, we will set $\mathbb{K} = \mathbb{B}$, the boolean semiring. We have adapted the definitions of this section to the boolean context to give a simplified version of the theory but most of them admits well-known generalizations. More precisely, we describe Set-operads that are operads whose underlying graded spaces are simply sets.

2.1 Operadic structures

Operads are algebraic structures which mimic the composition of n -ary operators. More explicitly, the construction starts with an underlying graded set $\mathfrak{P} = \bigcup_{n \in \mathbb{N}} \mathfrak{P}(n)$; elements of $\mathfrak{P}(n)$ are called n -ary operations. This set is endowed with functions

$$\circ : \mathfrak{P}(n) \times \mathfrak{P}(k_1) \times \cdots \times \mathfrak{P}(k_n) \rightarrow \mathfrak{P}(k_1 + \cdots + k_n)$$

called compositions and has a special element $\mathbf{1} \in \mathfrak{P}(1)$ called identity. This confers to \mathfrak{P} a structure of multicategory with one object [12]; more precisely the operations of the operad are the morphisms of the multicategory.

Furthermore, the compositions satisfy the two rules:

1. Associativity:

$$\mathbf{p} \circ (\mathbf{p}_1 \circ (\mathbf{p}_{1,1}, \dots, \mathbf{p}_{1,k_1}), \dots, \mathbf{p}_n \circ (\mathbf{p}_{n,1}, \dots, \mathbf{p}_{n,k_n})) = (\mathbf{p} \circ (\mathbf{p}_1, \dots, \mathbf{p}_n)) \circ (\mathbf{p}_{1,1}, \dots, \mathbf{p}_{1,k_1}, \dots, \mathbf{p}_{n,1}, \dots, \mathbf{p}_{n,k_n}) \quad (1)$$

2. Identity:

$$\mathbf{p} \circ (\mathbf{1}, \dots, \mathbf{1}) = \mathbf{1} \circ \mathbf{p} = \mathbf{p}. \quad (2)$$

The structure of operad is easier to manipulate if we introduce partial composition operations \circ_i which split the composition. These operations are defined by:

$$\begin{aligned} \circ_i : \mathfrak{P}(m) \times \mathfrak{P}(n) &\rightarrow \mathfrak{P}(m+n-1) \\ (\mathbf{p}_1, \mathbf{p}_2) &\rightarrow \mathbf{p}_1 \circ_i \mathbf{p}_2 := \mathbf{p}_1 \circ \left(\underbrace{\mathbf{1}, \dots, \mathbf{1}}_{i-1 \times}, \mathbf{p}_2, \underbrace{\mathbf{1}, \dots, \mathbf{1}}_{m-i-1 \times} \right) \end{aligned} \quad (3)$$

when $1 \leq i \leq n$.

Let $\mathbf{p}_1 \in \mathfrak{P}(m)$, $\mathbf{p}_2 \in \mathfrak{P}(n)$ and $\mathbf{p}_3 \in \mathfrak{P}(q)$. The partial compositions satisfy the two associative rules which are deduced from the definition (3) and the associativity of the compositions (1):

1. Associativity 1:

If $1 \leq j < i \leq n$ then

$$(\mathbf{p}_1 \circ_i \mathbf{p}_2) \circ_j \mathbf{p}_3 = (\mathbf{p}_1 \circ_j \mathbf{p}_3) \circ_{i+q-1} \mathbf{p}_2 \quad (4)$$

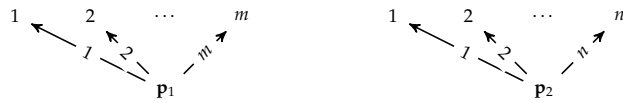
2. Associativity 2:

If $i \leq n$ then

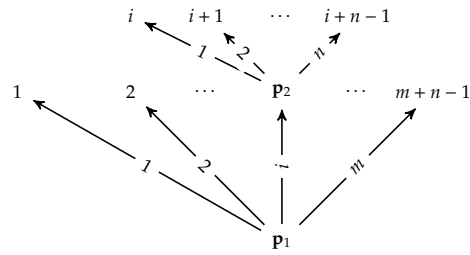
$$\mathbf{p}_1 \circ_j (\mathbf{p}_2 \circ_i \mathbf{p}_3) = (\mathbf{p}_1 \circ_j \mathbf{p}_2) \circ_{i+j-1} \mathbf{p}_3. \quad (5)$$

The composition is graphically interpreted by grafting trees together; a n -array operation is represented by a n -ary tree and a composition \circ_i consists in grafting the root of a tree onto the i th leaf of another tree, the resulting tree is associated to an element of \mathfrak{P} .

Consider two operations $\mathbf{p}_1 \in \mathfrak{P}(m)$ and $\mathbf{p}_2 \in \mathfrak{P}(n)$:

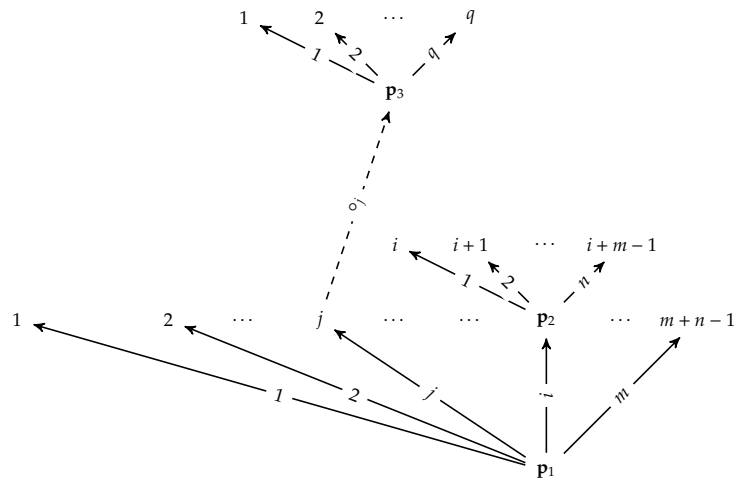


The composition $\mathbf{p}_1 \circ_i \mathbf{p}_2$ is represented by the tree:

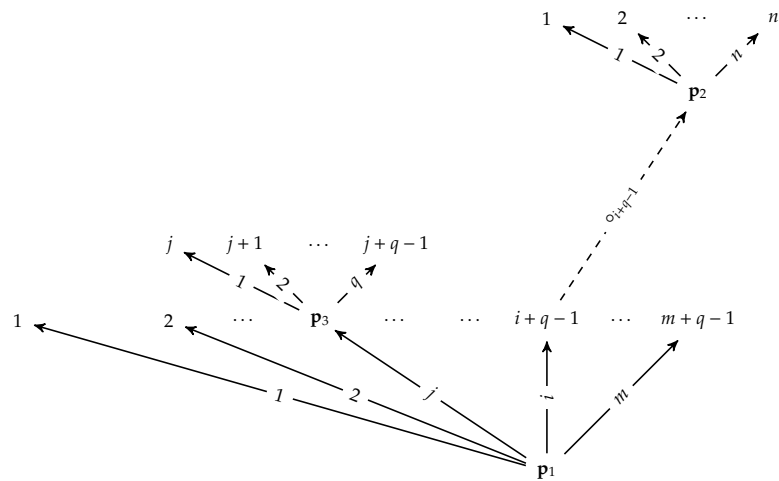


Now the two associativity rules are easily understood in terms of trees:

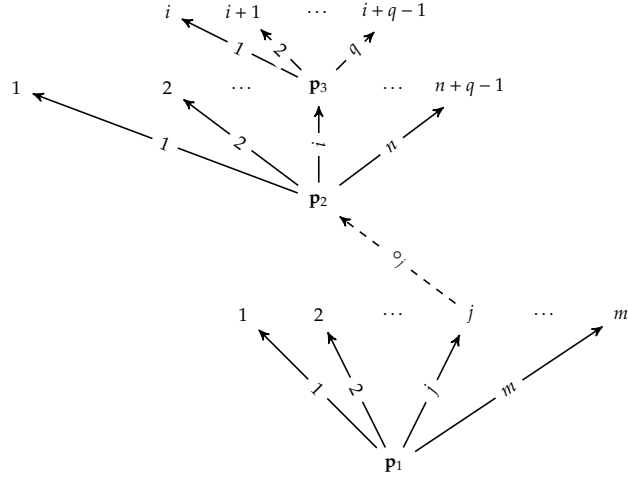
1. Associativity 1: The left hand side of (4) gives



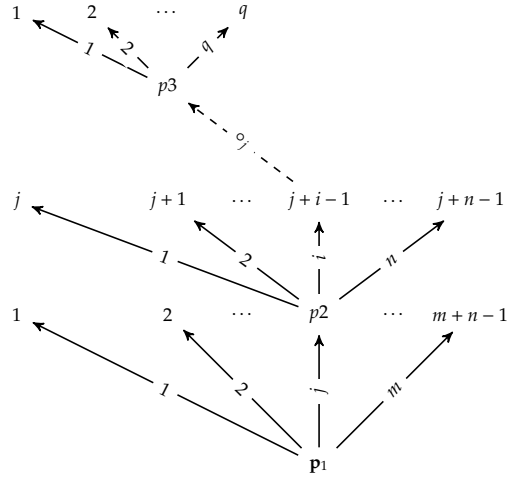
whilst the right hand side reads



2. Associativity 2: Drawing the left hand side of (5), we find



and the right hand side of (5) gives



In short, the compositions satisfy the same branching rules than the trees together with a correct relabeling of the leaves.

Let us give one of the simplest example of operads. Consider a set S and the set of the functions $\text{Map}_S(n)$ from S^n to S . The set $\text{Map}_S := \bigcup_{n \in \mathbb{N}} \text{Map}_S(n)$ endowed with the classical composition defined by

$$(f \circ (g_1, \dots, g_n))(s_{11}, \dots, s_{1k_1}, \dots, s_{n1}, \dots, s_{nk_n}) := f(g_1(s_{11}, \dots, s_{1k_1}), \dots, g_n(s_{n1}, \dots, s_{nk_n}))$$

for each $f \in \text{Map}_S(n)$, $g_i \in \text{Map}_S(k_i)$, $s_{ij} \in S$, is an operad.

