Acta Cybernetica **17** (2006) 799–810.

Rotational tree structures on binary trees and triangulations*

Jean Marcel Pallo†

Abstract

A rotation in a binary tree is a simple and local restructuring technique commonly used in computer science. We propose in this paper three restrictions on the general rotation operation. We study the case when only leftmost rotations are permitted, which corresponds to a natural flipping on polygon triangulations. The resulting combinatorial structure is a tree structure with the root as the greatest element. We exhibit an efficient algorithm for computing the join of two trees and the minimum number of leftmost rotations necessary to transform one tree into the other.

Keywords: Binary trees; Rotation; Distance; Lattice; Algorithms

1 Introduction

Rotation is one of the most common operations for restructuring binary trees. It has the advantage of altering the depths of some of the nodes in the tree while preserving the symmetric order of all the nodes. Thus rotation is commonly used in a variety of algorithms for maintaining binary search trees with a good amortized behavior [10, 24, 28].

The combinatorial properties of binary trees under the rotation operation have been studied for thirty years [27]. In [17] we have shown that a directed version of the rotation graph of binary trees with n nodes is a lattice, known as the nth Tamari lattice. This corresponds to the case when only left rotations are permitted in the binary tree transformation. Over the last ten years, Tamari lattices have often been used as examples to illustrate algebraic theories [1, 3, 16, 25].

Initially, Tamari lattices were orderings of parenthesizations of words. But nowadays they can be described in other ways via the well-known bijections between families of Catalan combinatorial objects. A system that is isomorphic to Tamari lattices is that of triangulations of a convex polygon related by the diagonal flip operation. This is the transformation that converts one triangulation into another

*The results reported in this paper were presented at the $11th$ International Conference AFL 2005 (Automata and Formal Languages) held at Dobogókő, Hungary, May 17–20.

†LE2I, UMR 5158, Université de Bourgogne, BP 47870, F21078 DIJON-Cedex, France, E-mail: pallo@u-bourgogne.fr

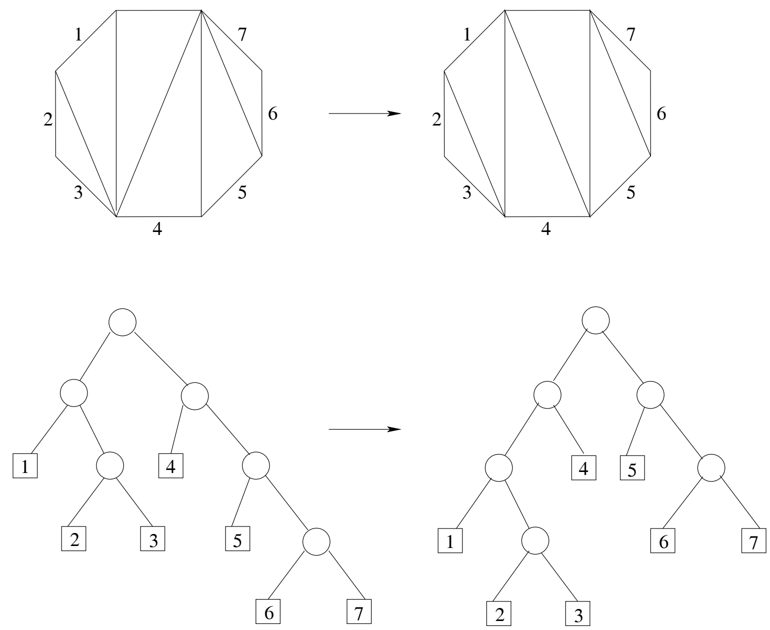

Figure 1: A triangulation diagonal flip and its corresponding binary tree rotation

by removing a diagonal in the triangulation and adding the diagonal that subdivides the resulting quadrilateral on the opposite way [8, 9, 26] (see Fig. 1).

In 1982, Culik and Wood defined the rotation distance between two binary trees with the same number of leaves as the minimum number of rotations necessary to transform one tree into the other [4]. Using the classical bijection between binary trees with n internal nodes and triangulations of $(n+2)$- gons, the previous distance is equivalent to the minimum number of diagonal-flip transformations needed to convert one triangulation of a polygon into another. There remains today an open problem whether the rotation distance can be computed in polynomial time.

