An analogue of the plactic monoid for binary search trees

Florent HIVERT* Jean-Christophe NOVELLI†
Jean-Yves THIBON‡

15th August 2003

Abstract

We introduce a monoid structure on the set of binary search trees, by a process similar to the construction of the plactic monoid, the Robinson-Schensted insertion is here replaced by the binary search insertion. This leads to a new construction and a new interpretation of the algebra of planar binary trees of Loday-Ronco. The non-commutative symmetric functions, the free symmetric functions and the algebra of planar binary trees are constructed is a unified way.

Contents

1 Introduction **1**

2 The Sylvester monoid **2**
2.1 The binary search insertion 3
2.2 The sylvester monoid 5

3 A new construction of the algebra of binary trees of Loday-Ronco **6**
3.1 Free quasi-symmetric functions and Free symmetric functions . . 6
3.2 The binary trees algebra 8
3.3 Duality 9

1 Introduction

There exist certain analogies between the combinatorics of binary search trees and young tableaux. For example, the increasing labeling of a given binary

*florent.hivert@univ-mlv.fr, Institut Gaspard Monge, Université de Marne-la-Vallée, 77454 Marne-la-Vallée cedex
†novelli@univ-mlv.fr
‡jyt@univ-mlv.fr

tree and the standard tableaux of a given shape are both counted by a "hook-formula" [7, 18, 8] which admit in both cases natural q-analogs [18, 1]. It is also known that one can construct by the means of the Robinson-Schensted correspondence a hopf algebra with basis the set of standard Young tableaux [16]. Moreover, a natural realization of this algebra by means of non-commutative polynomials gives a en-lighting proof of the Littlewood-Richardson rule [14, 3, 4]. In this realization, each tableau t of shape λ is an homogeneous polynomial of degree $n = |\lambda|$ whose commutative image is the Schur function s_λ.

Recently, Loday and Ronco have introduced a Hopf algebra with basis the set of planar binary trees [11, 12]. This algebra can, as the preceding one, be realized in a free associative algebra [3, 4]. Each complete binary trees with n internal nodes is represented by a homogeneous polynomial of degree n in some non-commutative indeterminates. Actually, these two algebras are by definition Hopf sub-algebras of the algebra of permutations of Malvenuto and Reutenauer [15], which has been realized in [3, 4] as the algebra of free quasi-symmetric functions.

This realization show that the algebra of tableaux and the algebra of binary trees are two particular case of the same construction, which is based on the existence of a Robinson-Schensted like correspondence and a plactic like monoid. This construction seems to be fundamental since there is a third example which fits in this setting: the pair of Hopf algebra in duality $(\mathbf{Sym}, QSym)$ (non-commutative symmetric functions and the quasi-symmetric functions), which corresponds to the hypoplactic monoid [9].

This paper is organized in two sections, the first one is devoted to the sylvester monoid and the binary search algorithm. In the second one, we give a new construction of the Loday-Ronco algebra as an algebra of non-commutative polynomials and a few consequences.

2 The Sylvester monoid

We start with some basic combinatorial definitions. Let $A = \{a_1 < a_2 < \cdots\}$ be a totally ordered infinite alphabet. To each word w of A^* of length n, we associate a permutations $\mathrm{std}\,(w) \in \mathfrak{S}_n$ called the standardized of w defined as the permutation obtained by iteratively scanning w from left to right, and labelling $1, 2, \ldots$ the occurrences of its smallest letter, then numbering the occurrences of the next one, and so on. Alternatively, $\sigma = \mathrm{std}\,(w)^{-1}$ can be characterized as the unique permutation of minimal length such that $w\sigma$ is a nondecreasing word. For example, $\mathrm{std}\,(abcadbcaa) = 157296834$ as shown here:

a	b	c	a	d	b	c	a	a
a_1	b_5	c_7	a_2	d_9	b_6	c_8	a_3	a_4
1	5	7	2	9	6	8	3	4

The goal of this first section is to construct a new monoid namely the sylvester monoid which is very similar to the plactic monoid. This later monoid

is closely related to the Robinson-Schensted correspondence. To stress the similarity with our construction, we just recall some basis fact, the reader which is not used to deals with tableaux and the plactic monoid should refer to [13, 14] or to the founding paper [10].

