J. Math. Study
doi: 10.4208/jms.v56n1.23.01

Vol. **56**, No. 1, pp. 1-17
March 2023

# Protected Branches in Ordered Trees

Lin Yang and Shengliang Yang*

*Department of Applied Mathematics, Lanzhou University of Technology, Lanzhou 730050, China.*

Received December 3, 2021; Accepted March 17, 2022;
Published online November 22, 2022.

**Abstract.** In this paper, we consider the class of ordered trees and its two subclasses, bushes and planted trees, which consist of the ordered trees with root degree at least 2 and with root degree 1 respectively. In these three classes, we study the number of trees of size $n$ with $k$ protected (resp. unprotected) branches, and the total number of branches (resp. protected branches, unprotected branches) among all trees of size $n$. The explicit formulas as well as the generating functions are obtained. Furthermore, we find that, in each class, as $n$ goes to infinity, the proportion of protected branches among all branches in all trees of size $n$ approaches $1/3$.

**AMS subject classifications**: 05A10, 05A15, 05A16, 05C05

**Key words**: ordered tree, bush, planted tree, protected branch, unprotected branch, Catalan number, generating function.

## 1 Introduction

An ordered tree is defined recursively as having a root and an ordered set of subtrees [14, 21]. We will draw ordered trees with the root on the top level, the root being connected with the roots of its subtrees by line segments, called edges. The size of a tree is defined to be the number of edges. For each vertex $v$, the number of subtrees rooted at $v$ is defined as the degree of $v$. In the graph theory, it is also named as outdegree. There are many different definitions for bush [8, 11, 12]. In this paper, we use the definition in [8]. A bush is an ordered tree in which the degree of the root is at least 2, while a planted tree is an ordered tree with root degree 1. The 14 ordered trees of size 4 are shown in Figure 1, in which the first 5 trees are planted trees and the remaining 9 trees are bushes.

A vertex of degree zero is called a leaf. A vertex of positive degree is called an internal node. A protected point is a vertex which is not a leaf and which is not distance 1 from a leaf. Cheon and Shapiro [5] started the study of protected points in ordered trees, and

*Corresponding author. *Email addresses:* `yanglinmath@163.com` (Yang L), `slyang@lut.cn` (Yang S L)

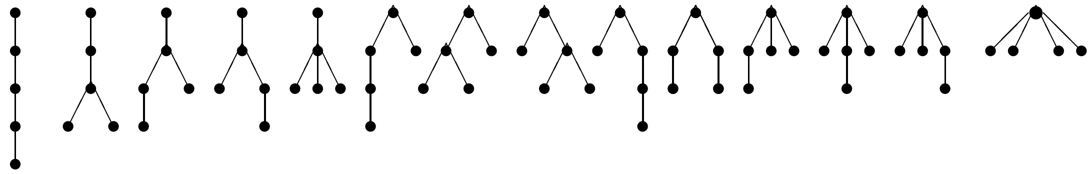

Figure 1: The 14 ordered trees of size 4 in which the last 9 are bushes.

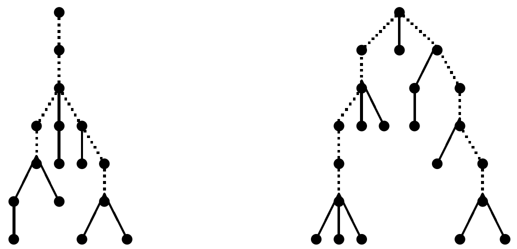

Figure 2: A planted ordered tree and a bush.

they showed that the average portion of protected points in ordered trees with $n$ edges approaches $1/6$ as $n$ goes to $\infty$. Since this pioneering paper, a large number of extensions have been studied [3,4,7,13,15,16]. We will focus on protected branches in ordered trees in this paper. An internal node of degree at least 2 is called a branch node. A tree with no branch nodes is called a path. By a branch we mean a path connecting either the root and a nearest branch node, or two nearest branch nodes, or a leaf and the nearest branch node. Riordan [18] enumerated the plane trees by number of branches and endpoints. Deutsch [12] introduced a new decomposition of ordered trees by branches, and using this decomposition he enumerated the ordered trees with prescribed root degrees, node degrees, and branch lengths.

We define a protected branch of a ordered tree is a branch which does not contain a leaf, and an unprotected branch is a branch ending at a leaf. For instance, a planted tree with 4 protected branches and a bush with 5 protected branches are displayed in Figure 2, in which the dashed branches are all protected branches.

In this paper, we will enumerate the number of ordered trees (resp. bushes, planted trees) of size $n$ with $k$ protected (resp. unprotected) branches, and the total number of branches (resp. protected branches, unprotected branches) among all ordered trees (resp. bushes, planted trees) of size $n$. In Section 2, we compute the numbers of branches in ordered trees, bushes, and planted trees. In Section 3, the enumerations of protected branches will be discussed. We will show that, as $n$ goes to infinity, the average proportion of protected branches among all branches of ordered trees (resp. bushes, planted trees) of size $n$ approaches $1/3$. In Section 4, the enumerations of unprotected branches will be considered. As by-products, we obtain three new combinatorial interpretations

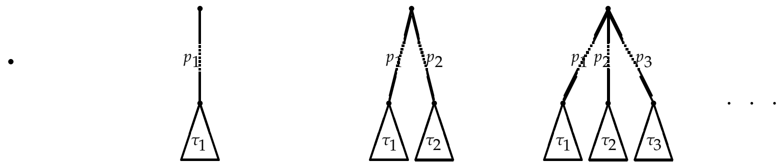


Figure 3: The decomposition of ordered trees.

for the Catalan numbers. The proofs in this paper are based mainly on functional equations which are obtained from the symbolic method of Flajolet [14], and alternative bijective proofs would be of interest.