Therefore it seems natural to consider special instances of rotation transformations in order to obtain simpler operations [12, 24]. In [2] the rotation operation is limited to the case where the leftmost subtree is constrained to be a leaf. In [5, 6, 11, 22] the authors only allow rotations at nodes along the right arm of a tree.

The current paper belongs to this appoach. We consider the problem by limiting the general rotation operation to the restricted version where only leftmost rotations on trees are allowed. We obtain a tree structure which is a join-semilattice with the root as the greatest element. An efficient algorithm computes the corresponding restricted rotation distance. This algorithm is constructive: it builds a sequence of leftmost rotations transforming one tree into the other.

Clearly, the restricted rotation distance defined above is bounded below by the usual rotation distance for which no efficient algorithm is known to compute it exactly. However, this restricted rotation distance is a weak approximation of the usual rotation distance. A better approximation can be found in [21, 23]. This new metric can be considered as a way of measuring the difference in shape between two binary trees.

2 Definitions and terminology

Let us denote by $\bigcirc$ (respectively $\square$) internal nodes (respectively leaves) of a binary tree. Let T_L (respectively T_R) denote the left (respectively right) subtree of a binary tree T (the order is significant). Thus we can write $T = \bigcirc T_L T_R$ in Polish notation, i.e. by traversing T in preorder (visit the root and then the left and right subtrees recursively). The weight $|T|$ of a binary tree T is the number of leaves of T. Let B_n denote the set of binary trees with n internal nodes (and thus with $n+1$ leaves). The leaves of $T \in B_n$ are numbered from 1 to $n+1$ by a preorder traversal of T (i.e. from left to right). The left (respectively right) arm of $T \in B_n$ is the path from the root of T to its first (respectively $(n+1)$th) leaf. The mirror image $\widetilde{T}$ of T is recursively defined by $\widetilde{T} = \bigcirc\widetilde{T_R}\widetilde{T_L}$ and $\widetilde{\square} = \square$. Let us define $\mathbf{0}_n = (\bigcirc\square)^n\square$ (respectively $\mathbf{1}_n = \bigcirc^n\square^{n+1}$) the tree of B_n where every internal node has a leaf as a left (respectively right) child.

In this paper we use the representation of binary trees via weight sequences introduced in [17]. This coding is defined as follows. Given $T \in B_n$, the weight sequence of T is the integer sequence $w_T = (w_T(1), \ldots, w_T(n))$ where $w_T(i)$ is the weight of the largest subtree of T whose last leaf is the ith leaf (see Fig. 2). The usual left-rotation $\rightarrow$ on B_n is defined as follows. A tree $T \in B_n$ being given, it associates a tree T' obtained by replacing some subtree $\bigcirc T_1 \bigcirc T_2 T_3$ of T by the

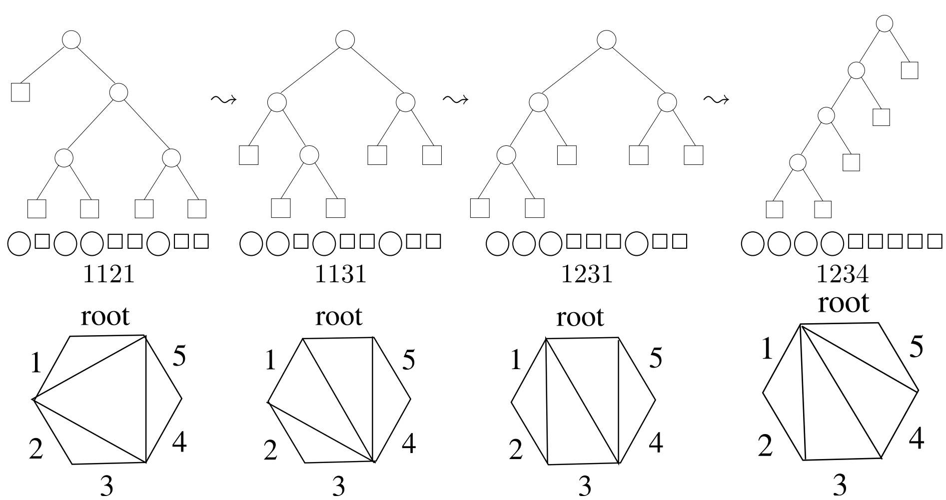

Figure 2: Three lefmost left-rotations in B_4 and the corresponding flips in T_6

subtree $\bigcirc\bigcirc T_1T_2T_3$. Let $\overset{-1}{\to}$ denote the right-rotation and let $\overset{*}{\to}$ denote the reflexive transitive closure of $\to$. The usual rotation distance between T and $T' \in B_n$ is the fewest number of left- and right-rotations required to convert T into T'. We have proved in [17] the following characterization: given $T, T' \in B_n$, we have $T \overset{*}{\to} T'$ iff for all $i \in [1, n]$: $w_T(i) \le w_{T'}(i)$.

