Chin. Ann. of Math.
4B (4) 1983

SOME COMBINATORIAL PROBLEMS ON ORDERED TREES

WANG ZHENGYU (王振宇)
(*Wuhan Digital Engineering Institute*)

Abstract

Some formulas are established to calculate the number of leaves on all structurally different ordered trees with n nodes, and the total leaf path length and the total node path length for those trees. Certainly, respective average numbers are obtained (the average path length may be considered as the average height of random ordered trees with n nodes), Actually, this paper presents a general method dealing with some classes of combinaorial problems on ordered trees.

1. Introduction

The present paper aims at discussing some combinatorial problems on ordered trees with n nodes. The author proved:[1] if we denote the numder of all structurally different, ordered tress with n nodes by d_n $\left(\text{it is known that } d_n=b_{n-1}=\frac{1}{n}\binom{2n-2}{n-1}\right)$, then the total leaf number of d_n ordered trees with n nodes is $\frac{n}{2}$, $d_n=\frac{1}{2}\binom{2n-2}{n-1}$ for $n\geqslant 2$, that is to say, the average leaf number is $\frac{n}{2}$ for $n\geqslant 2$. For the purpose of our discussion, we will first give a new proof for this result, and then discuss other combinatorial problems such as leaf path length, node path length, etc.

Throughout the paper, all discussions are concerning with d_n different ordered trees with n nodes, and all these trees are alike.

For each node N of an ordered tree, there is a node path from the root of the tree to N, consisting of a sequence of edges. If N is a leaf of the tree, then the node path is also called a leaf path. If N is an internal node of the tree, then the node path is also called an internal node path. The number of edges on the path is calleb its length.

2. Leaf Number

Theorem 1. *Let L_n denote the total leaf number, then*

Manuscipt received November 6, 1981. Revised October 16, 1982.

$$L_n=\frac{1}{2}\binom{2n-2}{n-1} \quad if\ n\geqslant 2, \qquad (1)$$
$$L_1=1.$$

Proof An ordered tree with n nodes has one subtree at least, $n-1$ subtrees at most. Thus, for $n\geqslant 2$, we have the following recursive relation

$$L_n=L_{n-1}$$
the contribution of the trees with 1 subtree
$$+2\sum_{i,j\geqslant 1,\, i+j=n-1} d_i L_j$$
the contribution of the trees with 2 subtrees
$$+3\cdot\sum_{i_1,i_2,i_3\geqslant 1,\, i_1+i_2+i_3=n-1} d_{i_1}d_{i_2}L_{i_3}$$
the contribution of the trees with 3 subtrees
$$+\cdots$$
$$+r\cdot\sum_{i_1,i_2,\cdots,i_r\geqslant 1,\, i_1+i_2+\cdots+i_r=n-1} d_{i_1}d_{i_2}\cdots d_{i_{r-1}}L_{i_r}$$
the contribution of the trees with r subtrees
$$+\cdots$$
$$+(n-1)d_1^{n-2}L_1$$
the contribution of the tree with $n-1$ subtrees
$$=L_{n-1}+2\cdot\sum_{i\geqslant 0,\, j\geqslant 1,\, i+j=n-2} b_i L_j$$
$$+3\cdot\sum_{i_1,i_2\geqslant 0,\, i_3\geqslant 1,\, i_1+i_2+\cdots+i_3=n-3} b_{i_1}b_{i_2}L_{i_3}$$
$$+\cdots$$
$$+r\cdot\sum_{i_1,i_2,\cdots,i_{r-1}\geqslant 0,\, i_r\geqslant 1,\, i_1+i_2+\cdots+i_r=n-r} b_{i_1}b_{i_2}\cdots b_{i_{r-1}}L_{i_r}$$
$$+\cdots$$
$$+(n-1)b_0^{n-2}L_1, \qquad (2)$$
$$L_1=1. \qquad (3)$$

Let
$$\mathscr{B}(z)=\frac{1-\sqrt{1-4z}}{2z}=\sum_{n=0}^{\infty} b_n z^n=\sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}z^n$$
and
$$\mathscr{L}(z)=\sum_{n=1}^{\infty} L_n z^n,$$
then the generating function $\mathscr{L}(z)$ satisfies
$$\mathscr{L}(z)=z\mathscr{L}(z)+2z^2\mathscr{B}(z)\mathscr{L}(z)+3z^3\mathscr{B}(z)^2\mathscr{L}(z)$$
$$+\cdots+nz^n\mathscr{B}(z)^{n-1}\mathscr{L}(z)+\cdots+z$$
$$=z+z\mathscr{L}(z)\cdot\frac{1}{(1-z\mathscr{B}(z))^2}=z+\frac{4z}{(1+\sqrt{1-4z})^2}\mathscr{L}(z), \qquad (4)$$
therefore
$$\mathscr{L}(z)=\frac{z}{2}+\frac{z}{2}\frac{1}{\sqrt{1-4z}}. \qquad (5)$$

Thus we have (1), and the theorem is proved.