The Robinson-Schensted correspondence is a bijection

$$w \longmapsto (\mathbf{P}(w), \mathbf{Q}(w)) \tag{1}$$

between the set of words and the set of pairs $(\mathbf{P}(w), \mathbf{Q}(w))$, where $\mathbf{P}(w)$ and $\mathbf{Q}(w)$ are two young tableaux of the same shape and $\mathbf{Q}(w)$ is standard. The equivalence relation on words u and v defined by

$$u \equiv v \iff \mathbf{P}(u) = \mathbf{Q}(v) \tag{2}$$

can be shown to coincide with the plactic congruence, which is generated by Knuth's plactic relations

$$\begin{cases} aba \equiv baa\ , \quad bba \equiv bab & \text{for } a < b, \\ acb \equiv cab\ , \quad bca \equiv bac & \text{for } a < b < c. \end{cases} \tag{3}$$

The quotient of the free monoid by the plactic relation is by definition the plactic monoid, the row reading of the tableaux form a section of the later (*i.e.* in each equivalence class there is only one word which is the row reading of a tableau).

2.1 The binary search insertion

The fundamental algorithm which will play here the role of Schensted's algorithm is a well known algorithm: *the binary search insertion*. After recalling it, we will give some of its basic properties.

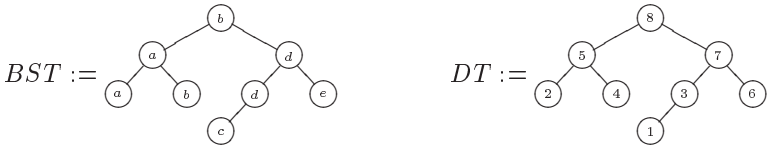
Through this paper, by *binary tree* we means (un-complete) planar rooted binary tree (*i.e.* a binary tree is either void $\emptyset$ or a pair of (possibly void) binary trees grafted on a node). The size of a binary tree is it's number of node. The number of binary trees of size n is the Catalan number $C_n := \frac{\binom{2n}{n}}{n+1}$. By labeled trees we means a tree with a label attached to each node, the label are take either from the alphabet A or from $\mathbb{N}$.

Definition 2.1 *A* binary search tree *T is a labeled binary tree such that for each node n the label of n is greater or equal than all the labels of the left subtree and strictly smaller that all the label of the right one.*

A decreasing trees *T is a labeled tree such that the label of each node is greater than the label of its child.*

A labeled trees which is labeled by the first natural number $1, \ldots, n$ where each number appear only once is called standard.

For example, here are a binary search tree and a standard decreasing tree.



Let us recall the well known binary search insertion algorithm [8]. It takes a word w of length n and return a binary search tree of size n (*i.e.* with n node). The reader have to be careful that for compatibility with the convention of [11], we decided to read the letter of the word from *right to left*. The last letter is attached to the root. Then each letter is added recursively in the left or right subtree whether it is $\leq$ or $>$ than the root. The obtained trees is denoted $\mathcal{P}(w)$.

The figure 1 shows the insertion of the word *cadbaedb*. The meaning of the second tree will be explained later.

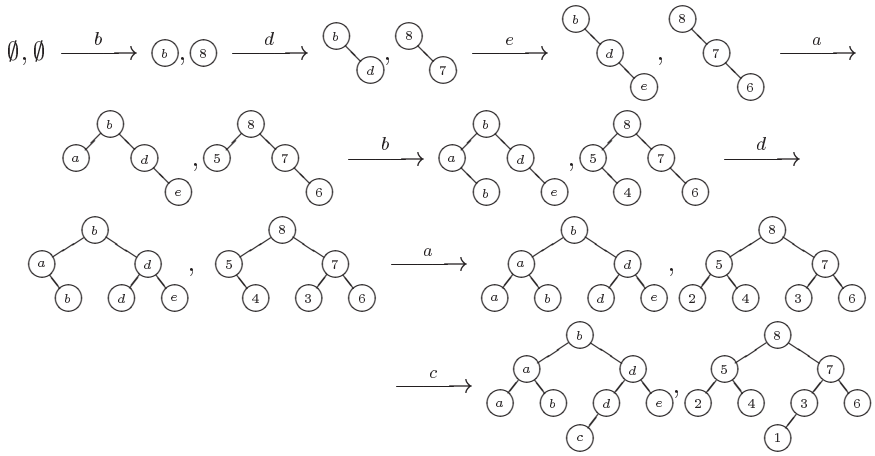


Figure 1: sylvester insertion of the word *cadbaedb*.

We want to describe the set of words that gives a given trees. Recall that the (left to right) *postfix reading* of a tree T is the word w_T obtained by reading the left subtree, then the right and finally the root. A word which is the postfix reading of a tree is called a tree-word.

