ON A SUBPOSET OF THE TAMARI LATTICE

SEBASTIAN A. CSAR, RIK SENGUPTA, AND WARUT SUKSOMPONG

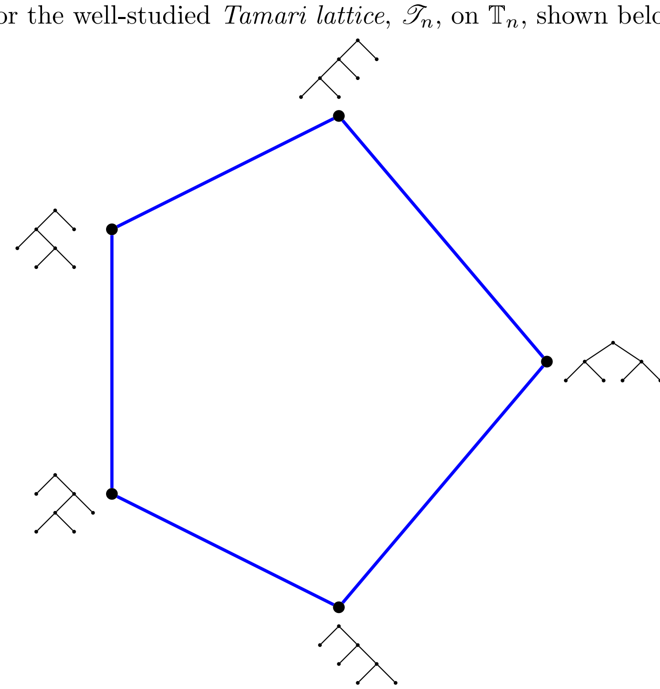
Abstract. We explore some of the properties of a subposet of the Tamari lattice introduced by Pallo, which we call the comb poset. We show that a number of binary functions that are not well-behaved in the Tamari lattice are remarkably well-behaved within an interval of the comb poset: rotation distance, meets and joins, and the common parse words function for a pair of trees. We relate this poset to a partial order on the symmetric group studied by Edelman.

1. Introduction

The set $\mathbb{T}_n$ of all full binary trees with n leaves, or parenthesizations of n letters, has been well-studied, and carries much structure. Its cardinality $|\mathbb{T}_n|$ is the $(n-1)^{\text{th}}$ Catalan number

$$C_{n-1} = \frac{1}{n}\binom{2n-2}{n-1}.$$

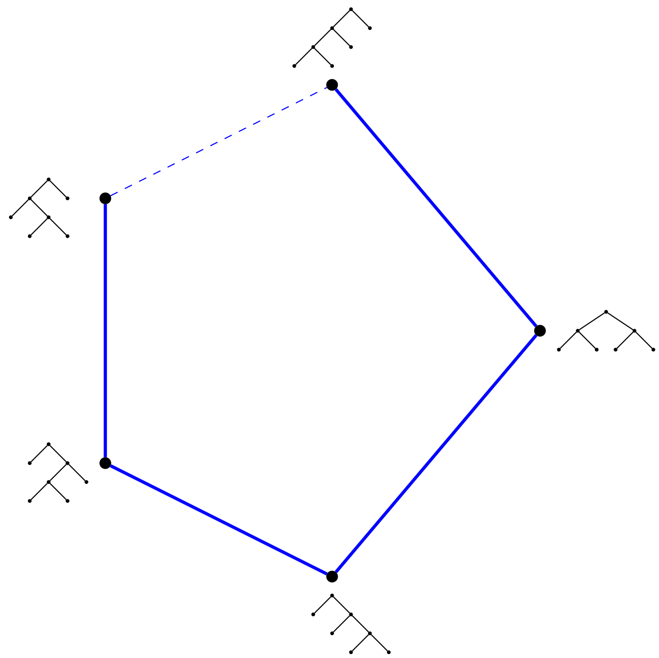
The *rotation graph*, $\mathscr{R}_n$, is the graph with vertex set $\mathbb{T}_n$, in which edges correspond to a local change in the tree called a *rotation*, corresponding to changing a single parenthesis pair in the parenthesization. This graph $\mathscr{R}_n$ forms the vertices and edges of an $(n-2)$-dimensional convex polytope called the *associahedron*, K_{n+1}. If we direct the edges of $\mathscr{R}_n$ in a certain fashion, we obtain the Hasse diagram for the well-studied *Tamari lattice*, $\mathscr{T}_n$, on $\mathbb{T}_n$, shown below for $n = 4$.



The Tamari lattice has many properties, but it has certain deficiencies. For instance, it is not ranked. Although one can encode the Tamari order by componentwise comparison of *weight vectors* $\langle T\rangle \in \{0, 1, \ldots, n-2\}^{n-1}$ for $T \in \mathscr{T}_n$, introduced by Huang and Tamari in [4] for the lattice dual to $\mathscr{C}_n$, only the meet is given by the componentwise minimum of these weight vectors; the join cannot be characterized similarly. Furthermore, computing the *rotation distance* $d_{\mathscr{R}_n}(T_1, T_2)$ between two trees T_1, T_2 in the graph $\mathscr{R}_n$ does not appear to follow easily from knowing their meet and join in the Tamari lattice.

Relying on work of Whitney in [13], Kauffman reformulated the Four Color Theorem using the vector cross product in [5]. More recently, in [1], Cooper, Rowland and Zeilberger transformed the Four Color Theorem into a question about another binary function on $\mathbb{T}_n$: the size of the set $\mathrm{ParseWords}(T_1, T_2)$ consisting of all words $w \in \{0, 1, 2\}^n$ which are *parsed* by both T_1 and T_2. Here, a word w is parsed by T if the labeling of the leaves of T by $w_1, w_2, \ldots, w_n$ from left to right extends to a proper 3-coloring with colors $\{0, 1, 2\}$ of *all* $2n-1$ vertices in T, such that no two children of the same vertex have the same label and such that no parent and child share the same label. The Four Color Theorem is equivalent to the statement that for all n and all $T_1, T_2 \in \mathbb{T}_n$, one has $|\mathrm{ParseWords}(T_1, T_2)| \geq 1$. Tamari offers a similar reformulation of the Four Color Theorem in [12].

This last application to the Four Color Theorem motivated us to investigate a poset $\mathscr{C}_n$ on the set $\mathbb{T}_n$, which we call the *(right) comb order*, a weakening of the Tamari order. Pallo first defined $\mathscr{C}_n$ in [8], where he proved that it is a meet-semilattice having the same bottom element as $\mathscr{T}_n$, called the right comb tree and denoted $\mathrm{RCT}(n)$. The solid edges in the diagram below form the Hasse diagram of $\mathscr{C}_4$. The dashed edge lies in $\mathscr{T}_4$ but not in $\mathscr{C}_4$.



While the comb order $\mathscr{C}_n$ is a meet-semilattice whose meet $\wedge_{\mathscr{C}_n}$ does not in general coincide with the Tamari meet $\wedge_{\mathscr{T}_n}$, it fixes several deficiencies of $\mathscr{T}_n$ noted above:

- $\mathscr{C}_n$ is ranked, with exactly $\binom{n+r-2}{r} - \binom{n+r-2}{r-1}$ elements of rank r, $0 \leq r \leq n-2$ (see Theorem 3.2).
- $\mathscr{C}_n$ is *locally distributive*; each interval forms a distributive lattice (see Corollary 2.12(i)).
- If T_1 and T_2 have an upper bound in $\mathscr{C}_n$ (or equivalently, if they both lie in some interval), the meet $T_1 \wedge_{\mathscr{C}_n} T_2$ and join $T_1 \vee_{\mathscr{C}_n} T_2$ are easily described combinatorially in two different ways (see Corollary 2.12(i) and Theorem 5.3). These operations also coincide with the Tamari meet $\wedge_{\mathscr{T}_n}$ and Tamari join $\vee_{\mathscr{T}_n}$ (see Corollary 5.4).
- When trees T_1, T_2 have an upper bound in $\mathscr{C}_n$, one has (see Theorem 4.4)
$$\begin{aligned} d_{\mathscr{R}_n}(T_1, T_2) &= \mathrm{rank}(T_1) + \mathrm{rank}(T_2) - 2 \cdot \mathrm{rank}(T_1 \wedge_{\mathscr{C}_n} T_2) \\ &= 2 \cdot \mathrm{rank}(T_1 \vee_{\mathscr{C}_n} T_2) - (\mathrm{rank}(T_1) + \mathrm{rank}(T_2)) \\ &= \mathrm{rank}(T_1 \vee_{\mathscr{C}_n} T_2) - \mathrm{rank}(T_1 \wedge_{\mathscr{C}_n} T_2), \end{aligned}$$
where, for any $T \in \mathbb{T}_n$, $\mathrm{rank}(T)$ refers to the rank of T in $\mathscr{C}_n$.
- Furthermore, for T_1, T_2 having an upper bound in $\mathscr{C}_n$, one has (see Theorem 7.9)
$$\mathrm{ParseWords}(T_1, T_2) = \mathrm{ParseWords}(T_1 \wedge_{\mathscr{C}_n} T_2, T_1 \vee_{\mathscr{C}_n} T_2),$$
with cardinality $3 \cdot 2^{n-1-k}$, where $k = \mathrm{rank}(T_1 \vee_{\mathscr{C}_n} T_2) - \mathrm{rank}(T_1 \wedge_{\mathscr{C}_n} T_2)$ (see Theorem 7.7).

Lastly, Section 6 discusses a well-known order-preserving surjection from the (right) weak order on the symmetric group $\mathfrak{S}_n$ to the Tamari poset $\mathscr{T}_{n+1}$ and its restriction to an order-preserving surjection from $\mathscr{E}_n$ to $\mathscr{C}_{n+1}$ (where $\mathscr{E}_n$ is a subposet of the weak order considered by Edelman in [2]). Furthermore, this surjection is a distributive lattice morphism on each interval of $\mathscr{C}_{n+1}$ (see Theorem 6.9).

Because we will be mainly confining our attention for the rest of this paper to the poset $\mathscr{C}_n$, we will drop the subscripts from $\wedge$, $\vee$, $>$ and $<$ when we mean meet, join, greater than, and less than in $\mathscr{C}_n$ respectively. Furthermore, we will use $\mathrm{rank}(T)$ to denote the rank of T in $\mathscr{C}_n$. Much of our notation in Section 7 is from [1].

2. The Comb Poset and Distributivity

In [4], Huang and Tamari describe the dual to the Tamari lattice in terms of *binary bracketings*, which are the usual parenthesizations of the leaves of a binary tree. However, the comb poset is most readily defined in terms of a variation on this parenthesization.

First, recall the definition of the parenthesization of a binary tree.

Definition 2.1. Suppose $T \in \mathbb{T}_n$ and its leaves are labeled $a_1, \dots, a_n$. The *parenthesization* of T is a set $\mathcal{P}$, whose elements are the subsets, J, of $\{a_1, \dots, a_n\}$ such that $J = \{a_i < \dots < a_j\}$ and $a_i < \dots < a_j$ label the leaves of a subtree of T.

