Theoretical Computer Science 409 (2008) 382–393

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

The pruning-grafting lattice of binary trees

J.L. Baril *, J.M. Pallo

LE2I, UMR 5158, Université de Bourgogne, B.P. 47870, F21078 DIJON-Cedex, France

ARTICLE INFO

Article history:
Received 6 April 2007
Received in revised form 4 June 2008
Accepted 25 August 2008
Communicated by M. Crochemore

Keywords:
Lattices
Binary trees
Feasible sequences
Distribution sequences
Catalan numbers

ABSTRACT

We introduce a new lattice structure B_n on binary trees of size n. We exhibit efficient algorithms for computing meet and join of two binary trees and give several properties of this lattice. More precisely, we prove that the length of a longest (resp. shortest) path between **0** and **1** in B_n equals to the Eulerian numbers $2^n-(n+1)$ (resp. $(n-1)^2$) and that the number of coverings is $\binom{2n}{n-1}$. Finally, we exhibit a matching in a constructive way. Then we propose some open problems about this new structure.

1. Introduction

This paper introduces a new operation on rooted, ordered, binary trees called the pruning-grafting transformation. We show that this transformation endows the set of binary trees with a new lattice structure. This is a new attempt to define a lattice structure on a Catalan set, i.e. a set of combinatorial objects enumerated by Catalan numbers.

The set B_n of binary trees with n internal nodes related by rotations is a lattice, the so-called n-th Tamari lattice [24]. These lattices have been extensively studied for algebraic and combinatorial purposes. A number of references on this subject are available in [28]. Like rotation, our pruning-grafting transformation is a simple and natural operation on binary trees. But as for rotation, the characterization of this transformation is unfortunately rather complex.

The idea developed in this paper is both similar to, and different from the tool introduced by Parker and Ram in [30]. Similar because our pruning-grafting operation resembles their balancing exchange. But different because binary trees used in [30] for constructing Huffman codes are *unordered*, whereas the binary trees in our paper are *ordered*.

Ordered binary trees of B_n are enumerated by the well-known n-th Catalan number $\binom{2n}{n}/(n+1)$. A large number of various classes of combinatorial objects are Catalan sets. It is the case, among others, of ballot sequences, planar trees, Young tableaux, nonassociative products, stack sortable permutations, and so on. A list of over 60 types of such combinatorial classes of independent interest has been compiled by Stanley [39, p. 219]. A certain number of explicit bijections between these Catalan classes can also be found in the literature.

Of much of interest are the following lattices. First the uppermentioned Tamari lattices can be obtained equivalently in two other ways. The coverings correspond to reparenthesizations of letters products [14] and to diagonal flips in triangulations [4,19,37]. Then the lattice of noncrossing partitions, the so-called Kreweras lattice, is equipped with the refinement order [21]. See [11,12,31,35] and numerous references in the exhaustive survey [36]. Also, Dyck words [3,17] have been endowed by lattice structures. In [1,13] the authors studied the Stanley lattices in term of Dyck paths.

The paper is organized as follows. First we recall some classical definitions on binary trees. Then we show that B_n equipped with the pruning-grafting transformation is a lattice. We exhibit an efficient algorithm for computing the meet and join of two binary trees. We study some properties of this lattice and conclude by some open problems.

* Corresponding author. Tel.: +33 0380393801.
E-mail addresses: barjl@u-bourgogne.fr (J.L. Baril), pallo@u-bourgogne.fr (J.M. Pallo).

0304-3975/$ – see front matter
doi:10.1016/j.tcs.2008.08.031

2. Distribution sequences of binary trees

A *binary tree* is a rooted, ordered, unlabeled tree in which every node is either a leaf $\square$ (i.e. a node without child) or an internal node $\bigcirc$ having two children. We denote by B_n the set of binary trees with n internal nodes (and thus with $n+1$ leaves). It is well-known that $\mathrm{card}(B_n) = |B_n|$ equals the Catalan number $\binom{2n}{n}/(n+1)$. The set B of binary trees is recursively defined by $B = \square + \bigcirc BB$ in Polish or linear notation. It is well known that a sequence with n circles and $(n+1)$ squares corresponds to the Polish notation of a tree of B_n iff in every proper prefix of this sequence, the number of circles is greater or equal to the number of squares, the last one being a square. Let T_L and T_R be the left and right subtrees of $T \in B_n$. Thus, we can write $T = \bigcirc T_L T_R$. The leaves of a binary tree T are numbered by a preorder traversal (i.e. from left to right). The order of leaves is *significant*. The *feasible sequence* of T (also called *path-length sequence*) is the integer sequence $\ell_T = (\ell_T(1), \ell_T(2), \dots, \ell_T(n+1))$ where $\ell_T(i)$ is the *level number* of the i-th leaf, i.e. the length of unique path between the i-th leaf and the root [34]. Let also $\bar{\ell}$ be the mirror sequence $\bar{\ell} = (\ell_T(n+1), \ell_T(n), \dots, \ell_T(1))$ which is the feasible sequence of the mirror tree $\bar{T}$ recursively defined by $\bar{T} = \bigcirc \bar{T}_R \bar{T}_L$ and $\bar{\square} = \square$. For example the tree defined in Polish notation by $\bigcirc\square\bigcirc\bigcirc\square\bigcirc\square\bigcirc\square\square\bigcirc\bigcirc\square\square\square$ has the feasible sequence $(1, 3, 4, 5, 5, 4, 4, 3)$. Its mirror is $\bigcirc\bigcirc\bigcirc\square\bigcirc\square\square\bigcirc\bigcirc\bigcirc\square\square\square\square\square$ and has the feasible sequence $(3, 4, 4, 5, 5, 4, 3, 1)$. Now let us consider a sequence $(a(1), a(2), \dots, a(n+1))$ with an integer $k \in [2, n+1]$ such that $a(k-1) = a(k) = q$. The process of replacing the pair $(a(k-1), a(k)) = (q, q)$ by $q-1$ in order to get the new sequence $(a(1), a(2), \dots, a(k-2), q-1, a(k+1), \dots, a(n+1))$ is called a *reduction*. Ruskey and Hu [34] give the following necessary and sufficient condition so that an $(n+1)$-length sequence represents a binary tree of B_n. A sequence $(a(1), a(2), \dots, a(n+1))$ is feasible iff a series of n reductions from the left or right (in any order) reduce the original sequence to the single integer 0.

Moreover, the feasible sequence of a binary tree obeys what we call Kraft equality, a special case of the Kraft inequality of noiseless coding theory (see [15], p. 45). A necessary condition [20, p. 404, ex. 3] for a sequence $(a(1), a(2), \dots, a(n+1))$ to be feasible is that:

$$\sum_{i=1}^{n+1} 2^{-a(i)} = 1.$$

Lemma 1 (*[20, p. 404, ex. 3]*). *Let $\ell_T = (\ell_T(1), \ell_T(2), \dots, \ell_T(n+1))$ be a feasible sequence of a binary tree T and let T' be a subtree of T where the root of T' is at level k in T. If $m_1 < m_2$ and $(\ell_T(m_1), \ell_T(m_1+1), \dots, \ell_T(m_2-1), \ell_T(m_2))$ is the subsequence of ℓ_T corresponding to the leaves of T' then*

$$\sum_{i=m_1}^{m_2} 2^{-\ell_T(i)} = 2^{-k}.$$

The sequence $2^{-\ell_T} = (2^{-\ell_T(1)}, 2^{-\ell_T(2)}, \dots, 2^{-\ell_T(n+1)})$ is called the *density sequence* of the binary tree (the sum of its entries is one). So we define its associated *distribution sequence* as the ascending sequence:

$$L_T = \left(2^{-\ell_T(1)}, 2^{-\ell_T(1)} + 2^{-\ell_T(2)}, \dots, \sum_{j=1}^{i} 2^{-\ell_T(j)}, \dots, 1\right).$$

In the sequel of this paper, the feasible sequences are denoted by lowercase letters (ℓ for instance) and their distribution sequences by uppercase letters (L). We assume that $\leq$ is the usual ordering on two n-length sequences, i.e. if $\ell = (\ell(1), \dots, \ell(n+1))$ and $\ell' = (\ell'(1), \dots, \ell'(n+1))$, we say that $\ell \leq \ell'$ if $\ell(i) \leq \ell'(i)$ for all $i \in [1, n+1]$.

3. The pruning-grafting transformation

In this section, we define a new transformation on binary trees which can be characterized using distribution sequences. This induces a new structure lattice on B_n.

Definition 1. The *pruning-grafting transformation* $\rightarrow$ on B_n with $n \geq 2$ is defined by the covering $T \rightarrow T'$ if there exist $k \geq 1$, $\tau_1 \in \bigcirc\{\bigcirc, \square\}^*$ and $\tau_2 \in \{\bigcirc, \square\}^*$ such that $T = \tau_1 \square \bigcirc^k \square\square \tau_2$ and $T' = \tau_1 \bigcirc \square\square \bigcirc^{k-1} \square \tau_2$. Let $\overset{*}{\rightarrow}$ be the reflexive transitive closure of $\rightarrow$ on B_n.