Proposition 2.2 *Let T be a binary search tree and w_T its postfix reading. Then $\mathcal{P}(w_T) = T$. Moreover w_T is the smallest word w such that $\mathcal{P}(w) = T$.*

The word corresponding to the previous tree is the word $w_{BT} = abacdedb$. On the contrary the infix reading gives the sorting of the word. Note that the biggest word w such that $\mathcal{P}(w) = T$ is obtained by a right to left postfix reading, in our example we get *ecddbaab*.

2.2 The sylvester monoid

Let us first define the sylvester monoid by means of congruencies without using any trees.

Definition 2.3 *Let w_1 and w_2 two words. They are said to be sylvester congruent if there exists three words u, v, w and three letters $A \leq B < C$ such that*

$$w_1 = u\,AC\,v\,B\,w \qquad \text{and} \qquad w_2 = u\,C\,A\,v\,B\,w\,. \tag{4}$$

The sylvester equivalence is the transitive closure of the sylvester congruence. That is, two words u, v are sylvester equivalent if there exists a chain

$$u = w_1, w_2, \ldots, w_k = v \tag{5}$$

of words such that w_i and w_{i+1} are congruent for a all i. In this case we write $u \equiv_{sylv} v$.

Definition 2.4 *The* sylvester monoid $\mathrm{Sylv}(A)$ *is the quotient of the free monoid A^* by the sylvester equivalence:* $\mathrm{Sylv}(A) := A^* / \equiv_{sylv}$.

For example the class of the word 21354 is the set

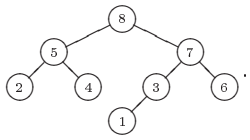
$$\{52134, 25134, 21534, 21354\}\,.$$

The fundamental theorem is the following:

Theorem 2.5 *Two words u and v are equivalent if and only if they corresponds to the same binary search tree: $\mathcal{P}(u) = \mathcal{P}(v)$.*

Consequently the postfix readings $(w_T)_T$ of the binary search trees gives a section of the sylvester monoid, that is there is one and only one tree-word in each equivalence class of $\equiv_{sylv}$.

Actually, it in possible to have not only the analogue of Schensted insertion, but also the full Robinson-Schensted correspondence. Suppose that $\sigma \in \mathfrak{S}_n$ is a permutation of $\{1, 2, \ldots, n\}$. To σ we associate a decreasing tree $\mathcal{T}(\sigma)$ as follows. The root is labeled by the biggest letter n of σ and if as a word $\sigma = unv$, then the left subtree is $\mathcal{T}(u)$ and the right subtree is $\mathcal{T}(v)$. For each $w \in A^*$, we set $\mathcal{Q}(w) = \mathcal{T}((\mathrm{std}\, w)^{-1})$. For example, with $w = cadbaedb$, one has $\mathrm{std}\,(w) = 51632874$, the inverse is $\mathrm{std}\,(w)^{-1} = 25481376$, the decreasing tree is therefore

$\mathcal{Q}(cadbaedb) =$ 

.

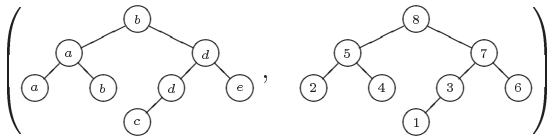
The following theorem is the analogue of the Robinson-Schensted correspondence $w \longmapsto (\mathbf{P}(w), \mathbf{Q}(w))$, where $\mathbf{P}(w)$ and $\mathbf{Q}(w)$ are two young tableaux of the same shape ($\mathbf{Q}(w)$ is standard). The trees play here the role of the tableaux [10, 14].

Theorem 2.6 *For each word $w \in A^*$, the tree $\mathcal{P}(w)$ is obtained by replacing in $\mathcal{Q}(w)$ each label i by the i-th letter of w.*

The tree $\mathcal{Q}(w)$ records the reverse order of creation of the node in the binary insertion of w.

The map $w \longmapsto (\mathcal{P}(w), \mathcal{Q}(w))$ is a bijection between A^* and the set of pairs of binary search trees and standard decreasing tree of the same shape (unlabeled tree).

The two first points are easy to see and can be checked on the following example:

cadbaedb $\longmapsto$ 

The third point is proved by giving the converse bijection: the word corresponding to a pair (BST, DT) is obtained by reading the labels of BST in the order of the label of DT. Note that, in opposition to the plactic insertion the labels does never move until they reach a leaf.