Proposition 2.2. *Suppose $T \in \mathbb{T}_n$. Then, for $E \in \mathcal{P}$, either $|E| = 1$ or $E = E_1 \sqcup E_2$, where $E_1, E_2 \in \mathcal{P}$, $E_1 \cap E_2 = \emptyset$ and E_1, E_2 are unique.*

Definition 2.3. For each $T \in \mathbb{T}_n$, the *reduced parenthesization* of T is denoted RP_T and $RP_T = \mathcal{P} \smallsetminus \{E \in \mathcal{P} | a_n \in E\}$. An element E of the set RP_T with $|E| > 1$ is called a *parenthesis pair.*

As the name suggests, one can write the parenthesization and reduced parenthesization as parenthesizations of the sequence $a_1 a_2 \cdots a_n$, with the singleton sets of $\mathcal{P}$ and RP_T not drawn. This convention will be used for the remainder of the paper.

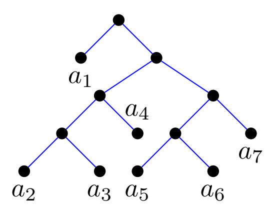


Figure 1.

Example. The full parenthesization of the tree in Figure 1 is $(a_1(((a_2a_3)a_4)((a_5a_6)a_7)))$ and its reduced parenthesization is $a_1((a_2a_3)a_4)(a_5a_6)a_7$.

Remark. All n-leaf binary trees have a unique reduced parenthesization, since there is a bijection between the full parenthesization of a tree T and its reduced parenthesization. The full parenthesization is recovered by pairing the two rightmost elements of RP_T successively.

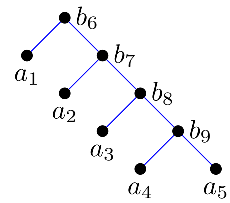
Proposition 2.4. *A collection $\mathcal{P}$ of subsets of $\{a_1, \dots, a_n\}$ is the reduced parenthesization of a tree $T \in \mathbb{T}_n$ if and only if the following conditions hold:*

(i) for each a_i, $1 \leq i < n$, there is a $P \in \mathcal{P}$ such that $a_i \in P$ and there is no $P \in \mathcal{P}$ such that $a_n \in P$
(ii) for each $P \in \mathcal{P}$, if $a_i, a_k \in P$ and $i < j < k$, then $a_j \in P$
(iii) for each $P \in \mathcal{P}$, either $|P| = 1$ or there are $P_1, P_2 \in \mathcal{P}$, with $P_1 \cap P_2 = \emptyset$, such that $P = P_1 \sqcup P_2$
(iv) if $P_1, P_2 \in \mathcal{P}$ with $P_1 \cap P_2 \neq \emptyset$, then either $P_1 \subset P_2$ or $P_2 \subset P_1$.

When $P \in RP_T$ can be written as $P_1 \sqcup P_2$ for $P_1, P_2 \in RP_T$, P_1 and P_2 are called the *factors* of P. Without loss of generality assume that for each $a_i \in P_1$ and $a_j \in P_2$ that $i < j$. Then P_1 is the *left factor* and P_2 is the *right factor* of P.

Definition 2.5. For $n \geq 2$, the *right comb tree of order* n, denoted by $\mathrm{RCT}(n) \in \mathbb{T}_n$, is the n-leaf binary tree with $RP_T = \emptyset$, corresponding to $a_1 a_2 \cdots a_n$. Similarly, the *left comb tree* of order n is defined as the n-leaf binary tree corresponding to the reduced parenthesization $RP_T = \{\{a_1, \ldots, a_i\} \colon i = 1, \ldots, n-1\}$, corresponding to $(((\cdots((a_1 a_2)a_3)\cdots)a_{n-2})a_{n-1})a_n$.

Example. RCT(5), the right comb tree of order 5, is shown below. The nodes labeled $a_1, \ldots, a_5$ are the leaves of the tree, and $b_6, \ldots, b_9$ are the internal vertices. Note that the structure of the left comb tree of order 5 is given by the reflection of the right comb tree about the vertical axis.



Definition 2.6. For $n \geq 2$, the *(right) comb poset* of order n is the poset whose elements are $\mathbb{T}_n$ and with $T_1 \leq T_2$ if $RP_{T_1} \subseteq RP_{T_2}$.

Remark. One sees immediately that $\mathrm{RCT}(n)$ is the unique minimal element of $\mathscr{C}_n$ since its reduced parenthesization is the empty set.

Example. The Hasse diagram of the right comb poset of order 5 is shown in Figure 2. For the sake of a cleaner diagram, the leaf a_i is labeled by i in Figure 2 for $i \in \{1, 2, 3, 4, 5\}$.

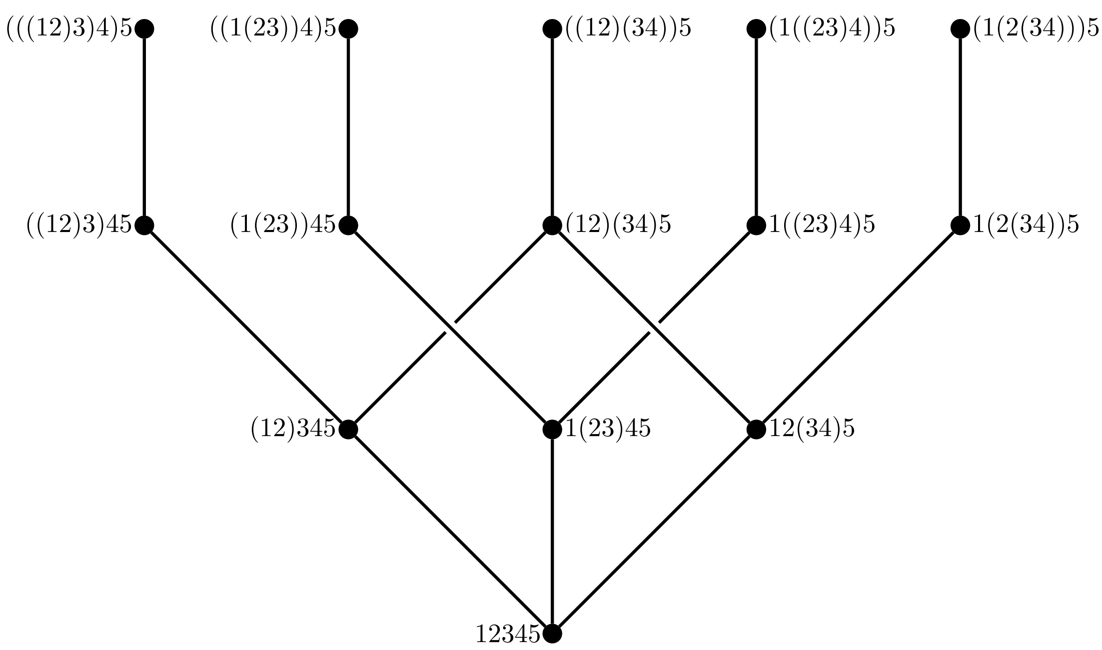


FIGURE 2. The Hasse diagram of $\mathscr{C}_5$

Proposition 2.7. *There is an order-preserving involution on $\mathscr{C}_n$.*

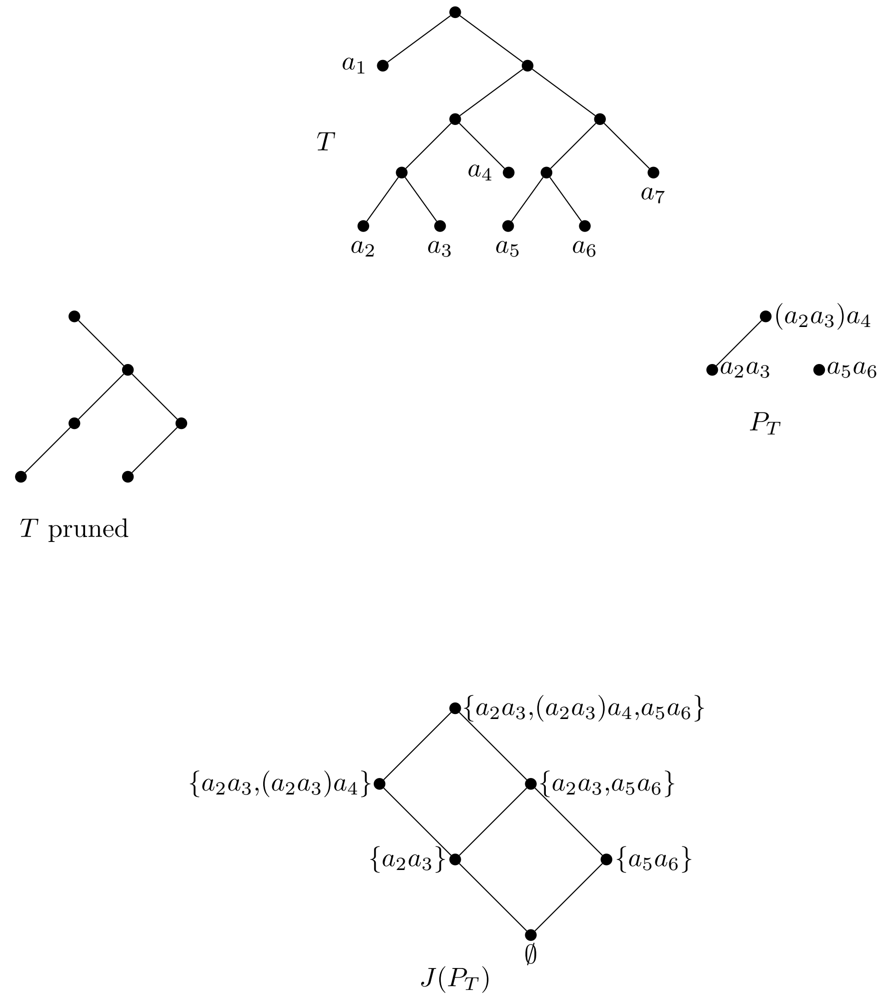


FIGURE 3.

Proof. For any tree T, take RP_T, and construct a new parenthesization $RP_{T'}$ as follows. For every parenthesis pair in RP_T which encloses leaves a_i through a_j, take $RP_{T'}$ to have a parenthesis pair enclosing leaves a_{n-j} through a_{n-i}. It is not hard to see (using Proposition 2.4) that $RP_{T'}$ corresponds to a tree T'. Define π to be the map that takes T to T' as described above. Then, π is an order preserving involution on $\mathscr{C}_n$. □

Definition 2.8. For n ≥ 2, the *right arm* of a tree $T \in \mathbb{T}_n$ is the path induced by the vertices of T that lie in the left subtree of no other vertex in T.

To understand the properties of the intervals of $\mathscr{C}_n$, one needs to define another poset using RP_T. It is well known that the operation of "pruning" a tree, i.e. deleting the leaves, is a bijection between n-leaf binary trees and (possibly incomplete) binary trees with $n-1$ vertices.

Definition 2.9. For a tree $T \in \mathbb{T}_n$, the *reduced pruned poset* of T, denoted P_T is the poset obtained by ordering by inclusion those elements of RP_T which are not singleton sets. Its Hasse diagram is obtained by pruning T, removing the right arm and removing those edges incident to the right arm.

Example. Consider the tree of Figure 1, given by reduced parenthesization $a_1((a_2a_3)a_4)(a_5a_6)a_7$. Figure 3 depicts its "pruned" form, the corresponding reduced pruned poset P_T and $J(P_T)$.

