

A refinement of the formula for k -ary trees and the Gould-Vandermonde's convolution

Ricky X. F. Chen

Center for Combinatorics, LPMC-TJKLC
Nankai University, Tianjin 300071, P. R. China
ricky_chen@mail.nankai.edu.cn

Submitted: Jul 10, 2007; Accepted: Mar 25, 2007; Published: Apr 3, 2008

Mathematics Subject Classification: 05A19, 05C05

Abstract

In this paper, we present an involution on some kind of colored k -ary trees which provides a combinatorial proof of a combinatorial sum involving the generalized Catalan numbers $C_{k,\gamma}(n) = \frac{\gamma}{kn+\gamma} \binom{kn+\gamma}{n}$. From the combinatorial sum, we refine the formula for k -ary trees and obtain an implicit formula for the generating function of the generalized Catalan numbers which obviously implies a Vandermonde type convolution generalized by Gould. Furthermore, we also obtain a combinatorial sum involving a vector generalization of the Catalan numbers by an extension of our involution.

1 Introduction

Recently, the author obtained the following identity involving the Catalan numbers $C_n = \frac{1}{n+1} \binom{2n}{n}$ by accident which is similar to the identity 2(a) in Riordan's book [6, p. 152–153]:

$$\sum_{i=0}^n (-1)^{n-i} \binom{i+1}{n-i} C_i = \delta_{0n}, \text{ for } n \geq 0, \quad (1)$$

where δ_{0n} is the Kronecker symbol. It is well known [10] that C_n counts the number of 2-ary trees (or complete binary trees) with n internal vertices (vertices with outdegree at least 1). Now suppose the number of ordered forests with γ β -ary trees and with total number of n internal vertices is $C_{\beta,\gamma}(n)$ which is a natural generalization of the Catalan numbers. It is also well known [4, 10] that

$$C_{\beta,\gamma}(n) = \frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n}.$$

Then by generalizing the case $C_n = C_{2,1}(n)$ to $C_{\beta,\gamma}(n)$, we obtain a generalization identity of (1):

Theorem 1.1. For $n \geq 0$, $\alpha, \beta, \gamma \in \mathbb{C}$,

$$\sum_{i=0}^n (-1)^{n-i} \binom{(\beta-1)i + \alpha}{n-i} \frac{\gamma}{\beta i + \gamma} \binom{\beta i + \gamma}{i} = (-1)^n \binom{\alpha - \gamma}{n}. \quad (2)$$

Actually, we can even generalize (2) by introducing following notations: For any vectors $\mathbf{a} = (a_1, \dots, a_t)$ and $\mathbf{b} = (b_1, \dots, b_t)$, we denote $\mathbf{a} \leq \mathbf{b}$ if $a_i \leq b_i$ for all $1 \leq i \leq t$; We also define $\mathbf{b} - \mathbf{a} = (b_1 - a_1, \dots, b_t - a_t)$ and $\mathbf{a} \cdot \mathbf{b} = \sum_{k=1}^t a_k b_k$; As usual, any dimension vector with constant entries k 's will be denoted by \mathbf{k} for short if there is no confusion; For $x \in \mathbb{C}$, $\sum_{i=1}^t m_i = x$ and $m_1, \dots, m_{t-1} \in \mathbb{N}$, we interpret the general multinomial coefficient as follows:

$$\binom{x}{m_1, \dots, m_{t-1}, m_t} = \binom{x}{m_1} \binom{x - m_1}{m_2} \cdots \binom{x - m_1 - \cdots - m_{t-2}}{m_{t-1}},$$

where $\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}$ for $k \in \mathbb{N}^+$ and $\binom{x}{0} = 1$.