Let us consider $(n+2)$-gons, i.e. convex polygons with $n+2$ sides and with a distinguished side as the top. We label the other sides from 1 to $n+1$ counterclockwise. Any triangulation of the $(n+2)$-gon has n triangles and $n-1$ non-crossing diagonals. Let T_{n+2} denote the set of triangulations of the $(n+2)$-gon. There is an explicit bijection τ between B_n and T_{n+2} [23,26]. The top of the $(n+2)$-gon $\tau(T)$ corresponds to the root of the tree T. The ith side of $\tau(T)$ corresponds to the ith leaf of T. Diagonals corresponds to internal nodes recursively as follows. If j is the last leaf of the left subtree T_L of T, then T_L corresponds to the $(j+1)$-gon having edge set $\{1, \ldots, j\}$ and the right subtree T_R corresponds to the $(n-j+2)$-gon having edge set $\{j+1, \ldots, n+1\}$ (see Fig. 1 and 2).

Given some $T \in B_n$ with $T \ne \mathbf{1}_n$, according to the Polish notation of T, consider the leftmost $\square$ followed by a $\bigcirc$ which respectively are the last leaf of a subtree T_1 and the root of a subtree $\bigcirc T_2T_3$. Thus the root of $\bigcirc T_1 \bigcirc T_2T_3$ is located on the left arm of T. Then define as the leftmost left-rotation on B_n the transformation $T \rightsquigarrow T'$ which consists in converting the leftmost subtree $\bigcirc T_1 \bigcirc T_2T_3$ of T into $\bigcirc\bigcirc T_1T_2T_3$ (see Fig. 2). Given $T \in B_n$, the leftmost left-rotation transformation is uniquely defined.

Let us describe the transformation on $\tau(T) \in T_{n+2}$ which corresponds to the leftmost left-rotation on $T \in B_n$ via the classical bijection τ between B_n and T_{n+2}. This transformation is the unique operation on $\tau(T)$ which consists in removing some diagonal and adding a new diagonal an end of which coincides with the vertex located between the root side and the side labelled 1 (see Fig. 2 and 3). This alternative formulation may seem more natural and intuitive. But the weight sequences of binary trees are more appropriate for calculations.

Let $\overset{*}{\rightsquigarrow}$ denote the reflexive transitive closure of $\rightsquigarrow$. The leftmost rotation graph LG_n is the directed graph which has a node for each tree of B_n. Two nodes are adjacent when their corresponding trees differ by a single leftmost left-rotation. Since the leftmost left-rotation operation on T is uniquely defined, LG_n enjoys a tree structure. The leftmost rotation distance $d(T, T')$ between T and $T' \in B_n$ is the length of the unique path between T and T' in the directed graph LG_n. LG_n is a subgraph of the graph G_n according to the usual rotation. Algebraic properties of G_n can be found in [1, 3, 16, 20, 25, 26].

3 Tree structure B_n

Given $T \in B_n$ with $T \ne \mathbf{1}_n$ and w_T, we obtain the weight sequence of the unique T' such that $T \rightsquigarrow T'$ in the following way. Let $i \ge 2$ be the smallest integer such that $w_T(i) = 1$. Let $j = max\{m \in [i, n] | i = m - w_T(m) + 1\}$, i.e. the greatest integer m such that the largest subtree with last leaf m has i as the first leaf. Then