The pruning-grafting transformation prunes by replacing a "little" subtree $T_1 = \bigcirc\square\square$ of T by a leaf and grafts this subtree $T_1 = \bigcirc\square\square$ instead of the leaf $\square$ just before T_1 in the Polish notation of T. See Fig. 1 for instance.

Proposition 1. *Let $T, T' \in B_n$ and $\ell_T, \ell_{T'}$ be their corresponding feasible sequences. Then we have $T \rightarrow T'$ iff there exist $p \geq 1$, $q \geq 2$ and $2 \leq i \leq n$ such that*

$$\begin{aligned} \ell_T &= (\ell_T(1), \ell_T(2), \dots, \ell_T(i-2), p, q, q, \ell_T(i+2), \dots, \ell_T(n+1)) \text{ and} \\ \ell_{T'} &= (\ell_T(1), \ell_T(2), \dots, \ell_T(i-2), p+1, p+1, q-1, \ell_T(i+2), \dots, \ell_T(n+1)). \end{aligned}$$

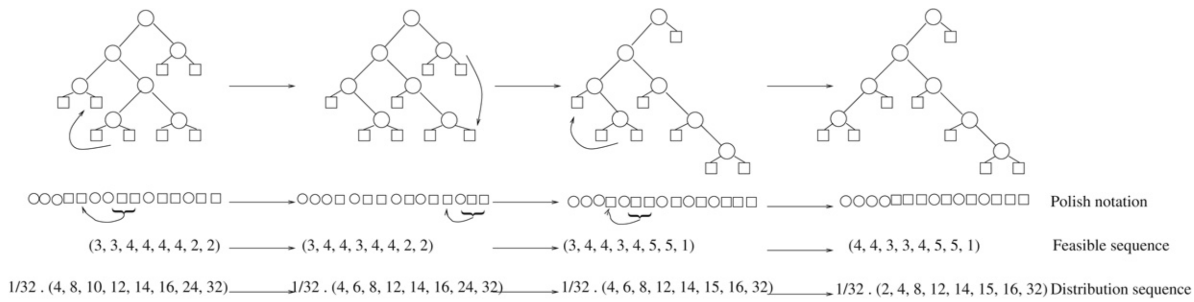

Fig. 1. Three pruning-grafting transformations in B_7.

Proof. One can easily prove that the condition is necessary. Conversely, let us assume that the feasible sequences of T and T' are $\ell_T = (\ell_T(1), \ell_T(2), \ldots, \ell_T(i-2), p, q, q, \ell_T(i+2), \ldots, \ell_T(n+1))$ and $\ell_{T'} = (\ell_T(1), \ell_T(2), \ldots, \ell_T(i-2), p+1, p+1, q-1, \ell_T(i+2), \ldots, \ell_T(n+1))$.
We distinguish between two cases.

(1) The i-th leaf of T is a left leaf. Since the i-th and $(i+1)$-th leaves have the same level number q, the Polish notation of T is $\tau \square \bigcirc^k \underset{i\text{-th}}{\square} \underset{(i+1)\text{-th}}{\square} \ldots$ where $\tau \in \bigcirc\{\bigcirc, \square\}^*$ and $k \geq 1$. Thus, T' is obtained from T by a pruning-grafting transformation.

(2) Assume that the i-th leaf of T is a right leaf. Since the i-th and $(i+1)$-th leaves have the same level number q, T can be written $T = \tau \square \underset{i\text{-th}}{\square} \bigcirc^k \underset{(i+1)\text{-th}}{\square} \ldots$ where $\tau \in \bigcirc\{\bigcirc, \square\}^*$ and $k \geq 1$. Since the $(i-1)$-th and i-th leaves of T' have level $p+1$, T' is of the form $\tau \bigcirc \square \underset{i\text{-th}}{\square} \ldots$. Moreover, the $(i+1)$-th leaf of T' is attached to a node of τ that also is the parent of the i-th leaf in T. This would mean that the level number of the $(i+1)$-th leaf of T' is at least q which contradicts the hypothesis. $\square$

Corollary 1. *Let $T, T' \in B_n$ and L_T, $L_{T'}$ their distribution sequences. Then we have $T \to T'$ iff there exist $p \geq 1$, $q \geq 2$ and $2 \leq i \leq n$ such that $L_T = (L_T(1), \ldots, L_T(i-2), L_T(i-2) + 2^{-p}, L_T(i-2) + 2^{-p} + 2^{-q}, L_T(i+1), \ldots, 1)$ and $L_{T'} = (L_T(1), \ldots, L_T(i-2), L_T(i-2) + 2^{-(p+1)}, L_T(i-2) + 2^{-p}, L_T(i+1), \ldots, 1)$.*

Definition 2. Let T and T' be two different trees of B_n. Let $\sigma_1 \in \bigcirc\{\bigcirc, \square\}^*$ be the longest common prefix of T and T' in their Polish notation. Let us assume that

$$T = \sigma_1 \sigma_2 \square \bigcirc^j \underbrace{\bigcirc \square \square}_{\tau} \sigma_3$$

where

- $j \geq 0$, and
- σ_2 is the empty word or $\sigma_2 \in \square\{\bigcirc, \square\}^*$ and $\sigma_2 \square$ does not contain any occurrence of $\tau = \bigcirc \square \square$ (i.e. $\tau = \bigcirc \square \square$ is the leftmost occurrence of $\bigcirc \square \square$ in T after σ_1), and
- $\sigma_3 \in \{\bigcirc, \square\}^*$.

Then we necessarily have $T' = \sigma_1 \sigma_2'$ with $\sigma_2' \in \bigcirc\{\bigcirc, \square\}^*$.
We define the tree $U(T, T') \in B_n$ (U for up) such that $T \to U(T, T')$ and

$$U(T, T') = \sigma_1 \sigma_2 \underbrace{\bigcirc \square \square}_{\tau} \bigcirc^j \square \sigma_3.$$

Moreover, we set $U(T, T) = T$.

For example, if:

$$T = \bigcirc \bigcirc \square \square \bigcirc \square \bigcirc \bigcirc \square \bigcirc \bigcirc \square \square \square \square$$
$$T' = \bigcirc \bigcirc \square \square \bigcirc \bigcirc \square \square \bigcirc \square \bigcirc \bigcirc \square \square \square \square$$

then $\sigma_1 = \bigcirc \bigcirc \square \square \bigcirc$ $\sigma_2 = \square \bigcirc \bigcirc$, $\sigma_3 = \square \square$ and $j = 1$. We obtain $U(T, T') = \bigcirc \bigcirc \square \square \bigcirc \square \bigcirc \bigcirc \bigcirc \square \square \bigcirc \square \square \square$.

Lemma 2. *Let T and T' be two trees in B_n such that their distribution sequences L_T and $L_{T'}$ verify $L_T > L_{T'}$. Thus, $U(T, T')$ exists and its distribution sequence $L_{U(T,T')}$ verifies $L_T > L_{U(T,T')} \geq L_{T'}$.*

Proof. Via Definition 2, $U(T, T')$ exists since $L_T > L_{T'}$ implies $T = \sigma \square \nu$ and $T' = \sigma \bigcirc \nu'$ with $\sigma \in \bigcirc\{\bigcirc, \square\}^*$ and $\nu, \nu' \in \{\bigcirc, \square\}^*$. Let us prove the inequality $L_U \geq L_{T'}$ where $L_U = L_{U(T,T')}$. Let $r = \min\{i \geq 1 | L_T(i) > L_{T'}(i)\}$ and let t be the smallest $i > r$ such that T is of the form $T = \ldots \bigcirc \underset{i\text{-th}}{\square} \underset{(i+1)\text{-th}}{\square} \ldots$. Let T_0 be the smallest subtree in T containing the r-th and t-th leaves. In Polish notation let $T_0 = \bigcirc T_1 T_2$. The r-th leaf belongs to T_1 and the t-th belongs to T_2. The rightmost leaf of T_1 is numbered by $s \geq r$ in T. With these hypotheses, we necessarily have the following properties:

(a) $\ell_T(r) \geq \ell_T(r+1)$ and $\ell_T(i) > \ell_T(i+1)$ for $r+1 \leq i \leq s-1$;
(b) $\ell_T(i) < \ell_T(i+1)$ for $s+1 \leq i \leq t-1$.

As ℓ_T and $\ell_{T'}$ have the same entries for $i \leq r-1$, we have $L_T(r-1) = L_{T'}(r-1)$. Moreover, $L_T(r) > L_{T'}(r)$ implies that $\ell_T(r) < \ell_{T'}(r)$ and Lemma 1 shows that there exists $r_0, r_0 \geq r+1$, such that $L_T(r-1) + 2^{-\ell_T(r)} = L_T(r-1) + \sum_{i=r}^{r_0} 2^{-\ell_{T'}(i)}$, i.e. $L_T(r) = L_{T'}(r_0)$.

In the case where r verifies $r+1 \leq s$, then the $(r+1)$-th leaf of T is attached to a node of the common prefix of T and T'. Thus, we necessarily have $\ell_T(r+1) \leq \ell_{T'}(r_0+1)$. Therefore, there exists r_1 with $r < r_0 < r_1$, such that $L_T(r+1) = L_{T'}(r_1)$. By repeating this process until reaching the s-th leaf, we prove that there exists $s' \geq s+1$ such that $L_T(s) = L_{T'}(s')$.

