

Decomposition of Triply Rooted Trees

William Y. C. Chen¹, Janet F. F. Peng² and Harold R. L. Yang³

^{1,2,3}Center for Combinatorics, LPMC-TJKLC
Nankai University, Tianjin 300071, P.R. China

¹Center for Applied Mathematics
Tianjin University, Tianjin 300072, P.R. China

¹chen@nankai.edu.cn, ²janet@mail.nankai.edu.cn
³yangruilong@mail.nankai.edu.cn

Submitted: Dec 28, 2012; Accepted: Apr 9, 2013; Published: Apr 17, 2013

Mathematics Subject Classifications: 05A15, 05A19

Abstract

We give a decomposition of triply rooted trees into three doubly rooted trees. This leads to a combinatorial interpretation of an identity conjectured by Lacasse in the study of the PAC-Bayesian machine learning theory, and proved by Younsi by using the Hurwitz identity on multivariate Abel polynomials. Let $[n] = \{1, 2, \dots, n\}$. We also give a bijection between the set of functions from $[n+1]$ to $[n]$ and the set of triply rooted trees on $[n]$, which leads to a symmetry property and a refined enumeration of functions from $[n+1]$ to $[n]$ with respect to the number of elements in the orbit of $n+1$ and the number of periodic points.

Keywords: doubly rooted tree, triply rooted tree, bijection

1 Introduction

Lacasse [5] introduced the functions $\xi(n)$ and $\xi_2(n)$ in his study of the classical PAC-Bayes theorem in the theory of machine learning, where

$$\xi(n) = \sum_{k=0}^n \binom{n}{k} \left(\frac{k}{n}\right)^k \left(1 - \frac{k}{n}\right)^{n-k}$$

and

$$\xi_2(n) = \sum_{j=0}^n \sum_{k=0}^{n-j} \binom{n}{j} \binom{n-j}{k} \left(\frac{j}{n}\right)^j \left(\frac{k}{n}\right)^k \left(1 - \frac{j}{n} - \frac{k}{n}\right)^{n-j-k}.$$

He showed that $\xi(n)$ can be used to give a tighter bound of the Kullback-Leibler divergence between the risk and the empirical risk on a sample space S of a hypothesis function in a hypothesis space, whereas $\xi_2(n)$ can be used to bound the Kullback-Leibler divergence between the risk and the empirical risk on S of the joint distribution of two hypothesis functions in a hypothesis space.

While $\xi_2(n)$ is a double sum, based on numerical evidence Lacasse [5] posed the following conjecture stating that $\xi_2(n)$ can be reduced to the single sum $\xi(n)$.

Conjecture 1.1 *For $n \in \mathbb{N}$, we have*

$$\xi_2(n) = \xi(n) + n. \tag{1.1}$$

By applying a multivariate Abel identity due to Hurwitz, Younsi [9] gave an algebraic proof of this conjecture. Let

$$A_n(x_1, x_2, \dots, x_m; p_1, p_2, \dots, p_m) = \sum_{k_1+k_2+\dots+k_m=n} \binom{n}{k_1, k_2, \dots, k_m} \prod_{j=1}^m (x_j + k_j)^{k_j+p_j},$$

where p_1, p_2, \dots, p_m are integers. When $p_1 = p_2 = \dots = p_m = 0$, Hurwitz proved that

$$A_n(x_1, \dots, x_m; 0, \dots, 0) = \sum_{k=0}^n \binom{n}{k} (x_1 + x_2 + \dots + x_m + n)^{n-k} \alpha_k(m-1), \tag{1.2}$$

where $\alpha_k(r) = r(r+1)\dots(r+k-1)$ is the rising factorial, see, for example, Riordan [7].