2.2 Free operad, morphisms, suboperads, quotients etc.

The definition of the morphism induces the existence of operads with universal properties called *free operads*. Let $G = (G_k)_k$ be a collection of sets, the set $\text{Free}_G(n)$ is the set of planar rooted trees with n leaves with labeled nodes where nodes with k branches are labeled by the elements of G_k . The free operad on G is obtained by endowing the set $\text{Free}_G = \bigcup_{n \in \mathbb{N}} \text{Free}_G(n)$ with the composition $p_1 \circ_i p_2$

which consists in grafting the i th leaf of \mathbf{p}_1 with the root of \mathbf{p}_2 . Note that Free_G contains a copy of G which is the set of the trees with only one inner node (the root) labeled by elements of G ; for simplicity we will identify it with G . The universality means that for any map $\varphi : G \rightarrow \mathfrak{F}$ it exists a unique morphism of operad $\phi : \text{Free}_G \rightarrow \mathfrak{F}$ such that $\phi(g) = \varphi(g)$ for each $g \in G$.

Consider the equivalence relation \equiv_ϕ on Free_G defined by $\mathbf{p}_1 \equiv_\phi \mathbf{p}_2$ if and only if $\phi(\mathbf{p}_1) = \phi(\mathbf{p}_2)$. Obviously this relation is compatible with the composition in the sense that $\mathbf{p}_1 \equiv_\phi \mathbf{p}'_1$ and $\mathbf{p}_2 \equiv_\phi \mathbf{p}'_2$ implies $\mathbf{p}_1 \circ_i \mathbf{p}_2 \equiv_\phi \mathbf{p}'_1 \circ_i \mathbf{p}'_2$. Hence, the operad \mathfrak{F} can be defined as a *quotient* (up to an isomorphism) of the free operad: $\mathfrak{F} = \text{Free}_G / \equiv_\phi$. More generally, when an equivalence relation \equiv on the elements of an operad \mathfrak{F} is compatible with the composition, this induces a structure of operad on the set \mathfrak{F} / \equiv ; the resulting operad is called the *quotient* of \mathfrak{F} by \equiv .

The dual notion of quotient is the notion of suboperad whose definition is very classical: a suboperad of an operad \mathfrak{F} is a subset of \mathfrak{F} which contains $\mathbf{1}$ and is stable by composition.

Consider a set \mathbf{S} together with an action of an operad \mathfrak{F} . That is: for each $\mathbf{p} \in \mathfrak{F}(n)$ we define a map $\mathbf{p} : \mathbf{S}^n \rightarrow \mathbf{S}$. We will say that \mathbf{S} is a \mathfrak{F} -module if the action of \mathfrak{F} is compatible with the composition in the following sense: for each $\mathbf{p}_1 \in \mathfrak{F}(m)$, $\mathbf{p}_2 \in \mathfrak{F}(n)$, $1 \leq i \leq m$, $s_1, \dots, s_{m+n-1} \in \mathbf{S}$ one has:

$$\mathbf{p}_1(s_1, \dots, s_{i-1}, \mathbf{p}_2(s_i, \dots, s_{i+n-1}), s_{i+n}, \dots, s_{m+n-1}) = (\mathbf{p}_1 \circ_i \mathbf{p}_2)(s_1, \dots, s_{m+n-1}),$$

i.e. there is a morphism of operads from $\text{Map}_{\mathbf{S}}$ to \mathfrak{F} .

3 Multitildes operad

Multitildes operators have been introduced by Caron *et al.* [3] in order to increase the number of regular languages represented by an expression of a fixed width. The only operation involved in the computation of the languages they denote is the addition of the empty word in several catenation factors in a catenation product. This operation is not very interesting as long as it is considered as unary, since for every regular expression E , the addition of the empty word in $L(E)$ is denoted by the expression $E + \varepsilon$, the width of which is the same as E . Multitildes extend it to k -ary operators allowing to define new regular expressions, the EMTREs (Extended to Multitilde Regular Expression) [3]. One of the main interest of multitilde operators is that, for any simple regular expression with no star, there exists an equivalent EMTRE with only one k -ary operator which is a multitilde. We will see that this equivalence involves a natural composition which defines a structure of operad.

In this section, we first recall the main definitions and results about multitildes then we describe in more details the operad structure acting on languages.

3.1 Extended to multitilde regular expression

Let Σ be an alphabet and ε be the empty word. A *regular expression* E over Σ is inductively defined by $E = \emptyset$, $E = \varepsilon$, $E = a$, $E = (F \cdot G)$, $E = (F + G)$, $E = (F^*)$ where $a \in \Sigma$ and F, G are regular expressions. The *language denoted by* the expression E is inductively defined by $L(\emptyset) = \emptyset$, $L(\varepsilon) = \{\varepsilon\}$, $L(a) = \{a\}$, $L(F + G) = L(F) \cup L(G)$, $L(F \cdot G) = L(F) \cdot L(G)$, $L(F^*) = L(F)^*$ where $a \in \Sigma$ and F, G are regular expressions.

Let i, j, n be three positive integers. Let (E_1, \dots, E_n) be a list of n expressions. The catenation $E_i \cdot E_{i+1} \cdots E_j$ is denoted by $E_{i..j}$ and for convenience the language $L(E_{i..j})$ equals $\{\varepsilon\}$ when $i > j$. A list $(E_i, E_{i+1}, \dots, E_j)$ is denoted by $E_{1..n}$. Given a set S , we denote by $\#S$ the number of elements of S . Suppose that $i \leq j$. We set $\llbracket i, j \rrbracket = \{i, i+1, \dots, j-1, j\}$, $\llbracket i, f \rrbracket_{\leq}^2 = \{(k, k') \mid k, k' \in \llbracket i, f \rrbracket \text{ and } k \leq k'\}$ and $\mathcal{S}_n = \mathcal{P}(\llbracket 1, n \rrbracket_{\leq}^2)$ the set of the subsets of $\llbracket 1, n \rrbracket_{\leq}^2$. The set of indices of $S \in \mathcal{S}_n$ is $I_S = \llbracket 1, \#S \rrbracket$ and it holds $S = \{(i_k, f_k)\}_{k \in I_S}$, with for all $k \in I_S$, $(i_k, f_k) \in \llbracket 1, n \rrbracket_{\leq}^2$.

A multitilde of arity n is a formal symbol $\widetilde{\sim}_T$ where $T \subset \mathcal{S}_n$. Remark for convenience, we use the same symbol to denote several operations with different arities. For instance, the multitilde $\widetilde{\sim}_{\{(1,3)\}}$

should be n -ary for any $n \geq 3$. These symbols together with the regular operations define a new family of expressions.

Definition 1 ([3]). An Extended to Multitilde Regular Expression (*EMTRE*) over an alphabet Σ is inductively defined by:

$$\begin{aligned} E &= \emptyset, & E &= a, \\ E &= (F + G), & E &= (F \cdot G), & E &= (F^*), \\ E &= (\sim\sim_T(E_{1,n})), \end{aligned}$$

with $a \in \Sigma$, F and G are two *EMTREs*, $E_{1,n}$ a list of *EMTREs* and T a list in \mathcal{S}_n .

Now let us recall how to extend the notion of languages denoted by regular expressions to *EMTREs*, by defining the action of multitilde operators.

A set $S \in \mathcal{S}_n$ is said *free* if and only if for all $(i, f), (i', f')$ in S such that $(i, f) \neq (i', f')$, the condition $\llbracket i, f \rrbracket \cap \llbracket i', f' \rrbracket = \emptyset$ holds. Let $S \in \mathcal{S}_n$, we denote by $\mathcal{F}(S)$ the set of all free subsets of S .

Let L_1, \dots, L_n be a list of languages over an alphabet Σ and let T be a list in \mathcal{S}_n . We set $\mathcal{W}_T(L_1, \dots, L_n) = L'_1 \cdots L'_n$ where

$$L'_k = \begin{cases} \{\varepsilon\} & \text{if } k \in \bigcup_{(i,f) \in T} \llbracket i, f \rrbracket \\ L_k & \text{otherwise.} \end{cases}$$

Definition 2. Let $L_{1,n}$ be a list of languages over an alphabet Σ . Let T be a list in \mathcal{S}_n . The language $\sim\sim_T(L_{1,n})$ is defined by:

$$\sim\sim_T(L_{1,n}) := \bigcup_{S \in \mathcal{F}(T)} \mathcal{W}_S(L_{1,n}).$$

Let $E_{1,n}$ be a list of *EMTREs*, the language denoted by $\sim\sim_T(E_{1,n})$ where $T \in \mathcal{S}_n$ is inductively defined by

$$L(\sim\sim_T(E_{1,n})) := \sim\sim_T(L(E_1), \dots, L(E_n)).$$

As usual two *EMTREs* E and F are said *equivalent* if they denote the same language; we write $E \equiv F$.

Note all these definitions are slight rewordings to those given in [3].

Every regular expression E that does not use any star operator can be turned into an equivalent expression $F = \sim\sim_T(e_1, \dots, e_n)$ where $\sim\sim_T(\in) \mathcal{S}_n$ and $\forall k \in \llbracket 1, n \rrbracket, e_k \in \Sigma \cup \{\emptyset\}$. Conversion formulas are given Table 1.

$$\begin{array}{ll} \emptyset \equiv \sim\sim_{\emptyset}(\emptyset) & E'_1 + E'_2 \equiv \sim\sim_{(1,2),(2,3)}(E'_1, \emptyset, E'_2) \\ \varepsilon \equiv \sim\sim_{(1,1)}(\emptyset) & E'_1 \cdot E'_2 \equiv \sim\sim_{\emptyset}(E'_1, E'_2) \\ a \equiv \sim\sim_{\emptyset}(a) & \end{array}$$

Table 1: Multitildes conversion.