Proposition 2.10. *For any tree $T \in \mathbb{T}_n$, the maximal elements of P_T correspond to the left subtrees of the vertices of the right arm of T.*

Proposition 2.11. *For any $T \in \mathbb{T}_n$, the interval $[\mathrm{RCT}(n), T]_{\mathscr{C}_n}$ is isomorphic to the lattice of order ideals in the reduced pruned poset of T, ordered by inclusion. In other words, for any tree T,*

$$[\mathrm{RCT}(n), T]_{\mathscr{C}_n} \cong J(P_T)$$

Proof. One has a natural map $J(P_T) \to [\mathrm{RCT}(n), T]$, given by $I \mapsto S$, where S is the tree with RP_S having precisely the parentheses in I. The definition of the order on $\mathscr{C}_n$ ensures this map is both well-defined and order-preserving. Furthermore, this map has an inverse $[\mathrm{RCT}(n), T] \to J(P_T)$ given by $S \mapsto \{E \in RP_S : |E| > 1\}$, which is again order-preserving. □

This proposition yields a number of immediate corollaries.

Corollary 2.12.

(i) Any interval in $\mathscr{C}_n$ is a distributive lattice, with the reduced parenthesizations of the join and meet of trees T_1 and T_2 in an interval given by the ordinary union and intersection of parenthesis pairs from RP_{T_1} and RP_{T_2}.

(ii) In $\mathscr{C}_n$, T_1 covers T_2 if and only if RP_{T_1} can be obtained from RP_{T_2} by adding one parenthesis pair.

(iii) $\mathscr{C}_n$ is a ranked poset, with the rank of any tree T in $\mathscr{C}_n$ given by the number of parenthesis pairs in RP_T.

(iv) For any two trees T_1 and T_2 that are in the same interval of $\mathscr{C}_n$, we have

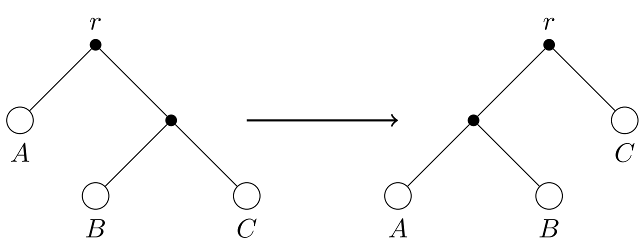
$$\mathit{rank}(T_1) + \mathit{rank}(T_2) = \mathit{rank}(T_1 \wedge T_2) + \mathit{rank}(T_1 \vee T_2)$$

(v) For any tree $T \in \mathbb{T}_n$ of rank k, the length of the right arm of T is $n - 1 - k$.

Remark. It is important to note that Corollary presumes pairs of parentheses have knowledge of their factors when talking about the "ordinary" union and intersection. For example, the trees with reduced parenthesizations $(a(bc))d$ and $((ab)c)d$ do not have a join, as, in one case, the factors of $\{a, b, c\}$ are $\{a\}$ and $\{b, c\}$, while in the other the factors are $\{a, b\}$ and $\{c\}$. Similarly, the meet of these two trees is the right comb tree, having reduced parenthesization $abcd$.

Note. In the remainder of this paper, we shall consider only the right comb poset of order n. Analogous results hold for the left comb poset by symmetry.

Remark. A *left rotation* is the following operation on a tree, which takes place in a subtree with root r:



Right arm rotations are those where r lies on the right arm of the tree. The covering relation described in Corollary 2.12(ii) corresponds right arm rotation, precisely the covering relation used by Pallo in [8] to define the poset $(B_n, \overset{*}{\rightsquigarrow})$, which he showed to be a meet-semilattice [8, Lemma 3].

3. Rank Sizes in the Comb Poset

In this section, we will prove some enumerative properties of the ranks of $\mathscr{C}_n$. To simplify notation, let Q_i denote the i^{th} rank of $\mathscr{C}_n$.

Proposition 3.1. *For $0 \leq i \leq n-2$, every tree in Q_i is covered by precisely $n - 2 - i$ trees.*

Proof. This fact follows from the definition of rotation, and the observation that a tree in rank Q_i has, by Corollary 2.12(v), a right arm of length $n - 1 - i$, i.e. $n - i$ vertices. □

Theorem 3.2. *For $n \geq 3$, $\mathscr{C}_n$ is a ranked poset. A tree is a maximal element of $\mathscr{C}_n$ if and only if it is of rank $n-2$ in $\mathscr{C}_n$ (or equivalently, from Corollary 2.12(v), if and only if its right arm has length 1). In particular, the left comb tree is in the maximal rank of $\mathscr{C}_n$. Furthermore, for $0 \leq r \leq n-2$, the number of elements in rank r of $\mathscr{C}_n$ is*

$$|Q_r| = \binom{n+r-2}{r} - \binom{n+r-2}{r-1}.$$

The authors thank an anonymous referee for pointing out the following combinatorial proof.

Proof. Suppose $n \geq 3$ and T is in rank r of $\mathscr{C}_n$. Then RP_T has r pairs of parentheses, which are completely determined by the open parentheses as a consequence of Proposition 2.4. Furthermore, when viewing RP_T as a parenthesization of $a_1 \cdots a_n$ and reading from right to left there are always at least two more a_i than open parentheses, as there can be no open parenthesis immediately preceding either a_{n-1} or a_n. One may read off a lattice path from $(0,0)$ to $(n-1,r)$ that touches the line $y=x$ only at $(0,0)$ as follows. One deletes the closed parentheses and a_n from RP_T, leaving a string consisting of a_i, $i \in \{1,\ldots,n-1\}$, and open parentheses. One obtains a lattice path by reading from right to left and recording an east step for each a_i and a north step for each (. (Having deleted a_n means that there will only be $n-1$ east steps and that there will always have been at least one more east step than north step.) The number of such paths is well-known (see, for example, [11, Exercise 6.20b]) and one has

$$\begin{aligned}
|Q_r| &= \frac{n-1-r}{n-1+r}\binom{n-1+r}{r} \\
&= ((n-1)-r)\frac{(n-2+r)!}{r!(n-1)!} \\
&= \frac{(n-2+r)!}{r!(n-2)!} - \frac{(n-2+r)!}{(r-1)!(n-1)!} \\
&= \binom{n+r-2}{r} - \binom{n+r-2}{r-1}.
\end{aligned}$$

□

Corollary 3.3. *The sizes of the ranks in $\mathscr{C}_n$ weakly increase. In fact, they strictly increase until the final rank Q_{n-2}, which has the same size, C_{n-2}, as the penultimate rank Q_{n-3}.*

Proof. From Theorem 3.2, it can be seen that $|Q_i| = \frac{(n+i-2)!}{i!(n-1)!} \cdot (n-i-1)$, and so, for consecutive ranks r and $r+1$, one has

$$\frac{|Q_{r+1}|}{|Q_r|} = \frac{(n+r-1)(n-2-r)}{(r+1)(n-1-r)}.$$

The rank size increases weakly whenever the numerator is at least as large as the denominator, and hence the condition for weakly increasing rank size is $(n+r-1)(n-r-2) \geq (r+1)(n-r-1)$. But this condition reduces after a few simple manipulations to the condition $n^2-4n+3-r(n-1) \geq 0$. The result can be verified easily. □

4. Distances in $\mathscr{C}_n$ and $\mathscr{R}_n$

We now prove some properties of the comb poset relating to the distance between pairs of trees in the rotation graph $\mathscr{R}_n$.

Proposition 4.1. *Any ascending chain in the right comb poset $\mathscr{C}_n$ is an ascending chain in $\mathscr{T}_n$.*

Proof. From Corollary 2.12(ii), one has that T_2 is a cover of T_1 in $\mathscr{C}_n$ if and only if RP_{T_2} can be obtained from RP_{T_1} by adding precisely one more parenthesis pair. Adding any parenthesis pair to RP_T is the same as shifting a pair of parentheses to the *left* in the corresponding full parenthesization of the leaves of T. □

Proposition 4.2. *Suppose T_1 and T_2 are two trees having a common upper bound in $\mathscr{C}_n$. Furthermore, suppose there are pairs of parentheses J_1 in RP_{T_1} and J_2 in RP_{T_2} such that J_1 and J_2 enclose a common factor. Then, $J_1 = J_2$.*

Proof. Suppose T_1 and T_2 have a common upper bound in $\mathscr{C}_n$ and suppose there are pairs of parentheses J_1 in RP_{T_1} and J_2 in RP_{T_2} enclosing a common factor, P. As T_1 and T_2 have a common upper bound, $T_1 \vee T_2$, from Corollary 2.12(i), one has that $P, J_1, J_2 \in RP_{T_1 \vee T_2}$. From Proposition 2.4, one has without loss of generality that $J_2 \subset J_1$, with $J_1 = P \sqcup P_1$ and $J_2 = P \sqcup P_2$. One then has that $P_1 = P_2 \sqcup P_3$, for some P_3. Both J_2 and P_1 are in $RP_{T_1 \vee T_2}$ and have nontrivial intersection, yet neither contains the other, contradicting Proposition 2.4. $\square$

Lemma 4.3. *Suppose T_1 and T_2 are two trees with a common upper bound in $\mathscr{C}_n$ and suppose $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ is a path from T_1 to T_2 in $\mathscr{R}_n$, with λ_i being the rotation between trees S_i and S_{i+1}, with $S_1 = T_1$ and λ_ℓ the rotation from S_ℓ to T_2. Let $RP_{T_1} \triangle RP_{T_2}$ denote the symmetric difference of RP_{T_1} and RP_{T_2}. Let $f\colon RP_{T_1} \triangle RP_{T_2} \to \{\lambda_1, \ldots, \lambda_s\}$ be the map defined by $f(J) = \lambda_j$, where j is the minimum index such that $J \in RP_{S_i} \setminus RP_{S_{i+1}}$ or $J \in RP_{S_{i+1}} \setminus RP_{S_i}$. Then f is injective and the shortest possible length of a path from T_1 to T_2 along the edges of the rotation graph $\mathscr{R}_n$ is $|RP_{T_1} \triangle RP_{T_2}|$.*

Proof. From Corollary 2.12(i) one has that $RP_{T_1 \wedge T_2}$ contains all the common parenthesis pairs of RP_{T_1} and RP_{T_2}. Hence, RP_{T_1} and RP_{T_2} are formed by adding, respectively, some r and s *extra* pairs of parentheses to $RP_{T_1 \wedge T_2}$, from Corollary 2.12(ii), where r and s are nonnegative integers, and $|RP_{T_1} \triangle RP_{T_2}| = r + s$.