Younsi [9] observed that $\xi(n) = A_n(0, 0; 0, 0)$ and $\xi_2(n) = A_n(0, 0, 0; 0, 0, 0)$, and derived the following expressions for $\xi(n)$ and $\xi_2(n)$ by using (1.2)

$$\xi(n) = \frac{1}{n^n} \sum_{j=0}^n n^j \frac{n!}{j!}, \tag{1.3}$$

$$\xi_2(n) = \frac{1}{n^n} \sum_{j=0}^n n^{n-j} \binom{n}{j} (j+1)!. \tag{1.4}$$

Conjecture 1.1 can be easily deduced from (1.3) and (1.4). Other proofs have been found by Prodinger [6] and Sun [8].

In this paper, we give a combinatorial explanation of relation (1.1). Rewriting (1.1) as

$$\sum_{j=0}^n \sum_{k=0}^{n-j} \binom{n}{j} \binom{n-j}{k} j^j k^k (n-j-k)^{n-j-k} = \sum_{k=0}^n \binom{n}{k} k^k (n-k)^{n-k} + n^{n+1}, \tag{1.5}$$

we see that it is equivalent to the following form

$$\sum_{j=1}^n \sum_{k=0}^{n-j} \binom{n}{j} \binom{n-j}{k} j^j k^k (n-j-k)^{n-j-k} = n^{n+1}. \quad (1.6)$$

The right hand side of (1.6) indicates that we need the notion of triply rooted trees, namely, labeled trees with an ordered list of three distinguished, but not necessarily distinct vertices. To be more specific, the three distinguished vertices of a triply rooted tree are called the first, the second and the third root, respectively. It can be easily seen that the summand on the left hand side of (1.6) can be interpreted as the number of triples of doubly rooted trees with a given number of vertices in each doubly rooted tree. Hence relation (1.6) can be deduced from a decomposition of a triply rooted tree into three doubly rooted trees.

The second result of this paper is a correspondence between the set of functions from $[n+1]$ to $[n]$ and the set of triply rooted trees on $[n]$, where n is a positive integer and $[n] = \{1, 2, \dots, n\}$. Let f be a function from $[n+1]$ to $[n]$ and let T be the corresponding triply rooted tree. We find that the orbit of $n+1$ on f is mapped to the set of ancestors of the second root in T , and the set of periodic points of f is mapped to the set of ancestors of the third root in T . Based on this property of our bijection, we deduce a symmetry property of functions from $[n+1]$ to $[n]$ with respect to the number of periodic points and the size of the orbit of $n+1$. Moreover, we obtain a formula for the number of triply rooted trees for which the second root has i ancestors and the third root has j ancestors. This formula also gives the number of functions from $[n+1]$ to $[n]$ such that the size of the orbit of $n+1$ is $i+1$ and the number of periodic points is j .

2 Decomposition of triply rooted tree

In this section, we give a combinatorial interpretation of Lacasse's identity by providing a decomposition of a triply rooted tree into three doubly rooted trees.

Recall that a rooted tree is defined to be a labeled tree with a specific vertex, which is called the root. Let R_n denote the set of rooted trees on $[n]$. It is well-known that the set R_n is counted by n^{n-1} , as shown by Cayley [2]. A doubly rooted tree is defined as a labeled tree with an ordered list of two distinguished vertices r_1 and r_2 , where we call r_1 the first root and r_2 the second root. Similarly, a triply rooted tree is a labeled tree with an ordered list of three distinguished vertices. Notice that the roots are not necessarily distinct. We denote by D_n (T_n) the set of doubly (triply) rooted trees on $[n]$. One sees that $|D_n| = n^n$ and $|T_n| = n^{n+1}$. A bijection from doubly rooted trees to functions from $[n]$ to $[n]$ was independently obtained by Joyal [4] and Goulden and Jackson [3]. Hence the right hand side of (1.6) can be interpreted as the number of triply rooted trees on $[n]$.