One of the authors has shown in [17] how to compose multitildes operators preserving languages: $L(\sim\sim_T(E_1, \dots, E_{k-1}, \sim\sim_{T'}(E'_1, \dots, E'_{k'}), E_{k+1}, \dots, E_n)) = L(\sim\sim_{T \circ_k T'}(E_1, \dots, E_{k-1}, E'_1, \dots, E'_{k'}, E_{k+1}, \dots, E_n))$.

We do not recall the original definition here but we describe this operation using new operators, namely \ggg and $\circ \rightarrow$ in the sequel of this section. We show that this composition endows the set of multitildes with a structure of operad.

3.2 The operators \ggg and $\circ\rightarrow$

The first operators to be defined are the \ggg operators, parametrized by any integer k . Any operator \ggg^k will increase both elements of a couple (x, y) of integers by k :

$$\ggg^k(x, y) := (x + k, y + k).$$

These operators are commutative ones:

$$\ggg^k(\ggg^\ell(x, y)) := \ggg^\ell(\ggg^k(x, y)) \quad (6)$$

Indeed, both of the expressions result in the same couple:

$$\ggg^k(\ggg^\ell(x, y)) = (x + k + \ell, y + k + \ell) = \ggg^\ell(\ggg^k(x, y)) \quad (7)$$

The second family of operators is the set of $\circ\rightarrow^{k,n}$ operators for any integers n, k , which shift elements in couples by inserting n elements in position k (see Example 1 for details):

$$\circ\rightarrow^{k,n}(x, y) := \begin{cases} (x, y) & \text{if } y < k, \\ (x, y + n - 1) & \text{if } x \leq k \leq y, \\ \ggg^{n-1}(x, y) & \text{otherwise.} \end{cases}$$

Example 1. Let n be a positive integer, and $x \leq y$ be two integers in $\llbracket 1, n \rrbracket$. The operators $\circ\rightarrow^{k,n}$ transforms the couple (x, y) performing an insertion of n elements at position k presented in Figure 1. After an insertion of 6 elements in position 5 (using the operator $\circ\rightarrow^{5,6}$), couples $(1, 3)$, $(3, 7)$ and $(7, 8)$ are respectively transformed into $(1, 3)$, $(3, 12)$ and $(12, 13)$.

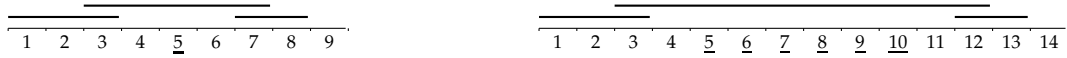


Figure 1: The $\circ\rightarrow^{5,6}$ operator.

For convenience, definition of $\circ\rightarrow^{k,n}$ operators can be rewritten using the characteristic function $\mathbb{1}_E(x)$ defined for any integer set E and any integer x by:

$$\mathbb{1}_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Since we will only use interval, we will denote by $\mathbb{1}_k(x)$ for any integer k the function $\mathbb{1}_{\llbracket k, \infty \rrbracket}(x)$.

Lemma 1. Let k, n, x, y be four integers. Then:

$$\circ\rightarrow^{k,n}(x, y) = (x + \mathbb{1}_{k+1}(x) \times (n - 1), y + \mathbb{1}_k(y) \times (n - 1)).$$

Next lemma presents the commutation of \ggg and $\circ\rightarrow$.

Lemma 2. Let x, y, i, m, n be five integers with $x \leq y$. Then:

$$\ggg^n(\circ\rightarrow^{i,m}(x, y)) = \circ\rightarrow^{i+n,m}(\ggg^n(x, y)).$$

Proof. From formulas for \ggg and $\circ\rightarrow$ operators:

$$\begin{aligned} \ggg^n(\circ\rightarrow^{i,m}(x, y)) &= \ggg^n(x + \mathbb{1}_{i+1}(x) \times (m - 1), y + \mathbb{1}_i(y) \times (m - 1)) \\ &= (x + n + \mathbb{1}_{i+1}(x) \times (m - 1), y + n + \mathbb{1}_i(y) \times (m - 1)) \\ &= (x + n + \mathbb{1}_{i+n+1}(x + n) \times (m - 1), y + n + \mathbb{1}_{i+n}(y + n) \times (m - 1)) \\ &= \circ\rightarrow^{i+n,m}(x + n, y + n) \\ &= \circ\rightarrow^{i+n,m}(\ggg^n(x, y)) \end{aligned}$$

□

The $\circ\rightarrow$ operators commute in a particular way, described as follows:

Lemma 3. Let x, y, m, n, k, ℓ be six integers with $x \leq y$ and $k < \ell$. Then:

$$\overset{k,m}{\circ\rightarrow}(\overset{\ell,n}{\circ\rightarrow}(x, y)) = \overset{\ell+m-1,n}{\circ\rightarrow}(\overset{k,m}{\circ\rightarrow}(x, y)).$$

Proof. From formulas for $\circ\rightarrow$ operators:

$$\begin{aligned} \overset{k,m}{\circ\rightarrow}(\overset{\ell,n}{\circ\rightarrow}(x, y)) &= \overset{k,m}{\circ\rightarrow}(x + \mathbb{1}_{\ell+1}(x) \times (n-1), y + \mathbb{1}_{\ell}(y) \times (n-1)) \\ &= (x', y') \end{aligned}$$

$$\begin{aligned} \text{with } x' &= x + \mathbb{1}_{\ell+1}(x) \times (n-1) + \mathbb{1}_{k+1}(x + \mathbb{1}_{\ell+1}(x) \times (n-1)) \times (m-1) \\ \text{and } y' &= y + \mathbb{1}_{\ell}(y) \times (n-1) + \mathbb{1}_k(y + \mathbb{1}_{\ell}(y) \times (n-1)) \times (m-1). \end{aligned}$$

Since $k < \ell$, then

$$\begin{aligned} \mathbb{1}_{\ell+1}(x) &= \mathbb{1}_{\ell+m}(x + m - 1) = \mathbb{1}_{\ell+m}(x + \mathbb{1}_{k+1}(x) \times (m-1)) \\ \text{and } \mathbb{1}_{k+1}(x) &= \mathbb{1}_{k+1}(x + \mathbb{1}_{\ell+1}(x) \times (n-1)). \end{aligned}$$

Consequently,

$$\begin{aligned} x' &= x + \mathbb{1}_{k+1}(x) \times (m-1) + \mathbb{1}_{\ell+m}(x + \mathbb{1}_{k+1}(x) \times (m-1)) \times (n-1), \\ \text{and } y' &= y + \mathbb{1}_k(x) \times (m-1) + \mathbb{1}_{\ell+m-1}(y + \mathbb{1}_k(y) \times (m-1)) \times (n-1). \end{aligned}$$

Finally,

$$\begin{aligned} (x', y') &= \overset{\ell+m-1,n}{\circ\rightarrow}((x + \mathbb{1}_{k+1}(x) \times (m-1), y + \mathbb{1}_k(y) \times (m-1))) \\ &= \overset{\ell+m-1,n}{\circ\rightarrow}(\overset{k,m}{\circ\rightarrow}(x, y)). \end{aligned}$$

□

Finally, another combination of $\circ\rightarrow$ operators:

Lemma 4. Let i, n, k, m, x, y be six integers such that $x \leq y$ and $0 \leq i < n$. Then:

$$\overset{k+i,m}{\circ\rightarrow}(\overset{k,n}{\circ\rightarrow}(x, y)) = \overset{k,m+n-1}{\circ\rightarrow}(x, y).$$

Proof. From formulas for $\circ\rightarrow$ operators:

$$\begin{aligned} \overset{k+i,m}{\circ\rightarrow}(\overset{k,n}{\circ\rightarrow}((x, y))) &= \overset{k+i,m}{\circ\rightarrow}(x + \mathbb{1}_{k+1}(x) \times (n-1), y + \mathbb{1}_k(y) \times (n-1)) \\ &= (x', y') \end{aligned}$$

$$\begin{aligned} \text{with } x' &= x + \mathbb{1}_{k+1}(x) \times (n-1) + \mathbb{1}_{k+i+1}(x + \mathbb{1}_{k+1}(x) \times (n-1)) \times (m-1) \\ \text{and } y' &= y + \mathbb{1}_k(y) \times (n-1) + \mathbb{1}_{k+i}(y + \mathbb{1}_k(y) \times (n-1)) \times (m-1). \end{aligned}$$

Since $0 \leq i < n$, then

$$\mathbb{1}_{k+1}(x) = \mathbb{1}_{k+i+1}(x + \mathbb{1}_{k+1}(x) \times (n-1)).$$

Finally,

$$\begin{aligned} (x', y') &= (x + \mathbb{1}_{k+1}(x) \times (m+n-1), y + \mathbb{1}_k(y) \times (m+n-1)) \\ &= \overset{k,m+n-1}{\circ\rightarrow}(x, y). \end{aligned}$$

□

The definition domain of $\circ\rightarrow$ operators can be extended to sets of couples by the following way. Let $T \subset \llbracket 1, n \rrbracket_{\leq}^2$ for any integer n :

$$\overset{k,n}{\circ\rightarrow}(T) = \{\overset{k,n}{\circ\rightarrow}((x, y)) : (x, y) \in T\}, \ggg^k(T) = \{\ggg^k((x, y)) : (x, y) \in T\}.$$

From an algebraic point of view, operators $\circ\rightarrow$ and \ggg are linearly extended. Relations are still valid:

Proposition 1. Let k, ℓ, n, i, m be five integers, and let T be a set of couples (x, y) with $x \leq y$. Then:

1. $\ggg^k(\ggg^\ell(T)) := \ggg^\ell(\ggg^k(T)),$
2. $\ggg^n(\overset{i,m}{\circ\rightarrow}(T)) = \overset{i+n,m}{\circ\rightarrow}(\ggg^n(T)),$
3. $\overset{k,m}{\circ\rightarrow}(\overset{\ell,n}{\circ\rightarrow}(T)) = \overset{\ell+m-1,n}{\circ\rightarrow}(\overset{k,m}{\circ\rightarrow}(T))$ if $k \leq \ell$.

$$4. \quad \circlearrowright^{k+i,m} (\circlearrowright^{k,n} (T)) = \circlearrowright^{k,m+n-1} (T) \text{ if } 0 \leq i < n.$$

3.3 The Operad of multitildes

Multitildes are k -ary operators for any integer $k > 0$ represented by a subset of $\llbracket 1, n \rrbracket_{\leq}^2$. The set of k -ary multitildes is denoted by \mathcal{T}_k . The set of multitildes (i.e. $\bigcup_{k \in \mathbb{N}} \mathcal{T}_k$) is denoted by \mathcal{T} .