Suppose $f(J) = f(K) = \lambda_j$. If $J \neq K$, then λ_j is a rotation sending J to K or vice versa. Without loss of generality, assume λ_j is a rotation sending J to K. Since $J, K \in RP_{T_1} \triangle RP_{T_2}$, both are in $RP_{T_1 \vee T_2}$. However, since λ_j is a rotation sending J to K, one must have that J and K share a factor. But then Proposition 4.2 forces $J = K$, a contradiction. Thus f must be injective, so the minimum length of a path from T_1 to T_2 in $\mathscr{R}_n$ is $|RP_{T_1} \triangle RP_{T_2}|$. $\square$

Theorem 4.4. *If T_1 and T_2 are two trees in some interval in $\mathscr{C}_n$, then the shortest distance between them* along the edges of the rotation graph $\mathscr{R}_n$ *is given by*

$$d_{\mathscr{R}_n}(T_1, T_2) = \mathit{rank}(T_1) + \mathit{rank}(T_2) - 2 \cdot \mathit{rank}(T_1 \wedge T_2).$$

Equivalently, from Corollary 2.12(iv), this shortest distance is also given by

$$d_{\mathscr{R}_n}(T_1, T_2) = 2 \cdot \mathit{rank}(T_1 \vee T_2) - \mathit{rank}(T_1) - \mathit{rank}(T_2).$$

Proof. If T_1 and T_2 are two trees in some interval in $\mathscr{C}_n$, $|RP_{T_1} \triangle RP_{T_2}| = \mathrm{rank}(T_1) + \mathrm{rank}(T_2) - 2 \cdot \mathrm{rank}(T_1 \wedge T_2)$. From Lemma 4.3, one knows that the minimal possible length of a path from T_1 to T_2 is $|RP_{T_1} \triangle RP_{T_2}|$. Furthermore, one knows a path of this length exists – the path in $\mathscr{C}_n$ from T_1 to $T_1 \wedge T_2$ obtained by deleting the pairs of parentheses in RP_{T_1} that do not appear in RP_{T_2}, followed by the path from $T_1 \wedge T_2$ to T_2 obtained by adding the pairs of parentheses in RP_{T_2} not appearing in RP_{T_1}. $\square$

Theorem 4.5. *For T_1, T_2 with an upper bound in $\mathscr{C}_n$, any shortest path in $\mathscr{R}_n$ from T_1 to T_2 also lies in $\mathscr{C}_n$.*

Proof. Suppose T_1 and T_2 lie in some interval of $\mathscr{C}_n$ and $(\lambda_1, \ldots, \lambda_{r+s})$ is a shortest path from T_1 to T_2 in $\mathscr{R}_n$, with λ_i being a rotation between trees S_i and S_{i+1}. Suppose λ_i is a rotation not centered on the right arm of S_i, i.e. it "shifts" a pair of parentheses, so $|RP_{S_i}| = |RP_{S_{i+1}}|$, and assume that λ_i is the first such rotation. In precise terms, this means there are two pairs of parentheses J, J' with $\{J\} = RP_{S_i} \setminus RP_{S_{i+1}}$ and $\{J'\} = RP_{S_{i+1}} \setminus RP_{S_i}$. Since $(\lambda_1, \ldots, \lambda_{r+s})$ is a path of shortest possible length, the map f in Lemma 4.3 is a bijection and one must have $J \in RP_{T_1} \setminus RP_{T_2}$ or $J' \in RP_{T_2} \setminus RP_{T_1}$, but not both.

Suppose $J' \in RP_{T_2} \setminus RP_{T_1}$. Then there must be some rotation λ_j with $j < i$ and $\{J\} = RP_{S_{j+1}} \setminus RP_{S_j}$. However, since $J \notin RP_{T_1}$, since f is a bijection, one must have that λ_j transformed some $J'' \in RP_{T_1} \cap RP_{S_j}$ into J, i.e. λ_j cannot be centered on the right arm of RP_{S_j}, contradicting that λ_i was the first such rotation.

Consequently, one must have $J \in RP_{T_1} \setminus RP_{T_2}$. Without loss of generality, one may assume that λ_i shifts J to the right, i.e. $J = A \sqcup B$ for factor $A, B \in RP_{S_i}$ and $J' = B \sqcup C$ for factors $B, C \in RP_{S_{i+1}}$. Moreover, for the rotation λ_i to take place, one must have that $A, B, C \in RP_{S_i} \cap RP_{S_{i+1}}$ and that there is some $K = A \sqcup B \sqcup C \in RP_{S_i} \cap RP_{S_{i+1}}$. Since $J \in RP_{T_1}$, one has that $J \in RP_{T_1 \vee T_2}$, so since $J \cap J' = B$, one must have that $J' \notin RP_{T_1 \vee T_2}$. Then there is a $j > i$ such that λ_j transforms J' to some $J'' \in RP_{T_1 \vee T_2}$. Moreover, since f is a bijection, one must have that $J \neq J''$.

Suppose J encloses the leaves $\{a_{m+1}, \dots, a_q\}$, J' encloses the leaves $\{a_{p+1}, \dots, a_t\}$ and K encloses $\{a_{m+1}, \dots, a_t\}$ for $0 \leq m < p < q < t < n$. J'' is obtained from J' by a rotation, so J'' must enclose either a_{p+1} or a_t. However, $a_{p+1} \in J$ and, since $J, J'' \in RP_{T_1 \vee T_2}$, $J \cap J'' = \emptyset$. Consequently, $a_t \in J''$. However, since a_t was the last leaf enclosed by J', it cannot be the last leaf enclosed by J''. Recall that a_t is the last leaf enclosed by K, so $J'' \not\subset K$. Since $K \cap J'' \neq \emptyset$ and $J'' \not\subset K$, one must have that $K \notin RP_{T_1 \vee T_2}$. Thus, K is in neither RP_{T_1} nor RP_{T_2}, so it must result from a rotation λ_k with $RP_{S_k} \triangle RP_{S_{k+1}} = \{K, K'\}$, i.e. with λ_k not centered on the right arm of S_k. But since $K \in RP_{S_i}$, $k < i$, a contradiction.

Thus, all rotations λ_j must be centered on the right arm of S_j, i.e. the path $(\lambda_1, \dots, \lambda_{r+s})$ lies entirely in $\mathscr{C}_n$.

□

Corollary 4.6. *The rank of any tree $T \in \mathbb{T}_n$ in $\mathscr{C}_n$ is its distance from the right comb tree along the edges of the rotation graph $\mathscr{R}_n$. Furthermore, from Corollary 2.12(iii), the distance of T from the right comb tree in $\mathscr{R}_n$ is given by the number of parenthesis pairs in RP_T.*

Remark. It can be easily shown from the result above that the *diameter* of the rotation graph $\mathscr{R}_n$, given by the maximum distance between any pair of trees in $\mathscr{R}_n$, is at most $2n-4$ for any $n \in \mathbb{N}$. In [9], Sleator, Tarjan and Thurston established the tighter bound of $2n-6$ on the diameter of the rotation graph for $n \geq 11$.

5. Tamari Meets and Joins for Two Trees in Some Interval

From Corollary 2.12(i), we know the meaning of the meet and join of a pair of trees having a common upper bound in our poset. It is natural to ask how these meets and joins relate to meets and joins in the Tamari lattice. As before, we will refers to meets and joins in the Tamari lattice $\mathscr{T}_n$ as the "Tamari meet" and "Tamari join".

The first observation is that, while two arbitrary trees in $\mathscr{C}_n$ *do* have a well-defined meet in $\mathscr{C}_n$, this meet does not necessarily correspond to the Tamari meet. For example, consider the pair of trees represented by $T_1 = (((a_1a_2)a_3)a_4)a_5$ and $T_2 = ((a_1(a_2a_3))a_4)a_5$. This pair has Tamari meet T_2, while their meet in $\mathscr{C}_n$ is just the right comb tree. Further, recall that $\mathscr{C}_n$ is a meet-semilattice rather than a lattice, so not all pairs of trees have a join.

However, something much stronger can be said if both the trees under consideration are in some interval in the comb poset; it turns out that their meet and join in $\mathscr{C}_n$ correspond to their Tamari meet and join.

In [4], Huang and Tamari consider the lattice dual to $\mathscr{T}_n$ and characterize the meet in that lattice as the componentwise minimum of the *bracketing vectors*. In [7], Pallo obtains an analogous result for $\mathscr{T}_n$ in terms of *weight vectors*, which will be of use here.

Definition 5.1. Suppose $T \in \mathbb{T}_n$. For each $i \in \{1, \dots, n-1\}$, let $w_T(i) = \max_{E \in RP_T\colon i = \max E} |E|$. The *weight vector* of T is $\langle T \rangle = \langle w_T(i) \rangle$.

Example. Consider the tree T having reduced parenthesization $((a_1a_2)(a_3a_4))a_5(a_6(a_7a_8))a_9$. For illustrative purposes, enclose each a_i in a pair of parentheses to represent the singleton sets in RP_T, giving

$$(((a_1)(a_2))((a_3)(a_4)))(a_5)((a_6)((a_7)(a_8)))(a_9).$$

Then $\langle T \rangle = (1, 2, 1, 4, 1, 1, 1, 3)$.

Theorem 5.2 (Pallo, [7, Theorem 2])**.** *For two n-leaf binary trees T and T', one has $T \leq T'$ if and only if the weight vector of T is component-wise less than or equal to the weight vector of T'. Furthermore, the bracketing vector for the meet of two trees in the Tamari lattice corresponds to the componentwise minimum of the weight vectors of the two trees.*

Theorem 5.3. *Let $\langle T \rangle$ denote the weight vector of $T \in \mathbb{T}_n$. Let T_1 and T_2 be arbitrary trees in the same interval of $\mathscr{C}_n$. Then, their meet and join in $\mathscr{C}_n$ are given by the trees corresponding respectively to the componentwise minimum and the componentwise maximum of $\langle T_1 \rangle$ and $\langle T_2 \rangle$.*

Proof. First, consider $\langle T_1 \vee T_2 \rangle$. Suppose the ith coordinate is k. Then, by definition, $k = \max_{E \in RP_{T_1 \vee T_2}\,:\, i \in E} |E|$. From Corollary 2.12(i), one has that $RP_{T_1 \vee T_2} = RP_{T_1} \cup RP_{T_2}$, so one must have that $k = \max(w_{T_1}(i), w_{T_2}(i))$. In other words, $\langle T_1 \vee T_2 \rangle$ is the componentwise maximum of $\langle T_1 \rangle$ and $\langle T_2 \rangle$. The proof for $T_1 \wedge T_2$ is analogous. $\square$

Corollary 5.4. *For T_1 and T_2 in some interval in $\mathscr{C}_n$, their meet and join in $\mathscr{C}_n$ correspond respectively to their meet and join in the Tamari lattice $\mathscr{T}_n$.*

Proof. The proof for the meet follows directly from Theorems 5.2 and 5.3. For the join, observe that the tree corresponding to the componentwise maximum of $\langle T_1 \rangle$ and $\langle T_2 \rangle$ would be the join of T_1 and T_2 in $\mathscr{T}_n$. However, in general, one does not know that such a tree exists. However, Theorem 5.3 gives that such a tree exists—it is the join of T_1 and T_2 in $\mathscr{C}_n$. $\square$

6. Relation with a Poset of Edelman

In [2], Edelman introduced a subposet of the right weak order on the symmetric group $\mathfrak{S}_n$. Although this poset is not a lattice, the intervals are known to each be distributive lattices, as is the case for the comb poset $\mathscr{C}_n$.

Definition 6.1. The *right weak order on* $\mathfrak{S}_n$ is a partial ordering of the elements of $\mathfrak{S}_n$ defined as the transitive closure of the following covering relation: a permutation σ covers a permutation τ if σ is obtained from τ by a transposition of $\tau(i)$ and $\tau(i+1)$, two adjacent elements of the one line notation of τ, such that $\tau(i) < \tau(i+1)$.

Edelman imposed an additional constraint on this ordering, under which σ covers τ, if, after the transposition of $\tau(j)$ and $\tau(j+1)$ as above, nothing to the left of $\tau(j+1)$ in σ is greater than $\tau(j+1)$. This restriction results in a subposet of the right weak ordering on $\mathfrak{S}_n$. Denote this poset by $\mathscr{E}_n$.