## 2 The total number of branches

Let $T(t,z)$ be the generating function of ordered trees of a given size and a given number of branches, let $P(z)$ be the generating function of all paths, and let $H(t,z)$ be the generating function of all trees except the planted ones.

Using the decomposition of ordered trees given by Deutsch [12], as shown in Figure 3, we have the following two equations:

$$H = 1 + \sum_{j=2}^{\infty} t^j P^j H^j, \tag{2.1}$$

$$T = 1 + \sum_{j=1}^{\infty} t^j P^j H^j, \tag{2.2}$$

where $P(z) = \frac{z}{1-z}$. Hence,

$$H = 1 - tPH + (tP + t^2P^2)H^2, \tag{2.3}$$

and eliminating $H$ and $P$ from Eqs. (2.2) and (2.3), we can obtain

$$T = 1 - z + zT + tz(1 - T + T^2). \tag{2.4}$$

**Proposition 2.1.** *For $n \geq 1$, let $a_n$ denote the number of branches in all ordered trees of size $n$. Then*

$$a_n = \frac{3n^2 - 2n + 1}{n^2 + n} \binom{2n-2}{n-1}.$$

*Proof.* Taking into account that $T(1,z) = C(z)$, by partial differentiation for both sides of

(2.4) with respect to $t$ gives

$$\begin{aligned} a_n &= [z^n]\frac{\partial T(t,z)}{\partial t}|_{t=1} \\ &= [z^n](1-z)zB(z)C(z)^2 \\ &= \binom{2n-1}{n-2}+\binom{2n-2}{n-1} \\ &= \frac{3n^2-2n+1}{n^2+n}\binom{2n-2}{n-1}. \end{aligned}$$

the generating function is $(1-z)zB(z)C(z)^2 = z+3z^2+11z^3+41z^4+154z^5+582z^6+\cdots$. □

Eq. (2.4) is also obtained by Riordan [18] in 1975 by using recurrence relations, he also showed that

$$T(t,z)=1+\sum_{n=1}^{\infty}\sum_{k=1}^{n}\binom{n-1}{k-1}m_{k-1}t^kz^n,$$

with $m_k=\sum_{j=0}^{k}\binom{k}{2j}C_j$ are Motzkin numbers. From here, we can get that the number of branches in all ordered trees of size $n$ equals

$$a_n=\sum_{k=1}^{n}k\binom{n-1}{k-1}m_{k-1}=\frac{3n^2-2n+1}{n+1}C_{n-1}.$$

This sequence appears as A076540 in the OEIS [20], but we did not find other explicit interpretations in literatures.

Let $a_{n,k}$ denote the number of ordered trees of size $n$ with $k$ branches, then $a_{n,k}=\binom{n-1}{k-1}m_{k-1}$ and the first few rows of the array $(a_{n,k})_{n\geq 1,k\geq 1}$ are

$$(a_{n,k})_{n\geq 1,k\geq 1}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 2 & 2 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 3 & 6 & 4 & 0 & 0 & 0 & \cdots \\ 1 & 4 & 12 & 16 & 9 & 0 & 0 & \cdots \\ 1 & 5 & 20 & 40 & 45 & 21 & 0 & \cdots \\ 1 & 6 & 30 & 80 & 135 & 126 & 51 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

which is the sequence A091187 in OEIS [20]. Baril and Kirgizov [1] showed that $a_{n,k}$ is the number of permutations of length $n$ with $k$ pure descents and avoiding all patterns in $\{132,213,312\}$.

Let $B(t,z)$ be the generating function of all bushes according to the number of edges and the number of branches. It is easy to see that $B(t,z)=H(t,z)-1$. Hence, from (2.3) we get

$$B(t,z)=(tP+t^2P^2)(B(t,z)+1)^2-tP(B(t,z)+1),$$

or equivalently

$$(B(t,z)+1)^2tz(1-z+tz)-(B(t,z)+1)(1-z)(1-z+tz)+(1-z)^2=0. \tag{2.5}$$

If $b_{n,k}=[t^kz^n]B(t,z)$ denotes the number of all bushes of size $n$ with $k$ branches, then we have the following array

$$(b_{n,k})_{n\geq 1,k\geq 1}=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 3 & 3 & 3 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 4 & 6 & 12 & 6 & 0 & 0 & 0 & \cdots \\ 0 & 5 & 10 & 30 & 30 & 15 & 0 & 0 & \cdots \\ 0 & 6 & 15 & 60 & 90 & 90 & 36 & 0 & \cdots \\ 0 & 7 & 21 & 105 & 210 & 315 & 252 & 91 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Proposition 2.2.** *For $n\geq 2$, let $b_n$ denote the number of branches in all bushes of size $n$. Then*