Let k, n be two integers such that $1 \leq k \leq n$. The partial composition \circ_k of multitildes can be expressed from \circlearrowright and \ggg operators as follows:

$$\circ_k : \begin{cases} \mathcal{T}_m \otimes \mathcal{T}_n & \rightarrow \mathcal{T}_{n+m-1} \\ T_1 \circ_k T_2 & = \circlearrowright^{k,n} (T_1) \cup \ggg^{k-1} (T_2) \end{cases}$$

Note this composition is straightforwardly the same than those defined by one of the authors in [17]. It satisfies the property of commutativity and associativity defining an operad.

Proposition 2 (Commutativity). *Let m, n, p, k, ℓ be five integers such that $1 \leq k < \ell \leq m$. Let $T_1 \in \mathcal{T}_m$, $T_2 \in \mathcal{T}_n$ and $T_3 \in \mathcal{T}_p$ be three multitildes. Then:*

$$(T_1 \circ_{\ell} T_2) \circ_k T_3 = (T_1 \circ_k T_3) \circ_{\ell+p-1} T_2.$$

Proof. From the definition of composition:

$$(T_1 \circ_{\ell} T_2) \circ_k T_3 = \circlearrowright^{k,p} (\circlearrowright^{\ell,n} (T_1)) \cup \circlearrowright^{k,p} (\ggg^{\ell-1} (T_2)) \cup \ggg^{k-1} (T_3).$$

Following Proposition 1 on the sets of couples computed in the right member:

1. From relation (3): $\circlearrowright^{k,p} (\circlearrowright^{\ell,n} (T_1)) = \circlearrowright^{\ell+p-1,n} (\circlearrowright^{k,p} (T_1))$;
2. Since $k < \ell$, from relation (2):

$$\circlearrowright^{k,p} (\ggg^{\ell-1} (T_2)) = \ggg^{\ell-1} (\circlearrowright^{k-\ell+1,p} (T_2)) = \ggg^{\ell+p-2} (T_2).$$

$$\text{Consequently: } (T_1 \circ_{\ell} T_2) \circ_k T_3 = \circlearrowright^{\ell+p-1,n} (\circlearrowright^{k,p} (T_1)) \cup \ggg^{\ell+p-2} (T_2) \cup \ggg^{k-1} (T_3).$$

Since $k < \ell$, the operator $\circlearrowright_{\ell+p-1,n}$ does not modify the set $\ggg^{k-1} (T_3)$ of couples the elements of couples of which are upper-bounded by $k + p - 1 < \ell + p - 1$. Hence: $\ggg^{k-1} (T_3) = \circlearrowright^{\ell+p-1,n} (\ggg^{k-1} (T_3))$.

$$\text{Then: } (T_1 \circ_{\ell} T_2) \circ_k T_3 = \circlearrowright^{\ell+p-1,n} (\circlearrowright^{k,p} (T_1)) \cup \circlearrowright^{\ell+p-1,n} (\ggg^{k-1} (T_3)) \cup \ggg^{\ell+p-2} (T_2).$$

$$\text{Finally: } (T_1 \circ_{\ell} T_2) \circ_k T_3 = (T_1 \circ_k T_3) \circ_{\ell+p-1} T_2.$$

□

The associativity of \circ_k is shown by the following proposition.

Proposition 3. *Let m, n, p, i, k be five integers such that $i < n$. Let $T_1 \in \mathcal{T}_m$, $T_2 \in \mathcal{T}_n$ and $T_3 \in \mathcal{T}_p$ be three multitildes. Then:*

$$T_1 \circ_k (T_2 \circ_i T_3) = (T_1 \circ_k T_2) \circ_{k+i-1} T_3.$$

Proof. According to definition of \circ_i :

$$\begin{aligned} T_1 \circ_k (T_2 \circ_i T_3) &= T_1 \circ_k \left(\circlearrowright^{i,p} (T_2) \cup \ggg^{i-1} (T_3) \right) \\ &= \circlearrowright^{k,n+p-1} (T_1) \cup \ggg^{k-1} (\circlearrowright^{i,p} (T_2)) \cup \ggg^{k-1} (\ggg^{i-1} (T_3)). \end{aligned}$$

Following Proposition 1 on the sets of couples computed in the right member:

1. from relation (1): $\ggg^{k-1} (\ggg^{i-1} (T_3)) = \ggg^{k+i-2} (T_3)$;
2. from relation (2): $\ggg^{k-1} (\circlearrowright^{i,p} (T_2)) = \circlearrowright^{k+i-1,p} (\ggg^{k-1} (T_2))$;

3. Since $i < n$, from relation (4): $\overset{k,n+p-1}{\circlearrowright} (T_1) = \overset{k+i-1,p}{\circlearrowright} (\overset{k,n}{\circlearrowright}(T_1))$.

Consequently: $T_1 \circ_k (T_2 \circ_i T_3) = \overset{k+i-1,p}{\circlearrowright} (\overset{k,n}{\circlearrowright}(T_1)) \cup \overset{k+i-1,p}{\circlearrowright} (\overset{k-1}{\ggg}(T_2)) \cup \overset{k+i-2}{\ggg}(T_3)$.

The right side of the expression is the definition of \circ_{k+i-1} and finally:

$$\begin{aligned} T_1 \circ_k (T_2 \circ_i T_3) &= (\overset{k,n-1}{\circlearrowright}(T_1) \cup \overset{k-1}{\ggg}(T_2)) \circ_{k+i-1} T_3 \\ &= (T_1 \circ_k T_2) \circ_{k+i-1} T_3. \end{aligned}$$

□

From the partial composition \circ_k , ones can define the composition \circ as follows:

$$\circ : \begin{cases} \mathcal{T}_m \otimes \mathcal{T}_{n_1} \otimes \cdots \otimes \mathcal{T}_{n_m} & \rightarrow \mathcal{T}_{n_1 + \cdots + n_m} \\ T \circ (T'_1, \dots, T'_m) & = (\cdots ((T_1 \circ_m T'_m) \circ_{m-1} T'_{m-1}) \cdots) \circ_1 T'_1. \end{cases}$$

Theorem 1. *The structure (\mathcal{T}, \circ) is an operad.*

4 Actions on languages

In this section, we recover the construction of language associated to a multitildes expression ([3]) by means of an operad based on boolean vectors. We show the connection between this operad and the operad \mathcal{T} .

4.1 An operad of boolean vectors and its action on languages

Denote by $\mathbb{B} = \{0, 1\}$ and $\mathcal{B}_n = \mathcal{P}(\mathbb{B}^n)$. We endow the set $\mathcal{B} = \bigcup_{n \in \mathbb{N}} \mathcal{B}_n$ with a structure of operad whose composition is as follows:

Let $E \in \mathcal{B}_m$ and $F \in \mathcal{B}_n$, the set $E \circ_k F$ is defined by

$$E \circ_k F := \{[e_1, \dots, e_{k-1}, e_k f_1, \dots, e_k f_n, e_{k+1}, \dots, e_m] : [e_1, \dots, e_m] \in E, [f_1, \dots, f_n] \in F\}.$$

Proposition 4. *The structure (\mathcal{B}, \circ) is an operad.*

Proof. The set $\mathbf{1} := \{[1]\}$ is the identity for the composition \circ_k . Indeed,

$$\mathbf{1} \circ_1 F = F$$

and

$$\begin{aligned} E \circ_k \{1\} &:= \{[e_1, \dots, e_{k-1}, 0, e_{k+1}, \dots, e_m] : [e_1, \dots, e_{k-1}, 0, e_{k+1}, \dots, e_m] \in E\} \cup \\ &\quad \{[e_1, \dots, e_{k-1}, 1, e_{k+1}, \dots, e_m] : [e_1, \dots, e_{k-1}, 1, e_{k+1}, \dots, e_m] \in E\} = E. \end{aligned}$$

It remains to prove the two associativity rules:

1. Let $E \in \mathcal{B}_m, F \in \mathcal{B}_n, G \in \mathcal{B}_p$ and $1 \leq j < i \leq m$. One has

$$\begin{aligned} (E \circ_i F) \circ_j G &= \{[e_1, \dots, e_{j-1}, e_j g_1, \dots, e_j g_n, e_{j+1}, \dots, e_{i-1}, e_i f_1, \dots, e_i f_n, e_{i+1}, \dots, e_m] \\ &\quad : [e_1, \dots, e_m] \in E, [f_1, \dots, f_n] \in F, [g_1, \dots, g_p] \in G\} \\ &= \{[e_1, \dots, e_{j-1}, e_j g_1, \dots, e_j g_p, e_{j+1}, \dots, e_m] : [e_1, \dots, e_m] \in E, \\ &\quad [g_1, \dots, g_p] \in G\} \circ_{j+p-1} F \\ &= (E \circ_j G) \circ_{j+p-1} F. \end{aligned}$$

2. Let $E \in \mathcal{B}_m, F \in \mathcal{B}_n, G \in \mathcal{B}_p, 1 \leq i \leq n$ and $1 \leq j \leq m$. One has

$$\begin{aligned} E \circ_j (F \circ_i G) &= \{[e_1, \dots, e_{j-1}, e_j f_1, e_j f_i g_1, \dots, e_j f_i g_p, e_j f_{i+1}, \dots, e_j f_n, e_{j+1}, \dots, e_m] : \\ &\quad [e_1, \dots, e_m] \in E, [f_1, \dots, f_n] \in F, [g_1, \dots, g_p] \in G\} \\ &= (E \circ_j F) \circ_{i+j-1} G. \end{aligned}$$

This proves that \mathcal{B} is an operad. □

We define the action of \mathcal{B}_n on the languages by

$$E(L_1, \dots, L_n) = \bigcup_{[e_1, \dots, e_n] \in E} L_1^{e_1} \cdot L_2^{e_2} \dots L_n^{e_n}.$$

for each $E \in \mathcal{B}$.