Now we first want to show that there exists $s' \geq t$ such that $L_{T'}(s') = L_T(s)$ and $s \leq t-1$.

If $s' < t$, the hypothesis gives us $L_{T'}(s'+1) \leq L_T(s'+1)$. This means that $2^{-\ell_{T'}(s'+1)} \leq 2^{-\ell_T(s+1)} + \cdots + 2^{-\ell_T(s'+1)}$. Since $s+1 \leq s' < t$, the inequalities follow $\ell_T(s+1) < \ell_T(s+2) < \cdots < \ell_T(s'+1)$ and we obtain $2^{-\ell_{T'}(s'+1)} \leq 2^{-\ell_T(s+1)} + 2^{-\ell_T(s+2)-1} < 2^{-\ell_T(s+1)+1}$. Thus, $\ell_{T'}(s'+1) \geq \ell_T(s+1)$ and there exists $s'' \geq s'+1$ such that $L_{T'}(s'') = L_T(s+1)$. If $s'' \leq t-1$, we repeat this process for the index $s+2$ and so on, and we stop it when s'' holds $s'' \geq t$.

So, we have obtain s' such that $s \leq t-1$ and $s' \geq t$ with $L_{T'}(s') = L_T(s)$. Notice that if $s = t-1$, then we necessarily have $L_T(t-2) = L_{T'}(s)$.

Since $L_{T'}(i) \leq L_T(i)$ for all i, we obtain $L_{T'}(i) \leq L_T(i) = L_U(i)$ for $i \leq t-2$ and $i \geq t+1$.

In the case where $i = t-1$, we obtain

$$\begin{aligned} L_U(t-1) - L_{T'}(t-1) &= L_T(t-1) - 2^{-\ell_T(t-1)+1} - L_{T'}(t-1) \\ &= L_T(t-1) - L_T(t-2) + L_T(t-2) - 2^{-\ell_T(t-1)+1} - L_{T'}(t-1) \\ &= 2^{-\ell_T(t-1)+1} + L_T(t-2) - L_{T'}(t-1) \\ &\geq 2^{-\ell_T(t-1)+1} + L_T(s) - L_{T'}(t-1) \\ &\geq 2^{-\ell_T(t-1)+1} + L_T(s) - L_{T'}(s') \geq 2^{-\ell_T(t-1)+1}. \end{aligned}$$

If $i = t$, we have

$$\begin{aligned} L_U(t) - L_{T'}(t) &= L_T(t-1) - L_{T'}(t) \\ &\geq L_T(s) - L_{T'}(t) \\ &\geq L_T(s) - L_{T'}(s') \geq 0. \quad \square \end{aligned}$$

Definition 3. Let T and T' be two different trees of B_n with $n \geq 2$. Let $\sigma_1 \in \bigcirc\{\bigcirc, \square\}^*$ be their longest common prefix in their Polish notation. Let us assume that $T = \sigma_1 \square \sigma_2$ with $\sigma_2 \in \{\bigcirc, \square\}^*$. Then we necessarily have

$$T' = \sigma_1 \sigma_2' \underbrace{\bigcirc \square \square}_{\tau} \bigcirc^j \square \sigma_3'$$

where

- $j \geq 0$, and
- σ_2' is the empty word or $\sigma_2' \in \bigcirc\{\bigcirc, \square\}^*$ such that σ_2' does not contain any occurrence $\tau = \bigcirc \square \square$ (i.e. τ is the leftmost occurrence of $\bigcirc \square \square$ in T' after σ_1), and
- $\sigma_3' \in \{\bigcirc, \square\}^*$.

We define the tree $D(T, T') \in B_n$ (D for down) such that $D(T, T') \to T'$ and

$$D(T, T') = \sigma_1 \sigma_2' \square \bigcirc^j \underbrace{\bigcirc \square \square}_{\tau} \sigma_3'.$$

Moreover, we set $D(T, T) = T$.

For example, if:

$$\begin{aligned} T &= \bigcirc \bigcirc \square \square \bigcirc \square \bigcirc \square \bigcirc \bigcirc \square \bigcirc \square \square \square \\ T' &= \bigcirc \bigcirc \square \square \bigcirc \bigcirc \square \bigcirc \square \square \bigcirc \square \bigcirc \square \square \end{aligned}$$

then $\sigma_1 = \bigcirc \bigcirc \square \square \bigcirc$, $\sigma_2' = \bigcirc \square$, $\sigma_3' = \bigcirc \square \square$ and $j = 1$. We obtain $D(T, T') = \bigcirc \bigcirc \square \square \bigcirc \bigcirc \square \square \bigcirc \bigcirc \square \square \bigcirc \square \square$.

Lemma 3. *Let T and T' be two trees in B_n such that their distribution sequences L_T and $L_{T'}$ verify $L_T > L_{T'}$. Thus, $D(T, T')$ exists and its distribution sequence $L_{D(T,T')}$ verifies $L_T \geq L_{D(T,T')} > L_{T'}$.*

Proof. Via Definition 3, $D(T, T')$ exists since $L_T > L_{T'}$ implies $T = \sigma \square \nu$ and $T' = \sigma \bigcirc \nu'$ with $\sigma \in \bigcirc\{\bigcirc, \square\}^*$ and $\nu, \nu' \in \{\bigcirc, \square\}^+$. Let us prove the inequality $L_D \leq L_T$ where $L_D = L_{D(T,T')}$. Let $r = \min\{i \geq 1 | L_T(i) > L_{T'}(i)\}$ and let s be the smallest $i \geq r$ such that T' is of the form $T' = \sigma \bigcirc \ldots \bigcirc \underset{i\text{-th}}{\square} \underset{(i+1)\text{-th}}{\square} \ldots$ where σ is the common prefix of T and T'. Since L_D differs from $L_{T'}$ at indices s and $s+1$, the inequality $L_D(i) = L_{T'}(i) \leq L_T(i)$ holds for $i \notin \{s, s+1\}$. Now let us examine $L_D(i) - L_T(i)$ for $i \in \{s, s+1\}$. Using Lemma 1, we deduce easily that $L_T(r) \geq L_{T'}(s+1)$. This provides $L_D(s) - L_T(s) = L_{T'}(s+1) - L_T(s) \leq L_T(r) - L_T(s) \leq 0$.

For the index $s+1$, $L_D(s+1) - L_T(s+1) = L_{T'}(s+2) - L_T(s+1) - 2^{-(\ell_{T'}(s+2)+1)}$. We distinguish two cases:

(1) if there exists $k \le s+1$ such that $L_{T'}(s+2) \le L_T(k)$ then we clearly have $L_D(s+1) - L_T(s+1) \le 0$;

(2) otherwise (i.e. $k \ge s+2$), $L_{T'}(s+2) > L_T(s+1)$. Thus, $2^{-\ell_{T'}(s+2)} > \sum_{i=r+1}^{s+1} 2^{-\ell_T(i)}$ and $\forall\, i \in [r+1, s+1]$, $2^{-\ell_T(i)} < 2^{-\ell_{T'}(s+2)}$ which means $\ell_T(i) \ge \ell_{T'}(s+2)+1$, $\forall\, i \in [r+1, s+1]$. This implies that there exists $k \ge r+1$ such that $L_T(k) = L_{T'}(s+2)$. The hypothesis $L_{T'} < L_T$ induces $k \le s+2$ and with $k \ge s+2$ we obtain $k = s+2$, i.e. $L_{T'}(s+2) = L_T(s+2)$.

Therefore, $L_{T'}(s+1) + 2^{-\ell_{T'}(s+2)} = L_T(s+1) + 2^{-\ell_T(s+2)}$ and since $L_{T'}(s+1) \le L_T(r) \le L_T(s) < L_T(s+1)$, we obtain $2^{-\ell_T(s+2)} < 2^{-\ell_{T'}(s+2)}$ (i.e. $\ell_T(s+2) \ge \ell_{T'}(s+2)+1$). Thus,

$$\begin{aligned} L_D(s+1) - L_T(s+1) &= L_{T'}(s+2) - L_T(s+1) - 2^{-(\ell_{T'}(s+2)+1)} \\ &= L_T(s+2) - L_T(s+1) - 2^{-(\ell_{T'}(s+2)+1)} \\ &= 2^{-\ell_T(s+2)} - 2^{-(\ell_{T'}(s+2)+1)} \le 0. \quad \square \end{aligned}$$

Theorem 1. *Giving T and $T' \in B_n$. Let L_T and $L_{T'}$ be their distribution sequences. Then we have $T \xrightarrow{*} T'$ iff $L_T \ge L_{T'}$.*

Proof. Let us assume that $T \xrightarrow{*} T'$, i.e. there exists a path $T = T_0 \to T_1 \to T_2 \to \cdots \to T_k = T'$ ($k \ge 1$). Via Corollary 1, the distribution sequences L_{T_i} of T_i verify: $L_T \ge L_{T_1} \ge L_{T_2} \ge \cdots \ge L_{T_{k-1}} \ge L_{T'}$. Thus, we obtain $L_T \ge L_{T'}$.