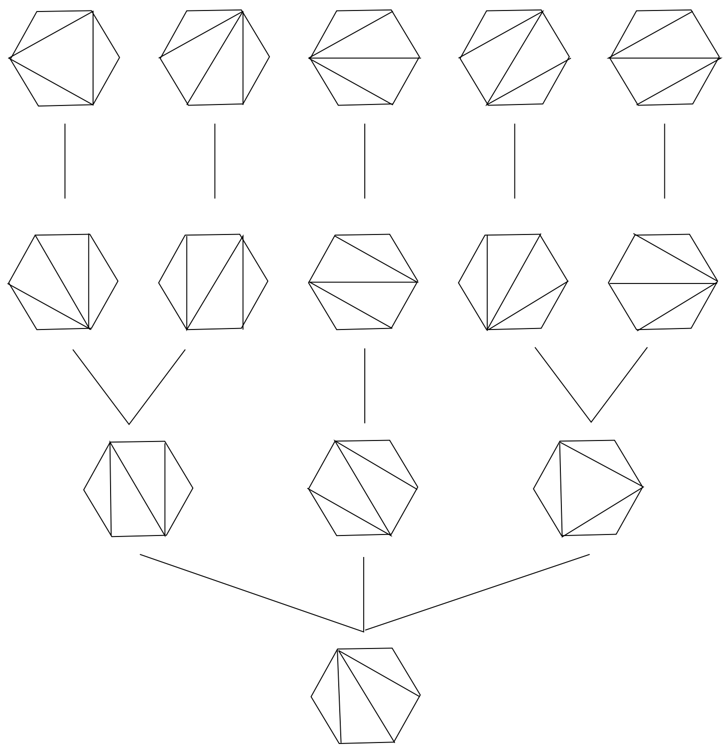

Figure 3: The flipping tree structure T_6

$w_{T'} = w_T$ except for the integer j: $w_{T'}(j) = j$. It is worth noting that this integer j cannot be modified further since we have $1 \leq w_T(k) \leq k$ for all $T \in B_n$ and $k \in [1, n]$.

The poset $(B_n, \overset{*}{\rightsquigarrow})$ enjoys some properties which can be easily obtained. $(B_n, \overset{*}{\rightsquigarrow})$ is a poset with greatest element $\mathbf{1}_n$ for which $w_{\mathbf{1}_n} = (1, 2, 3, \ldots, n)$. This poset has a tree structure (with the greatest element $\mathbf{1}_n$ as root) and thus is a join-semilattice (see Fig. 4 and 5). The poset $(B_n, \overset{*}{\rightsquigarrow})$ is graded, i.e. there exists an integer-valued function r defined on B_n by $r(T) = card\{i \in [1, n] | w_T(i) = i\}$ such that $T \overset{*}{\rightsquigarrow} T'$ and $r(T') = 1 + r(T)$ iff $T \rightsquigarrow T'$. $r(T)$ is equal to the number of internal nodes that are on the left arm of T. We have $r(\mathbf{1}_n) = n - 1$.

Let us remark that B_n is isomorphic to two subtrees of B_{n+1}. One is obtained by sustituting $\bigcirc\square\square$ for the last leaf $\square$ in all the B_n trees. If $(w_1, \ldots, w_n) \in B_n$, then $(w_1, \ldots, w_n, 1)$ is the weight sequence of a tree in the corresponding subtree of B_{n+1}. The other is obtained by sustituting $\bigcirc\square\square$ for the one before last leaf $\square$ in all the B_n trees. If $(w_1, \ldots, w_n) \in B_n$, then $(w_1, \ldots, w_{n-1}, 1, 1 + w_n)$ is the weight sequence of a tree in the corresponding subtree of B_{n+1}. For example in Fig. 5, the left and right subtrees of B_5 are isomorphic to B_4 (Fig. 4).