$$b_n=\frac{9n^2-15n+6}{n^2+n}\binom{2n-4}{n}.$$

*Proof.* By partial differentiation for both sides of (2.5) with respect to $t$ gives

$$\begin{aligned}&\frac{\partial B(t,z)}{\partial t}|_{t=1}\\ =&(C(z)-z-1)(1-z)B(z)=(zC(z)^2-z)(1-z)B(z)=(1-z)zB(z)(C(z)^2-1),\end{aligned}$$

where the fact $B(1,z)=C(z)-zC(z)-1$ is used. Hence

$$\begin{aligned}b_n&=[z^n]\frac{\partial B(t,z)}{\partial t}|_{t=1}\\ &=[z^n](1-z)zB(z)(C(z)^2-1)\\ &=\binom{2n-1}{n-2}+\binom{2n-4}{n-2}\\ &=\frac{9n^2-15n+6}{n^2+n}\binom{2n-4}{n}.\end{aligned}$$

The generating function is $(1-z)zB(z)(C(z)^2-1)=2z^2+7z^3+27z^4+104z^5+400z^6+\cdots$. $\square$

The sequence $(b_n)_{\geq 2}$ in the above proposition equals the sequence obtained by multiplying the infinite vector $(2,3,4,5,\cdots)^T$ on the right of the array $(b_{n,k})_{n\geq 2,k\geq 2}$. They are not registered in OEIS.

Let $Q(t,z)$ be the generating function for the planted trees with respect to the size and the number of branches, i.e., the coefficient $[t^kz^n]Q(t,z)$ is the number of planted trees with $n$ edges and $k$ branches. Since any planted tree is either a path or a path connecting a bush, we have that $Q(t,z)=tP+tPB(t,z)$. Applying (2.5),

$$Q(t,z)^2(1-z+tz)-Q(t,z)(1-z+tz)+tz=0. \tag{2.6}$$

Riordan [18] showed that

$$Q(t,z)=\sum_{n=1}^{\infty}\sum_{k=1}^{n}\binom{n-1}{k-1}r_{k-1}t^kz^n,$$

with $r_k=\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}C_j$ are Riordan numbers (A005043).

If $c_{n,k}$ denotes the number of all planted trees of size $n$ with $k$ branches, then we have the following array

$$(c_{n,k})_{n\geq 1,k\geq 1}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 3 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 6 & 4 & 3 & 0 & 0 & \cdots \\ 1 & 0 & 10 & 10 & 5 & 6 & 0 & \cdots \\ 1 & 0 & 15 & 20 & 45 & 36 & 15 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The reversal array $(c_{n,k})_{n,n-k\geq 1}$ is the sequence A091867.

**Proposition 2.3.** *For $n\geq 1$, let $c_n$ denote the number of branches in all planted tree of size $n$. Then*

$$c_n=\binom{2n-2}{n-1}-\binom{2n-4}{n-2},$$

*and the generating function is $(1-z)zB(z)=z+z^2+4z^3+14z^4+50z^5+182z^6+\cdots$.*

The sequence $(1,4,14,50,182,\cdots)$ appears as A051924 in the OEIS [20].

## 3 Number of protected branches

In this section, we will study the number of protected branches in ordered trees, bushes and planted trees, respectively.

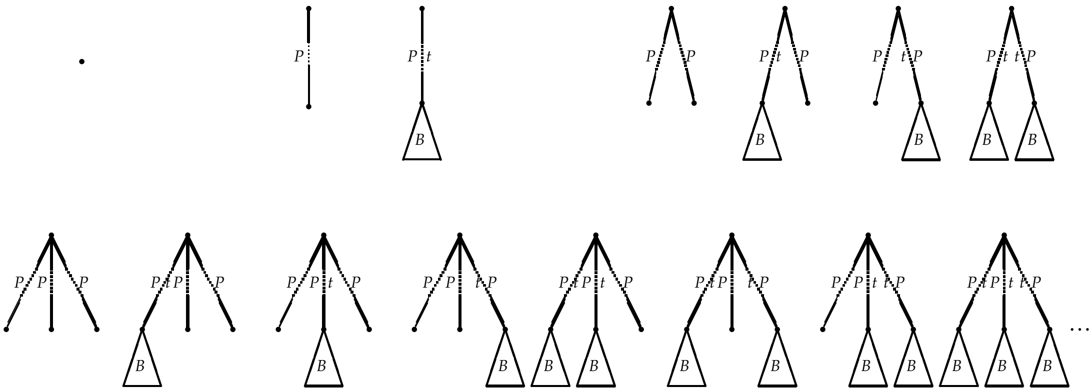


Figure 4: Marking protected branches.

**Theorem 3.1.** *Let $\bar{T}(t,z)$ be the generating function for the number of ordered trees of size n according to the number of protected branches, where z marks edges and t marks protected branches, and let $\bar{B}(t,z)$ be the analogous generating function of all bushes. Then*

$$\bar{B}=\frac{P^2(1+t\bar{B})^2}{1-P(1+t\bar{B})}, \tag{3.1}$$

$$\bar{T}=1-z(1-t)+2z(1-t)\bar{T}+tz\bar{T}^2, \tag{3.2}$$

*where $P(z)=\frac{z}{1-z}$.*

*Proof.* For any ordered tree, let $j$ be the degree of the root.