Conversely, assume that L_T and $L_{T'}$ verify $L_T \ge L_{T'}$ with $T \ne T'$. Let $i_0 \ge 1$ be the smallest integer i such that $L_T(i) > L_{T'}(i)$. There are two cases to consider:

(1) there doesn't exist any internal node $\bigcirc$ after the i_0-th leaf in T. Since $\ell_T(i) = \ell_{T'}(i)$ for $i < i_0$ and $\ell_T(i_0) < \ell_{T'}(i_0)$, the Polish notation of T and T' can be written as $T = \sigma \square^k$ for $k \ge 1$, $\sigma \in \bigcirc\{\bigcirc, \square\}^*$ and $T' = \sigma \bigcirc \ldots$ where σ is the longest common prefix of T and T'. Thus, the number of nodes of T' is greater than those of T which is a contradiction

(2) there exists an internal node $\bigcirc$ after the i_0-th leaf in T. This means that there exists an occurrence of $\bigcirc\square\square$ after the i_0-th leaf in T. Via Lemma 2, we have $T \to U(T, T')$ and if L_U is the distribution sequence of $U(T, T')$ then $L_T > L_U \ge L_{T'}$.

By iterating this process, we obtain $T \xrightarrow{*} T'$. $\square$

Theorem 2. *For all n, the poset $(B_n, \xrightarrow{*})$ is a lattice.*

Proof. It suffices to show that any two elements of B_n have a least upper bound. The existence of greatest lower bound then follows automatically since B_n is finite with as least element $\mathbf{0}$ and as greatest element $\mathbf{1}$ defined by $\ell_{\mathbf{0}} = (1, 2, \ldots, n-1, n, n)$ and $\ell_{\mathbf{1}} = (n, n, n-1, \ldots, 2, 1)$. Let T and T' be two different trees in B_n and L_T and $L_{T'}$ their distribution sequences. In this proof, we construct a tree $S \in B_n$ that is candidate to be the join element of T and T' and we show that this element is really the join.

Given τ and τ' in B_n, we define the following function **join**(τ,τ'):

```
Function join(τ,τ')
begin
    while τ ≠ τ' do
        i0 := min{i ∈ [1, n] | Lτ(i) ≠ Lτ'(i)}
        if Lτ(i0) < Lτ'(i0) then
            ν' := U(τ', τ)  (thus τ' → ν')
            τ' := ν'
        else
            ν := U(τ, τ')  (thus τ → ν)
            τ := ν
        endif
    enddo
    return τ
end
```

For example, if we perform this function for $\tau = (2, 2, 3, 3, 3, 3)$ and $\tau' = (1, 5, 5, 4, 3, 2)$ (i.e. $L_\tau = \frac{1}{32}(8, 16, 20, 24, 28, 32)$ and $L_{\tau'} = \frac{1}{32}(16, 17, 18, 20, 24, 32)$), then $i_0 = 1$ and $L_\tau(1) < L_{\tau'}(1)$; we replace τ' by $U(\tau', \tau) = (2, 2, 4, 4, 3, 2)$ that have the distribution sequence $\frac{1}{32}(8, 16, 18, 20, 24, 32)$; $i_0 = 3$ and $L_\tau(3) > L_{\tau'}(3)$, we replace τ by $U(\tau, \tau') = (2, 2, 3, 4, 4, 2)$ whose distribution sequence is $\frac{1}{32}(8, 16, 20, 22, 24, 32)$. Yet, $i_0 = 3$, and $L_\tau(3) > L_{\tau'}(3)$, we replace τ by $U(\tau, \tau') = (2, 2, 4, 4, 3, 2)$ which implies that $\tau = \tau'$. The function returns $\tau = (2, 2, 4, 4, 3, 2)$.

In order to construct the tree $S = T \vee T'$, we apply the function **join**(τ,τ') with $\tau = T$ and $\tau' = T'$. At the end of this process, τ and τ' are the same tree: we obtain $\tau = \tau' = S$. So we have constructed a path $\mathcal{P}$ between T and S such that $\mathcal{P} : T \to \tau_1 \to \cdots \to \tau_k = S$ and a path $\mathcal{P}'$ between T' and S such that $\mathcal{P}' : T' \to \tau'_1 \to \cdots \to \tau'_{k'} = S$. Moreover, if L_S is the distribution sequence of S, we obviously have $L_S(i) = L_T(i) = L_{T'}(i)$ for $i < i_0$ and $L_S(i_0) = \min(L_T(i_0), L_{T'}(i_0))$.

Now, we show that the tree S is really the least upper bound of T and T'. For this, let T'' be a tree in B_n such that $T \xrightarrow{*} T''$ and $T' \xrightarrow{*} T''$ and prove us that $S \xrightarrow{*} T''$ or equivalently $L_{T''} \le L_S$ in term of distribution sequences.

In the case where there exists $i_1 \in [1, i_0[$ such that $L_{T''}(i) = L_S(i)$ for $1 \le i < i_1$ and $L_{T''}(i_1) \ne L_S(i_1)$, we necessarily have $L_{T''}(i_1) < L_T(i_1) = L_{T'}(i_1)$. Using Lemma 3, there exists $S^{(1)} = D(T'', T) = D(T'', T') \in B_n$ such that $S^{(1)} \to T''$, $T \xrightarrow{*} S^{(1)}$ and

$T' \xrightarrow{*} S^{(1)}$. If we denote by $L_{S^{(1)}}$ the distribution sequence of $S^{(1)}$ then we have $L_{S^{(1)}} \leq L_T$ and $L_{S^{(1)}} \leq L_{T'}$. By repeating this process with $S^{(1)}$, we obtain a path $S^{(0)} = T'' \leftarrow S^{(1)} \leftarrow \cdots \leftarrow S^{(m)}$ such that $T \xrightarrow{*} S^{(m)}$, $T' \xrightarrow{*} S^{(m)}$ and the distribution sequence $L_{S^{(m)}}$ of $S^{(m)}$ verifies $L_{S^{(m)}}(i) = L_S(i) = L_T(i) = L_{T'}(i)$ for $1 \leq i < i_0$. Thus, $L_{S^{(m)}} \leq L_T$ and $L_{S^{(m)}} \leq L_{T'}$.

Moreover, if we have $L_T(i_0) < L_{T'}(i_0)$ (resp. $L_T(i_0) > L_{T'}(i_0)$) then, via Lemma 2, $\tau_1 \xrightarrow{*} S^{(m)}$ (resp. $\tau_1' \xrightarrow{*} S^{(m)}$).

We repeat all this process as follows: if $L_T(i_0) < L_{T'}(i_0)$, we replace T by τ_1 and T'' by $S^{(m)}$; or, if $L_T(i_0) > L_{T'}(i_0)$, we replace T' by τ_1' and T'' by $S^{(m)}$, and so on.

At the end of this process, $S \xrightarrow{*} S^{(m)}$ which proves that S corresponds to the least upper bound of T and T'. $\square$

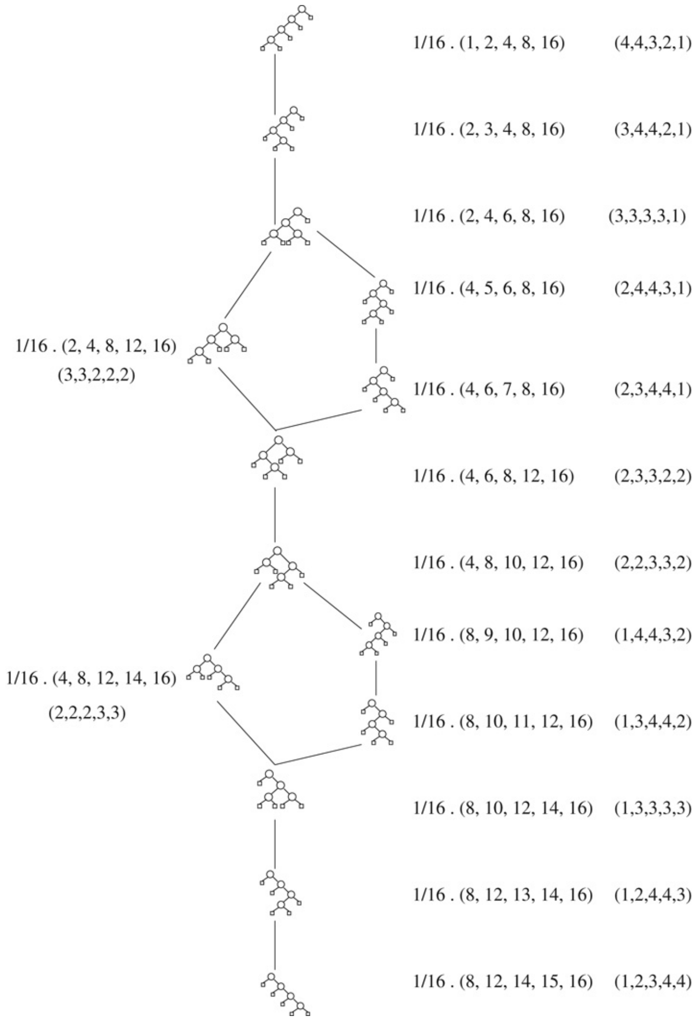

Fig. 2. The pruning-grafting lattice B_4. Each tree is encoded with its feasible and distribution sequences.