Theorem 1.2. For $\mathbf{n} = (n_1, \dots, n_t) \geq \mathbf{0}$, $\mathbf{p} = (p_1, \dots, p_t) \geq \mathbf{1} \in \mathbb{N}^t$, $t \geq 1$, $\gamma \geq 0 \in \mathbb{N}$ and $\alpha \in \mathbb{C}$, define $Q(\mathbf{n}; \mathbf{p}; \gamma)$ which can be viewed as a vector generalization of the Catalan numbers as follows:

$$Q(\mathbf{n}; \mathbf{p}; \gamma) = \frac{\gamma}{\mathbf{n} \cdot \mathbf{p} + \gamma} \binom{\mathbf{n} \cdot \mathbf{p} + \gamma}{n_1, \dots, n_t, \mathbf{n} \cdot (\mathbf{p} - \mathbf{1}) + \gamma}.$$

Then, there holds

$$\begin{aligned} \sum_{\mathbf{i}=(i_1, \dots, i_t) \geq \mathbf{0}} (-1)^{i_1 + \cdots + i_t} \binom{(\mathbf{p} - \mathbf{1}) \cdot (\mathbf{n} - \mathbf{i}) + \alpha}{i_1, \dots, i_t, (\mathbf{p} - \mathbf{1}) \cdot (\mathbf{n} - \mathbf{i}) + \alpha - \sum_{j=1}^t i_j} Q(\mathbf{n} - \mathbf{i}; \mathbf{p}; \gamma) \\ = (-1)^{n_1 + \cdots + n_t} \binom{\alpha - \gamma}{n_1, \dots, n_t, \alpha - \gamma - \sum_{j=1}^t n_j}. \quad (3) \end{aligned}$$

In fact, $Q(\mathbf{n}; \mathbf{p}; \gamma)$ also counts some kind of ordered forests which we will say in details later. When looking up (2) and (3) in the literature, we first found a special case of (2) for $\alpha = 0, \gamma = 1$ in a very recent paper of Bajunaid et al. [1, 2005]. They claimed that they did not find their result in the literature and they did not know any obvious combinatorial proof for their special case, even for (1). But, shortly later, we found that a special case of (2) for $\alpha = \beta$ also appeared in an early paper of Gould [3] though he did not state it explicitly. Hence, the main purpose of this paper is to prove (2) and (3), especially combinatorially.

Our combinatorial proofs of (2) and (3) are via an involution on some kind of colored ordered trees. We note that ordered trees and k -ary trees have been extensively studied and refer to [4, 10] and references therein. In addition, in representing ordered trees we will put the root on the top and put other vertices on different levels: the root on level 0, the children of vertices of level i on level $i + 1$ (right below level i).

In the rest of this paper, we first present an involution on colored k -ary trees which leads to a combinatorial proof of (2). Next, we comment briefly on how to prove (3) by a similar argument as (2). Finally we give a noncombinatorial proof of (2) by a modified Riordan array theorem derived from theorems in Sprugnoli et al. [5]. Also we refine the formula for $C_{\beta,\gamma}(n)$ with aids of a Gould classes of inverse relation and obtain an implicit formula for the generating function of $C_{\beta,\gamma}(n)$ which obviously implies a Vandermonde type convolution generalized by Gould [2, 3].

2 Proofs of (2), (3) via an involution on colored trees

We rewrite identity (2):

$$\sum_{i=0}^n (-1)^i \binom{(\beta-1)(n-i)+\alpha}{i} \frac{\gamma}{\beta(n-i)+\gamma} \binom{\beta(n-i)+\gamma}{n-i} = (-1)^n \binom{\alpha-\gamma}{n}. \quad (4)$$

It is known that the left side of (4) is a polynomial in α, β, γ . Hence it suffices to prove (4) for all positive integers α, β, γ . Therefore, we first give a combinatorial proof for the case $\gamma = 1, \alpha = 1, \beta \geq 1$. Then we will show that with a little more work the general cases will follow.