The sylvester congruence share also with the plactic congruence the following important property;

Proposition 2.7 *Suppose that u and v are to sylvester congruent word. Let $I = [a_k\, dots, a_l]$ an interval of the alphabet. Let u/I (resp. v/I) be the word obtained by erasing the letter of u (resp. v) that are not in I. Then the two word u/I and v/I are sylvester congruent.*

This is easily checked on the congruence relations.

3 A new construction of the algebra of binary trees of Loday-Ronco

The goal of this section is to give a new construction of the algebra of Loday and Ronco. This construction is very similar to the construction of the algebra of free-symmetric functions **FSym** of [4], replacing the plactic monoid by the sylvester monoid. The basis ingredient is the algebra of free quasi-symmetric functions **FQSym**.

3.1 Free quasi-symmetric functions and Free symmetric functions

The free algebra on the alphabet A is denoted by $\mathbb{Z}\langle A\rangle$. Recall that the shuffle $w_1 \sqcup\!\sqcup w_2$ of to words is defined recursively by

- $w_1 \sqcup\!\sqcup \epsilon = w_1 \qquad \text{and} \qquad \epsilon \sqcup\!\sqcup w_2 = w_2,$
- $au \sqcup\!\sqcup bv = a(u \sqcup\!\sqcup bv) + b(au \sqcup\!\sqcup v),$

where a, b are letters, u, v words and ϵ is the empty word.

Definition 3.1 *The* free quasi-ribbon $\mathbf{F}_\sigma$ *labelled by a permutation* $\sigma \in \mathfrak{S}_n$ *is the noncommutative polynomial*

$$\mathbf{F}_\sigma = \sum_{\mathrm{std}\,(w)=\sigma^{-1}} w \quad \in \mathbb{Z}\langle A \rangle \tag{6}$$

where $\mathrm{std}\,(w)$ *denotes the standardized of the word* w*.*

For example :
$\mathbf{F}_{123} = aaa + aab + abb + bbb + aac + abc + acc + bbc + bcc + ccc + \dots$
$\mathbf{F}_{213} = bab + bac + cac + cbc.$

For a word $w = x_1 x_2 \cdots x_n$ in the letters $1, 2, \dots$ and an integer k, denote by $w[k]$ the shifted word $(x_1 + k)(x_2 + k) \cdots (x_n + k)$, e.g., $312[4] = 756$. The shifted concatenation of two words u, v is defined by

$$u \bullet v = u \cdot v[k] \tag{7}$$

where k is the length of u.

Proposition 3.2 *Let* $\alpha \in \mathfrak{S}_k$ *and* $\beta \in \mathfrak{S}_l$*. Then,*

$$\mathbf{F}_\alpha \mathbf{F}_\beta = \sum_{\sigma \in \alpha \,\sqcup\!\sqcup\, \beta[k]} \mathbf{F}_\sigma \tag{8}$$

Therefore, the free quasi-ribbons span a $\mathbb{Z}$*-subalgebra of the free associative algebra called the algebra* **FQSym** *of free quasi-symmetric functions.*

This algebra is in fact a Hopf algebra, the coproduct is defined as follows: Let A' and A'' be two mutually commuting ordered alphabets. Identifying $F \otimes G$ with $F(A')G(A'')$, we set $\Delta(F) = F(A' \oplus A'')$, where $\oplus$ denotes the ordered sum. Clearly, this is an algebra homomorphism and thus this define a coproduct which is compatible with the product.

FQSym is isomorphic (and in fact dual) to the algebra of permutations of Malvenuto and Reutenauer [15]. We set also $\mathbf{G}_\sigma = \mathbf{F}_{\sigma^{-1}}$ and we define the scalar product $\langle \mathbf{F}_\sigma, \mathbf{G}_\tau \rangle = \delta_{\sigma\tau}$. This make **FQSym** a self dual Hopf algebra.

To stress the parallel, we recall the construction of the algebra of free symmetric function. Let t be a standard tableau of shape λ. Let

$$\mathbf{S}_t := \sum_{P(\sigma)=t} \mathbf{F}_\sigma = \sum_{Q(w)=t} w\,, \tag{9}$$

where $w \mapsto (P(w), Q(w))$ is the usual Robinson-Schensted map.

As pointed out in [13], Schützenberger's version of the Littlewood-Richardson rule is equivalent to the following statement, which shows that the free Schur functions span a subalgebra of **FQSym**. We will call it the algebra of *free symmetric functions* and denote it by **FSym**. It provides a realization of the algebra of tableaux introduced by Poirier and Reutenauer [16] as a subalgebra of the free associative algebra.