Notice that some relations do exist between the already known Catalan lattices. Indeed, recall that if T and T' are two elements in the Kreweras lattice such that $T \leq T'$, then $T \leq T'$ occurs in the Tamari lattice. If $T \leq T'$ in the Tamari lattice, then $T \leq T'$ occurs in the Stanley lattice. If $T \leq T'$ in the phagocyte lattice defined in [3], then $T \leq T'$ occurs in the Kreweras lattice. The covering set of the phagocyte lattice is included in the covering set of Kreweras lattice. However, the pruning-grafting lattice does not appear to have some similar relations with these lattices, see Figs. 2 and 3.

4. Some properties of $(B_n, \xrightarrow{*})$

The lattice $(B_n, \xrightarrow{*})$ has a greatest **1** and a least element **0** and their feasible sequences are respectively, $\ell_{\mathbf{1}} = (n, n, n-1, \ldots, 2, 1)$ and $\ell_{\mathbf{0}} = (1, 2, \ldots, n-1, n, n)$. $(B_n, \xrightarrow{*})$ is symmetric because $T \xrightarrow{*} T'$ iff $\bar{T'} \xrightarrow{*} \bar{T}$. The lattice $(B_n, \xrightarrow{*})$ is not modular since it contains pentagons (see Figs. 2 and 3).

1.	(1, 2, 3, 4, 5, 5)	1/32.(16,24,28,31,31,32)	22.	(2, 3, 3, 3, 3, 2)	1/32.(8,12,16,20,24,32)
2.	(1, 2, 3, 5, 5, 4)	1/32.(16,24,28,29,30,32)	23.	(3, 3, 2, 3, 3, 2)	1/32.(4,8,16,20,24,32)
3.	(1, 2, 4, 4, 4, 4)	1/32.(16,24,26,28,30,32)	24.	(2, 3, 4, 4, 2, 2)	1/32.(8,12,14,16,24,32)
4.	(1, 2, 4, 5, 5, 3)	1/32.(16,24,26,27,28,32)	25.	(2, 3, 4, 5, 5, 1)	1/32.(8,12,14,15,16,32)
5.	(1, 3, 3, 3, 4, 4)	1/32.(16,20,24,28,30,32)	26.	(2, 4, 4, 3, 2, 2)	1/32.(8,10,12,16,24,32)
6.	(1, 2, 5, 5, 4, 3)	1/32.(16,24,25,26,28,32)	27.	(2, 3, 5, 5, 4, 1)	1/32.(8,12,13,14,16,32)
7.	(1, 3, 3, 4, 4, 3)	1/32.(16,20,24,26,28,32)	28.	(3, 3, 3, 3, 2, 2)	1/32.(4,8,12,16,24,32)
8.	(2, 2, 2, 3, 4, 4)	1/32.(8,16,24,28,30,32)	29.	(2, 4, 4, 4, 4, 1)	1/32.(8,10,12,14,16,32)
9.	(2, 2, 2, 4, 4, 3)	1/32.(8,16,24,26,28,32)	30.	(2, 4, 5, 5, 3, 1)	1/32.(8,10,11,12,16,32)
10.	(1, 3, 4, 4, 3, 3)	1/32.(16,20,22,24,28,32)	31.	(3, 3, 3, 4, 4, 1)	1/32.(4,8,12,14,16,32)
11.	(1, 3, 4, 5, 5, 2)	1/32.(16,20,22,23,24,32)	32.	(2, 5, 5, 4, 3, 1)	1/32.(8,9,10,12,16,32)
12.	(1, 4, 4, 3, 3, 3)	1/32.(16,18,20,24,28,32)	33.	(3, 3, 4, 4, 3, 1)	1/32.(4,8,10,12,16,32)
13.	(1, 3, 5, 5, 4, 2)	1/32.(16,20,21,22,24,32)	34.	(3, 4, 4, 2, 2, 2)	1/32.(4,6,8,16,24,32)
14.	(1, 4, 4, 4, 4, 2)	1/32.(16,18,20,22,24,32)	35.	(4, 4, 3, 2, 2, 2)	1/32.(2,4,8,16,24,32)
15.	(2, 2, 3, 3, 3, 3)	1/32.(8,16,20,24,28,32)	36.	(3, 4, 4, 3, 3, 1)	1/32.(4,6,8,12,16,32)
16.	(1, 4, 5, 5, 3, 2)	1/32.(16,18,19,20,24,32)	37.	(3, 4, 5, 5, 2, 1)	1/32.(4,6,7,8,16,32)
17.	(2, 2, 3, 4, 4, 2)	1/32.(8,16,20,22,24,32)	38.	(4, 4, 3, 3, 3, 1)	1/32.(2,4,8,12,16,32)
18.	(1, 5, 5, 4, 3, 2)	1/32.(16,17,18,20,24,32)	39.	(3, 5, 5, 4, 2, 1)	1/32.(4,5,6,8,16,32)
19.	(2, 2, 4, 4, 3, 2)	1/32.(8,16,18,20,24,32)	40.	(4, 4, 4, 4, 2, 1)	1/32.(2,4,6,8,16,32)
20.	(2, 3, 3, 2, 3, 3)	1/32.(8,12,16,24,28,32)	41.	(4, 5, 5, 3, 2, 1)	1/32.(2,3,4,8,16,32)
21.	(3, 3, 2, 2, 3, 3)	1/32.(4,8,16,24,28,32)	42.	(5, 5, 4, 3, 2, 1)	1/32.(1,2,4,8,16,32)

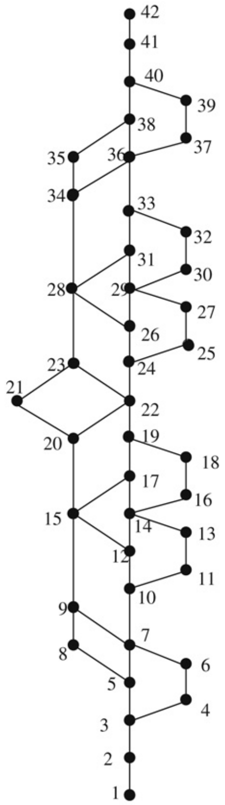

Fig. 3. The pruning-grafting lattice B_5. Each tree is encoded with its feasible and distribution sequences.

Proposition 2. *Let T and T' be two trees of B_n such that $T \overset{*}{\to} T'$, then*

(i) $\overline{U(\bar{T}, \bar{T'})} \to T'$,

(ii) $T \to \overline{D(\bar{T}, \bar{T'})}$,

(iii) *the path $T' \leftarrow D(T, T') \leftarrow D(T, D(T, T')) \leftarrow D(T, D(T, D(T, T'))) \leftarrow \cdots \leftarrow T$ is a longest path between T and T',*

(iv) *the path $T \to U(T, T') \to U(U(T, T'), T') \to U(U(U(T, T'), T'), T') \to \cdots \to T'$ is a shortest path between T and T'.*

Proof. The two first items are deduced from the symmetry of the lattice.

Let us prove (iii). First, notice that if T' is obtained from T by one pruning-grafting transformation, T' cannot be obtained from T by two or more consecutive pruning-grafting transformations (we call this remark R_1).

We proceed by induction on the length of a longest path between T and T'. The property holds when the length of the longest path is two. Indeed, let $T \to T_1 \to T'$ be a longest path between T and T'. If T' has not two lower covers then $T_1 = D(T, T')$ and we obtain directly the result. Otherwise, T' has at least two lower covers, T_1 and T_2 and assume that $T_1 \neq D(T, T')$. As the longest length between T and T' is two, we easily see that there exists a two length path between T and T' such that $T \to D(T, T') \to T'$ which gives the results. This is true because we do not have $T = D(T, T')$ with the remark R_1. Now, let us assume that the property holds when the length of the longest path between T and T' is $\ell \geq 2$ and let us examine the case for a longest path of length $\ell + 1$. Let also $T = T_0 \to T_1 \to \cdots \to T_{\ell+1} = T'$ be a longest path of length $\ell + 1$ between T and T'.

If $T_\ell = D(T, T')$ and using the recurrence hypothesis for the path $T \to \cdots \to T_\ell$ we conclude directly.

Now, we assume that $T_\ell \neq D(T, T')$. This induces that T' has at least two lower covers T_ℓ and $D(T, T')$. We apply the recurrence hypothesis for the path between T and T_ℓ. This means that the path can be of the form $T \to T_1 \to \cdots \to T_{\ell-1} = D(T, T_\ell) \to T_\ell \to T'$.

Moreover, the remark R_1 allows us to prove that $D(T, T') \neq T_{\ell-1}$. Then we distinguish two cases: (I) $D(T, D(T, T')) = T_{\ell-1}$, (II) $D(T, D(T, T')) \neq T_{\ell-1}$ and we will prove that the second case does not occur.