Firstly, we know that $\frac{1}{\beta(n-i)+1} \binom{\beta(n-i)+1}{n-i}$ counts the number of β -ary trees with $n-i$ internal vertices and $(\beta-1)(n-i)+1$ is the number of leaves in such a tree. Hence,

$$\binom{(\beta-1)(n-i)+1}{i} \frac{1}{\beta(n-i)+1} \binom{\beta(n-i)+1}{n-i}$$

counts the number of β -ary trees with $n-i$ internal vertices and i colored leaves. (Note: When a β -ary tree has only one vertex, this vertex will be treated as a leaf.)

However, we can classify all those β -ary trees with $n-i$ internal vertices and i colored leaves into two classes for $n > 0, i \geq 0$:

- The first class consists of trees in which there is at least one colored leaf in the lowest two levels (the two levels furthest to the root) and to the left of which all vertices in the same level have no children. The leftmost and lowest such a colored leaf will be called the *candidate leaf*.
- The second class consists of the rest of trees, i.e., the trees in which there is no colored leaf in the lowest level (the level furthest to the root) and there is an internal vertex in the second lowest level to whose left in the same level there is no colored leaf. The leftmost such an internal vertex will be called the *incumbent internal vertex*.

Now, we define an operation on these colored β -ary trees which will give an involution:

- For each β -ary tree with $n-i$ internal vertices and i colored leaves of the first class, we attach β leaves to its candidate leaf and erase the color of the candidate leaf (now an internal vertex). Thus we obtain a β -ary tree with $n-i+1$ internal vertices and $i-1$ colored leaves.

- For each β -ary tree of the second class, we delete all children of its incumbent internal vertex and color the incumbent internal vertex (now a leaf). Thus, we obtain a β -ary tree with $n - i - 1$ internal vertices and $i + 1$ colored leaves.

Figure 1 shows an example of this involution. From above involution, when $n > 0$, we obtain a perfect matching over all β -ary trees with $n - i$ internal vertices and i colored leaves for all $i \geq 0$. Moreover, if each colored ordered tree T has weight $(-1)^{\#\text{of colored leaves in } T}$, then the two trees matched to each other are weighted by $+1$ and -1 respectively. Hence, the total weight over these colored β -ary trees is 0. The case for $n = 0$ is obvious. This completes the proof of the case $\alpha = 1, \gamma = 1$ of (2)

$$\sum_{i=0}^n (-1)^i \binom{(\beta-1)(n-i)+1}{i} \frac{1}{\beta(n-i)+1} \binom{\beta(n-i)+1}{n-i} = \delta_{0n}. \quad \square$$

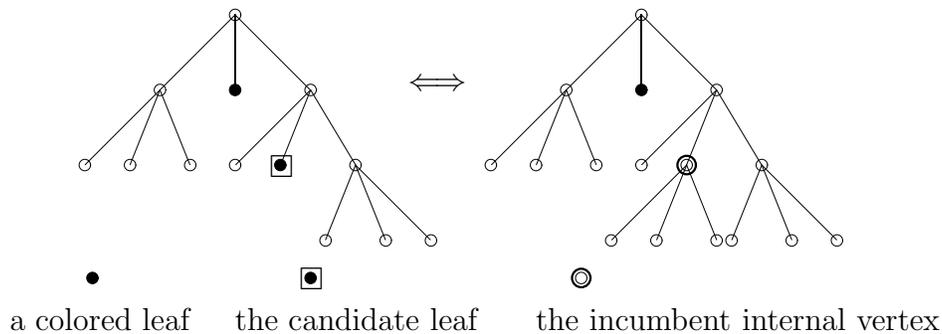


Figure 1: An example of above involution for $n = 6, \alpha = 1, \beta = 3$, the tree on the left side is of the first class while the other belongs to the second class.

For simplicity of the remaining presentment, we describe the core idea of the above involution in the following obvious lemma.