Proposition 5. *The sets 2^{Σ^*} (the set of the languages over Σ) and $\text{Reg}(\Sigma^*)$ (the set of regular languages over Σ) are both \mathcal{B} -modules.*

Proof. Let us show first the compatibility of the action previously defined with the composition. One has

$$E(L_1, \dots, L_{i-1}, F(L_i, \dots, L_{i+n-1}), L_{i+n}, \dots, L_{m+n-1}) = \bigcup_{[e_1, \dots, e_m] \in E} L_1^{e_1} \dots L_{i-1}^{e_{i-1}} (L_i^{f_1} \dots L_{i+n-1}^{f_n})^{e_i} L_{i+n}^{e_{i+1}} \dots L_{m+n-1}^{e_m}.$$

Since $e_i \in \{0, 1\}$,

$$\begin{aligned} E(L_1, \dots, L_{i-1}, F(L_i, \dots, L_{i+n-1}), L_{i+n}, \dots, L_{m+n-1}) &= \bigcup_{\substack{[e_1, \dots, e_m] \in E, \\ [f_1, \dots, f_n] \in F}} L_1^{e_1} \dots L_{i-1}^{e_{i-1}} L_i^{e_i f_1} \dots L_{i+n-1}^{e_i f_n} L_{i+n}^{e_{i+1}} \dots L_{m+n-1}^{e_m} \\ &= \bigcup_{[g_1, \dots, g_{n+m-1}] \in E \circ_i F} L_1^{g_1} \dots L_{n+m-1}^{g_{n+m-1}} \\ &= E \circ_i F(L_1, \dots, L_{n+m-1}). \end{aligned}$$

Hence, 2^{Σ^*} is a \mathcal{B} -module. Note a finite union of catenation of regular languages is regular, so each operator $E \in \mathcal{B}$ maps $\text{Reg}(\Sigma^*)$ on $\text{Reg}(\Sigma^*)$. This implies that it is also a \mathcal{B} -module. □

4.2 Action of \mathcal{T} on languages

We define a map $V : \mathcal{T}_k \rightarrow \mathcal{B}_k$ by

$$V(T) = \{v(S) : S \in \mathcal{F}(T)\}$$

with

$$v(S) = (v_1, \dots, v_k) \text{ where } v_j = \begin{cases} 0 & \text{if } j \in \bigcup_{(x,y) \in S} \llbracket x, y \rrbracket \\ 1 & \text{otherwise.} \end{cases}$$

Example 2. Consider $T = \{(1, 2), (2, 3), (3, 4), (4, 4)\} \in \mathcal{T}_4$. The images of the elements of $\mathcal{F}(T)$ are

$$\begin{aligned} v(\{(1, 2), (3, 4)\}) &= (0, 0, 0, 0) \\ v(\{(1, 2), (4, 4)\}) &= (0, 0, 1, 0) \\ v(\{(2, 3), (4, 4)\}) &= (1, 0, 0, 0) \\ v(\{(1, 2)\}) &= (0, 0, 1, 1) \\ v(\{(2, 3)\}) &= (1, 0, 0, 1) \\ v(\{(3, 4)\}) &= (1, 1, 0, 0) \\ v(\{(4, 4)\}) &= (1, 1, 1, 0) \\ v(\emptyset) &= (1, 1, 1, 1) \end{aligned}$$

Hence

$$V(T) = \{(0, 0, 0, 0), (0, 0, 1, 0), (1, 0, 0, 0), (0, 0, 1, 1), (1, 0, 0, 1), (1, 1, 0, 0), (1, 1, 1, 0), (1, 1, 1, 1)\}$$

Theorem 2. *The map V is a morphism of operads.*

Proof. First the image of $\emptyset \in \mathcal{T}_1$ by V is the set 1.

Now let us examine the image of a composition. Let $T_1 \in \mathcal{T}_m, T_2 \in \mathcal{T}_n$ and $1 \leq k \leq m$.

We need the following lemma which explains how the composition modifies the free subsets.

Lemma 5.

$$\begin{aligned} \mathcal{F}(T_1 \circ_k T_2) = & \{ \overset{n,k}{\circ} \rightarrow S : S \in \mathcal{F}(T_1) \} \cup \\ & \{ \overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T : S \in \mathcal{F}(T_1), T \in \mathcal{F}(T_2), \\ & \forall (x, y) \in S, x \geq k \text{ or } y < k - 1 \} \end{aligned} \quad (8)$$

Proof. For simplicity, we will denote by \mathcal{R} the right hand side of equality (8). Note $\mathcal{R} \subseteq \mathcal{F}(T_1 \circ_k T_2)$ is obtained remarking that each element of \mathcal{R} is free.

Now let us prove the inclusion $\mathcal{F}(T_1 \circ_k T_2) \subseteq \mathcal{R}$. Straightforwardly from the definition of \mathcal{F} , we have the inclusion

$$\mathcal{F}(T_1 \circ_k T_2) \subset \{ \overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T : S \in \mathcal{F}(T_1), T \in \mathcal{F}(T_2) \}.$$

Hence, if $R \in \mathcal{F}(T_1 \circ_k T_2)$, there exist $S \in \mathcal{F}(T_1)$ and $T \in \mathcal{F}(T_2)$ such that $R = \overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T$. If $T = \emptyset$ then $R \in \{ \overset{n,k}{\circ} \rightarrow S : S \in \mathcal{F}(T_1) \} \subset \mathcal{R}$. Suppose $T \neq \emptyset$ and $R \notin \mathcal{R}$. Consequently there exist two couples $(x, y + n), (x' + k - 1, y' + k - 1) \in R$ with $(x, y) \in S, x \leq k, y > k - 1$ and $(x', y') \in T$. Hence, $\llbracket x, y + n \rrbracket \cap \llbracket x' + k - 1, y' + k - 1 \rrbracket = \llbracket x' + k - 1, y' + k - 1 \rrbracket \neq \emptyset$. This contradicts the fact that R is free, proves the second inclusion and then the equality of the two sets. \square

End of the proof From equation (8) we have

$$\begin{aligned} V(T_1 \circ_k T_2) = & \{ v(\overset{n,k}{\circ} \rightarrow S) : S \in \mathcal{F}(T_1) \} \cup \\ & \{ v(\overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T) : S \in \mathcal{F}(T_1), T \in \mathcal{F}(T_2), \\ & \forall (x, y) \in S, x \geq k \text{ or } y < k - 1 \} \end{aligned}$$

Set $V_1 := \{ v(\overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T) : S \in \mathcal{F}(T_1), T \in \mathcal{F}(T_2), \forall (x, y) \in S, x \geq k \text{ or } y < k - 1 \}$ and $V_2 = \{ v(\overset{n,k}{\circ} \rightarrow S) : S \in \mathcal{F}(T_1) \}$. We have

$$V_1 = \{ v(\overset{n,k}{\circ} \rightarrow S \cup \overset{k-1}{\ggg} T) : S \in \mathcal{F}(T_1), T \in \mathcal{F}(T_2), v(s)_k = 1 \}.$$

Hence,

$$V_1 = \{ (s_1, \dots, s_{k-1}, s_k t_1, \dots, s_k t_n, s_{k+1}, \dots, s_m) : (s_1, \dots, s_m) \in V(T_1), (t_1, \dots, t_n) \in V(T_2), s_k = 1 \}$$

On another hand V_2 splits as

$$V_2 = V_2^0 \cup V_2^1$$

where

$$V_2^i = \{ v(\overset{n,k}{\circ} \rightarrow S : S \in \mathcal{F}(T_1) \text{ and } v(s)_k = i \}.$$

We have $V_2^1 \subset V_1$ (it suffices to put $T = \emptyset$ in V_1) and

$$V_2^0 = \{ (s_1, \dots, s_{k-1}, s_k t_1, \dots, s_k t_n, s_{k+1}, \dots, s_m) : (s_1, \dots, s_m) \in V(T_1), (t_1, \dots, t_n) \in V(T_2), s_k = 0 \}$$

It follows that $V(T_1 \circ_k T_2) = V(T_1) \circ_k V(T_2)$. \square

Corollary 1. *The sets 2^{Σ^*} and $\text{Reg}(\Sigma^*)$ are both \mathcal{T} -module.*

Proof. It suffices to set $T(L_1, \dots, L_k) := V(T)(L_1, \dots, L_k)$. \square

Note the action of \mathcal{T} on languages matches with the definition of \sim_T .