Case (I). In this case, we can conclude directly that there exists a longest path of the form $T \to T_1 \to \cdots \to T_{\ell-1} \to D(T, T') \to T'$. See the following diagram for an illustration of this case.

$$\begin{tikzcd}
& T' & \\
D(T,T') \arrow[ur] & & T_\ell \arrow[ul] \\
& T_{\ell-1} \arrow[ul] \arrow[ur] &
\end{tikzcd}$$

with

$$\begin{aligned}
T' &= \sigma_1 \bigcirc \square\square \bigcirc^{r} \square\sigma_2 \bigcirc \square\square \bigcirc^{s} \square\sigma_3 \\
D(T,T') &= \sigma_1\square \bigcirc^{r+1} \square\square\sigma_2 \bigcirc \square\square \bigcirc^{s} \square\sigma_3 \\
T_\ell &= \sigma_1 \bigcirc \square\square \bigcirc^{r} \square\sigma_2\square \bigcirc^{s+1} \square\square\sigma_3 \\
T_{\ell-1} &= \sigma_1\square \bigcirc^{r+1} \square\square\sigma_2\square \bigcirc^{s+1} \square\square\sigma_3
\end{aligned}$$

where $r, s \geq 0$, $\sigma_1 \in \bigcirc\{\square, \bigcirc\}^*$ and $\sigma_2, \sigma_3 \in \{\square, \bigcirc\}^*$.

Case (II). This induces the following diagram which gives a contradiction with the hypothesis that $T \to T_1 \to \cdots \to T_{\ell-1} \to T_\ell \to T'$ is a longest path between T and T'.

$$\begin{tikzcd}
& T' & \\
D(T,T') \arrow[ur] & & T_\ell \arrow[ul] \\
S \arrow[u] & & \\
& T_{\ell-1} \arrow[ul] \arrow[uur] &
\end{tikzcd}$$

with

$$\begin{aligned}
T' &= \sigma_1 \bigcirc \square\square \bigcirc^{r} \square\square \bigcirc^{s} \square\sigma_2 \\
D(T,T') &= \sigma_1\square \bigcirc^{r+1} \square\square\square \bigcirc^{s} \square\sigma_2 \\
S &= \sigma_1\square \bigcirc^{r} \square \bigcirc \square\square \bigcirc^{s} \square\sigma_2 \\
T_\ell &= \sigma_1 \bigcirc \square\square \bigcirc^{r-1} \square \bigcirc^{s+1} \square\square\sigma_2 \\
T_{\ell-1} &= \sigma_1\square \bigcirc^{r} \square\square \bigcirc^{s+1} \square\square\sigma_2
\end{aligned}$$

where $r \geq 1$, $s \geq 0$, $\sigma_1 \in \bigcirc\{\square, \bigcirc\}^*$ and $\sigma_2 \in \{\square, \bigcirc\}^*$.

This case does not occur since this would mean that the path $T \to T_1 \to \cdots \to T_{\ell-1} \to T_\ell \to T'$ is not a longest path between T and T'.

In conclusion, only the first case appears and we have constructed a path of length $(\ell + 1)$ between T and T' such that $T \to \cdots \to D(T, T_\ell) \to D(T, T') \to T'$ which proves the results by induction.

The item (iv) is obtained *mutatis mutandis*. Indeed, we also proceed by induction on the length of the path between T and T'. With the remark R_1, the result holds when the length of the shortest path is less or equal than two. Let us assume that the property holds when the length of the shortest path between T and T' is less or equal than $\ell \geq 2$ and let us examine the case for a shortest path of length $\ell + 1$. Let also $T = T_0 \to T_1 \to \cdots \to T_{\ell+1} = T'$ be a shortest path of length $\ell + 1$ between T and T'.

If $T_1 = U(T, T')$ and using the recurrence hypothesis for the path $T_1 \to \cdots \to T_\ell \to T'$, we conclude directly.

Now, let us assume that $T_1 \neq U(T, T')$. This induces that T has at least two upper covers T_1 and $U(T, T')$. We apply the recurrence hypothesis for the path between T_1 and T'. This means that the path can be of the form $T \to T_1 \to U(T_1, T') = T_2 \to \cdots \to T'$.

Thus, we distinguish two cases: (I) $T_2 = U(U(T, T'), T')$, (II) $T_2 \neq U(U(T, T'), T')$ and we will prove that the second case does not occur.

Case (I). $T_2 = U(U(T, T'), T')$; as previously, we can conclude directly that there exists a shortest path of the form $T \to U(T, T') \to U(T_1, T') = T_2 \to \cdots \to T'$.

Case (II). $T_2 \neq U(U(T,T'),T')$; we have the following diagram:

$$
\begin{tikzcd}
 & T_3 & \\
U(T,T') \arrow[ur] & & T_2 \arrow[ul] \\
 & & T_1 \arrow[u] \\
 & T \arrow[uul] \arrow[ur] &
\end{tikzcd}
$$

with

$$
\begin{aligned}
T &= \sigma_1 \square \bigcirc^r \square\square \bigcirc^s \square\square \sigma_2\\
T_1 &= \sigma_1 \square \bigcirc^r \square \bigcirc \square\square \bigcirc^{s-1} \square \sigma_2\\
T_2 &= \sigma_1 \square \bigcirc^{r+1} \square\square\square \bigcirc^{s-1} \square \sigma_2\\
T_3 &= \sigma_1 \bigcirc \square\square \bigcirc^r \square\square \bigcirc^{s-1} \square \sigma_2\\
U(T,T') &= \sigma_1 \bigcirc \square\square \bigcirc^{r-1} \square \bigcirc^s \square\square \sigma_2
\end{aligned}
$$

where $r, s \geq 1$ $\sigma_1 \in \bigcirc\{\square, \bigcirc\}^*$ and $\sigma_2 \in \{\square, \bigcirc\}^*$. This case produces a contradiction with the hypothesis that the path $T \to T_1 \to T_2 \to \cdots \to T'$ is a shortest path between T and T'. Thus, only the case (I) occurs and there is a shortest path between T and T' of the form $T \to U(T,T') \to \cdots \to T'$ which gives the result by induction. $\square$

Corollary 2. *Let $\mathcal{C}$ be the path defined by:*

$$\mathbf{1} \leftarrow D(\mathbf{0},\mathbf{1}) \leftarrow D(\mathbf{0}, D(\mathbf{0},\mathbf{1})) \leftarrow D(\mathbf{0}, D(\mathbf{0}, D(\mathbf{0},\mathbf{1}))) \leftarrow \cdots \leftarrow \mathbf{0}.$$

T belongs to $\mathcal{C}$ if and only if T verifies the condition (A) defined by: T has either a unique occurrence of $\bigcirc\square\square$, or two occurrences of $\bigcirc\square\square$ such that there does not exist any leaf between them in the Polish notation of T. Moreover, if T and T' belong to the path $\mathcal{C}$ and verify $T \to T'$, then T' is obtained from T by a pruning-grafting transformation of the rightmost occurrence of $\bigcirc\square\square$ in T.

Proof. $\mathbf{1}$ has only one occurrence of $\bigcirc\square\square$ thus $\mathbf{1}$ verifies condition (A). Assume that $T' \in \mathcal{C}$ verifies (A). We put $T = D(\mathbf{0}, T')$ (thus $T' \leftarrow T$). If T' has only one occurrence of $\bigcirc\square\square$ then $T = D(\mathbf{0}, T')$ has at most two occurrences of $\bigcirc\square\square$, and there does not exist any leaf between them. Thus, T' is necessarily obtained from T by a pruning-grafting of the rightmost occurrence of $\bigcirc\square\square$. If T' has exactly two occurrences of $\bigcirc\square\square$ such that there does not exist any leaf between them. Then $T = D(\mathbf{0}, T')$ has either a unique occurrence of $\bigcirc\square\square$ or two occurrences such that there does not exist any leaf between them. By induction all trees $T \in \mathcal{C}$ hold (A). Conversely, let T be a tree verifying (A). It is straightforward to obtain T from $\mathbf{0}$ by successive pruning-grafting transformations of rightmost occurrences of $\bigcirc\square\square$. $\square$

Corollary 3. *Let $\mathcal{C}'$ be the path defined by:*

$$\mathbf{0} \to U(\mathbf{0},\mathbf{1}) \to U(U(\mathbf{0},\mathbf{1}),\mathbf{1}) \to U(U(U(\mathbf{0},\mathbf{1}),\mathbf{1}),\mathbf{1}) \to \cdots \to \mathbf{1}.$$

If $T \to T'$ is on the path $\mathcal{C}'$ then T' is obtained from T by the leftmost possible pruning-grafting transformation in T.