Example. Figure 4 depicts the Hasse diagram of $\mathscr{E}_3$, with an additional dashed edge indicating the extra order relation in the right weak order on $\mathfrak{S}_3$.

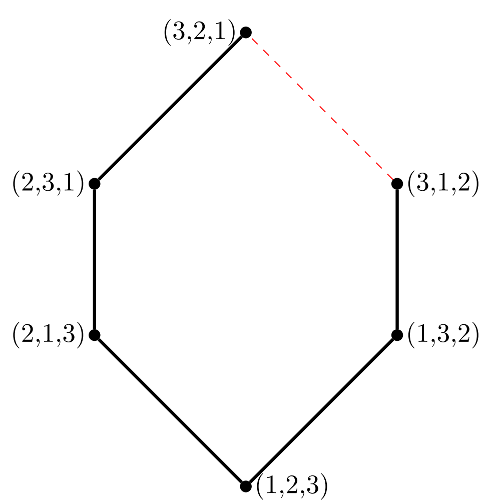
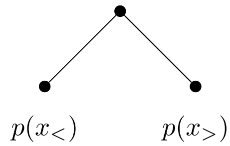


Figure 4. Edelman's Poset $\mathscr{E}_3$.

Definition 6.2. The *pruned tree map*, $p\colon \mathfrak{S}_n \to \{\text{pruned trees on } n \text{ vertices}\}$, is defined recursively as follows. For $x \in \mathfrak{S}_1$, $p(x)$ is the tree with a single vertex. Then, for $n > 1$ and $x \in \mathfrak{S}_n$, define

$p(x) =$ 

where $x_< = (x_{i_1}, \dots, x_{i_k})$ where $i_1 < \cdots < i_k$ are the indices of all elements of x less than x_1 and $x_>$ is defined similarly for elements of x greater than x_1. Extend p to a map $\beta\colon \mathfrak{S}_n \to \mathbb{T}_{n+1}$ by attaching leaves to $p(x)$ to give a binary tree (in other words, "unpruning" $p(x)$).

Remark. Amending the definition of p slightly so that the root of $p(x)$ is labeled by x_1 results in the pruned tree having the *in-order labeling*, where a vertex's label is greater than those of the vertices in its left subtree and smaller than those of the vertices in its right subtree. This labeled tree is, in fact, the unbalanced binary search tree for the permutation. (See [6].) The pruned tree map is also related to the bijection between permutations and increasing binary trees on n vertices (see [10, p. 24]): the pruned tree associated to w is the increasing binary tree associated to w^{-1} with the labels removed. Consequently, the pruned tree map is a surjection.

Example. Figure 5 shows $p\colon \mathfrak{S}_4 \to \{\text{pruned trees with 4 vertices}\}$. Permutations having the same image are circled.

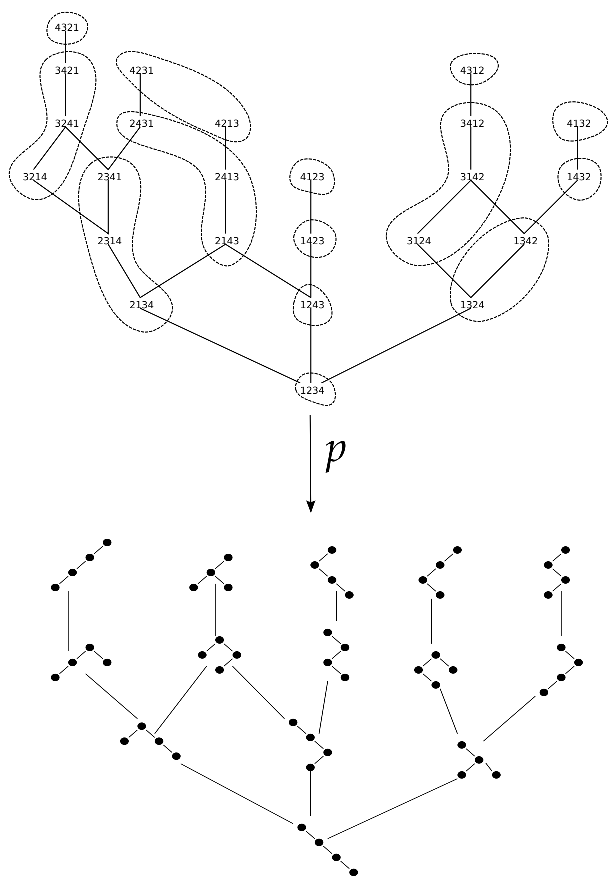


FIGURE 5.

Theorem 6.3. *The map $p\colon \mathfrak{S}_n \to \mathbb{T}_{n+1}$ gives an order-preserving surjection from $\mathscr{E}_n$ to $\mathscr{C}_{n+1}$.*

Proof. As noted above, it is well-known that p is a surjection. It suffices to show that if $\sigma \lessdot \tau$ in $\mathscr{E}_n$ and $T_1 = p(\sigma)$ and $T_2 = p(\tau)$, then either $T_1 = T_2$, or $T_1 \lessdot T_2$ in $\mathscr{C}_{n+1}$.

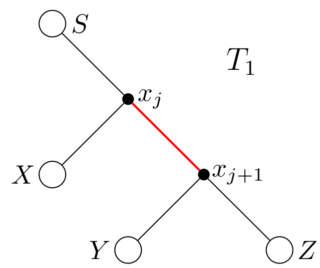
Suppose

$$\sigma = (x_1, x_2, \ldots, x_j, x_{j+1}, \ldots, x_n) \in \mathfrak{S}_n \quad \text{and} \quad \tau = (x_1, x_2, \ldots, x_{j+1}, x_j, x_{j+2}, \ldots, x_n) \in \mathfrak{S}_n,$$

with τ covering σ in $\mathscr{E}_n$. One then has that $x_s < x_{j+1}$ for all $s < j+1$. Now, if $j = 1$, then the transposition changing σ to τ corresponds, in the image of p, to a left rotation centered on the root, and therefore $T_1 \lessdot T_2$ in $\mathscr{C}_{n+1}$. So assume $j \neq 1$; in other words, x_j is not the root x_1 of the tree.

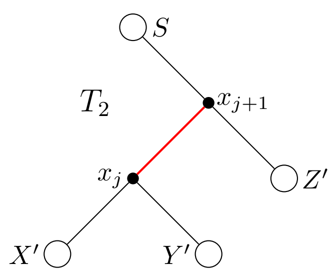
Recall that $x_s < x_{j+1}$ for all $s < j+1$. Suppose there is an $s < j$ such that $x_j < x_s < x_{j+1}$. Then, from the definition of p, one knows that the vertex labeled x_j lies in the left subtrees of that labeled x_s and x_{j+1} lies in the right subtree. When x_j and x_{j+1} are exchanged to obtain τ, their positions in the image of p do not change and $T_1 = T_2$.

Now suppose there is no $s < j$ such that $x_j < x_s < x_{j+1}$. In such a case, T_1 has the form



Here the white circle S denotes the parent tree of the entire subtree shown, with the condition that x_j and x_{j+1} lie on the right arm. The white circles X, Y and Z denote arbitrary subtrees, whose interpretations in terms of the elements in σ are as follows: X is the image under P of the ordered sequence of elements appearing *after* x_j which are less than x_j, while Z is the ordered sequence of elements appearing *after* x_{j+1} which are greater than x_{j+1}, and Y is the ordered sequence of elements appearing after x_j that lie between x_j and x_{j+1}.

Now, consider what happens to T_2, when x_j and x_{j+1} are exchanged. The tree T_2 is depicted below.



Here, S is going to be unchanged, and x_j and x_{j+1} must move as shown. In addition, there will be subtrees X', Y', and Z' as drawn above. However, notice that, if one considers what these subtrees must be with respect to the permutation τ, the fact that x_j and x_{j+1} are adjacent forces the conclusion that the subtrees are unchanged from σ, or in other words that $X = X'$, $Y = Y'$ and $Z = Z'$. So then, T_2 is obtained by a left rotation centered on a vertex on the right arm of T_1. Therefore, T_2 covers T_1 in $\mathscr{C}_{n+1}$, completing the proof. □

To relate the intervals of $\mathscr{C}_{n+1}$ to those of $\mathscr{E}_n$ more deeply, a formal discussion of $\mathscr{E}_n$ is needed. In [2], Edelman defined the following order on the inversion set of a permutation σ.

Definition 6.4. Define $I(\sigma) := \{(j, i) : j > i \text{ and } \sigma^{-1}(j) < \sigma^{-1}(i)\}$. Order $I(\sigma)$, with $(k, \ell) \geq (j, i)$ if and only if $k \geq j$ and $\sigma^{-1}(\ell) \leq \sigma^{-1}(i)$. In a slight abuse of notation, the poset $(I(\sigma), <)$ shall be referred to as $I(\sigma)$ as well.

Theorem 6.5 (Edelman, [2, Theorem 2.13])**.** *$[e, w]_{\mathscr{E}_n} \simeq J(I(w))$, where $[e, w]_{\mathscr{E}_n} = \{v \in \mathfrak{S}_n : v \leq_{\mathscr{E}_n} w\}$, via $v \mapsto I(v)$.*

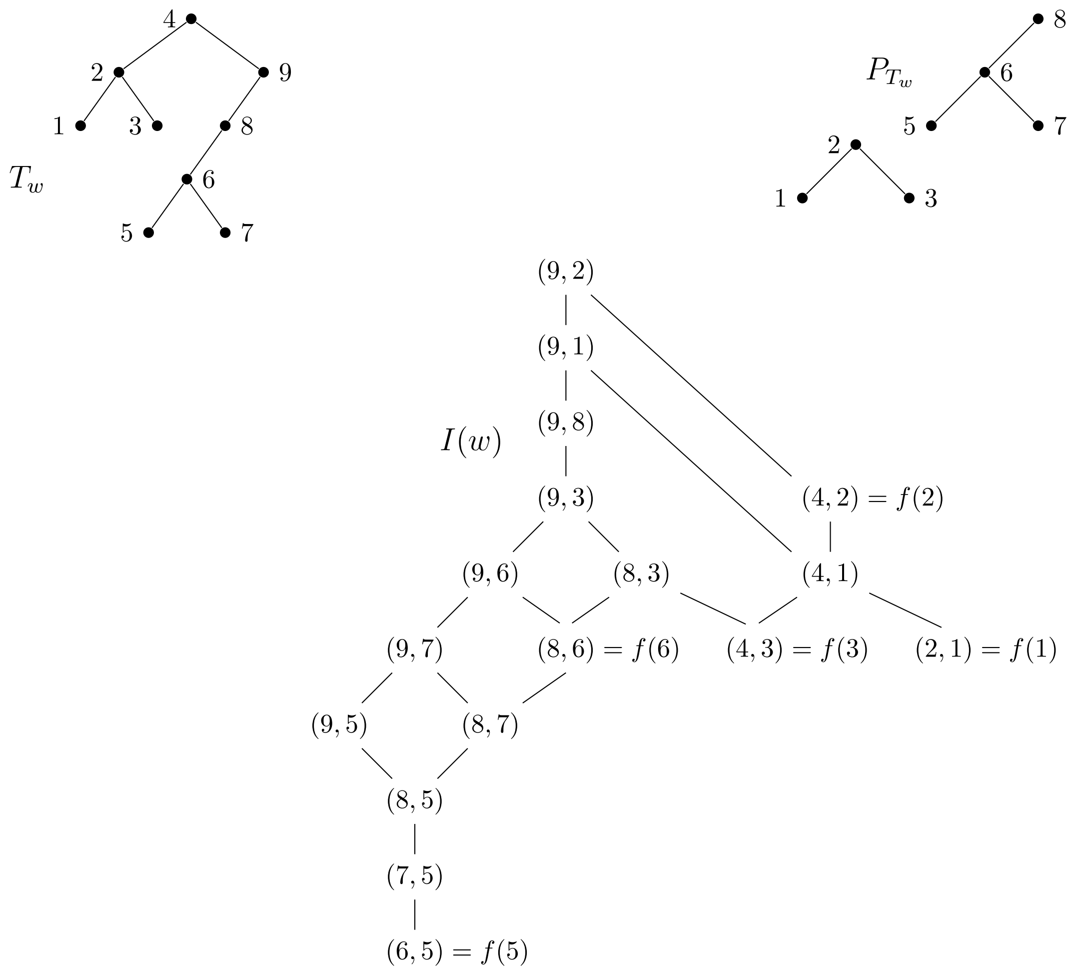


FIGURE 6. T_w, P_{T_w} and $I(w)$ with the image of f indicated.