Lemma 2.1. *Suppose S is a finite set of structures each of which has two finite integer indices (or called parameters, etc.), say x and y , and has weight $(-1)^x$. All these structures can be partitioned into two classes: the first and the second. Suppose every structure of the first class in S corresponds uniquely to a structure of the second class in S with index x increasing by 1 and index y decreasing by 1 and vice versa. Then the total weight of structures in S is 0.*

Now, when $\alpha > 1$, we just view

$$\binom{(\beta-1)(n-i)+\alpha}{i} \frac{1}{\beta(n-i)+1} \binom{\beta(n-i)+1}{n-i}$$

as the number of $(\alpha - 1)$ -planted β -ary trees (analog to usual planted tree except that there are $\alpha - 1$ planted roots, see Figure 2 for an example) with $n - i$ internal vertices and j colored leaves and $i - j$ colored planted roots for all $0 \leq i - j \leq \alpha - 1$. (Note: The planted roots are neither internal vertices nor leaves.)

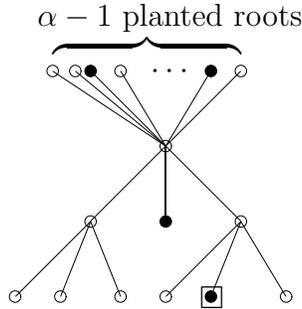


Figure 2: An example of $(\alpha - 1)$ -planted 3-ary tree in which some planted roots are also colored.

Following Lemma 2.1, the two indices here are respectively the number of internal vertices (y) and the sum of colored leaves together with the number of colored planted roots (x), and the two classes are defined to be the same as in the above proof for $\alpha = 1, \gamma = 1$.

- (i) When $n > \alpha - 1 > 0$, for all possible $0 \leq i < n$, any planted β -ary tree with $n - i$ internal vertices and j colored leaves and $i - j$ colored planted roots is either of the first class or of the second class. So their total weight is 0 from our involution.
- (ii) When $n \leq \alpha - 1$, except for those structures which belong to the defined two classes, there still remain structures for $i = n$ and with only the planted roots colored. See Figure 3 for an example of the case $i = n$. Since the total weight of structures belonging to the defined two class is 0, we only need to calculate the total weight of those exceptional structures. This is equivalent to choosing n out of all $\alpha - 1$ planted roots. Hence, the total weight is $(-1)^n \binom{\alpha - 1}{n}$.

Thus the proof for all $\alpha \geq 1, \beta \geq 1$ follows. □

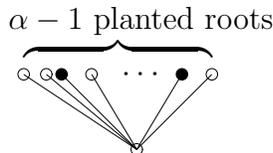


Figure 3: An example of $(\alpha - 1)$ -planted β -ary tree with 0 internal vertices and n colored planted roots, i.e., $i = n$.

For $\gamma \geq 1$, it is well known that

$$\frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n}$$

counts the number of ordered forests of β -ary trees with totally n internal vertices and γ components and in which there are just $(\beta - 1)n + \gamma$ leaves. Thus, we can view

$$\binom{((\beta - 1)(n - i) + \alpha)}{i} \frac{\gamma}{\beta(n - i) + \gamma} \binom{\beta(n - i) + \gamma}{n - i}$$

as the number of $(\alpha - \gamma)$ -planted β -ary forests with γ components where the planted roots are planted on the first ordered tree in the forest. We put all roots (except planted roots) of the forest on the same level and define the two classes of those forests in the same way as trees. So, with same argument, the general case $\gamma \geq 1$ follows. See an illustration in Figure 4. Thus, we have completed a combinatorial proof of (2) for all positive integers α, β, γ , which implies that (2) holds for all $\alpha, \beta, \gamma \in \mathbb{C}$. \square

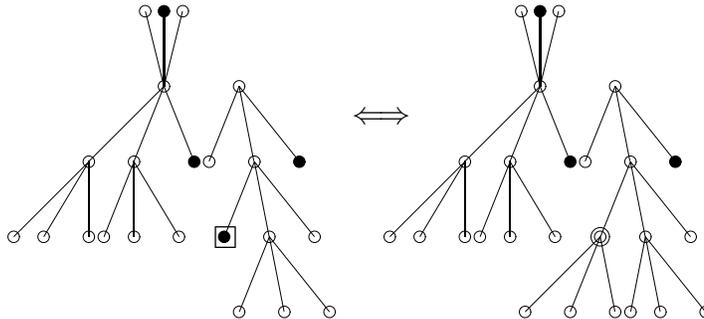


Figure 4: An example of the involution for $\alpha = 4, \beta = 3, \gamma = 2, n = 10$.