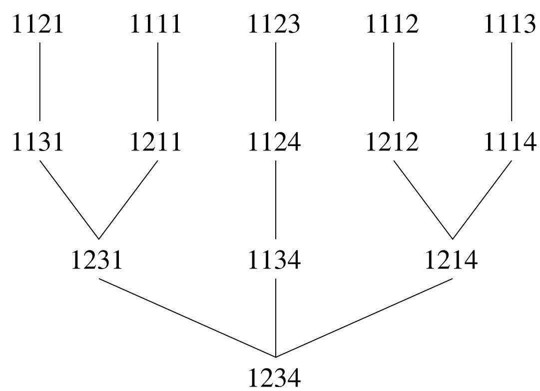

Figure 4: The leftmost tree structure B_4

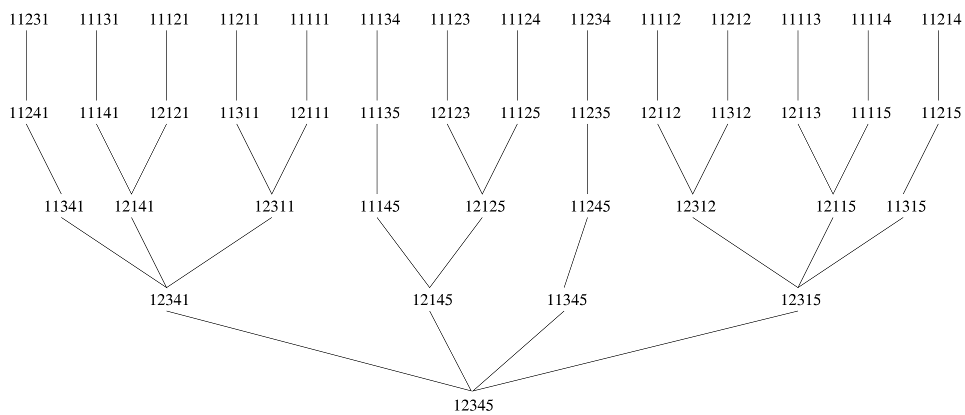

Figure 5: The leftmost tree structure B_5

The leftmost rotation distance between T and T' can be computed by the formula $d(T,T') = 2r(T \vee T') - r(T) - r(T')$. Thus we are led to compute the join $T \vee T'$ of any couple of trees T and T'.

4 Computing joins and leftmost rotation distance

We already have observed that in applying the leftmost rotation $T \rightsquigarrow T'$ the unique integer which has been transformed reaches its maximal possible value and thus cannot increase. Now, for every $T \in B_n$, compute from w_T an ordered array a_T which keeps track of the sequence of all the integer transformations for designing the unique path between T and $\mathbf{1}_n$.

Algorithm (Computation of a_T from w_T)
Given $T \in B_n$ and its weight sequence w_T
$k := 1$
for $i := 1$ **to** n **do**
  if $w_T(i) = 1$ **then**
    for $j := n$ **downto** i **do**
      if $i = j - w_T(j) + 1$ **then** $a_T(k) := j; k := k + 1$ **endif**
    enddo
  endif
enddo

This algorithm requires $O(n^2)$ time in the worst case and $O(n)$ space.

The join $T \vee T'$ of T and T' is located at the intersection of the two paths connecting T and T' to $\mathbf{1}_n$. Thus we compute $w_{T\vee T'}$ in the following way.

Let us consider the greatest suffix which is common to a_T and $a_{T'}$ (if it exists). The corresponding prefixes of a_T and $a_{T'}$ contain the same integers i (possibly in different order) for which $w_{T\vee T'}(i) = i$. The remaining integers j verify $w_{T\vee T'}(j) = w_T(j) = w_{T'}(j)$. Therefore it is easy to compute $w_{T\vee T'}$, and then $r(T)$, $r(T')$, $d(T,T') = 2r(T \vee T') - r(T) - r(T')$ using the rank function $r(T) = card\{i \in [1,n] | w_T(i) = i\}$. See some examples in Table 1 where suffixes are shown in bold type.

Table 1:

w_T	$w_{T'}$	a_T	$a_{T'}$	$w_{T\vee T'}$	$d(T,T')$
11112	11315	1235**4**	5312**4**	12315	4
11234	11345	1543**2**	5431**2**	11345	3
11214	11215	15**324**	51**324**	11215	1
11111	12345	12345	54321	12345	4
11111	11114	12345	15234	12345	8
11315	12112	5312**4**	2135**4**	12315	3
11212	11114	1325**4**	1523**4**	12315	6

a_T (respectively $a_{T'}$) allows to build the unique path between T and $T \vee T'$ (respectively T' and $T \vee T'$). Thus we obtain the unique path $(T, T \vee T', T')$ between T and T'.