Proof. Indeed, if T and T' belong to $\mathcal{C}'$ such that $T \to T'$, then $T' = U(T, \mathbf{1})$ and T' is clearly obtained from T by the leftmost possible pruning-grafting transformation on T. $\square$

Remark 1. Via Corollary 2, the longest path between $\mathbf{0}$ and $\mathbf{1}$ has length $2^n - (n+1)$. These are the Eulerian numbers given by the sequence **A000295** of [38]. Indeed, if $T \to T'$ belongs to this path then T' is obtained from T by a pruning-grafting transformation of the rightmost occurrence $\bigcirc\square\square$ in T. Let us denote by $\ell(n)$ the length of the longest path in B_n between $\mathbf{1}$ and $\mathbf{0}$. Then there are $\ell(n-1)$ pruning-grafting transformations between $\mathbf{0}$ and T_1 with the feasible sequence $(1, n, n, n-1, \ldots, 2)$; there are n transformations between T_1 and T_2 with feasible sequence $(2, 3, 4, \ldots, n-1, n, n, 1)$; there are $\ell(n-1)$ transformations between T_2 and $\mathbf{1}$. This means that $\ell(n)$ verifies the recurrence relation $\ell(n) = 2 \cdot \ell(n-1) + n$ with $\ell(2) = 1$, which implies that $\ell(n) = 2^n - (n+1)$.

On the other hand, the shortest path has the length $(n-1)^2$. Indeed, with Corollary 3 we obtain the successor (in the shortest path) of a tree T by the leftmost possible pruning-grafting transformation in T. Thus, there are $(n-1)$ pruning-grafting transformations between $\mathbf{0}$ and the tree T_1 with feasible sequence $(2, 2, 2, 3, 4, \ldots, n-2, n-1, n-1)$. After iterating this process $(n-1)$ times from T_1, we obtain directly that $\ell(n) = (n-1)^2$.

This result allows us to compute the complexity of the above **join** algorithm for computing the least upper bound $T \vee T'$ of two binary trees $T, T' \in B_n$. The *while* loop is performed at most twice the length of a shortest path between $\mathbf{0}$ and $\mathbf{1}$, i.e. $2(n-1)^2$. A loop iteration takes $\mathcal{O}(n)$ time. Thus, the time complexity is $\mathcal{O}(n^3)$ in the worst case.

Now, we prove that the join (resp. meet) irreducible elements in B_n are enumerated by the Eulerian numbers. Recall that $x \in B_n$ is a join (resp. meet)-irreducible element if $x = a \vee b$ (resp. $x = a \wedge b$) implies $x = a$ or $x = b$. In other words, join (resp. meet)-irreducible elements are the elements that have a unique lower (resp. upper) cover.

Theorem 3. *The number of join (resp. meet)-irreducible elements in the pruning-grafting lattice* $(B_n, \xrightarrow{*})$ *is given by the Eulerian number* $2^n - (n+1)$.

Proof. A join-irreducible element T ($T \neq \mathbf{0}$) of B_n can be characterized as follows: (1) it contains exactly one node having two leaves as children and these leaves are not the last two of the tree; or (2) it contains exactly two nodes with two leaves as children and one of these nodes is the parent of the two last leaf of T. There exists a one-to-one map between theses elements and the Dyck paths of semi-length n having exactly one long descent (i.e. descent of length at least two). Indeed, we associate with the Polish notation of T the Dyck path defined as follows: a $\bigcirc$ is replaced by $U = (1, 1)$ and a $\square$ is replaced by $D = (1, -1)$ except the last leaf $\square$ which is not considered. For example, if $n = 4$, the Dyck path associated with the join-irreducible element $\bigcirc\square\bigcirc\bigcirc\square\square\bigcirc\square\square$ is $UDUU\mathbf{DD}UD$. If $T \neq \mathbf{0}$ verifies (1), then its Polish notation verifies $\sigma\bigcirc\square\square\square^k$ where $k \geq 1$ and $\sigma \in \bigcirc\{\bigcirc, \square\}^*$ does not contain any occurrence of $\bigcirc\square\square$. Thus, its corresponding Dyck path has a unique long descent (i.e. a unique maximal sequence $D \dots D$ of length at least two). If T verifies (2), then its Polish notation verifies $\sigma\bigcirc\square\square\sigma'\bigcirc\square\square$ where $\sigma, \sigma' \in \bigcirc\{\bigcirc, \square\}^*$ do not contain any occurrence of $\bigcirc\square\square$. Thus, its corresponding Dyck path also has a unique long descent. Conversely, a Dyck path with a unique long descent can be associated with a Polish notation of a tree verifying (1) or (2). Thus, this map induces a one-to-one map between join-irreducible elements of B_n and Dyck paths of semi-length n having exactly one long descent. These elements are enumerated by the Eulerian numbers [20,33]. Using the symmetry of the lattice, we deduce the same result for the meet-irreducible elements. $\square$

Theorem 4. *The number* $cov(n)$ *of coverings in* B_n *is equal to* $\binom{2n}{n-1}$.

Proof. Recall that a covering of B_n corresponds to an edge of B_n, i.e. a pruning-grafting transformation in B_n. Thus, there exists a bijection between the covering set of B_n and the long descents (i.e. descents of length at least two) in all Dyck paths of semilength n. Indeed, we associate with the Polish notation of a tree $T \in B_n$ a Dyck path $\mathscr{P}$ as follows: each $\bigcirc$ is replaced by $U = (1, 1)$ and each $\square$ is replaced by $D = (1, -1)$ except the last $\square$ which is not considered. Thus, the number of possible pruning-grafting transformations from T is equal to the number of occurrences $\bigcirc\square\square$ that are not on the left of T. By symmetry, if $T' = \overline{T}$, it is also the number of occurrences $\bigcirc\square\square$ that are not on the right of T'. Thus, this corresponds to the number of long descents of the Dyck path $\mathscr{P}$. Then $cov(n)$ is the number of long descents in all Dyck paths of semi-length n. Deutsch [9,38] proved that this number is $\binom{2n}{n-1}$. We obtain the sequence **A001791** of [38]. For example, $cov(3) = 4$ because in the five Dyck paths of semilength 3, namely $UDUDUD$, $UDUU\mathbf{DD}$, $UU\mathbf{DD}UD$, $UUDU\mathbf{DD}$, and $UUU\mathbf{DDD}$, we have four long descents (shown in bold-face). $\square$

Let $J(B_n)$ (resp. $M(B_n)$) denote the set of nonzero join-irreducible (resp. nonunit meet-irreducible) elements of B_n. We say that B_n has a matching g if g is a map of $J(B_n) \cup \{\mathbf{0}\}$ to $M(B_n) \cup \{\mathbf{1}\}$ which is one-to-one and verifies $\tau \leq g(\tau)$ for each join-irreducible τ [2,10,22,32]. Kung [22] has proved that every consistent lattice has a matching. In the sequel, we build a matching of B_n in a *constructive* way.

Theorem 5. *For all n, there exists a matching in* $(B_n, \xrightarrow{*})$.

Proof. Let $\mathcal{C}$ be the particular maximal chain defined as follows: $T_0 = \mathbf{1}$ and $T_i = D(\mathbf{0}, T_{i-1})$. Following Remark 1, we have $T_{2^n-(n+1)} = \mathbf{0}$. By Corollary 2, a tree on the path $\mathcal{C}$ verifies: $T \in \mathcal{C}$ iff in the Polish notation of T we have: (a) there are at most two occurrences of $\bigcirc\square\square$, and (b) there does not exist any leaf between two different occurrences $\bigcirc\square\square$.

We will build an one-to-one map f between $M(B_n) \cup \{\mathbf{1}\}$ and $\mathcal{C}$ such that $\sigma \geq f(\sigma)$ for $\sigma \in M(B_n) \cup \{\mathbf{1}\}$. The symmetry of the lattice induces a one-to-one map $\overline{f}$ between $J(B_n) \cup \{\mathbf{0}\}$ and $\mathcal{C}$ such that $\tau \leq \overline{f}(\tau)$ for $\tau \in J(B_n) \cup \{\mathbf{0}\}$. Using these two bijections, we obtain a matching.

Recall that (cf. proof of Theorem 3) an element σ belongs to $M(B_n) \cup \{\mathbf{1}\}$ if and only if (1) it contains exactly one node having two leaves as children; or (2) it contains exactly two nodes with two leaves as children and one of these nodes is the parent of the two first leaf of T. Let M_1 (resp. M_2) be the set of meet-irreducible in the case (1) (resp. (2)). We have $M(B_n) \cup \{\mathbf{1}\} = M_1 \cup M_2$ and we define the map f as follows:

- if $\sigma \in M_1, f(\sigma) = \sigma$;
- if $\sigma \in M_2$, σ can be written in terms of feasible sequences:

$\sigma = (\sigma_2, \sigma_2, \sigma_3, \dots, \sigma_{k-1}, \sigma_k, \sigma_k, \sigma_{k+2}, \dots, \sigma_n, \sigma_{n+1})$ with $k \geq 3$ where there exists r, $2 \leq r \leq k-1$, such that $\sigma_i \geq \sigma_{i+1}$ for $i \leq r-1$ and $\sigma_i \leq \sigma_{i+1}$ for $r \leq i \leq k$. Thus, if $k \geq 4$, we set $f(\sigma) = (\sigma_r - 1, \sigma_{r+1}, \dots, \sigma_{k-2}, \sigma_{k-1} + 1, \dots, \sigma_{k-1} + r - 2, \sigma_{k-1} + r - 1, \sigma_{k-1} + r - 1, \sigma_k, \sigma_k, \sigma_{k+2}, \dots, \sigma_n, \sigma_{n+1})$; if $k = 3$ then $\sigma \in \mathcal{C}$ and we set $f(\sigma) = \sigma$. For example, if $\sigma = (4, 4, 3, 2, 3, 4, 7, 7, 6, 5, 2)$ then $k = 7$, $r = 4$ and $f(\sigma) = (1, 3, 5, 6, 7, 7, 7, 7, 6, 5, 2)$.