To prove (3), we first note the following facts:

Fact 2.2. (Stanley [10, p.34]) *The number of ordered forests with n_i internal vertices of outdegree p_i and γ component trees is*

$$Q(\mathbf{n}; \mathbf{p}; \gamma) = \frac{\gamma}{\mathbf{n} \cdot \mathbf{p} + \gamma} \binom{\mathbf{n} \cdot \mathbf{p} + \gamma}{n_1, \dots, n_t, \mathbf{n} \cdot \mathbf{p} + \gamma - \sum_{j=1}^t n_j}.$$

In addition, the number of leaves in such a forest is $\mathbf{n} \cdot (\mathbf{p} - \mathbf{1}) + \gamma$.

Thus, the summand in the left hand side of (3) counts the total weight of $(\alpha - \gamma)$ -planted ordered forests with leaves (and planted roots) colored with colors in $\{1, \dots, t\}$, and in which there are $n_j - i_j$ internal vertices of outdegree p_j and i_j leaves plus planted roots have color j for $1 \leq j \leq t$. We also assume that all planted roots are planted on the first ordered tree when $r > 1$ and the two classes are still the same. But we modify the involution a little bit:

- (i) If the candidate leaf has color j , then we attach p_j leaves to it and erase its color.
- (ii) Correspondingly, if the incumbent internal vertex has p_j leaves, we delete all its leaves and color it with color j .

We omit figures for illustration of this case since it is so clear and simple. After applying this modified involution and by the same reasoning as ordered β -ary forests, all cancel out except for the case $i_j = n_j$ with only planted roots colored. It reduces to choosing out n_i planted roots from $\alpha - \gamma$ and coloring them with color i for all $1 \leq i \leq t$, for which there are

$$\binom{\alpha - \gamma}{n_1, \dots, n_t, \alpha - \gamma - \sum_{i=1}^t n_i}$$

different ways. Thus we obtain (3). However, because the left side of (3) is a polynomial in α but not necessarily a polynomial in \mathbf{p} and γ , we can not generalize it to $\mathbf{p} \in \mathbb{C}^t$ and $\gamma \in \mathbb{C}$ in the manner of (2).

3 Refinement of $C_{\beta,\gamma}(n)$ and the Gould-Vandermonde's convolution

In this section, we will prove (2) and refine the formula for k -ary trees and the Gould-Vandermonde's convolution through Riordan array theory. The Riordan array theory is an effective tool to prove and invert combinatorial identities [5–9]. For clarity, here we present the related results which will be used later.

Definition 3.1. A lower triangular infinite matrix, T , is a Riordan array, if the generating function of the k th column is

$$g(x)f(x)^k,$$

for $k = 0, 1, 2, 3, \dots$, where

$$\begin{aligned} g(x) &= g_0 + g_1x + g_2x^2 + g_3x^3 + \dots, & g_0 &\neq 0, \\ f(x) &= f_1x + f_2x^2 + f_3x^3 + \dots, & f_1 &\neq 0. \end{aligned}$$

In addition, we will denote this Riordan array as $[g(x), f(x)]$ and denote its (n, k) entry as $[g(x), f(x)]_{n,k}$.