On the other hand, let Q_n denote the set of triples of doubly rooted trees (D, D', D'') such that the vertex sets of D, D', D'' form a weak partition of $[n]$ with D being nonempty. To be more specific, a triple (X, Y, Z) of subsets of a set S is said to be a weak partition of S if X, Y , and Z are disjoint and their union equals S . It is obvious that Q_n is counted by

$$\sum_{j=1}^n \sum_{k=0}^{n-j} \binom{n}{j} \binom{n-j}{k} j^j k^k (n-j-k)^{n-j-k},$$

which is the left hand side of (1.6). Hence identity (1.6) follows from the following bijection.

Theorem 2.1 *For $n \geq 1$, there is a bijection between Q_n and T_n .*

To present the proof of the above theorem, we recall some terminology. Given two vertices i and j of a rooted tree T , we say that j is a *descendant* of i , or i is an *ancestor* of j , if i lies on the unique path from the root to j . In particular, each vertex is a descendant as well as an ancestor of itself. A *child* of i means a descendant j of i such that (i, j) is an edge of T . The depth of i is defined to be the number of edges of the unique path from the root to i . Given two vertices v_1 and v_2 of T , there is a unique vertex v that is the common ancestor of v_1 and v_2 with the largest depth. This vertex is called the *least common ancestor* of v_1 and v_2 , see Aho, Hopcroft and Ullman [1]. For example, for the tree in Figure 2.1, the least common ancestor of 1 and 3 is 5, while the least common ancestor of 1 and 6 is the root 4.

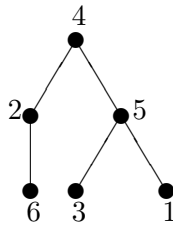


Figure 2.1: A rooted tree on $[6]$.

Throughout this paper, we use $r_1(D)$ and $r_2(D)$ to denote the first root and the second root of a doubly rooted tree D , respectively, and we use $r_1(T)$, $r_2(T)$ and $r_3(T)$ to denote the first root, the second root and the third root of a triply rooted tree T , respectively.

Proof of Theorem 2.1. We define a map φ from Q_n to T_n . Given a triple (D, D', D'') of doubly rooted trees in Q_n , we first consider the case when neither D' nor D'' is empty.

We merge D and D' by setting $r_1(D')$ to be a child of $r_2(D)$, and we merge D and D'' by setting $r_1(D'')$ to be a child of $r_2(D)$. By setting $r_2(D')$ and $r_2(D'')$ to be the second root and the third root of the resulting tree, we obtain a triply rooted tree T .

For example, Figure 2.2 gives an illustration of a triple of doubly rooted trees and the corresponding triply rooted tree, where the second root is represented by a solid square, and the third root is represented by a hollow square.

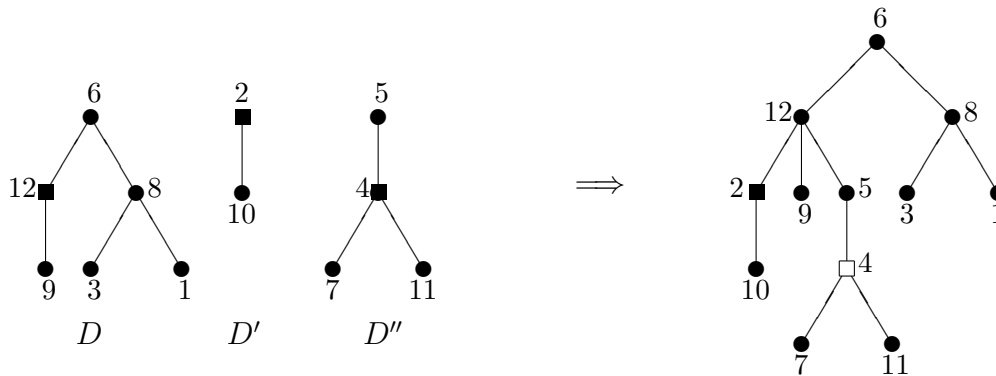


Figure 2.2: The merging process when $D' \neq \emptyset$ and $D'' \neq \emptyset$.