Definition 6.6. Fix a permutation $w \in \mathfrak{S}_n$. Let T_w be the image of w under the pruned tree map, p. Recall the reduced pruned poset from Definition 2.9. Here it will be useful to label its vertices by the labels they have in T_w, rather than by pairs of parentheses as in the definition of P_T. Define a map $f : P_{T_w} \to I(w)$ as follows: $f(j) = (i, j)$, where i is the smallest label of a vertex of T_w such that j lies in the left subtree of i.

Example. Suppose $w = (4, 9, 2, 1, 8, 3, 6, 7, 5) \in \mathfrak{S}_9$. Figure 6 depicts T_w, P_{T_w} and $I(w)$, with the image of f indicated in $I(w)$.

Proposition 6.7. *The map f is order-preserving.*

Proof. It suffices to show that if $j > k$, and j covers k in P_{T_w}, then $f(j) > f(k)$. Since j covers k, one has that k is a child of j, and there are two cases.

(1) If k is a left child of j, then $f(k) = (j, k)$. By the definition of the pruned tree map (Definition 6.2), one knows $w^{-1}(j) < w^{-1}(k)$. Suppose $f(j) = (i, j)$. Then, by definition, $j < i$, which means that $(j, k) < (i, j)$ in $I(w)$, as desired.
(2) If k is a right child of j, then $f(j) = (i, j)$ means that $f(k) = (i, k)$. Now, $w^{-1}(j) < w^{-1}(k)$, and so $(i, j) > (i, k)$, as desired.

These cover all the cases, proving the result. $\square$

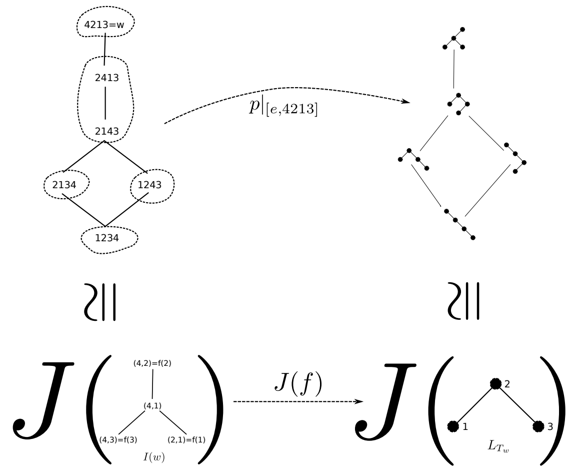


FIGURE 7. $p|_{[e,4213]}$ and $J(f)$ for the interval $[e, 4213]_{\mathscr{E}_n}$.

Definition 6.8. Let P_1, P_2 be two posets and suppose $\phi\colon P_1 \to P_2$ is order-preserving. Then ϕ induces a map $J(\phi)\colon J(P_2) \to J(P_1)$ defined by $J(\phi)(I) = \phi^{-1}(I)$. One calls $J(\phi)$ the *Birkhoff-Priestley dual* to ϕ. In fact, $J(\phi)$ is a lattice morphism.

For further details on Birkhoff-Priestley duality, see [10, Theorem 3.4.1].

Theorem 6.9. *For each $w \in \mathfrak{S}_n$, the map $f\colon P_{T_w} \to I(w)$ defined in Definition 6.6 is Birkhoff-Priestley dual to the pruned tree map $p\colon [e,w]_{\mathscr{E}_n} \to J(P_{T_w})$. In particular, $p\colon \mathscr{E}_n \to \mathscr{C}_n$ becomes a lattice morphism when restricted to any interval in $\mathscr{E}_n$. As a commutative diagram, one has*

$$\begin{tikzcd}
\mathscr{E}_n \arrow[r, "p"] & \mathscr{C}_{n+1} \\
{[e,w]_{\mathscr{E}_n}} \arrow[u, hook] \arrow[r, "p|_{[e,w]_{\mathscr{E}_n}}"] & {[\mathrm{RCT}(n+1), T_w]_{\mathscr{C}_{n+1}}} \arrow[u, hook] \\
J(I(w)) \arrow[u, "\wr"] \arrow[r, "J(f)"] & J(P_{T_w}) \arrow[u, "\wr"]
\end{tikzcd}$$

Example. Figure 7 depicts Theorem 6.9 on the interval $[e, 4213]_{\mathscr{E}_n}$.

Proof. Begin by noting that, strictly speaking, the Birkhoff-Priestley dual to f, $J(f)$, is not a map from $[e,w]_{\mathscr{E}_n} \to J(P_{T_w})$ as p is, but $J(f)\colon J(I(w)) \to J(P_{T_w})$. However, from Theorem 6.5, $J(I(w)) \simeq [e,w]_{\mathscr{E}_n}$, so one can use p in place of such a $J(f)$.

Fix $w \in \mathfrak{S}_n$. From Theorem 6.3 one has that $p\colon [e,w]_{\mathscr{E}_n} \to J(P_{T_w})$ is order-preserving. Then one must show that $p(I)$ is, in fact, $f^{-1}(I)$.

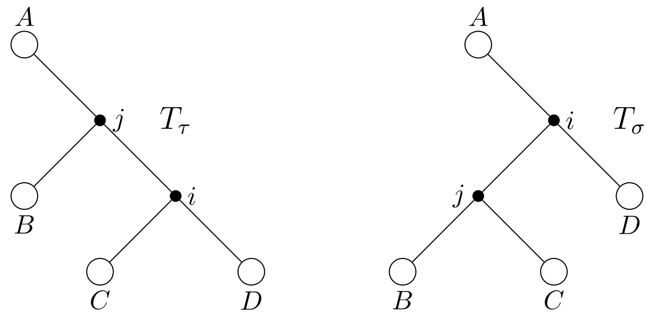
Induct on the number of inversions in a permutation in $[e,w]_{\mathscr{E}_n} \simeq J(I(w))$. Note that the claim is trivially true for $(1, 2, \ldots, n)$, the identity permutation, which corresponds to $\varnothing \in J(I(w))$.

Now consider the permutation $\tau \in [e,w]_{\mathscr{E}_n}$, and $f^{-1}(I(\tau)) = T_\tau$. Suppose σ covers τ. Then, σ has precisely one more inversion than τ; call this inversion (i,j), with $j < i$.

Suppose there is an inversion (ℓ, j) in both τ and σ with $\tau^{-1}(\ell) = \sigma^{-1}(\ell) < \sigma^{-1}(i)$. Then, j is in the left subtree of ℓ in T_τ and T_σ, meaning that j is not the parent of i in T_τ and so adding the inversion (i,j) does not change T_τ, forcing $T_\tau = T_\sigma$, as desired. Thus, one may concentrate on the case where there is no such inversion (ℓ, j), so j is a left-right maximum in τ.

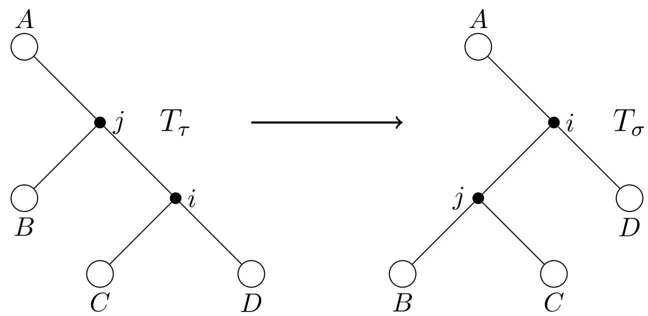
There are two cases:

(1) If (i,j) is not in the image of f, then $f^{-1}(I(\sigma)) = f^{-1}(I(\tau))$, and so one must show that $T_\sigma = T_\tau$. One knows i and j are adjacent in τ, and i is a left-right maximum. In particular, this means that neither τ nor σ has an inversion (k,i), meaning i lies in the right arm of both T_τ and T_σ. Recalling that j is a left-right maximum in τ one has

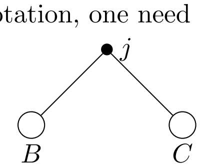


Since (i,j) is not in the image of f, one cannot have that j is the left child of i in T_w. However, j is the left child of i in T_σ and $\sigma < w$, meaning j must also be the left child of i in T_w, a contradiction. Thus (i,j) must be in the image of f.

(2) In the second case, suppose (i,j) is in the image of f. Then, the addition of the inversion of (i,j) to τ results in σ, and in the following rotation from T_τ to T_σ.



One needs to show that $f^{-1}(I(\sigma))$ is the order ideal in P_{T_w} that corresponds to T_σ. Now, since T_τ and T_σ differ only in this rotation, one need only show that



appears in P_{T_w}. Left-right maxima occur only on the right arm of the image of a permutation under the pruned tree map, and so subsequent inversions on the way from σ to w result in rotations in the pruned tree that cannot affect the children of j. Hence, the above subtree appears in P_{T_w}, and so, T_σ is the pruned tree associated to $I(\sigma)$.

These two cases cover all possibilities, concluding the proof. □

7. The ParseWords Function for the Comb Poset

The number of common parsewords for any two trees having a common upper bound in $\mathscr{C}_n$ can be computed precisely. Recall that $w \in \mathrm{ParseWords}(T)$ means that T admits a labeling of its vertices by $0, 1, 2$ such that the leaves are labeled by the word w, the children of each vertex have distinct labels and no vertex has the same label as either of its children. Recent work by Cooper, Rowland and Zeilberger in [1] led us to first consider the comb poset. They showed that a statement equivalent to the Four Color Theorem due to Kaufmann in [5] is, in turn, equivalent to $\mathrm{ParseWords}(T_1, T_2) \neq \emptyset$ for all $T_1, T_2 \in \mathbb{T}_n$ for any $n \in \mathbb{N}$.

Example. An example of a tree parsing the word 2202 is shown in Figure 8.

Example. An example of two trees parsing the same word 010 is shown in Figure 9.

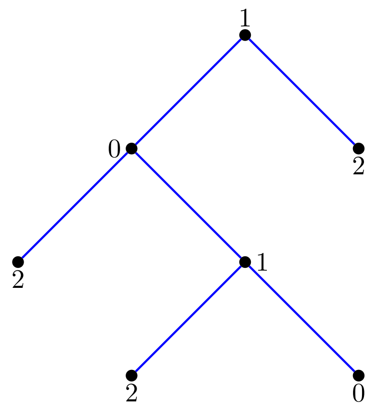


FIGURE 8.