5 Mirror leftmost rotation distance

Let us define the mirror leftmost rotation $\hookrightarrow$ on B_n by $T \hookrightarrow T'$ iff $\widetilde{T'} \rightsquigarrow \widetilde{T}$. Then $(B_n, \overset{*}{\hookrightarrow})$ is a poset with least element $\mathbf{0}_n$ for which $w_{\mathbf{0}_n} = (1,1,1,\ldots,1)$. This poset has a tree structure (with the least element $\mathbf{0}_n$ as root) and thus is a meet-semilattice. This poset $(B_n, \overset{*}{\hookrightarrow})$ is ranked by the rank function $\rho(T) = n - k_T + 1$ where k_T is the number of internal nodes on the right arm of $T \in B_n$. We have $\rho(\mathbf{0}_n) = 1$ The following algorithm computes $\rho(T)$ using the weight sequence of T:

Rank algorithm (Computation of $\rho(T)$ from w_T)
Given $T \in B_n$ and its weight sequence w_T;
$k_T := 1$; $i := n$;
while $i > 1$ **do**
  if $w_T(i) = 1$ **then** $k_T := k_T + 1$;$i := i - 1$
  else $i := i - w_T(i) + 1$ **endif**
enddo
$\rho(T) = n - k_T + 1$

See $(B_4, \overset{*}{\hookrightarrow})$ in Fig. 6. Observe that $\hookrightarrow$ is a particular case of the right-arm rotation transformation defined in [22]. As illustration, compare for example Fig. 3 of [22, p. 176] and Fig. 6 of this paper. The edge which links 1112 and 1212 in Fig. 3 of [22] has disappeared in Fig. 6. The graph drawn in Fig. 3 of [22] does not enjoy the tree structure property.

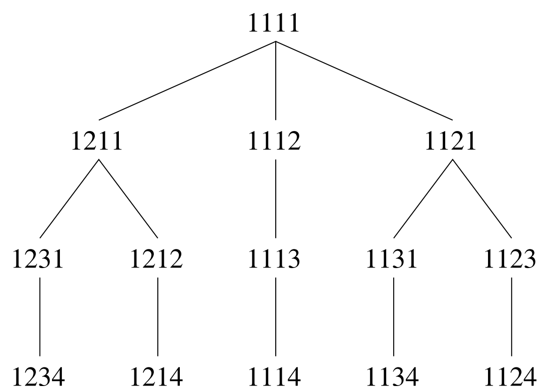

Figure 6: The mirror image of B_4

Let us define the mirror leftmost rotation distance $\widetilde{d}(T, T')$ between T and $T' \in B_n$ as the length of the unique path between T and T' in the graph of $(B_n, \overset{*}{\hookrightarrow})$. Therefore we have: $\widetilde{d}(T, T') = d(\widetilde{T}, \widetilde{T'})$.

Since $w_{\widetilde{T}}$ can be easily computed recursively from w_T, the mirror leftmost rotation distance $\widetilde{d}(T, T') = d(\widetilde{T}, \widetilde{T'})$ is computed using Section 4. Then $\delta(T, T') = min(d(T, T'), d(\widetilde{T}, \widetilde{T'}))$ is bounded below by the usual rotation distance for which no polynomial time algorithm is known to compute it exactly today. See some examples in B_8 (Table 2).

Table 2:

w_T	$w_{T'}$	$d(T,T')$	$w_{\widetilde{T}}$	$w_{\widetilde{T'}}$	$d(\widetilde{T},\widetilde{T'})$	$\delta(T,T')$
11121511	12123611	8	12312148	12311141	9	8
11121518	11234112	11	11212147	11341118	11	11
11312312	11214111	13	11311612	12341218	6	6
11115123	12311312	11	11141234	11312611	11	11
11111123	11231237	14	11145678	11114118	8	8
11235112	12115111	11	11341114	12341231	9	9
11211612	11111312	5	11312315	11312678	9	5
11311245	11341678	6	11113612	12112678	9	6