We now consider the case when either D' or D'' is empty. When $D' = \emptyset$, there is no need to merge D and D' . Instead, we set $r_2(D)$ to be the second root of the resulting tree in the merging process. Similarly, when $D'' = \emptyset$, we just set $r_3(D)$ to be the third root of the resulting tree in the merging process.

In summary, $(r_1(T), r_2(T), r_3(T))$ is given as follows:

$$\begin{cases} (r_1(D), r_2(D'), r_2(D'')), & \text{if } D' \neq \emptyset, \text{ and } D'' \neq \emptyset; \\ (r_1(D), r_2(D), r_2(D'')), & \text{if } D' = \emptyset, \text{ and } D'' \neq \emptyset; \\ (r_1(D), r_2(D'), r_2(D)), & \text{if } D' \neq \emptyset, \text{ and } D'' = \emptyset; \\ (r_1(D), r_2(D), r_2(D)), & \text{if } D' = \emptyset, \text{ and } D'' = \emptyset. \end{cases}$$

Figure 2.3 gives an illustration of the merging process when $D' = \emptyset$ and $D'' \neq \emptyset$.

To show that the above process is invertible, we give a description of the inverse procedure. Given a triply rooted tree T with three roots r_1, r_2 and r_3 , assume that w is the least common ancestor of r_2 and r_3 .

We first consider the case when $w \neq r_2$ and $w \neq r_3$. Find the child x of w such that r_2 is a descendant of x , and the child y of w such that r_3 is a descendant of y . Removing the edges (w, x) and (w, y) , we get three trees with three roots r_1, x and y . Let D be the doubly rooted tree with two roots r_1 and w , let D' be the doubly rooted tree with two roots x and r_2 , and let D'' be the doubly rooted tree with two roots y and r_3 .

When $w = r_2$, then in the previous process, x does not exist, and instead we set $D' = \emptyset$. Similarly, when $w = r_3$, we set $D'' = \emptyset$.

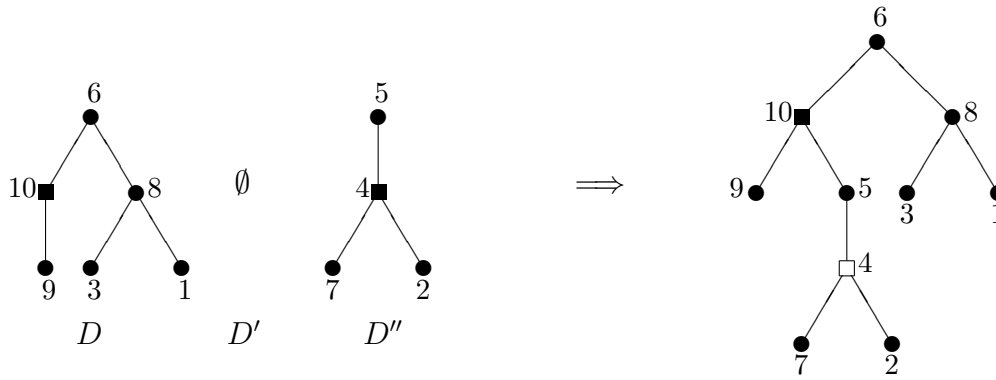


Figure 2.3: The merging process when $D' = \emptyset$ and $D'' \neq \emptyset$.

Now we see that in any case the three doubly rooted trees D , D' and D'' can be merged into the triply rooted tree T . That is, the above merging process is invertible. This completes the proof. \blacksquare

3 Functions from $[n + 1]$ to $[n]$

In this section, we establish a correspondence between functions from $[n + 1]$ to $[n]$ and triply rooted trees on $[n]$, which maps the orbit of $n + 1$ to the set of ancestors of the second root, and maps the set of periodic points to the set of ancestors of the third root. By the symmetry between the second and third roots, we deduce a symmetry property of the number of functions from $[n + 1]$ to $[n]$ with respect to the number of periodic points and the size of the orbit of $n + 1$.