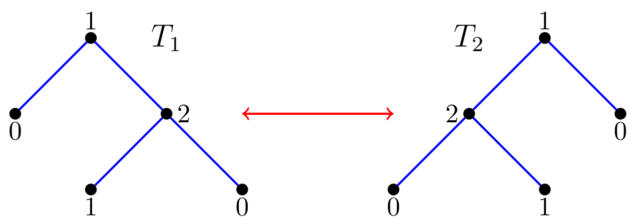


FIGURE 9.

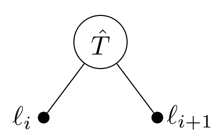
Example. The common parsewords for the trees in Figure 9 are 101, 202, 010, 212, 020, 121.

To simplify notation, let $T_{\leq b}$ be the subtree of a tree T having the vertex b as its root.

Proposition 7.1 (Common root property, [1, Proposition 2]). *If two trees $T_1, T_2 \in \mathbb{T}_n$ parse the same word, then their roots receive the same label when the trees are labeled with a common parseword. Hence, if for $T_1, T_2 \in \mathbb{T}_n$, there are vertices b_i in T_1 and b_j in T_2 such that ${T_1}_{\leq b_i}$ and ${T_2}_{\leq b_j}$ have precisely the same leaves (i.e. both the dangling subtrees contain precisely the leaves m_1 through m_2, for some natural numbers $m_1 < m_2 \leq n$), then b_i and b_j receive the same label if one labels the trees with a common parse word.*

If a tree T parses a word w and X is a subtree of T, let $w(X)$ be the label received by the root of X parsing w and let w_X be the segment of w parsed by the subtree X.

Definition 7.2. Say $T \in \mathbb{T}_n$ has a *leaf reduction* at $(i, i+1)$ for $i \in \{1, \ldots, n-1\}$ if the leaves $i, i+1$ have a common parent:



Define $\hat{T}$ as in the above diagram, i.e. remove ℓ_i and ℓ_{i+1} from T. For $i \in \{1, \ldots, n-1\}$, define two maps $f_i^<, f_i^> \colon \mathrm{ParseWords}(\hat{T}) \to \mathrm{ParseWords}(T)$ sending $\hat{w}$ to the w in $\mathrm{ParseWords}(T)$ that uniquely extends $\hat{w}$ in such a way that $w_i < w_{i+1}$ or $w_i > w_{i+1}$, respectively.

Proposition 7.3. *If $\{T_1 \ldots, T_m\} \subset \mathbb{T}_n$ share a leaf reduction at $(i, i+1)$, then*

$$\mathrm{ParseWords}(T_1, \ldots, T_m) = f_i^<(\mathrm{ParseWords}(\hat{T}_1, \ldots, \hat{T}_m)) \sqcup f_i^>(\mathrm{ParseWords}(\hat{T}_1, \ldots, \hat{T}_m)).$$

Proposition 7.3 is most frequently used several times in succession, to "collapse" a subtree common to two or move trees. In particular, it often allows a reduction to the case $T_1 \wedge T_2 \wedge \cdots \wedge T_m = \mathrm{RCT}(n)$.

Corollary 7.4. *For $T \in \mathbb{T}_n$, one has $|\mathrm{ParseWords}(T)| = 3 \cdot 2^{n-1}$.*

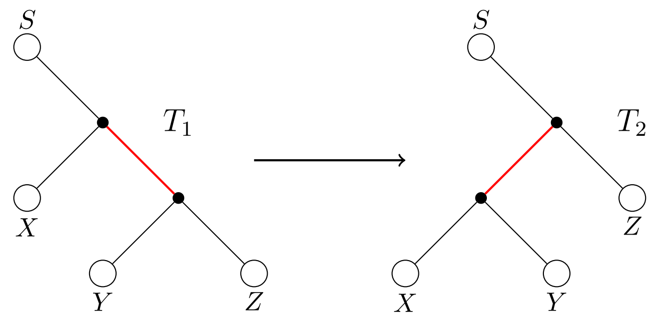
Proof. The case $n = 1$ is trivial. One can then induct on n, applying Proposition 7.3 to T alone. $\square$

Proposition 7.5. *For any $T_1, T_2 \in \mathbb{T}_n$ differing by a single rotation (not necessarily a right arm rotation),*

$$\begin{aligned}\text{ParseWords}(T_1, T_2) &= \{w \in \text{ParseWords}(T_1) \colon w(X) \neq w(Y)\}\\ &= \{w \in \text{ParseWords}(T_2) \colon w(Y) \neq w(Z)\}\\ &= \{w \in \text{ParseWords}(T_1) \colon w(X) = w(Z)\},\end{aligned}$$

where X, Y, Z are subtrees as indicated below. Furthermore, $|\text{ParseWords}(T_1, T_2)| = 3 \cdot 2^{n-2}$.

Proof. The conditions on w in the first part of the claim can be checked by inspection. For the second part, the case $n = 3$ can be checked directly. Taking this as a base case, one can induct on n. A rotation looks like



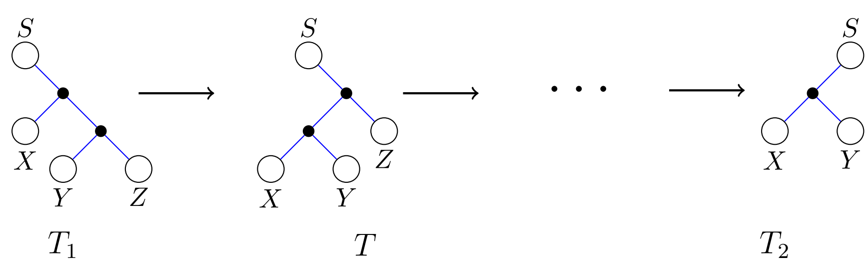
Applying Proposition 7.3 to any of the subtrees X, Y, Z not consisting of a single leaf, or to the subtrees taken together as a single subtree if all three are leaves, allows one to invoke the result for a smaller n and obtain the desired result. $\square$

Theorem 7.6. *Suppose $T_1 < T < T_2$ in $\mathscr{C}_n$. Then* $\text{ParseWords}(T_1, T_2) = \text{ParseWords}(T_1, T_2, T)$.

Proof. It suffices to prove the statement for $T_1 \lessdot T < T_2$ in $\mathscr{C}_n$. Assume the theorem holds in this case and obtain the general case by induction on the length of a chain between T_1 and T.

Suppose one has $T_1 \lessdot T_1' \lessdot T_2' \lessdot \cdots \lessdot T_k' \lessdot T < T_2$. Then, by induction, $\text{ParseWords}(T_1, T_k', T_2) = \text{ParseWords}(T_1, T_2)$. Suppose $w \in \text{ParseWords}(T_1, T_2) = \text{ParseWords}(T_1, T_k', T_2)$. Furthermore, $w \in \text{ParseWords}(T_k', T_2) = \text{ParseWords}(T_k', T, T_2)$, so $w \in \text{ParseWords}(T_1, T, T_2)$ as desired. By definition, $\text{ParseWords}(T_1, T, T_2) \subset \text{ParseWords}(T_1, T_2)$, so $\text{ParseWords}(T_1, T, T_2) = \text{ParseWords}(T_1, T_2)$, as desired.

To prove the initial case, now suppose $T_1 \lessdot T < T_2$. One has a sequence of right-arm rotations



Since the rotation between T_1 and T moves the subtrees labeled by X and Y off the right arm, they must remain in the same position relative to one another in T_2.

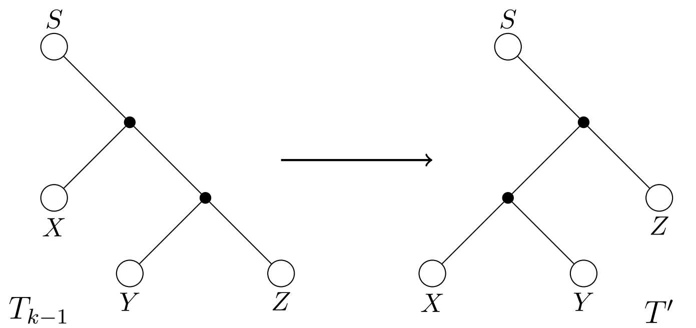
Suppose $w \in \text{ParseWords}(T_1, T_2)$. Since T_2 parses w, one must have $w(X) \neq w(Y)$ and, hence, by Proposition 7.5, T parses w. Thus $\text{ParseWords}(T_1, T_2) \subset \text{ParseWords}(T_1, T, T_2)$. By definition, $\text{ParseWords}(T_1, T, T_2) \subset \text{ParseWords}(T_1, T_2)$, so $\text{ParseWords}(T_1, T_2) = \text{ParseWords}(T_1, T, T_2)$, as desired. $\square$

Theorem 7.7. *Suppose $T < T'$ in $\mathscr{C}_n$ and* $\text{rank}(T') - \text{rank}(T) = k$. *Then* $|\text{ParseWords}(T, T')| = 3 \cdot 2^{n-1-k}$.

Proof. One proceeds by induction. Proposition 7.5 addresses the case $k = 1$. Via repeated leaf reductions, one may assume $T = \mathrm{RCT}(n)$. Now suppose the statement holds for $k-1$, that $T < T'$ and $\mathrm{rank}(T') - \mathrm{rank}(T) = k$. One has a chain in $\mathscr{C}_n$, $T \lessdot T_1 \lessdot T_2 \lessdot \cdots \lessdot T_{k-1} \lessdot T'$. By induction $|\mathrm{ParseWords}(T, T_{k-1})| = 3 \cdot 2^{n-k}$. One constructs a bijection

$$\mathrm{ParseWords}(T, T_{k-1}, T') \to \mathrm{ParseWords}(T, T_{k-1}) \smallsetminus \mathrm{ParseWords}(T, T_{k-1}, T').$$

First, one characterizes those parsewords in $\mathrm{ParseWords}(T, T_{k-1}, T')$. One knows that T_{k-1} and T' differ by a right arm rotation.



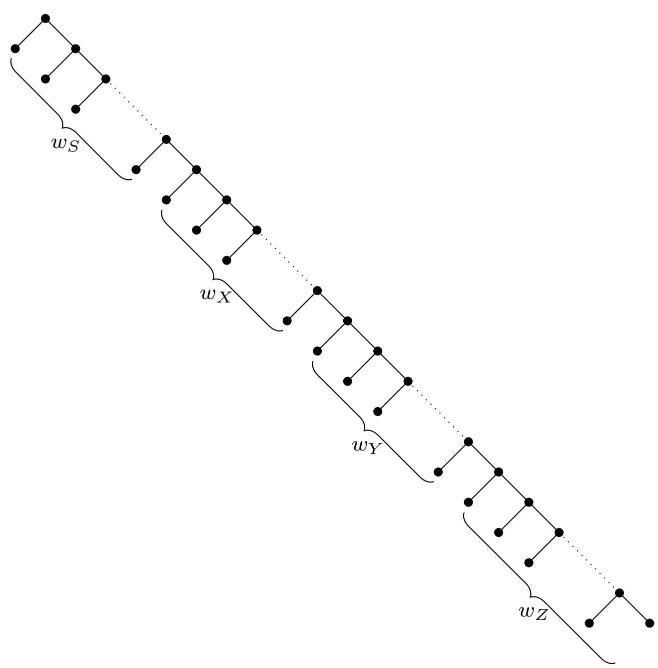
From Proposition 7.5, one has that T' also parses $w \in \mathrm{ParseWords}(T, T_{k-1})$ (i.e. $w \in \mathrm{ParseWords}(T, T_{k-1}, T')$) if and only if $w(X) \neq w(Y)$.