If $j=0$, then the tree has no edges, and its contribution to the generating function $\bar{T}(t,z)$ is 1.

If $j=1$, then there is only one path starting from the root, and the path connects to either the empty tree or a bush. If it is the empty tree, then its contribution to the generating functions is $P$. If it is a bush, then its contribution to the generating functions is $tP\bar{B}$. Hence, the contribution to the generating function $\bar{T}(t,z)$ is $P(1+t\bar{B})$.

If $j>1$, then there are $j$ paths starting from the root. If there exist $i$ $(0\le i\le j)$ paths connecting bushes, then its contribution to the generating functions is $P^j t^i \bar{B}^i$. Thus, we obtain that the contribution of the ordered trees whose root has degree $j$ is $P^j\sum_{i=0}^{j}\binom{j}{i}t^i\bar{B}^i=P^j(1+t\bar{B})^j$. We illustrate the generating functions for small values of $j$ in Figure 4.

Summing over all $j\ge 0$, we have

$$\bar{B}=\sum_{j=2}^{\infty}P^j(1+t\bar{B})^j, \tag{3.3}$$

$$\bar{T}=1+\sum_{j=1}^{\infty}P^j(1+t\bar{B})^j, \tag{3.4}$$

where $P(z)=\frac{z}{1-z}$. Hence, we obtain the equations (3.1) and (3.2). $\square$

**Proposition 3.1.** *We have*

$$\bar{T}(t,z)=\frac{1-2z+2tz-\sqrt{1-4z+4z^2-4tz^2}}{2tz},$$

*and the number $\bar{a}_{n,k}$ of ordered trees of size n with k protected branches is equal to*

$$\bar{a}_{n,k}=[t^kz^n]\bar{T}(t,z)=\frac{1}{n}\sum_{i=0}^{n}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\binom{2i-j}{n-1}\binom{n-i+j}{n-k}(-1)^{n+j-k}2^j.$$

*Proof.* Let $y(t,z)=\bar{T}(t,z)-1$. Then from (3.2), we have

$$y=z((t-1)+2(1-t)(y+1)+t(y+1)^2).$$

Using the Lagrange inversion formula [17] we obtain

$$\begin{aligned}[z^n]y(t,z)=&\frac{1}{n}[z^{n-1}]\left((t-1)+2(1-t)(z+1)+t(z+1)^2\right)^n\\=&\frac{1}{n}\sum_{i=0}^{n}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\binom{2i-j}{n-1}(-2)^j(t-1)^{n-i+j}t^{i-j}.\end{aligned}$$

Thus, we have

$$\begin{aligned}\bar{a}_{n,k}=&[t^kz^n]\bar{T}(t,z)=[t^kz^n]y(t,z)\\=&\frac{1}{n}\sum_{i=0}^{n}\sum_{j=0}^{i}\binom{n}{i}\binom{i}{j}\binom{2i-j}{n-1}\binom{n-i+j}{n-k}(-1)^{n+j-k}2^j.\end{aligned}$$

$\square$

**Remark 3.1.** The corresponding triangle formed by the coefficients of $\bar{T}(t,z)$ is

$$(\bar{a}_{n,k})_{n\geq 1,k\geq 0}=\begin{pmatrix}1&0&0&0&0&0&\cdots\\2&0&0&0&0&0&\cdots\\4&1&0&0&0&0&\cdots\\8&6&0&0&0&0&\cdots\\16&24&2&0&0&0&\cdots\\32&80&20&0&0&0&\cdots\\64&240&120&5&0&0&\cdots\\\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\ddots\end{pmatrix}.$$

This is denoted by A091894 in OEIS [20], is known as Touchard distribution [19, 22]. From [20], we know that $\bar{a}_{n,k}$ is the number of ordered trees on $n$ edges with $k$ prolific edges. A prolific edge is a edge whose child vertex has at least two children. For a given ordered tree, there is a bijection between the set of protected branches and the set of prolific edges: the last edge of a protected branch is a prolific edge, and each prolific edge belongs to only one protected branch. Baril, Kirgizov, and Vajnovszki [2] showed that $\bar{a}_{n,k}$ is also the number of length $n$ Catalan words with $k$ descents.

**Corollary 3.1.** *For $m \geq 0$, the number of ordered trees of size $2m+1$ with $m$ protected branches equals the mth Catalan number*

$$C_m = \frac{1}{m+1}\binom{2m}{m}.$$

*Proof.* Setting $n = 2m+1$ and $k = m$. □

**Proposition 3.2.** *For $n \geq 3$, let $\bar{a}_n$ be the number of protected branches in all ordered trees of size $n$. Then*

$$\bar{a}_n = \binom{2n-2}{n-3}.$$

*Proof.* Since $\bar{T}(1,z) = C(z)$, by partial differentiation for both sides of (3.2) with respect to $t$ gives

$$\frac{\partial \bar{T}(t,z)}{\partial t}\Big|_{t=1} = \frac{z(1-2C(z)+C(z)^2)}{1-2zC(z)} = \frac{z(1-C(z))^2}{1-2zC(z)} = z^3B(z)C(z)^4,$$

and hence

$$\bar{a}_n = [z^n]\frac{\partial \bar{T}(t,z)}{\partial t}\Big|_{t=1} = [z^n]z^3B(z)C(z)^4 = \binom{2n-2}{n-3}.$$

The generating function is $z^3B(z)C(z)^4 = z^3+6z^4+28z^5+120z^6+495z^7+2002z^8+\cdots$. □

**Remark 3.2.** The sequence

$$(\bar{a}_n)_{n\geq 3} = (1,6,28,120,495,2002,8008,31824,125970,497420,1961256,7726160,\cdots)$$

is denoted by A002694 in OEIS [20].