6 Open problems

We propose below two other new definitions of restricted rotations which lead to computing open problems.
First we can restrict the general definition of the rotation transformation by choosing $\bigcirc T_1 \bigcirc T_2 T_3$ as the rightmost subtree in the Polish notation of T. More precisely, let us consider in the Polish notation of T the rightmost pattern $\Box\bigcirc$ made up of a $\Box$ followed by a $\bigcirc$. This $\bigcirc$ is the root of a subtree denoted by $\bigcirc T_2 T_3$, and thus T_3 is always equal to a leaf $\Box$. Let us denote by T_1 the largest subtree of T whose last leaf is the leaf $\Box$ involved in the previous pattern $\Box\bigcirc$. The uniquely defined rotation which transforms $\bigcirc T_1 \bigcirc T_2 \Box$ of T into $\bigcirc \bigcirc T_1 T_2 \Box$ is called rightmost left-rotation on the tree T. B_n endowed with this transformation has a tree structure (with the root as the greatest element $\mathbf{1}_n$) and thus is a join-semilattice (see Fig. 7). Despite this tree structure, the direct computation of the joins of two trees seems to be more arduous. The definition of an efficient algorithm for computing the corresponding rightmost rotation distance d' seems difficult, too. However, we can easily exhibit the unique paths connecting T and T' with $\mathbf{1}_n$. The weight sequence of the unique tree $succ(T)$ obtained from T by a rightmost rotation is such that $w_{succ(T)} = w_T$ except for the integer $i = max\{k \in [j,n] | w_T(k) = k - j + 1\}$ where $j = max\{l \in [1,n] | w_T(l) = 1\}$. For this integer i, we have $w_{succ(T)}(i) = w_T(i) + w_T(i - w_T(i))$. The join $T \vee T'$ of T and T' is located at the intersection of the two paths $(T, \mathbf{1}_n)$ and $(T', \mathbf{1}_n)$. Unfortunately, this rough construction requires $O(n^3)$ time and $O(n^2)$ space.

It is worth noting that leftmost d and rightmost d' rotation distances cannot be compared. For example: $d(1112, 1114) = 3 < d'(1112, 1114) = 4$ and $d'(1113, 1121) = 2 < d(1113, 1121) = 6$ (see Fig. 4 and 7).

Second, we have limited in [2] the rotation operation to the case where the leftmost subtree T_1 of the subtree $\bigcirc T_1 \bigcirc T_2 T_3$ is always constrained to be a leaf $\Box$. This transformation $\bigcirc\Box\bigcirc T_2 T_3 \xrightarrow{L} \bigcirc\bigcirc\Box T_2 T_3$ induces a graded lower semimodular meet-semilattice structure on B_n. We can define a new restricted rotation by compelling, this time, the central subtree T_2 of the subtree $\bigcirc T_1 \bigcirc T_2 T_3$ to be always

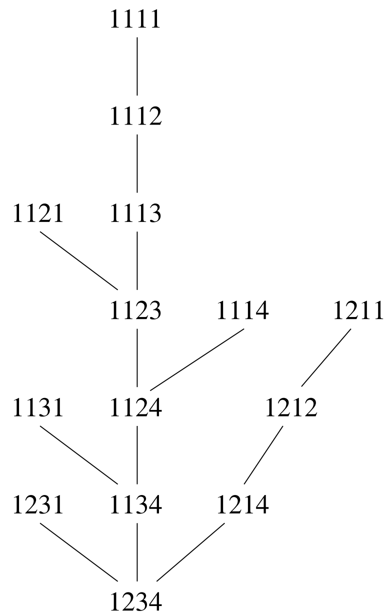

Figure 7: The rightmost tree structure B_4

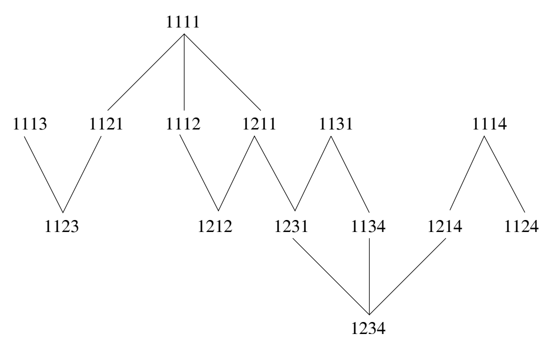

Figure 8: The central poset B_4

equal to a leaf $\Box$. This transformation $\bigcirc T_1 \bigcirc \Box T_3 \xrightarrow{C} \bigcirc\bigcirc T_1 \Box T_3$ induces a graded poset structure on B_n, but does not have as good algebraic properties as before. However, this "central" rotation operation $\xrightarrow{C}$ has a nice characterization: $T \xrightarrow{C} T'$ iff $w_T = w_{T'}$ except for an integer i such that $w_T(i) = 1 < w_{T'}(i)$ (see Fig. 8). The rank of $T \in B_n$ is easily computed by $r(T) = n + 1 - card\{i \in [1, n] | w_T(i) = 1\}$. Here too, it seems difficult to exhibit an efficient algorithm for computing the corresponding central rotation distance.