Proposition 3.3 (Littlewood-Richardson-Schützenberger rule)
Let t', t'' be standard tableaux, and let k be the number of cells of t'. Then,

$$\mathbf{S}_{t'}\mathbf{S}_{t''} = \sum_{t\in \mathrm{Sh}(t',t'')} \mathbf{S}_t \tag{10}$$

where $\mathrm{Sh}(t',t'')$ *is the set of standard tableaux in the shuffle of t' (regarded as a word via its row reading) with the plactic class of $t''[k]$.*

Finally is we replace the plactic monoid by the hypoplactic monoid we obtain by the very same formula the algebra of noncommutative symmetric function [9].

3.2 The binary trees algebra

Note that there is only one standard binary search tree of each shape. Therefore, any unlabeled trees T can be considered as a **P**-symbol corresponding to a permutation.

For each unlabeled trees T, we define the polynom

$$\mathbf{P}_T = \mathbf{P}_T(A) = \sum_{\mathcal{P}(\sigma)=T} \mathbf{F}_\sigma\,. \tag{11}$$

By the very same reasoning than [4], prop. 3.12, using the fact that the sylvester congruence is compatible with the restriction to intervals, one deduce the following identity

Theorem 3.4 *Let T' and T'' be two unlabeled binary trees. Then*

$$\mathbf{P}_{T'}\mathbf{P}_{T''} = \sum_{T\in \mathrm{Sh}(T',T'')} \mathbf{P}_T\,, \tag{12}$$

where $\mathrm{Sh}(T',T'')$ *is the set of trees T such that w_T occur in the shuffle product $u \sqcup\!\sqcup v$ where $u = w_{T'}$ and $v = w_{T''}[k]$ is the tree-word of T'' shifted by the number k of node of T'.*

This statement analogs of the "free Littlewood-Richardson rule" [14, 4] gives a new construction of the binary trees algebra of Loday and Ronco. In [12] the product $\mathbf{P}_{T'}\mathbf{P}_{T''}$ is describe by means of an order on the binary trees. This order can be obtained from the weak ordering of the symmetric group" it is the restriction of it to trees-words. It should be noticed that the results of Björner and Wachs [2] shows that the sylvester class are intervals for the weak order.

Hence, the set $\mathrm{Sh}(T', T'')$ is an interval for the order of Loday and Ronco. It the case of the tableaux, it is also possible to define such an order quotient of the weak order by Knuth relation. The sets $\mathrm{Sh}(t', t'')$ are also intervals for this order.

The cardinal of a plactic class of permutations is equal to the number of standard tableaux of a certain shape which is given by the celebrated hook formula [7]. This formula admits a q-analog which count the permutations by their major index [18]. In the same way, the counting of the sylvester class correponding to the tree T with n nodes is given by the specialization $(q)_n \mathbf{P}_T(1, q, q^2, \cdots)$ which is equal from [1] to

$$\sum_{\mathcal{P}(\sigma)=T} q^{\mathrm{maj}\,(\sigma)} = (q)_n \mathbf{P}_T(1, q, q^2, \cdots) = \frac{[n]_q!}{\prod_{\circ \in T} q^{\delta_\circ} [h_\circ]_q}\,, \tag{13}$$

where for a node $\circ$ of T, the coefficient $h_\circ$ is the size of the subtree rooted at $\circ$ and $\delta_\circ$ the size of its right subtree. Here we use the standard notations of the q-calculus, that is, for any integer n, the q-integer $[n]_q$ is equal to $1 + q + \cdots + q^{n-1}$, the q-factorial $[n]!_q$ is the product of the corresponding q-integer: $[n]!_q := [1]_q [2]_1 \ldots [n]_q$, and finally

$$(q)_n := \frac{[n]!_q}{(1-q)^n} = (1-q)\left(1-q^2\right)\ldots\left(1-q^n\right)\,. \tag{14}$$

3.3 Duality

The algebra of binary trees in not only a sub-algebra but also a sub Hopf algebra of **FQSym**. This is a consequence of the fact that the sylvester congruence is compatible with the standardization, (*i.e.* two words u and v with the same commutative image are congruent if and only if their standardized are congruent).