By Propositions 2.1 and 3.4, we know that the proportion of protected branches among all branches for all ordered trees of size $n$ is equal to

$$\frac{\bar{a}_n}{a_n} = \frac{\binom{2n-2}{n-3}}{\frac{3n^2-2n+1}{n^2+n}\binom{2n-2}{n-1}} = \frac{(n-2)(n-1)}{3n^2-2n+1}.$$

Hence , we have the following result.

**Theorem 3.2.** *The proportion of protected branches among all branches in all ordered trees of size $n$ approaches $1/3$ as $n \to \infty$.*

We have the average number of protected branches in all ordered trees of size $n$ is

$$\frac{\bar{a}_n}{C_n}=\frac{\binom{2n-2}{n-3}}{\frac{1}{n+1}\binom{2n}{n}}\sim\frac{n}{4}.$$

This means that, when $n$ is large, for every four new edges, there is about one protected branch.

**Proposition 3.3.** For $n\geq 2$, the number $\bar{b}_{n,k}$ of bushes of size $n$ with $k$ protected branches equals

$$\bar{b}_{n,k}=\frac{1}{k+1}\sum_{i=0}^{n}\binom{k+i}{i}\binom{n-1}{i+2k+1}\binom{i+2k+2}{k}.$$

*Proof.* By Theorem 3.1, we get

$$\bar{B}(t,z)=\frac{\frac{z^2}{(1-z)^2}(1+t\bar{B})^2}{1-\frac{z}{1-z}(1+t\bar{B})}.$$

Using the Lagrange inversion formula we know

$$\bar{B}(t,z)=\sum_{m=0}^{\infty}\frac{1}{m}\sum_{i=0}^{\infty}\binom{m+i-1}{i}\sum_{j=0}^{\infty}\binom{2m+i+j-1}{j}z^{2m+i+j}\binom{2m+i}{m-1}t^{m-1}.$$

It follows that

$$\bar{b}_{n,k}=[z^nt^k]\bar{B}(t,z)=\frac{1}{k+1}\sum_{i=0}^{n}\binom{k+i}{i}\binom{n-1}{i+2k+1}\binom{i+2k+2}{k}.$$

□

**Remark 3.3.** Arranging these coefficients in matrix form gives (this array do not appear in [20])

$$(\bar{b}_{n,k})_{n\geq 1,k\geq 0}=\begin{pmatrix} 0&0&0&0&0&0&0&0&\cdots\\ 1&0&0&0&0&0&0&0&\cdots\\ 3&0&0&0&0&0&0&0&\cdots\\ 7&2&0&0&0&0&0&0&\cdots\\ 15&13&0&0&0&0&0&0&\cdots\\ 31&54&5&0&0&0&0&0&\cdots\\ 63&183&51&0&0&0&0&0&\cdots\\ 127&552&308&14&0&0&0&0&\cdots\\ \vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\vdots&\ddots \end{pmatrix}.$$

**Corollary 3.2.** *For $m\geq 1$, the number of bushes of size $2m$ having $m-1$ protected branches equals the mth Catalan number $C_m=\frac{1}{m+1}\binom{2m}{m}$.*

*Proof.* Setting $n=2m$ and $k=m-1$. $\square$

**Proposition 3.4.** *For $n\geq 4$, let $\bar{b}_n$ be the number of protected branches in all bushes of size $n$. Then*

$$\bar{b}_n=\frac{3n-2}{n+1}\binom{2n-4}{n}.$$

*Proof.* Differentiating both sides of Eq. (2.6) with respect to $t$ and taking into account that $\bar{B}(1,z)=C(z)-zC(z)-1$ leads to

$$\frac{\partial\bar{B}(t,z)}{\partial t}|_{t=1}=\frac{2z^4C(z)^5-z^5C(z)^6}{1-2zC(z)}=2z^4B(z)C(z)^5-z^5B(z)C(z)^6.$$

Thus, we can obtain

$$\bar{b}_n=[z^n]\frac{\partial\bar{B}(t,z)}{\partial t}|_{t=1}=[z^n](2z^4B(z)C(z)^5-z^5B(z)C(z)^6)=\frac{3n-2}{n+1}\binom{2n-4}{n},$$

and the generating function

$$2z^4B(z)C(z)^5-z^5B(z)C(z)^6=2z^4+13z^5+64z^6+285z^7+\cdots.$$ $\square$

**Remark 3.4.** The sequence

$$(\bar{b}_n)_{n\geq 4}=(2,13,64,285,1210,5005,20384,82212,329460,1314610,\dots),$$

is denoted by A127531 in OEIS [20].

By Propositions 2.2 and 3.4,

$$\frac{\bar{b}_n}{b_n}=\frac{\frac{3n-2}{n+1}\binom{2n-4}{n}}{\frac{9n^2-15n+6}{n^2+n}\binom{2n-4}{n}}=\frac{n(3n-2)}{9n^2-15n+6}.$$

Hence , we have the following result.