Obviously $f(\sigma) \in \mathcal{C}, f(\sigma) \leq \sigma$ and by construction $f(\sigma) = f(\sigma')$ implies $\sigma = \sigma'$. Notice that $f(\sigma)$ is the smallest element of $\mathcal{C}$ such that $\sigma \geq f(\sigma)$. Via the symmetry of B_n, we obtain a bijection $\overline{f}$ from $J(B_n) \cup \{\mathbf{0}\}$ to $\mathcal{C}$ defined by: for $\tau \in J(B_n) \cup \{\mathbf{0}\}, \overline{f}(\tau) = \overline{f(\overline{\tau})}$, and such that $\tau \leq \overline{f}(\tau)$. Therefore, the composition $g = f^{-1} \circ \overline{f}$ is a matching of B_n since g is a map from $J(B_n) \cup \{\mathbf{0}\}$ to $M(B_n) \cup \{\mathbf{1}\}$ which is one-to-one and verifies $\tau \leq g(\tau)$ for every $\tau \in J(B_n) \cup \{\mathbf{0}\}$. $\square$

5. Conclusion and open problems

In this paper, a new lattice structure has been defined on the Catalan sets of binary trees via a natural transformation. The simple and natural definition of the pruning-grafting transformation is unfortunately at odds with the rather complex theorem which characterizes this transformation.

Some problems remain to be solved.

Is there an efficient and nonrecursive algorithm to compute the Möbius function of the pruning-grafting lattice of binary trees as in [26]? We conjecture that this Möbius function takes its values in $\{-1, 0, +1\}$.

Is there a polynomial time algorithm to compute the minimal path length distance between two binary trees in the pruning-grafting lattice [23]? If so, a new shortest-path-type metric could be obtained, and could be added to the existing metrics on Catalan sets [17,25,27,29,37]. Let us recall that we still do not know if the rotation distance on binary trees can be computed in polynomial time.

Let us define the integer-valued function v on $(B_n, \xrightarrow{*})$ by $v(T)$ be the length of a longest path between T and **1**. We conjecture that v is an anti-monotone supervaluation, i.e. v verifies for all T and T': $v(T) + v(T') \leq v(T \vee T') + v(T \wedge T')$ and $T < T' \implies v(T) > v(T')$. Were it the case, therefore $d(T, T') = v(T) + v(T') - 2v(T \vee T')$ would be a metric on $(B_n, \xrightarrow{*})$ [5,7].

Guttmann, Krattenthaler and Viennot [18] enumerate the k-chains in the Stanley lattices, i.e, the number of k-noncrossing Dyck paths. Kreweras [21] enumerates the k-chains in the lattices of noncrossing partitions. More generally, Chapoton proves in [8] that the sequence **A000260** of [38] enumerates the number of intervals in the Tamari lattices, i.e. the number of ordered pairs (T, T') such that $T \leq T'$: see also [16, p. 27]. Bernardi and Bonichon [6] construct bijections between the set of intervals of these lattices and the realizers of triangulations. Here, we obtain experimentally the numbers of intervals for the pruning-grafting lattices for small sizes:

$$1, 3, 15, 101, 818, 7486, 74648, 793005, 8843056, 102464586, \ldots$$

This sequence does not appear in [38]. Is it possible to obtain the generating function of this sequence?

Acknowledgements

We would like to thank the anonymous referees and Emeric Deutsch for their constructive remarks which have greatly improved this paper.

References

[1] E. Barcucci, A. Bernini, L. Ferrari, M. Poneti, A distributive lattice structure connecting Dyck paths, noncrossing partitions and 312-avoiding permutations, Order 22 (4) (2005) 311–328.
[2] J.L. Baril, J.M. Pallo, Matchings in three Catalan lattices, Internat. J. Comput. Math. 8 (2003) 193–200.
[3] J.L. Baril, J.M. Pallo, The phagocyte lattice of Dyck words, Order 23 (2–3) (2006) 97–107.
[4] J.L. Baril, J.M. Pallo, Efficient lower and upper bounds of the diagonal-flip distance between triangulations, Inform. Process. Lett. 100 (2006) 131–136.
[5] J.P. Barthélemy, B. Leclerc, B. Monjardet, On the use of ordered sets on problems of comparison and consensus of classifications, J. Classification 3 (1986) 187–224.
[6] O. Bernardi, N. Bonichon, Catalan's intervals and realizers of triangulations, in: FPSAC, Tianjin, China, 2007.
[7] S.A. Boorman, D.C. Olivier, Metrics on spaces of finite trees, J. Math. Psychol. 10 (1973) 26–59.
[8] F. Chapoton, Sur le nombre d'intervalles dans les treillis de Tamari, Séminaire Lotharingien de combinatoire 55 (2006).
[9] E. Deutsch, Private communication, 2008.
[10] D. Duffus, Matching in modular lattice, J. Combin. Theory 32 (1982) 303–314.
[11] P.H. Edelman, Chain enumeration and noncrossing partitions, Discrete Math. 31 (1980) 171–180.
[12] P.H. Edelman, R. Simion, Chains on the lattice of noncrossing partitions, Discrete Math. 126 (1994) 107–119.
[13] L. Ferrari, R. Pinzani, Lattices of lattice paths, J. Stat. Plan. Inference 135 (2005) 77–92.
[14] H. Friedman, D. Tamari, Problèmes d'associativité: une structure de treillis fini induite par une loi de demi-associative, J. Comb. Theory 2 (1967) 215–242.
[15] R.G. Gallager, Information Theory and Reliable Communications, J. Wiley, New York, 1968.
[16] C. Germain, J. Pallo, The number of coverings in four Catalan lattices, Internat. J. Comput. Math. 61 (1996) 19–28.
[17] C. Germain, J.M. Pallo, Two shortest path metrics on well-formed parentheses strings, Inform. Process. Lett. 60 (1996) 283–287.
[18] A.J. Guttmann, C. Krattenthaler, X.G. Viennot, Vicious walkers, friendly walkers and Young tableaux II: Between two walls, J. Statist. Phys. 110 (2003) 1069–1086.
[19] S. Hanke, T. Ottmann, S.J. Schuierer, The edge-flipping distance of triangulations, J. Univers. Comp. Sci. 2 (1996) 570–579.
[20] D.E. Knuth, Fondamental Algorithms, vol. 1, Addison-Wesley, 1997.
[21] G.J. Kreweras, Sur les partitions noncroisées d'un cycle, Discrete Math. 1 (1972) 570–579.
[22] J.P.S. Kung, Matching and Radon transforms in lattices, Order 2 (1985) 105–112.
[23] B. Monjardet, Metrics on partially ordered sets — A survey, Discrete Math. 35 (1981) 173–184.
[24] J.M. Pallo, Enumerating, ranking and unranking binary trees, The Computer J. 29 (1986) 171–175.
[25] J.M. Pallo, On the rotation distance in the lattice of binary trees, Inform. Process. Lett. 25 (1987) 369–373.
[26] J.M. Pallo, An algorithm to compute the Möbius function of the rotation lattice of binary trees, RAIRO Inform. Théorique et Appl. 27 (1993) 341–348.
[27] J.M. Pallo, An efficient upper bound of the rotation distance of binary trees, Inform. Process. Lett. 73 (2000) 87–92.
[28] J.M. Pallo, Generating binary trees by Glivenko classes on Tamari lattices, Inform. Process. Lett. 85 (2003) 235–238.
[29] J.M. Pallo, Right-arm rotation distance between binary trees, Inform. Process. Lett. 87 (2003) 173–177.
[30] D.S. Parker, P. Ram, The construction of Huffman codes is a submodular ("convex") optimization problem over a lattice of binary trees, SIAM J. Comput. 28 (1999) 1875–1905.
[31] V. Reiner, Noncrossing partitions for classical reflection groups, Discrete Math. 177 (1997) 195–222.

[32] K. Reuter, Matching for linearly indecomposable modular lattices, Discrete Math. 63 (1987) 245–247.
[33] D.P. Roselle, Permutations by number of rises and successions, Proc. Amer. Math. Soc. 20 (1968) 8–16.
[34] F. Ruskey, T.C. Hu, Generating binary trees lexicographically, SIAM J. Comput. 6 (4) (1977) 745–758.
[35] R. Simion, D. Ullman, On the structure of lattice of noncrossing partitions, Discrete Math. 98 (1991) 193–206.
[36] R. Simion, Noncrossing partitions, Discrete Math. 217 (2000) 367–409.
[37] D.D. Sleator, R.E. Tarjan, W.P. Thurston, Rotation distance, triangulations and hyperbolic geometry, J. Am. Math. Soc. 1 (1988) 647–681.
[38] N.J.A. Sloane, The On-line Encyclopedia of Integer Sequences, Available electronically at http://www.research.att.com/~njas/sequences.
[39] R.P. Stanley, Enumerative Combinatorics, vol. 2, Cambridge University Press, 1999.