Now define the map

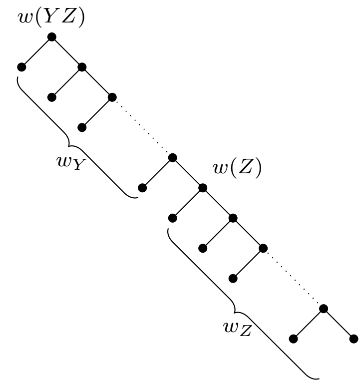
$$\phi\colon \mathrm{ParseWords}(T, T_{k-1}) \to \mathrm{ParseWords}(T, T_{k-1})$$

as follows. Suppose $w \in \mathrm{ParseWords}(T, T_{k-1})$. Then $w = w_S w_X w_Y w_Z$, where w_J is the word parsed by the leaves of the subtree J. Define a transposition in $\mathfrak{S}_{\{0,1,2\}}$ by $\sigma = (w(Y), w(Z))$. Then define $\phi(w) = w_S w_X \sigma(w_Y)\sigma(w_Z)$. One needs to show that $\phi(w) \in \mathrm{ParseWords}(T, T_{k-1})$.

On the one hand, σ permutes the alphabet within the smallest subtree of T_{k-1} containing both Y and Z, while leaving the label of its root unchanged, so $\phi(w)$ is certainly parsed by T_{k-1}. Recall that T was assumed to be $\mathrm{RCT}(n)$. Labeling T with w gives



Proposition 7.1 means that when the subtree of T containing the leaves of T_{k-1}'s Y and Z subtrees is fully labeled, the root of this subtree receives the same label as the root of the smallest subtree of T_{k-1} containing both Y and Z, call it $w(YZ)$ and another right arm vertex receives the label $w(Z)$:



Since $w(YZ)$ is equal to neither $w(Y)$ nor $w(Z)$, it is fixed by σ. Consequently, applying σ to the subtree of T consisting of those vertices in Y and Z has the same effect on parsewords as applying σ to Y and Z in T_{k-1}. In other words, $\phi(w)$ is parsed by T, so the map is well-defined.

Then ϕ is transparently a bijection

$$\mathrm{ParseWords}(T, T_{k-1}) \to \mathrm{ParseWords}(T, T_{k-1})$$

and, moreover, exchanges $\mathrm{ParseWords}(T, T_{k-1}, T')$ and $\mathrm{ParseWords}(T, T_{k-1}) \setminus \mathrm{ParseWords}(T, T_{k-1}, T')$: if $w \in \mathrm{ParseWords}(T, T_{k-1}, T')$, then $\phi(w)$ has $\phi(w)(X) = \phi(w)(Y)$, meaning it cannot parse T'. Thus, ϕ is a bijection

$$\mathrm{ParseWords}(T, T_{k-1}, T') \to \mathrm{ParseWords}(T, T_{k-1}, T') \setminus \mathrm{ParseWords}(T, T_{k-1}, T').$$

Consequently, $\mathrm{ParseWords}(T, T_{k-1}, T')$ contains precisely half the parsewords of $\mathrm{ParseWords}(T, T_{k-1})$, i.e. there are $3 \cdot 2^{n-1-k}$ of them. From Theorem 7.6, one has $\mathrm{ParseWords}(T, T_{k-1}, T') = \mathrm{ParseWords}(T, T')$, so $|\mathrm{ParseWords}(T, T')| = 3 \cdot 2^{n-1-k}$, as desired. $\square$

Corollary 7.8. *Since $k \leq n-2$ by Proposition 3.2, any pair of trees comparable in $\mathscr{C}_n$ has a common parse word.*

Theorem 7.9. *Suppose T_1 and T_2 have an upper bound in $\mathscr{C}_n$. Then,*

$$\mathrm{ParseWords}(T_1, T_2) = \mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2).$$

Proof. The statement is clear when T_1 and T_2 are comparable, so assume T_1 and T_2 are not comparable. By Theorem 7.6,

$$\mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2, T_1) = \mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_1) = \mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2, T_2).$$

One then immediately has that $\mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2) \subset \mathrm{ParseWords}(T_1, T_2)$.

All that remains is to show inclusion the other way. Suppose the theorem holds for trees with $k < n$ leaves. Without loss of generality, one may assume that $T_1 \wedge T_2 = \mathrm{RCT}(n)$ by making repeated leaf reductions in T_1 and T_2. There are two cases:

(1) Suppose T_1 and T_2 share a leaf reduction at, say, i. Then $T_1 \wedge T_2$ and $T_1 \vee T_2$ must also share this leaf reduction. Then,

$$\begin{aligned}\mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2) &= f_i^{<}(\mathrm{ParseWords}(\widehat{T_1 \wedge T_2}, \widehat{T_1 \vee T_2})) \\ &\quad \sqcup f_i^{>}(\mathrm{ParseWords}(\widehat{T_1 \wedge T_2}, \widehat{T_1 \vee T_2})) \\ &= f_i^{<}(\mathrm{ParseWords}(\hat{T}_1, \hat{T}_2)) \sqcup f_i^{>}(\mathrm{ParseWords}(\hat{T}_1, \hat{T}_2)) \\ &= \mathrm{ParseWords}(T_1, T_2),\end{aligned}$$

as desired, with the inductive hypothesis being used in the second equality.

(2) Suppose, on the other hand, that no such common leaf reduction exists. Then one must have that $T_1 \wedge T_2 = \mathrm{RCT}(n)$. Since $T_1 \vee T_2$ exists, one must then have that all parenthesis pairs in RP_{T_1} and RP_{T_2} are disjoint. Suppose $w = w_1 w_2 \cdots w_n \in \mathrm{ParseWords}(T_1, T_2)$. Then, since $T_1 \wedge T_2$ is $\mathrm{RCT}(n)$, either RP_{T_1} or RP_{T_2} contains a parenthesis pair enclosing a_{n-1}, else both trees would have a leaf reduction at $(n-1, n)$. Without loss of generality, assume RP_{T_1} contains a parenthesis pair enclosing a_{n-1}. Moreover, RP_{T_1} has a maximal parenthesis pair enclosing the leaves $a_j, \ldots, a_{n-1}$. Then, since all parenthesis pairs in RP_{T_1} and RP_{T_2} are disjoint, none of $a_j, \ldots, a_{n-1}$ are enclosed by a parenthesis pair in RP_{T_2}. Consequently, the subtrees of T_2 and $T_1 \wedge T_2$ with leaf set $a_j, \ldots, a_n$ are both isomorphic to $\mathrm{RCT}(n-j+1)$. Call this subtree X_1. By the maximality of the parenthesis pair containing $a_j, \ldots, a_{n-1}$, one has that T_1 and $T_1 \vee T_2$ have isomorphic subtrees whose leaf sets are $a_j, \ldots, a_n$, call this subtree X_2. Consequently, $w_j \cdots w_n$ is parsed by the subtree containing $a_j, \ldots, a_n$ in $T_1, T_2, T_1 \wedge T_2$ and $T_1 \vee T_2$. Then, from Proposition 7.1, one has that $w(X_1) = w(X_2)$. Collapse the subtrees X_1 and X_2 to obtain $T_1', T_2', T_1' \wedge T_2'$ and $T_1' \vee T_2'$. By induction, $\mathrm{ParseWords}(T_1', T_2') = \mathrm{ParseWords}(T_1' \wedge T_2', T_1' \vee T_2')$, meaning $w_1 \cdots w_{j-1} w(X_1) = w_1 \cdots w_{j-1} w(X_2)$ is parsed by $T_1' \wedge T_2'$ and $T_1' \vee T_2'$. It is then easy to see that this implies w lies in $\mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2)$, as desired.

□

Remark. If T_1 and T_2 have an upper bound in $\mathscr{C}_n$, and $\mathrm{rank}(T_1 \vee T_2) - \mathrm{rank}(T_1 \wedge T_2) = k$, combining Theorem 7.9 and Theorem 7.7, one has

$$|\mathrm{ParseWords}(T_1, T_2)| = |\mathrm{ParseWords}(T_1 \wedge T_2, T_1 \vee T_2)| = 3 \cdot 2^{n-1-k}.$$

Acknowledgements

This research was carried out at the School of Mathematics, University of Minnesota - Twin Cities, under the supervision of Vic Reiner and Dennis Stanton with funding from NSF grant DMS-1001933. We would like to thank Vic Reiner and Dennis Stanton for their invaluable guidance and support, Bobbe Cooper for giving a talk introducing us to this problem, Nathan Williams for his words of advice and continual assessment of our progress, as well as Alan Guo and an anonymous referee for their input in the proof of Theorem 3.2. We would also like to thank the referee for their very thoughtful and detailed comments.

References

[1] B. Cooper, E. Rowland, and D. Zeilberger. Toward a language theoretic proof of the Four Color Theorem, 2010, `arXiv:1006.1324v1 [math.CO]`.

[2] P. H. Edelman. Tableaux and chains in a new partial order of S_n. *J. Combin. Theory Ser. A*, 51(2):181–204, 1989.

[3] H. Friedman and D. Tamari. Problèmes d'associativité: Une structure de treillis finis induite par une loi demi-associative. *J. Combinatorial Theory*, 2:215–242, 1967.

[4] S. Huang and D. Tamari. Problems of associativity: A simple proof for the lattice property of systems ordered by a semi-associative law. *J. Combinatorial Theory Ser. A*, 13:7–13, 1972.

[5] L. H. Kauffman. Map coloring and the vector cross product. *J. Combin. Theory Ser. B*, 48(2):145–154, 1990.

[6] D. E. Knuth. *The Art of Computer Programming*, volume 3: Sorting and Searching. Addison-Wesley Publishing Company, Reading, Massachusetts, 1973.

[7] J. M. Pallo. Enumerating, ranking and unranking binary trees. *Comput. J.*, 29(2):171–175, 1986.

[8] J. M. Pallo. Right-arm rotation distance between binary trees. *Inform. Process. Lett.*, 87(4):173–177, 2003.

[9] D. D. Sleator, R. E. Tarjan, and W. P. Thurston. Rotation distance, triangulations, and hyperbolic geometry. *J. Amer. Math. Soc.*, 1(3):647–681, 1988.

[10] R. P. Stanley. *Enumerative combinatorics. Vol. 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[11] R. P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999.

[12] D. Tamari. Associativity theory and the theory of lists. Their applications from abstract algebra to the four-colour-map problem. In *Proceedings of the seventeenth Southeastern international conference on combinatorics, graph theory, and computing (Boca Raton, Fla., 1986)*, volume 54, pages 39–53, 1986.

[13] H. Whitney. A theorem on graphs. *Ann. of Math. (2)*, 32(2):378–390, 1931.

School of Mathematics, University Of Minnesota, Minneapolis, MN 55455, USA
E-mail address: csar@math.umn.edu

Department of Mathematics, Massachusetts Institute of Technology, MA 02139, USA
E-mail address: rsengupt@mit.edu

Department of Mathematics, Massachusetts Institute of Technology, MA 02139, USA
E-mail address: warutsuk@mit.edu