**Theorem 3.3.** *The proportion of protected branches among all branches for all bushes of size $n$ approaches $1/3$ as $n\to\infty$.*

By Propositions 2.2 and 3.4, we have average number of protected branches in all bushes of size $n$ is

$$\frac{\bar{b}_n}{C_n-C_{n-1}}=\frac{\frac{3n-2}{n+1}\binom{2n-4}{n}}{\frac{3}{n+1}\binom{2n-2}{n}}\sim\frac{n}{4}.$$

Informally, when $n$ is large, for every four new edges, there is about one protected branch.

Let $\bar{Q}(t,z)$ be the generating function for the planted trees with respect to the size and the number of protected branches, i.e., the coefficient $[t^kz^n]Q(t,z)$ is the number of planted trees with $n$ edges and $k$ protected branches. Since any planted tree is either a path or a path connecting a bush, we have that $\bar{Q}(t,z)=P+tP\bar{B}(t,z)$. Applying (3.1),

$$\bar{Q}(t,z)^2(1-z+tz)-\bar{Q}(t,z)+z=0. \tag{3.5}$$

If $\bar{c}_{n,k}$ denotes the number of all planted trees of size $n$ with $k$ protected branches, then

$$\bar{c}_{n,k}=\frac{1}{n+1}\binom{n+1}{k}\sum_{j=0}^{n-2k}\binom{k+j-1}{k-1}\binom{n+1-k}{n-2k-j},$$

and we get the following array (A091156).

$$(\bar{c}_{n,k})_{n\geq 1,k\geq 0}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 4 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 11 & 2 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 26 & 15 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 57 & 69 & 5 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Corollary 3.3.** *Planted trees of size $2m+1$ with $m$ protected branches are enumerated by the mth Catalan number $C_m=\frac{1}{m+1}\binom{2m}{m}$.*

**Proposition 3.5.** *For $n\geq 3$, let $\bar{c}_n$ be the number of protected branches in all planted trees of size $n$. Then $\bar{c}_n=\binom{2n-4}{n-3}$, and the generating function is $z^3B(z)C(z)^2=z^3+4z^4+15z^5+56z^6+210z^7+792z^8+3003z^9+\cdots$ (A001791).*

By Propositions 2.3 and 3.5,

$$\frac{\bar{c}_n}{c_n}=\frac{\binom{2n-4}{n-3}}{\binom{2n-2}{n-1}-\binom{2n-4}{n-2}}=\frac{n-2}{3n-5},$$

$$\frac{\bar{c}_n}{C_{n-1}}=\frac{\binom{2n-4}{n-3}}{\frac{1}{n}\binom{2n-2}{n-1}}=\frac{n-2}{2n(2n-3)}\sim\frac{n}{4}.$$

Hence, we can conclude the following results.

**Theorem 3.4.** *The proportion of protected branches among all branches for all planted trees of size $n$ approaches $1/3$ as $n\to\infty$, and the average number of protected branches (=number of protected branches per planted tree ) is about $\frac{n}{4}$.*

# 4 Number of unprotected branches

In this section, we will study the number of unprotected branches. Let $\widetilde{T}(t,z)$ be the generating function for the number of ordered trees of size $n$ according to the number of unprotected branches, where $z$ marks edges and $t$ marks unprotected branches, and let $\widetilde{B}(t,z)$ be the analogous generating function of all bushes. By a similar argument to Theorem 3.1, we can obtain

$$\widetilde{B}=\sum_{j=2}^{\infty}P^j(t+\widetilde{B})^j, \tag{4.1}$$

$$\widetilde{T}=1+\sum_{j=1}^{\infty}P^j(t+\widetilde{B})^j, \tag{4.2}$$

where $P(z)=\frac{z}{1-z}$. Thus, we can get,

$$\widetilde{B}=\frac{P^2(t+\widetilde{B})^2}{1-P(t+\widetilde{B})}, \tag{4.3}$$

$$\widetilde{T}=1+tz\widetilde{T}+z\widetilde{T}(\widetilde{T}-1). \tag{4.4}$$

By (4.4), we know the number $\widetilde{a}_{n,k}$ of ordered trees of size $n$ having $k$ unprotected branches is equal to

$$\widetilde{a}_{n,k}=\frac{1}{n}\binom{n}{k}\binom{n}{k-1}.$$

This is the celebrated formula of Narayana number (A001263), which also is the number of ordered trees with $n$ edges and $k$ leaves [10].

$$(\widetilde{a}_{n,k})_{n\geq 1,k\geq 1}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 3 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 6 & 6 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 10 & 20 & 10 & 1 & 0 & 0 & \cdots \\ 1 & 15 & 50 & 50 & 5 & 1 & 0 & \cdots \\ 1 & 21 & 105 & 175 & 105 & 21 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Proposition 4.1.** *For $n\geq 1$, let $\widetilde{a}_n$ be the number of unprotected branches in all ordered trees of size n. Then $\widetilde{a}_n=\binom{2n-1}{n}$, and the generating function is $zB(z)C(z)=z+3z^2+10z^3+35z^4+126z^5+462z^6+\cdots$.*