The sylvester congruence give also a nice characterization of the Hopf-dual of the algebra of binary trees:

Theorem 3.5 *The (graded) dual of the algebra of binary trees is isomorphic to the image of* **FQSym** *by the canonical projection*

$$\pi : \ \mathbb{C}\langle\, A\,\rangle \longrightarrow \mathbb{C}[\mathrm{Sylv}\,(A)] \simeq \mathbb{C}\langle\, A\,\rangle / \equiv\ . \tag{15}$$

the dual basis $\mathbf{Q}_T$ *of* $\mathbf{P}_T$ *is* $\mathbf{Q}_T = \pi(\mathbf{G}_{w_T})$*, where* w_T *is the permutation canonically associated to the tree* T*.*

It is easy to deduce that the algebra is in fact self dual (which is also a consequence of [17]). The dual of **FSym** admit also a similar description [16, 5]. These result are consequence of the "Free cauchy formula"

$$\sum_{\mathrm{std}\,(u)=\mathrm{std}\,(v)^{-1}} u \otimes v = \sum_{\sigma} \mathbf{U}_\sigma \otimes \mathbf{V}_\sigma$$

for all pair $(\mathbf{U}, \mathbf{V})$ of adjoint base of **FQSym**.

Finally, one can build a pair of graded graphs (Γ, Γ^*) in duality in Fomin's sense [6], whose vertex of degree n are the binary trees with n nodes. In Γ there is an edge between T and T' if $\mathbf{P}_{T'}$ belong to the expansion $\mathbf{P}_T\mathbf{P}_\bullet$ (the dot $\bullet$ is the tree with only one node). In Γ^*, this edge appears if $\mathbf{Q}_{T'}$ belongs to $\mathbf{Q}_T\mathbf{Q}_\bullet$. The sylvester correspondence is the Fomin correspondence associated to this pair of graphs.

References

[1] A. Björner, M. Wachs, q-Hook length formulas for forests, J. Combin. Theory Ser. A 52 (1989), 165–187.

[2] A. Björner, M. Wachs, Permutation statistics and linear extensions of posets, J. Combin. Theory Ser. A 58 (1991), 85–114.

[3] G. Duchamp, F. Hivert, J.-Y. Thibon, Une généralisation des fonctions quasi-symétriques et des fonctions symétriques non commutatives, C. R. Acad. Sci. Paris Sér. I Math. 328 (1999), no. 12, 1113–1116.

[4] G. Duchamp, F. Hivert, J.-Y. Thibon, Noncommutative symmetric functions VI: free quasi-symmetric functions and related algebras, Internat. J. Alg. Comput. (à paraître).

[5] G. Duchamp, F. Hivert, J.-Y. Thibon, Noncommutative symmetric functions VII (en préparation).

[6] S. Fomin, Duality of graded graphs, J. Alg. Combinatorics 3 (1994), 357–404.

[7] J.S. Frame, G. de B. Robinson, R.M. Thrall, The hook graphs of the symmetric groups, Canadian J. Math. 6, (1954). 316–324.

[8] D.E. Knuth, The art of computer programming, vol.3: Sorting and searching, Addison-Wesley, 1973.

[9] D. Krob, J.-Y. Thibon, Noncommutative symmetric functions IV: Quantum linear groups and Hecke algebras at $q = 0$, J. Algebraic Combin. 6 (1997), no. 4, 339–376.

[10] A. Lascoux, M.-P. Schützenberger, Le monoïde plaxique, Noncommutative structures in algebra and geometric combinatorics (Naples, 1978), pp. 129–156, Quad. Ricerca Sci., 109, CNR, Rome, 1981.

[11] J.-L. Loday, M.O. Ronco, Hopf algebra of the planar binary trees, Adv. Math. 139 (1998), no. 2, 293–309.

[12] J.-L. Loday, M.O. Ronco, Order structure on the algebra of permutations and of planar binary trees, J. Algebraic Combin. 15 (2002), no. 3, 253-270.

[13] A. Lascoux, B. Leclerc and J.-Y. Thibon, *The plactic monoid*, Chapter 5 of M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, to appear.

[14] M. Lothaire, Algebraic combinatorics on words, Cambridge University press, Cambridge, 2002.

[15] C. Malvenuto, C. Reutenauer, Duality between quasi-symmetric functions and Solomon descent algebra, J. Algebra 177 (1995), 967–892.

[16] S. Poirier, C. Reutenauer, Algèbres de Hopf de tableaux, Ann. Sci. Math. Québec 19 (1995), no. 1, 79–90.

[17] M.O. Ronco, Primitive elements in a free dendriform Hopf algebra, Contemp. Maths. Vol. 267 (2000), 245–264.

[18] R.P. Stanley, Ordered structures and partitions, Memoirs of the American Mathematical Society, No. 119, American Mathematical Society, Providence, R.I., 1972.