Theorem 3.2 (Riordan array theorem). Let $A(x) = a_0 + a_1x + a_2x^2 + \dots$, and $L(x) = l_0 + l_1x + l_2x^2 + l_3x^3 + \dots$, and $[g(x), f(x)]$ is a Riordan array. Then

$$\sum_{k=0}^n [g(x), f(x)]_{n,k} a_k = l_n, \quad \text{for } n = 0, 1, 2, \dots,$$

if and only if

$$g(x)A(f(x)) = L(x). \tag{5}$$

Sprugnoli et al. [5] studied how to invert combinatorial sums through Riordan array theory and Lagrange inversion formula. We present the following theorem easily derived from Theorem 2.1–2.4 in the paper [5].

Theorem 3.3 (modified Riordan array theorem). Let $A(x) = a_0 + a_1x + a_2x^2 + \dots$, and $L(x) = l_0 + l_1x + l_2x^2 + l_3x^3 + \dots$, and $[g(x), f(x)]$ is a Riordan array. Then

$$\sum_{k=0}^n [g(x), f(x)]_{n,k} a_k = l_n, \quad \text{for } n = 0, 1, 2, \dots,$$

if and only if

$$n[x^n]A(x) = [x^{n-1}] \left(\frac{x}{f(x)} \right)^n \left(\frac{L(x)}{g(x)} \right)', \quad \text{for } n > 0, \tag{6}$$

and $a_0 = \frac{L(0)}{g(0)}$.

Now, we come to a proof of (2) via the modified Riordan array theorem: Firstly, for (2) we have $A(x) = \sum_{n \geq 0} \frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n} x^n$, $f(x) = x(1-x)^{\beta-1}$, $g(x) = (1-x)^\alpha$, $L(x) = \sum_{n \geq 0} (-1)^n \binom{\alpha - \gamma}{n} x^n = (1-x)^{\alpha - \gamma}$. Thus, for $n \geq 1$,

$$\begin{aligned} [x^{n-1}] \left(\frac{x}{x(1-x)^{\beta-1}} \right)^n \left(\frac{(1-x)^{\alpha-\gamma}}{(1-x)^\alpha} \right)' &= [x^{n-1}] \gamma (1-x)^{(1-\beta)n - \gamma - 1} \\ &= \gamma \binom{\beta n + \gamma - 1}{n-1} \\ &= n [x^n] \sum_{n \geq 0} \frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n} x^n. \end{aligned}$$

Additionally, $\frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n} = \frac{(1-0)^{\alpha-\gamma}}{(1-0)^\alpha}$. By Theorem 3.3, the proof follows. \square

Let $C_{\beta, \gamma}(x) = \sum_{n \geq 0} C_{\beta, \gamma}(n) x^n$. By Riordan array theorem (5) and (2), we also obtain an implicit formula for $C_{\beta, \gamma}(x)$ which appeared in Gould [3].

Corollary 3.4.

$$C_{\beta, \gamma}(x(1-x)^{\beta-1}) = (1-x)^{-\gamma}. \quad (7)$$

From (7), we immediately obtain $C_{\beta, \alpha_1 + \alpha_2}(x) = C_{\beta, \alpha_1}(x) C_{\beta, \alpha_2}(x)$ which leads to the following famous convolution of Vandermonde type generalized by Gould [2, 3]. (We refer to it as Gould-Vandermonde's convolution in the title of this paper.)

Corollary 3.5. *For any $\alpha, \beta, \gamma \in \mathbb{C}$ and $n \geq 0$, there holds*

$$\begin{aligned} \frac{\alpha_1 + \alpha_2}{\beta n + \alpha_1 + \alpha_2} \binom{\beta n + \alpha_1 + \alpha_2}{n} \\ = \sum_{i=0}^n \frac{\alpha_1}{\beta i + \alpha_1} \binom{\beta i + \alpha_1}{i} \frac{\alpha_2}{\beta(n-i) + \alpha_2} \binom{\beta(n-i) + \alpha_2}{n-i}. \end{aligned} \quad (8)$$