Proposition 6.

$$\sim_T(L_1, \dots, L_k) = T(L_1, \dots, L_k)$$

Proof. We have

$$\mathcal{W}_S(L_1, \dots, L_k) = L'_1 \cdots L'_k$$

with

$$L'_i = \begin{cases} \{\epsilon\} & \text{if } i \in \bigcup_{(x,y) \in S} \llbracket x, y \rrbracket \\ L_k & \text{otherwise.} \end{cases}$$

Hence,

$$\mathcal{W}_S(L_1, \dots, L_k) = L_1^{v(S)_1} \cdots L_k^{v(S)_k}.$$

And the result follows from

$$\sim_T(L_1, \dots, L_k) = \bigcup_{S \in \mathcal{F}(T)} \mathcal{W}_S(L_1, \dots, L_k) = \bigcup_{v \in V(T)} L_1^{v_1} \cdots L_k^{v_k} = T(L_1, \dots, L_k).$$

□

5 \mathcal{T} , RAS and POSet

In this section, we give a combinatorial description of the operad \mathcal{T} . In particular we prove that it is isomorphic to an operad whose underlying set is the set RAS^{\leq} of reflexive and (necessarily) antisymmetric subrelations of the order \leq on $\mathbb{N} \setminus \{0\}$. In our context, we define a quotient POSet^{\leq} of RAS^{\leq} whose elements are indexed by partial ordered sets. This construction is based on the transitive closure equivalence relation on RAS ; the equivalent classes are indexed by POSet s. We prove that the operad POSet^{\leq} is isomorphic to a quotient of \mathcal{T} which is compatible in a natural way with the action on languages.

5.1 From \mathcal{T} to RAS^{\leq}

Since multitildes are defined by a set of couples (x, y) such that $x \leq y$, they can be seen as antisymmetric relations compatible with the natural order on \mathbb{N} , *i.e.* $(x, y) \in T$ implies $x \leq y$.

On the other hand some multitildes expressions are equivalent. For instance, \widetilde{abcd} and $\widetilde{\widetilde{abcd}}$ denote the same language : $\{\epsilon, ab, cd, abcd\}$. This phenomenon is not very natural when stated in terms of relations. Nevertheless, up to a slight transformation, this can be seen as a transitive closure.

Let us be more precise and define a graduated bijection ϕ between \mathcal{T}_k and $\text{RAS}_k^{\leq} := \text{RAS}^{\leq} \cap \{1, \dots, k+1\}^2$ for any integer k :

$$\phi : \begin{cases} \mathcal{T}_k & \longrightarrow \text{RAS}_k^{\leq} \\ T & \longrightarrow \{(x, y+1) \mid (x, y) \in T\} \cup \{(x, x) \mid x \in \llbracket 1, n \rrbracket\} \end{cases}$$

The fact that ϕ is a bijection is obvious and the inverse bijection ϕ^{-1} is given by:

$$\phi^{-1} : \begin{cases} \text{RAS}_k^{\leq} & \longrightarrow \mathcal{T}_k \\ R & \longrightarrow \{(x, y) \mid (x, y+1) \in R \wedge x \neq y\} \end{cases}$$

We endow the set RAS^{\leq} with a structure of operad, setting:

$$R_1 \circ'_i R_2 = \phi(\phi^{-1}(R_1) \circ_i \phi^{-1}(R_2)).$$

Moreover, the bijection ϕ is in fact an isomorphism from the operad (\mathcal{T}, \circ) to $(\text{RAS}^{\leq}, \circ')$.

We define the operator $\overset{n,k}{\diamond} \rightarrow$ acting on pairs of integers (x, y) by

$$\overset{n,k}{\diamond}(x, y) = \begin{cases} (x, y) & \text{if } y \leq k \\ (x, y+n-1) & \text{if } x \leq k < y \\ (x+n-1, y+n-1) & \text{otherwise.} \end{cases}$$

We extend the definition to the set of pairs by $\overset{n,k}{\diamondrightarrow}(S) = \{\overset{n,k}{\diamondrightarrow}(x, y) : (x, y) \in S\}$. Note $\overset{n,k}{\diamondrightarrow}$ acts by relabelling $x \rightarrow x + n - 1$ when $x > k$. Hence, we set as for the multitildes:

$$R_1 \diamond_k R_2 := \overset{n,k}{\diamondrightarrow}(R_1) \cup \overset{k-1}{\ggg}(R_2),$$

for $R_1 \in \text{RAS}_m^{\leq}$, $R_2 \in \text{RAS}_n^{\leq}$ and $k \leq m$.

The following lemma shows that RAS^{\leq} is closed under the action of the compositions \diamond_k and that these binary operators are compatible with the graduation.

Lemma 6. *Let $R_1 \in \text{RAS}_m^{\leq}$ and $R_2 \in \text{RAS}_n^{\leq}$ then $R_1 \diamond_k R_2 \in \text{RAS}_{m+n-1}^{\leq}$.*

Proof. The antisymmetry is straightforward from the definition of \diamond_k . So it suffices to prove the reflexivity, that is to check $(x, x) \in R_1 \diamond_k R_2$ for each $1 \leq x \leq m + n$. But, from the definitions, $(x, x) \in \overset{n,k}{\diamondrightarrow}(R_1)$ when $1 \leq x \leq k$ or $n + k \leq x \leq n + m$ and $(x, x) \in \overset{k-1}{\ggg}(R_2)$ is $k \leq x \leq n + k$. So for each $1 \leq x \leq m + n$ we obtain $(x, x) \in R_1 \diamond_k R_2$. Hence, $R_1 \diamond_k R_2 \in \text{RAS}_{m+n-1}^{\leq}$. \square

From the partial composition \diamond_k we define

$$\diamond : \begin{cases} \text{RAS}_m^{\leq} \otimes \text{RAS}_{n_1}^{\leq} \otimes \cdots \otimes \text{RAS}_{n_m}^{\leq} & \longrightarrow & \text{RAS}_{n_1 + \cdots + n_m}^{\leq} \\ R \diamond (R_1, \dots, R_m) & = & (\dots((R \diamond_m R_1) \diamond_{m-1} R_{m-1}) \dots) \diamond_1 R_1. \end{cases}$$

And more precisely:

Proposition 7. *We have $\circ'_k = \diamond_k$ and $\circ' = \diamond$.*

Proof. We have $\phi \overset{n,k}{\circ} = \overset{n,k}{\diamondrightarrow} \phi$ and $\phi \overset{k}{\ggg} = \overset{k}{\ggg} \phi$. So

$$\phi(T_1 \circ_k T_2) = \phi(T_1) \diamond_k \phi(T_2),$$

for each $T_1 \in \mathcal{T}_m$, $T_2 \in \mathcal{T}_n$ and $1 \leq k \leq m$. Since $\phi : (\mathcal{T}, \circ) \rightarrow (\text{RAS}^{\leq}, \circ')$ is an isomorphism of operads, we prove the result. \square

It follows that:

Theorem 3. *The structure $(\text{RAS}^{\leq}, \diamond)$ defines an operad isomorphic to (\mathcal{T}, \circ) .*

5.2 From RAS^{\leq} to POSet^{\leq}

Any reflexive and antisymmetric relation R can be turned into a partial order when applying the transitive closure denoted by $\gamma(R)$.

Note, since $\overset{n,k}{\diamondrightarrow}$ and $\overset{k}{\ggg}$ are just relabelling, they commute with the operator γ :

$$\gamma \overset{n,k}{\diamondrightarrow} = \overset{n,k}{\diamondrightarrow} \gamma, \quad \gamma \overset{k}{\ggg} = \overset{k}{\ggg} \gamma. \quad (9)$$

We need the following lemma

Lemma 7.

$$\gamma(\gamma R_1 \diamond_k \gamma R_2) = \gamma(R_1 \diamond_k R_2)$$

Proof. It suffices to prove that $\gamma R_1 \diamond_k \gamma R_2 \subset \gamma(R_1 \diamond_k R_2)$. Indeed by transitivity, this implies $\gamma(\gamma R_1 \diamond_k \gamma R_2) \subset \gamma(R_1 \diamond_k R_2)$ and the reverse inclusion is obvious.

We have $\gamma R_1 \diamond_k \gamma R_2 = \overset{k,n}{\diamondrightarrow}(\gamma R_1) \cup \overset{k-1}{\ggg}(\gamma R_2)$. From eq (9) we obtain $\gamma R_1 \diamond_k \gamma R_2 = (\gamma \overset{n,k}{\diamondrightarrow}(R_1)) \cup (\gamma \overset{k-1}{\ggg}(R_2)) \subset \gamma(R_1 \diamond_k R_2)$, as expected. \square

As an immediate consequence, the transitive closure is compatible with the structure of operads that is:

Lemma 8. *If $\gamma(R_1) = \gamma(R'_1)$ and $\gamma(R_2) = \gamma(R'_2)$ then $\gamma(R_1 \diamond_k R_2) = \gamma(R'_1 \diamond_k R'_2)$.*

Denote by \equiv_γ the equivalence relation on RAS^\leq defined by $R_1 \equiv_\gamma R_2$ if and only if $\gamma(R_1) = \gamma(R_2)$. The quotient $\text{RAS}^\leq / \equiv_\gamma$ is automatically endowed with a structure of operad whose composition is deduced from γ ; we denote by \diamond' the associated composition.

The equivalence classes are indexed by the element $\text{POSet}^\leq := \gamma(\text{RAS}^\leq)$. Let us set $\diamond_k := \gamma \diamond_k$ and $\diamond := \gamma \diamond$.

Immediately we obtain:

Proposition 8. *The structure $(\text{POSet}^\leq, \diamond)$ is an operad isomorphic to $(\text{RAS}^\leq / \equiv_\gamma, \diamond')$.*

5.3 Pseudotransitive tildes

A multitilde is said *pseudotransitive* if it is the image of a POSet by ϕ^{-1} . The set of pseudotransitive multitildes will be denoted by

$$\text{PTT} := \phi^{-1}(\text{POSet}^\leq).$$

This set is endowed with an operad structure induced by the partial products

$$T_1 \circ_k T_2 := \phi^{-1}(\phi(T_1) \diamond_k \phi(T_2)) = \phi^{-1} \gamma \phi(T_1 \circ_k T_2).$$

For simplicity, set $\tilde{\gamma} := \phi^{-1} \gamma \phi$. The equivalence relation \equiv on \mathcal{T} defined by $T_1 \equiv T_2$ if and only if $\tilde{\gamma}(T_1) = \tilde{\gamma}(T_2)$ is compatible with the operad structure. Indeed,

$$\begin{aligned} \tilde{\gamma}(T_1 \circ_k T_2) &= \phi^{-1} \gamma \phi(T_1 \circ_k T_2) \\ &= \phi^{-1} \gamma (\phi(T_1) \diamond_k \phi(T_2)) \end{aligned}$$

But, $\tilde{\gamma}(T) = \tilde{\gamma}(T')$ implies $\phi(T) \equiv_\gamma \phi(T')$. Since \equiv_γ is compatible with the composition in RAS^\leq we obtain

$$\begin{aligned} \tilde{\gamma}(T_1 \circ_k T_2) &= \phi^{-1} \gamma (\phi(T'_1) \diamond_k \phi(T'_2)) \\ &= \tilde{\gamma}(T'_1 \circ_k T'_2). \end{aligned}$$

Hence, the quotient \mathcal{T} / \equiv is isomorphic to (PTT, \circ) .