Given a function f from $[n + 1]$ to $[n]$, the *orbit* of x on f is defined to be the set $\{x, f(x), f^2(x), \dots\}$. If there exists some $j \geq 1$, such that $f^j(x) = x$, then x is called a *periodic point* of f . We have the following correspondence.

Theorem 3.1 *There is a bijection ϕ between the set of functions f from $[n + 1]$ to $[n]$ and the set of triply rooted trees on $[n]$ such that the orbit of $n + 1$ on f excluding $n + 1$ itself is mapped to the set of ancestors of the second root of $\phi(f)$ and the set of periodic points of f is mapped to the set of ancestors of the third root of $\phi(f)$.*

Proof. The map ϕ can be described as follows. Let f be a function from $[n + 1]$ to $[n]$. We proceed to construct a triply rooted tree T on $[n]$ based on f . We begin with the functional digraph G_f of f , that is, a digraph on $[n + 1]$ with arcs $(i, f(i))$ for $1 \leq i \leq n + 1$. Let C_1 be the connected component of G_f containing the vertex $n + 1$. Consider the longest path P starting from $n + 1$, say,

$$P: n + 1 = u_0 \rightarrow u_1 \rightarrow u_2 \rightarrow \dots \rightarrow u_k.$$

In other words, k is the smallest integer such that $f(u_k)$ equals u_j for some $j \leq k$. Assume that $f(u_k) = u_j$. Removing the arc (u_k, u_j) and the vertex $n + 1$ from C_1 , we get a tree H rooted at u_k .

Let C_2 be the digraph $G_f \setminus C_1$. When $C_2 = \emptyset$, we set u_k, u_1 and u_j to be the three roots of H to obtain a triply rooted tree T .

When $C_2 \neq \emptyset$, we see that C_2 is a functional digraph. By applying the bijection between functions and doubly rooted trees, obtained by Joyal [4] and Goulden and Jackson [3], C_2 corresponds to a doubly rooted tree D . Let w_1 and w_2 be the two roots of D . We merge the rooted tree H and the doubly rooted tree D by setting the first root w_1 of D as a child of the vertex u_j of H . Setting u_k, u_1 and w_2 to be the first, the second and the third root, respectively, we get a triply rooted tree T , and we set $\phi(f) = T$.

For example, let f be the following function from $[13]$ to $[12]$

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 8 & 6 & 8 & 5 & 4 & 12 & 4 & 6 & 12 & 2 & 4 & 2 & 3 \end{pmatrix}.$$

The functional digraph of f is given in Figure 3.4, where C_1 is the functional digraph on $\{1, 2, 3, 6, 8, 9, 10, 12, 13\}$ and C_2 is the functional digraph on $\{4, 5, 7, 11\}$. The longest

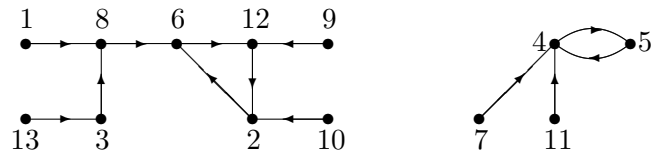


Figure 3.4: The functional digraph G_f .

path starting from 13 is

$$P: 13 \rightarrow 3 \rightarrow 8 \rightarrow 6 \rightarrow 12 \rightarrow 2,$$

with $f(2) = 6$, that is, $u_1 = 3$, $u_k = 2$ and $u_j = 6$ as in the proof. Deleting the arc $(2, 6)$ and vertex 13, we get a rooted tree H as illustrated in Figure 3.5. By applying the bijection between functional digraphs and doubly rooted trees, C_2 can be mapped to a doubly rooted tree D with roots 5 and 4 as shown in Figure 3.5, where $w_1 = 5$ and $w_2 = 4$ as in the proof. Merging H and D by adding an edge $(6, 5)$ and setting 2, 3 and 4 to be the three roots, we get a triply rooted tree T in Figure 3.5.