Acknowledgement

I would like to thank the anonymous referees for their constructive remarks and recommendations which have greatly helped to improve this paper.

References

[1] M.K. Bennet, G. Birkhoff, Two families of Newman lattices, Alg. Universalis **32**(1994), 115-144.

[2] A. Bonnin, J. Pallo, A shortest path metric on unlabeled binary trees, Pattern Recognition Lett. **13**(1992), 411-415.

[3] C. Chameni-Nembua, B. Monjardet, Les treillis pseudocomplments finis, European J. Combin. **13**(1992), 89-107.

[4] K. Cukik, D. Wood, A note on some tree similarity measures, Inform. Process. Lett. **15**(1982), 39-42.

[5] S. Cleary, Restricted rotation distance between binary trees, Inform. Process. Lett. **84**(2002), 333-338.

[6] S. Cleary, J. Taback, Bounding restricted rotation distance, Inform. Process. Lett. **88**(2003), 251-256.

[7] R.D. Dutton, R.O. Rogers, Properties of the rotation graph of binary trees, Congr. Numer. **109**(1995), 51-63.

[8] S. Hanke, T. Ottmann, S. Schuierer, The edge-flipping distance of triangulations, J. Universal Comput. Sci. **2**(1996), 570-579.

[9] F. Hurtado, M. Noy, J. Urrutia, Flipping edges in triangulations, Discrete Comput. Geom. **22**(1999), 333-346.

[10] K.S. Larsen, E. Soisalon-Soininen, P. Widmayer, Relaxed balance using standard rotations, Algorithmica **31**(2001), 501-512.

[11] J.M. Lucas, A direct algorithm for restricted rotation distance, Inform. Process. Lett. **90**(2004), 129-134.

[12] J.M. Lucas, Localized rotation distance in binary trees, Congr. Numer. **169**(2004), 161-178.

[13] J.M. Lucas, Untangling binary trees via rotations, Comput. J. **47**(2004), 259-269.

[14] F. Luccio, L. Pagli, On the upper bound of the rotation distance of binary trees, Inform. Process. Lett. **31**(1989), 57-60.

[15] E. Mäkinen, On the rotation distance of binary trees, Inform. Process. Lett. **26**(1987/88), 271-272.

[16] G. Markowski, Primes, irreducibles and extremal lattices, Order **9**(1992), 265-290.

[17] J.M. Pallo, Enumerating, ranking and unranking binary trees, Comput. J. **29**(1986), 171-175.

[18] J.M. Pallo, On the rotation distance in the lattice of binary trees, Inform. Process. Lett. **25**(1987), 369-373.

[19] J.M. Pallo, Some properties of the rotation lattice of binary trees, Comput. J. **31**(1988), 564-565.

[20] J.M. Pallo, An algorithm to compute the Mbius function of the rotation lattice of binary trees, RAIRO Theoret. Inform. Appl. **27**(1993), 341-348.

[21] J.M. Pallo, An efficient upper bound of the rotation distance of binary trees, Inform. Process. Lett. **73**(2000), 87-92.

[22] J.M. Pallo, Right-arm rotation distance between binary trees, Inform. Process. Lett. **87**(2003), 173-177.

[23] R.O. Rogers, On finding shortest paths in the rotation graph of binary trees, Congr. Numer. **137**(1999), 75-95.

[24] A.A. Ruiz, F. Luccio, A.M. Enriquez, L. Pagli, k-restricted rotation with an application to search tree rebalancing, $9th$ WADS, Lecture Notes in Computer Science, Springer, vol. **3608**(2005), 2-13.

[25] B.E. Sagan, A generalization of Rota's NBC theorem, Adv. Math. **111**(1995), 195-207.

[26] D.D. Sleator, R.E. Tarjan, W. Thurston, Rotation distance, triangulations and hyperbolic geometry, J. Amer. Math. Soc. **1**(1988), 647-681.

[27] D. Tamari, Monodes prordonns et chanes de Malcev, Bull. Soc. Math. France **82**(1954), 53-96.

[28] R. Wilber, Lower bounds for accessing binary search trees with rotations, SIAM J. Comput. **18**(1989), 56-67.