We explicitly describe $\tilde{\gamma}$.

Lemma 9. *The set $\tilde{\gamma}(T)$ is the smallest set S such that the two following assertions are satisfied:*

1. $T \subset S$
2. $(i, k), (k + 1, j) \in S$ implies $(i, j) \in S$.

Proof. Since $\gamma(\phi(T))$ is the transitive closure of $\phi(T)$, it is the smallest set R such that

1. $\phi(T) \in R$
2. $(i, k), (k, j) \in R$ implies $(i, j) \in R$.

Hence, $\phi^{-1}(\gamma(\phi(T)))$ is such that if $(i, k), (k + 1, j) \in \phi^{-1}(\gamma(\phi(T)))$ implies $(i, j) \in \phi^{-1}(\gamma(\phi(T)))$.

Indeed let $(i, k), (k + 1, j) \in \tilde{\gamma}(T)$ then $(i, k + 1), (k + 1, j + 1) \in \gamma(\phi(T))$. Since $\gamma(\phi(T))$ is transitive, we also have $(i, j + 1) \in \gamma(\phi(T))$. The image of this pair is $\phi^{-1}(i, j + 1) = (i, j) \in \tilde{\gamma}(T)$.

Let S be a subset of $\phi^{-1}(\gamma(\phi(T)))$ verifying

1. $T \subset S$
2. $(i, k), (k + 1, j) \in S$ implies $(i, j) \in S$.

Suppose that $(i, j) \notin \phi^{-1}(\gamma(\phi(T)))$, since ϕ is a bijection we obtain $\phi(i, j) \notin \gamma(\phi(T))$. This contradicts the minimality of $\gamma(\phi(T))$ and prove the result. \square

The set $\tilde{\gamma}(T)$ will be referred as the pseudotransitive closure of T .

Lemma 10. *For each tilde $T \in \mathcal{T}$, $V(T) = V(\tilde{\gamma}(T))$ holds.*

Proof. Suppose that $(i, k), (k+1, j) \in T$. Let us prove $V(T \cup \{(i, j)\}) = V(T)$. Suppose $(i, j) \notin T$ (otherwise the result is obvious). We compare the set $E := \{v(S) : S \in \mathcal{F}(T)\}$ and $F := \{v(S) : S \in \mathcal{F}(T \cup \{(i, j)\})\}$. Obviously, $\mathcal{F}(T) \subset \mathcal{F}(T \cup \{(i, j)\})$ hence $E \subset F$. Let us prove the reverse inclusion and let $S \in \mathcal{F}(T \cup \{(i, j)\})$. If $(i, j) \notin S$ then $S \in \mathcal{F}(T)$ and then $v(S) \in E$. Now suppose $(i, j) \in S$ and set $S' = S \cup \{(i, k), (k+1, j)\} \setminus \{(i, j)\}$. Obviously $S' \in \mathcal{F}(T)$. Set $v(S) = (s_1, \dots, s_n)$ and $v(S') = (s'_1, \dots, s'_n)$. We have $s_l = s'_l$ if $l \notin \llbracket i, j \rrbracket$ and 0 otherwise. But $s'_l = 0$ when $l \in \llbracket i, k \rrbracket \cup \llbracket k+1, j \rrbracket = \llbracket i, j \rrbracket$. So $v(S) = v(S') \in E$. \square

As a consequence, we define an action of POSet^{\leq} and PTT on languages:

Theorem 4. *The sets 2^{Σ^*} and $\text{Reg}(\Sigma^*)$ are POSet^{\leq} -module and PTT -module.*

Proof. From Corollary 1, 2^{Σ^*} and $\text{Reg}(\Sigma^*)$ are \mathcal{T} -module.

Note if $T_1 \equiv T_2$ are two tildes of \mathcal{T} then, by Lemma 10, T_1 and T_2 have the same action on 2^{Σ^*} and $\text{Reg}(\Sigma^*)$. These set are PTT -modules. Since PTT is isomorphic to POSet^{\leq} as an operad, they are also POSet^{\leq} -modules. \square

6 Consequences and perspectives

6.1 Enumeration

The purpose of this section is to compute an upper bound of the number $n(L_1, \dots, L_k)$ of languages that can be obtained by the action of tilde on the k -tuple of languages (L_1, \dots, L_k) .

First, Lemma 10 implies that $n(L_1, \dots, L_k)$ equals to the number of languages obtained by applying a PTT on (L_1, \dots, L_k) .

From the previous section, the number of PTT with arity k equals the numbers p_k of POSet s on $\{1, \dots, k+1\}$ that are contained in the usual linear order on integers. So,

$$n(L_1, \dots, L_k) \leq p_k. \quad (10)$$

The sequence of p_k has no known closed form but the first values can be found on [18] (A006455); the first values are

$$2, 7, 40, 357, 4824, 96428, 2800472, 116473461, 6855780268, \\ 565505147444, 64824245807684, \dots$$

Note this is also the number of $(k+1) \times (k+1)$ upper triangular idempotent boolean matrices with all diagonal entries 1.

Let us show that the bound is reached for $L_1 = \{a_1\}, \dots, L_k = \{a_k\}$ where a_1, \dots, a_k are distinct letters.

Let $P_1 \neq P_2$ be two PTT s with arity k and $(i, j) \in P_1 \setminus P_2$. We have $\{(i, j)\} \in \mathcal{F}(P_1)$ and then $(1^{i-1}, 0^{j-i}, 1^{n-j+1}) \in V(P_1)$.

Suppose $(1^{i-1}, 0^{j-i}, 1^{n-j+1}) \in V(P_2)$. This implies that there exists a set $\{(i, k_1), (k_1+1, k_2), \dots, (k_m+1, j)\} \in \mathcal{F}(P_2)$ for some m . So $\{(i, k_1), (k_1+1, k_2), \dots, (k_m+1, j)\} \subset P_2$ and from Lemma 9 we obtain $(i, j) \in P_2$. This contradicts our hypothesis. Then,

Corollary 2. *Let a_1, \dots, a_k be k distinct letters. We have*

$$n(\{a_1\}, \dots, \{a_k\}) = p_k.$$

Example 3. There are 7 languages obtained by applying PTT on a pair of letters ($\{a\}, \{b\}$):

tildes	languages
\emptyset	ab
$(1, 1)$	ab, b
$(1, 1), (1, 2)$	ab, b, ϵ
$(2, 2)$	ab, a
$(2, 2), (1, 2)$	ab, a, ϵ
$(1, 1), (2, 2), (1, 2)$	ab, a, b, ϵ
$(1, 2)$	ab, ϵ

6.2 Finite languages

Every finite languages can be generated by a multitilde acting on a sequences of elements in $\Sigma_0 := \{\{a\}\}_{a \in \Sigma} \cup \{\emptyset\}$.

Proposition 9. For any finite language L over an alphabet Σ , it exists a multitilde $T \in \mathcal{T}_k$ and a k -tuples (L_1, \dots, L_k) for some integer k , with $L_1, \dots, L_k \in \Sigma_0$ such that

$$L = T(L_1, \dots, L_k).$$

Proof. Note first, the set of finite languages is a \mathcal{T} -module, since it is contained in 2^{Σ^*} , which by Corollary 1 is a \mathcal{T} -module, and is closed by the actions of \mathcal{T} . Straightforwardly, we have $\emptyset = \emptyset_1(\emptyset)$, $\{\epsilon\} = \{(1, 1)\}(\emptyset)$ and $\{a\} = \emptyset_1(\{a\})$ for any $a \in \Sigma$.

Hence, since \mathcal{T} is an operad (Theorem 1), it suffices to prove that the union and the catenation of two languages can be obtained by using a multitilde. We verify easily that

$$L_1 L_2 = \emptyset_2(L_1, L_2) \text{ and } L_1 \cup L_2 = \{(1, 2), (2, 3)\}(L_1, \emptyset, L_2).$$

This ends the proof. □

Let us give some examples.

Example 4. The set of all subwords of a given word $a_1 \dots a_k$, is easily expressed as a multitilde by

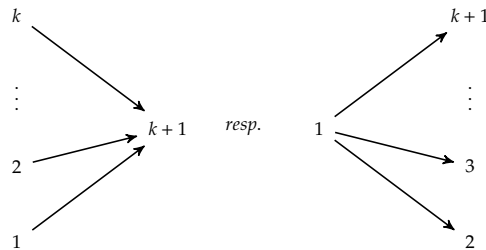
$$\{(i, i) : i \in \llbracket 1, k \rrbracket\} (a_1, \dots, a_k).$$

Note the corresponding POSet is the natural order relation on $\llbracket 1, k + 1 \rrbracket$.