To prove that ϕ is a bijection, we give a description of the inverse map. For a triply rooted tree $T \in T_n$ with the three roots r_1, r_2 and r_3 , we first find the least common ancestor of r_2 and r_3 , and we denote it by u_0 . Suppose that the unique path P from u_0 to r_3 in T is $u_0 u_1 \cdots u_i = r_3$. Removing the edge (u_0, u_1) from T , we get two components T_1 and T_2 , where T_1 is rooted at r_1 and T_2 is rooted at u_1 . Adding $n + 1$ to T_1 by setting it

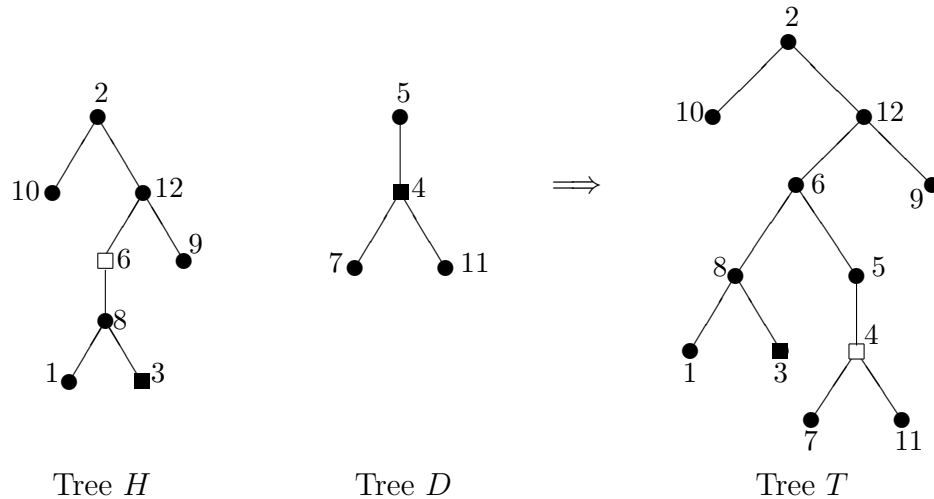


Figure 3.5: An example of the bijection ϕ .

as a child of r_2 . Now, T_1 can be viewed as a directed graph by making each edge point to the father. Then we add the arc (r_1, u_0) to T_1 to obtain a connected functional digraph C_1 .

Next, we transform T_2 rooted at u_1 into a doubly rooted tree by setting r_3 to be the second root. Then we get a functional digraph C_2 by applying the inverse map of the bijection of Joyal [4] and Goulden and Jackson [3].

Finally, let $G = C_1 \cup C_2$. Observe that G is a directed graph on $[n + 1]$ such that each vertex has outdegree one and the vertex $n + 1$ has indegree zero. In other words, G is the functional digraph of a function from $[n + 1]$ to $[n]$. We see that the above procedure is the inverse of the map ϕ .

It remains to prove the properties of ϕ as stated in the theorem. For a function f from $[n + 1]$ to $[n]$, an element x is a periodic point in f if and only if it is a vertex in a cycle in the functional digraph G_f . We see that x is in a cycle if and only if it is an ancestor of the third root in the triply rooted tree $\phi(f)$. Moreover, we observe that each element y in the orbit of $n + 1$ on f other than $n + 1$ itself corresponds to an ancestor of the second root in the triply rooted tree $\phi(f)$. This completes the proof. ■

For example, for the function f in Figure 3.4, there are five periodic points 2, 12, 6, 5, 4, which are the vertices in the path from the root 2 to the third root 4 in $\phi(f)$ as demonstrated in Figure 3.5. The orbit of 13 consists of 13, 3, 8, 6, 12, 2. The elements 3, 8, 6, 12, 2 correspond to the vertices in the path from the root 2 to the second root 3 in $\phi(f)$.