*Proof.* Taking into account that $\widetilde{T}(1,z)=C(z)$, by partial differentiation for both sides of (3.4) with respect to $t$ gives

$$\frac{\partial\widetilde{T}(t,z)}{\partial t}|_{t=1}=\frac{zC(z)}{1-2zC(z)}=zB(z)C(z).$$

This gives us

$$\widetilde{a}_n = [z^n]\frac{\partial \widetilde{T}}{\partial t}|_{t=1} = [z^n]zB(z)C(z) = \binom{2n-1}{n}. \qquad \square$$

**Remark 4.1.** These numbers appear in OEIS [20] as the sequence A001700.

Since $(a_n)_{n\geq 1} = (\bar{a}_n)_{n\geq 1} + (\widetilde{a}_n)_{n\geq 1}$, by Proposition 2.1, Theorem 3.2 and Proposition 4.1, we get the following result.

**Corollary 4.1.** *The proportion of unprotected branches among all branches in all ordered trees of size $n$ approaches $2/3$ as $n\to\infty$.*

The average number of unprotected branches in all ordered trees of size $n$ is $\frac{\binom{2n-1}{n}}{\frac{1}{n+1}\binom{2n}{n}} \sim \frac{n}{2}$. Informally, when $n$ is large, for every two new edges, there is about one protected branch.

**Proposition 4.2.** *For $n,k\geq 2$, the number $\widetilde{b}_{n,k}$ of bushes of size $n$ with $k$ unprotected branches is given by*

$$\widetilde{b}_{n,k} = \sum_{i=0}^{k-3}\frac{1}{k-i-1}\binom{k-2}{i}\binom{n-1}{2k-i-3}\binom{2k-i-2}{k-i-2}.$$

*Proof.* From Eq. (3.3), we get

$$\widetilde{B}(t,z) = z\frac{\frac{z}{(1-z)^2}(t+\widetilde{B})^2}{1-\frac{z}{1-z}(t+\widetilde{B})}.$$

Thus, by the Lagrange inversion formula we can show that

$$\widetilde{B}(t,z) = \sum_{m=0}^{\infty}\frac{1}{m}\sum_{i=0}^{\infty}\binom{m+i-1}{i}\sum_{j=0}^{\infty}\binom{2m+i+j-1}{j}z^{2m+i+j}\binom{2m+i}{m-1}t^{m+i+1}.$$

It follows that

$$\widetilde{b}_{n,k} = [z^n t^k]\widetilde{B}(t,z) = \sum_{i=0}^{k-3}\frac{1}{k-i-1}\binom{k-2}{i}\binom{n-1}{2k-i-3}\binom{2k-i-2}{k-i-2}. \qquad \square$$

**Remark 4.2.** The above array is denoted by A119308 in OEIS [20]. The first few rows and columns are as follows

$$(\widetilde{b}_{n,k})_{n\geq 1,k\geq 1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 3 & 5 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 4 & 14 & 9 & 1 & 0 & 0 & \cdots \\ 0 & 5 & 30 & 40 & 14 & 1 & 0 & \cdots \\ 0 & 6 & 55 & 125 & 90 & 20 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Proposition 4.3.** *For $n \geq 2$, let $\widetilde{b}_n$ be the number of unprotected branches in all bushes of size $n$. Then*

$$\widetilde{b}_n = \frac{3n-2}{n}\binom{2n-3}{n-1}.$$

*Proof.* Differentiating both sides of Eq. (3.3) with respect to $t$ and taking into account that $\widetilde{B}(1,z) = C(z) - zC(z) - 1$ leads to

$$\frac{\partial \widetilde{B}(t,z)}{\partial t}|_{t=1} = \frac{z^2C(z)^2 + z^2C(z)}{1-2zC(z)} = z^2B(z)C(z)^2 + z^2B(z)C(z).$$

Hence,

$$\widetilde{b}_n = [z^n]\frac{\partial \widetilde{B}(t,z)}{\partial t}|_{t=1} = [z^n](z^2B(z)C(z)^2 + z^2B(z)C(z)) = \frac{3n-2}{n}\binom{2n-3}{n-1}$$

the generating function is

$$z^2B(z)C(z)^2 + z^2B(z)C(z) = 2z^2 + 7z^3 + 25z^4 + 91z^5 + 336z^6 + \cdots. \qquad \square$$

**Remark 4.3.** The sequence of the above numbers

$$(0,0,2,7,25,91,336,1254,4719,17875,68068,260338,\dots)$$

is denoted by A097613 in OEIS [20].

Since $(b_n)_{n\geq 1} = (\bar{b}_n)_{n\geq 1} + (\widetilde{b}_n)_{n\geq 1}$, by Proposition 2.2, Theorem 3.3 and Proposition 4.3, we get the following result.

**Corollary 4.2.** *The proportion of unprotected branches among all branches for all bushes of size $n$ approaches $2/3$ as $n \to \infty$.*

The average number of unprotected branches in all bushes of size $n$ is $\frac{\frac{3n-2}{n}\binom{2n-3}{n-1}}{\frac{3}{n+1}\binom{2n-3}{n-1}} \sim \frac{n}{2}$. Thus, when $n$ is large, 2 new edges get you about one unprotected branches.