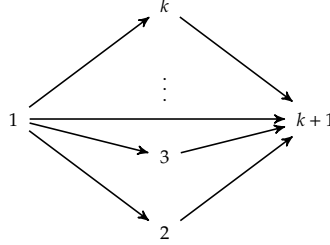
Example 5. The set of all the prefixes (resp. suffices) of a given word $a_1 \dots a_k$ admits a nice expression in terms of multitildes:

$$P_k := \{(i, k) : i \in \llbracket 1, k \rrbracket\} (a_1, \dots, a_k) \text{ (resp. } S_k := \{(1, i) : i \in \llbracket 1, k \rrbracket\} (a_1, \dots, a_k)).$$

The corresponding POSets are



Example 6. As a consequence, the set of all the factors of the word $a_1 \dots a_k$ is obtained by the action of $F_k := P_k \cup S_k$. Graphically, the corresponding POSet is:



For instance, consider the multiset

$$F_3 := \{(1, 1), (1, 2), (1, 3), (2, 3), (3, 3)\}$$

its free subsets are

$$\{\emptyset\} \cup \{c : c \in F_3\} \cup \{(1, 1), (2, 3)\}, \{(1, 2), (3, 3)\}, \{(1, 1), (3, 3)\}.$$

The corresponding boolean vectors are

$$(1, 1, 1), (0, 1, 1), (0, 0, 1), (0, 0, 0), (1, 0, 0), (1, 1, 0), (0, 1, 0).$$

When acting on $(\{a\}, \{b\}, \{c\})$ they respectively give $\{abc\}, \{bc\}, \{c\}, \{\varepsilon\}, \{a\}, \{ab\}$ and $\{b\}$.

Example 7. More generally, we have the following property:

If $L = T(L_1, \dots, L_k)$ with $L_i \in \{\{a\} : a \in \Sigma\} \cup \{\emptyset\}$ we have

1. The set of the prefixes of any words in L , $\text{Pref}(L)$ is a subset of $(T \cup P_k)(L_1, \dots, L_k)$
2. The set of the suffices of any words in L , $\text{Suff}(L)$ is a subset of $(T \cup S_k)(L_1, \dots, L_k)$.
3. The set of the factors of any words in L , $\text{Fact}(L)$ is a subset of $(T \cup F_k)(L_1, \dots, L_k)$.

One has only to show the first assertion. Indeed, the second point is obtained by symmetry from the first and the third is simply the composition of the first and the second.

The free subsets of $T \cup P_k$ are

$$\mathcal{F}(T \cup P_k) = \mathcal{F}(T) \cup \{S' = S \cup \{(i, k)\} : i \in \llbracket 1, k \rrbracket, S \in \mathcal{F}(T), S' \text{ is free}\}.$$

Consider a proper prefix p of word $w = ps \in L$. Let $I \subset \llbracket 1, k \rrbracket$ such that $i \in I$ if and only if $L_i \neq \emptyset$. Without loss of generality, we suppose that all the L_i for $i \in I$ are distinct. Set $s = a_i s'$ with $a_i \in \Sigma$ and let j be the unique integer such that $L_j = \{a_j\}$. Since $w \in L$, it exists a free list S such that $v(S)(L_1, \dots, L_k) = \{w\}$. Furthermore, $S' = \{(j_1, j_2) : (j_1, j_2) \in S, \llbracket j_1, j_2 \rrbracket \subset \llbracket 1, j \rrbracket\} \subset \mathcal{F}(T)$. Indeed, there is no pair (j_1, j_2) in S such that $j \in \llbracket j_1, j_2 \rrbracket$ (otherwise the letter a_j does not appear in w). Hence, S is composed only by pairs $(j_1, j_2) \in T$ such that $j_2 < j$ or $j < j_1$. It follows that $S' = \{(j_1, j_2) : j_2 < j\}$ is a free list of T because S is free. So $S' \cup \{(j+1, k)\} \subseteq T \cup P_k$ is free and $v(S' \cup \{(j+1, k)\})(L_1, \dots, L_k) = v(S')(L_1, \dots, L_j) = p$. It follows that $p \in (T \cup P)(L_1, \dots, L_k)$ and our claim.

The inclusions are strict in the general case. For instance, consider $\emptyset_4(\{a\}, \{b\}, \{c\}, \emptyset) = \emptyset$ and $(\emptyset_4 \cup P_4)(\{a\}, \{b\}, \{c\}, \emptyset) = P_4(\{a\}, \{b\}, \{c\}, \emptyset) = \{\varepsilon, a, ab, abc\}$.

Note if each $L_i \neq \emptyset$ the inclusions are equalities. Although this is a sufficient condition, it is not necessary as shown by $\{(1, 3), (3, 5)\}(\{a\}, \{b\}, \emptyset, \{c\}, \{d\}) = \{ab, cd\}$ and

$$(\{(1, 3), (3, 5)\} \cup P_5)(\{a\}, \{b\}, \emptyset, \{c\}, \{d\}) = \{\varepsilon, \{a\}, \{ab\}, \{c\}, \{cd\}\} = \text{Pref}(\{ab, cd\}).$$

6.3 Regular languages

In the previous section, we showed that each finite language can be obtained by the action of a multitilde on a k -tuple in Σ_0^k . As a consequence, any regular languages can be written as a combination of multitildes and stars acting on a k -tuple in Σ_0^k .

Consider the smallest operad \mathcal{T}^\star containing \mathcal{T} together a new formal 1-ary operator denoted \star . The elements of \mathcal{T}^\star are trees whose nodes are multitildes or \star and are such that the roots of the subtrees immediately issued of a multitilde are not multitildes (indeed, the composition \circ_i of two multitildes gives an other multitilde).

To define properly an action of \mathcal{T}^\star on languages, it suffices to define the action of the operator \star which is naturally $\star(L) = L^*$. Note $\star(\star(L)) = (L^*)^* = L^* = \star(L)$. To simplify the operators we introduce the smallest equivalence relation \equiv compatible with the composition such that $\star \circ \star \equiv \star$, and the operad $T^\star = \mathcal{T}^\star / \equiv$. Straightforwardly, the following result holds:

Proposition 10. 1. The sets 2^{Σ^*} and $\text{Reg}(\Sigma^*)$ are both T^\star and \mathcal{T}^\star -modules.

2. Each regular language can be obtained by the action of an element in T^\star on a k -tuple in Σ_0^k .

Similarly to the operad \mathcal{T} , there is a very interesting underlying combinatoric structure to the operad T^\star involving pairs of relations and a generalization of the dissections of polygons. This study is differed to a forthcoming paper.

7 Conclusion

In this paper, we introduced several operads which allow us to represent multitilde operations defined by one of the authors in [3]. Each of these operads plays a specific role in the understanding of these tools. For instance, the operad \mathcal{T} encodes the structure of composition of the multitilde operators. The operad \mathcal{B} explains how the multitildes act on languages. Finally, the operads RAS^{\leq} and POSet^{\leq} establish a relation between these objects and well known combinatorial structures. This formalizes the equivalence of two multitildes and enumerates the different inequivalent multitildes.

This study is a first step towards the construction of a new combinatorial approach of the theory of languages. First, regular languages can be seen as the result of the action of an element of T^\star on a k -tuples of letters. Second, another action of \mathcal{T} on languages have been defined in [3] and combined with the action of multitildes in [2]. Similar operadic constructions can be defined for these operators. These two points will be studied in forthcoming papers.

The multitilde have been introduced in the aim to enlarge the class of automata which can be represented as an expression using a number of symbols which is a linear function of the number of states. It remains to study the algorithmic improvement which can be brought by the different structures of operads.

References

- [1] V. Antimirov. Partial derivatives of regular expressions and finite automaton constructions. *Theoret. Comput. Sci.*, 155:291–319, 1996.
- [2] P. Caron, J.-M. Champarnaud, and L. Mignot. Acyclic automata and small expressions using multi-tilde-bar operators. *Theoret. Comput. Sci.*, 411(38–39):3423–3435, 2010.
- [3] P. Caron, J.-M. Champarnaud, and L. Mignot. Multi-bar and multi-tilde regular operators. *Journal of Automata, Languages and Combinatorics*, 16(1):11–26, 2011.
- [4] P. Caron and D. Ziadi. Characterization of Glushkov automata. *Theoret. Comput. Sci.*, 233(1–2):75–90, 2000.

- [5] J.-M. Champarnaud and D. Ziadi. From c-continuations to new quadratic algorithms for automata synthesis. *Internat. J. Algebra Comput.*, 11(6):707–735, 2001.
- [6] F. Chapoton. Operads and algebraic combinatorics of trees. *Séminaire Lotharingien de Combinatoire*, B58c:1–27, 2007.
- [7] A. Ehrenfeucht and H.-P. Zeiger. Complexity measures for regular expressions. *J. Comput. Syst. Sci.*, 12(2):134–146, 1976.
- [8] V.L. Ginzburg and M. Kapranov. Koszul duality for operads. *Duke Math. J.*, 76(1):203–272, 1994.
- [9] V. M. Glushkov. The abstract theory of automata. *Russian Mathematical Surveys*, 16:1–53, 1961.
- [10] S. Kleene. Representation of events in nerve nets and finite automata. *Automata Studies*, Ann. Math. Studies 34:3–41, 1956. Princeton U. Press.
- [11] M. Lazard. Lois de groupes et analyseurs. *Annales scientifiques de l’E.N.S. 3ième série*, 72(4):299–400, 1955.
- [12] T. Leinster. *Higher operads, Higher categories*. Cambridge University Press, 2004.
- [13] J.-L. Loday and B. Vallette. *Algebraic Operads*. draft available at <http://www-irma.u-strasbg.fr/loday/PAPERS/LodayVallette.pdf>, 2010.
- [14] J.P. May. Operads, algebras and modules. *Operads Proceedings of Renaissance Conferences Hartford CTLuminy 1995*, 202:1–14, 1971.
- [15] J.P. May. *The geometry of iterated loop spaces*. Number 271 in Lecture Notes in Mathematics. Springer-Verlag, 1972.
- [16] R. F. McNaughton and H. Yamada. Regular expressions and state graphs for automata. *IEEE Transactions on Electronic Computers*, 9:39–57, March 1960.
- [17] L. Mignot. *Des Codes Barres pour les Langages Rationnels*. PhD thesis, LITIS, Université de Rouen, France, 2010. available online, URL: <http://ludovicmignot.free.fr>.
- [18] N. J. A. Sloane. The on-line encyclopedia of integer sequences. URL: <http://oeis.org/>.