Note that even if we did not know the exact formula of $C_{\beta, \gamma}(n)$, we still have by the involution argument in section 2 that

$$\sum_{i=0}^n (-1)^{n-i} \binom{(\beta-1)i + \alpha}{n-i} C_{\beta, \gamma}(i) = (-1)^n \binom{\alpha - \gamma}{n}.$$

Hence, from the Gould classes of inverse relation [6, p.52]

$$b_n = \sum_{k=0}^n \binom{m + ak}{n-k} z^{n-k} a_k, \quad a_n = \sum_{k=0}^n \frac{-ak - m}{-an - m} \binom{-an - m}{n-k} b_k z^{n-k}, \quad (9)$$

we obtain

Corollary 3.6. *For $n > 0$,*

$$C_{\beta, \gamma}(n) = \sum_{k=0}^n (-1)^n \frac{(1-\beta)k - \alpha}{(1-\beta)n - \alpha} \binom{(1-\beta)n - \alpha}{n-k} \binom{\alpha - \gamma}{k}. \quad (10)$$

By setting $\alpha = 0$ in (10) and applying Vandermonde's convolution, it is easy to obtain a closed formula:

$$\begin{aligned}
C_{\beta,\gamma}(n) &= \sum_{k=0}^n (-1)^n \frac{n - (n - k)}{n} \binom{(1 - \beta)n}{n - k} \binom{-\gamma}{k} \\
&= \sum_{k=0}^n (-1)^n \binom{(1 - \beta)n}{n - k} \binom{-\gamma}{k} - \sum_{k=0}^n (-1)^n \frac{n - k}{n} \binom{(1 - \beta)n}{n - k} \binom{-\gamma}{k} \\
&= (-1)^n \binom{(1 - \beta)n - \gamma}{n} + (-1)^n (\beta - 1) \binom{(1 - \beta)n - 1 - \gamma}{n - 1} \\
&= \frac{\gamma}{\beta n + \gamma} \binom{\beta n + \gamma}{n}.
\end{aligned}$$

We finally remark that our approach to obtain the formula for k -ary trees may be applied to other combinatorial structures, especially recursive ones.

Acknowledgement. The author would like to thank Prof. Christian Reidys for help in his study. The author is also very grateful to the referees for their valuable suggestions and comments. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, the Ministry of Science and Technology and the National Science Foundation of China.

References

- [1] I. Bajunaid, J. M. Cohen, F. Colonna, D. Singman, Function Series, Catalan Numbers, and Random Walks on Trees, *Amer. Math. Monthly* **112** (2005) 765–785.
- [2] H. W. Gould, Final analysis of Vandermonde's convolution, *Amer. Math. Monthly* **64** (1957) 409–415.
- [3] H. W. Gould, Some Generalization of Vandermonde's Convolution, *Amer. Math. Monthly* **63** (1956) 84–91.
- [4] Ira M. Gessel, Seunghyun Seo, A Refinement of Cayley's Formula for Trees, *Electron. J. Combin.* **11(2)** (2006) #R27.
- [5] D. Merlini, R. Sprugnoli, M. C. Verri, Combinatorial inversions and implicit Riordan arrays, *Electron. Notes Discrete Math.* **26** (2006) 103–110.
- [6] J. Riordan, *Combinatorial Identities*, Wiley and Sons, New York, 1968.
- [7] L. W. Shapiro, S. Getu, W. J. Woan, L. Woodson, The Riordan group, *Discrete Appl. Math.* **34** (1991) 229–239.
- [8] R. Sprugnoli, Riordan arrays and combinatorial sums, *Discrete Math.* **132** (1994) 267–290.
- [9] R. Sprugnoli, Riordan arrays and the Abel–Gould identity, *Discrete Math.* **142** (1995) 213–233.
- [10] R. P. Stanley, *Enumerative Combinatorics*, vol. 2, Cambridge Univ. Press, Cambridge, 1997.