We notice that the bijection ϕ leads to a symmetry property of functions from $[n + 1]$ to $[n]$. Let $W_{n,i,j}$ denote the set of triply rooted trees on $[n]$ such that the depth of the

second root is i and the depth of the third root is j . Let $F_{n,i,j}$ denote the set of functions from $[n+1]$ to $[n]$ such that the size of the orbit of $n+1$ is i and the number of periodic points is j . By Theorem 3.1, we see that

$$|F_{n,i+1,j}| = |W_{n,i-1,j-1}|.$$

By the symmetry between the second roots and the third roots of triply rooted trees, that is,

$$|W_{n,i,j}| = |W_{n,j,i}|,$$

we arrive at the following symmetry property of functions from $[n+1]$ to $[n]$,

$$|F_{n,i+1,j}| = |F_{n,j+1,i}|. \quad (3.1)$$

The following theorem gives a formula for $W_{n,i,j}$.

Theorem 3.2 *For $n \geq 1$ and $0 \leq i, j \leq n-1$, we have*

$$|W_{n,i,j}| = \sum_{d=0}^{\min\{i,j\}} \frac{(i+j-d+1)n!}{(n-i-j+d-1)!} n^{n-i-j+d-2}, \quad (3.2)$$

where $\min\{i, j\}$ denotes the smaller value between i and j .

Proof. Let $W_{n,i,j}(d)$ denote the set of triply rooted trees T in $W_{n,i,j}$ such that d is the depth of the least common ancestor of the second root and the third root of T . We proceed to show that $W_{n,i,j}(d)$ is enumerated by the summand on the right hand side of (3.2).

Let T be a triply rooted tree in $W_{n,i,j}(d)$. We denote by P_1 the path from the first root to the second root and denote by P_2 the path from the first root to the third root. Observing that there are exactly $k = i + j - d + 1$ vertices on P_1 and P_2 , we see the number of ways to form P_1 and P_2 equals $\frac{n!}{(n-k)!}$. Moreover, as we know, there are kn^{n-k-1} forests consisting of k rooted trees on $[n]$ with k given roots. It follows that $W_{n,i,j}(d)$ is enumerated by the summand on the right hand side of (3.3). This completes the proof. ■

Combining Theorem 3.1 and Theorem 3.2, we obtain the following formula for $|F_{n,i+1,j}|$.

Theorem 3.3 *For $n \geq 1$ and $1 \leq i, j \leq n$, we have*

$$|F_{n,i+1,j}| = \sum_{s=0}^{\min\{i,j\}-1} \frac{(i+j-s-1)n!}{(n-i-j+s+1)!} n^{n-i-j+s}. \quad (3.3)$$

Acknowledgments. We wish to thank Catherine H.F. Yan and the referee for valuable suggestions. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, and the National Science Foundation of China.

References

- [1] A.V. Aho, J.E. Hopcroft and J.D. Ullman, On finding lowest common ancestors in trees, *SIAM J. Comput.* 5 (1976), 115–132.
- [2] A. Cayley, Note sur une formule pour la réversion des séries, *J. reine angew. Math.* 52 (1856), 276–284.
- [3] I.P. Goulden and D.M. Jackson, *Combinatorial Enumeration*, John Wiley, New York, 1983.
- [4] A. Joyal, Une théorie combinatoire des séries formelles, *Adv. Math.* 42 (1981), 1–82.
- [5] A. Lacasse, Bornes PAC-Bayes et algorithmes d’apprentissage, Ph.D. Thesis, Université Laval, Quebec, 2010.
- [6] H. Prodinger, An identity conjectured by Lacasse via the tree function, arXiv:1301.3669.
- [7] J. Riordan, *Combinatorial Identities*, Robert E. Krieger Publishing Co., New York, 1968.
- [8] Y. Sun, A simple proof of an identity of Lacasse, *Electron. J. Combin.* 20(2) (2013), #P11.
- [9] M. Younsi, Proof of a combinatorial conjecture coming from the PAC-Bayesian machine learning theory, arXiv:1209.0824.