Let $\widetilde{Q}(t,z)$ be the generating function for the planted trees with respect to the size and the number of unprotected branches, i.e., the coefficient $[t^kz^n]Q(t,z)$ is the number of planted trees with $n$ edges and $k$ unprotected branches. Since any planted tree is either a path or a path connecting a bush, we have that $\widetilde{Q}(t,z) = tP + P\widetilde{B}(t,z)$. Applying (4.3),

$$\widetilde{Q}(t,z)^2 - (1-z+tz)\widetilde{Q}(t,z) + tz = 0. \tag{4.5}$$

If $\widetilde{c}_{n,k}$ denotes the number of all planted trees of size $n$ with $k$ unprotected branches, then $\widetilde{c}_{n,k}=\frac{1}{n}\binom{n}{k}\binom{n-2}{k-1}$, and we have the following modified Narayana array .

$$(\widetilde{c}_{n,k})_{n\geq 1,k\geq 1}=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 3 & 1 & 0 & 0 & 0 & \cdots \\ 1 & 6 & 6 & 1 & 0 & 0 & \cdots \\ 1 & 10 & 20 & 10 & 1 & 0 & \cdots \\ 1 & 15 & 50 & 50 & 5 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

**Proposition 4.4.** *For $n\geq 2$, let $\widetilde{c}_n$ be the number of unprotected branches in all planted trees of size $n$. Then $\widetilde{c}_n=\frac{1}{2}\binom{2n-2}{n-1}$, and the generating function is $zB(z)(1-zC(z))=z+z^2+3z^3+10z^4+35z^5+126z^6+\cdots$ ( A088218).*

Since $(c_n)_{n\geq 1}=(\bar{c}_n)_{n\geq 1}+(\widetilde{c}_n)_{n\geq 1}$, by Proposition 2.3, Theorem 3.4 and Proposition 4.4, we get the following result.

**Corollary 4.3.** *The proportion of unprotected branches among all branches in all planted trees of size $n$ approaches $2/3$ as $n\to\infty$.*

## Acknowledgments

Our research was supported by the National Natural Science Foundation of China (Grant Nos. 11861045 and 12101280), and Gansu Province Science Foundation for Youths (Grant No. 20JR10RA187).

### References

[1] Baril J-L, Kirgizov S. The Pure Descent Statistic on Permutations, Preprint, 2016.
[2] Baril J-L, Kirgizov S, Vajnovszki V. Descent distribution on Catalan words avoiding a pattern of length at most three. Discrete Math., 2018, 341: 2608–2615.
[3] Bóna M. k-protected vertices in binary search trees. Adv. Appl. Math., 2014, 53: 1–11.
[4] Chen W Y C, Deutsch E, Elizalde S. Old and young leaves on plane trees. European J. Combin., 2006, 27: 414–427.
[5] Cheon G-S, Shapiro L W. Protected points in ordered trees. Appl. Math. Lett., 2008, 21: 516–520.
[6] Cheon G-S, Shapiro L W. The uplift principle for ordered trees. Appl. Math. Lett., 2012, 21: 1010–1015.
[7] Copenhaver K. k-protected vertices in unlabeled rooted plane trees. Graphs Combin., 2017, 33: 347-355

[8] Davenport D E, Shapiro L W, Pudwell L K, *et al.* The boundary of ordered trees. J. Integer Seq., 2015, 18: Article 15.5.8.

[9] Denise A, Simion R. Two combinatorial statistics on Dyck paths. Discrete Math., 1995, 137: 155–176.

[10] Dershowitz N, Zaks S. Enumerations of ordered trees. Discrete Math., 1980, 31: 9–28.

[11] Deutsch E, Shapiro L W. A bijection between ordered trees and 2-Motzkin paths and its many consequences. Discrete Math., 2002, 256: 655–6700

[12] Deutsch E. Ordered trees with prescribed root degrees, node degrees, and branch lengths. Discrete Math., 2004, 282: 89–94.

[13] Du R R X, Prodinger H. Notes on protected nodes in digital search trees. Appl. Math. Lett., 2012, 25: 1025–1028.

[14] Flajolet P, Sedgewick R. Analytic Combinatorics. Cambridge University Press, Cambridge, 2009.

[15] Heuberger C, Prodinger H. Protection number in plane trees. Appl. Anal. Discrete Math., 2017, 11: 314–326.

[16] Mansour T. Protected points in $k$-ary trees. Appl. Math. Lett., 2011, 24: 478–480.

[17] Merlini D, Sprugnoli R, Verri M C. Lagrange inversion: when and how. Acta Appl. Math., 2006, 94: 233–249.

[18] Riordan J. Enumeration of plane trees by branches and endpoints. J. Combin. Theory Ser. A, 1975, 19: 214–222.

[19] Shapiro L W. A short proof of an identity of Touchards concerning Catalan numbers. J. Combin. Theory Ser. A, 1976, 20: 375–376.

[20] Sloane N J A. On-Line Encyclopedia of integers sequences. Published electronically at http://oeis.org, 2019.

[21] Stanley R P. Enumerative Combinatorics(volume 2). Cambridge Univ. Press, Cambridge/New York, 1999.

[22] Touchard J. Sur certaines équations fonctionnelles. In: Proc. International Congress on Mathematics, University of Toronto Press, Toronto, 1928: 465–472.