Corollary 1. *Let l_n denote the average leaf number, then*

$$l_n = \frac{n}{2} \quad \text{if } n \geqslant 2,$$
$$l_1 = 1.$$

3. Leaf Path Length

Theorem 2. *Let P_n denote the total leaf path length of all structurally different ordered trees with n nodes, then*

$$P_n = 4^{n-2} \quad \text{if } n \geqslant 2, \tag{6}$$
$$P_1 = 0.$$

Proof The proof is analogous to the proof of Theorem 1, except that (2) has to be replaced by

$$\begin{aligned} P_n = P_{n-1} &+ 2 \cdot \sum_{i,j \geqslant 1, i+j=n-1} d_i P_j \\ &+ 3 \cdot \sum_{i_1, i_2, i_3 \geqslant 1, i_1+i_2+i_3=n-1} d_{i_1} d_{i_2} P_{i_3} \\ &+ \cdots \\ &+ r \cdot \sum_{i_1, i_2, \cdots, i_r \geqslant 1, i_1+i_2+\cdots+i_r=n-1} d_{i_1} d_{i_2} \cdots d_{i_{r-1}} P_{i_r} \\ &+ \cdots \\ &+ (n-1) d_1^{n-2} P_1 \\ &+ L_n, \end{aligned}$$
$$P_1 = 0,$$

and (4) becomes

$$\mathscr{P}(z) = \frac{4z}{(1+\sqrt{1-4z})^2} \mathscr{P}(z) + \mathscr{L}(z) - z,$$

where $\mathscr{P}(z)$ is the generating function of P_n' s.

Corollary 2. *Let p_n be the average leaf path length, $p_n = P_n / L_n$. then*

$$p_n = 4^{n-2} \Big/ \frac{1}{2} \binom{2n-2}{n-1} \quad \text{if } n \geqslant 2,$$
$$p_1 = 0.$$

4. Node Path Length

Theorem 3. *Let T_n be the total node path length of all structurally different ordered trees with n nodes, then*

$$T_n = \frac{1}{2} \left(4^{n-1} - \binom{2n-2}{n-1} \right), \quad n \geqslant 1.$$

Proof The proof is also analogous to the proof of Theorem 1. In this case, (2) becomes

$$
\begin{aligned}
T_n = T_{n-1} &+ 2\cdot \sum_{i,j\ge 1,\, i+j=n-1} d_i T_j \\
&+ 3\cdot \sum_{i_1,i_2,i_3\ge 1,\, i_1+i_2+i_3=n-1} d_{i_1} d_{i_2} T_{i_3} \\
&+ \cdots + r\cdot \sum_{i_1,i_2,\cdots,i_r\ge 1,\, i_1+i_2+\cdots+i_r=n-1} d_{i_1} d_{i_2} \cdots d_{i_{r-1}} T_{i_r} \\
&+ \cdots + (n-1) d_1^{n-2} T_1 \\
&+ (n-1) d_n,
\end{aligned}
$$

($T_1=0$) and (4) becomes

$$\mathscr{T}(z) = \frac{4z}{(1+\sqrt{1-4z})^2} \mathscr{T}(z) + z^2 \mathscr{B}'(z),$$

where $\mathscr{T}(z)$ is the generating function of T_n' s.

Corollary 3. *Let t_n be the average node path length, then*

$$t_n = \frac{2\cdot 4^{n-2}}{\binom{2n-2}{n-1}} - \frac{1}{2}.$$

Let $E_n = T_n - P_n$, i. e., the total internal node path length, and $e_n = E_n/(nd_n - L_n)$, i. e., the average internal node path length, then we can verify the following interesting result.

Theorem 4.

$$e_n + 1 = p_n.$$

References

[1] Wang Zhenyu (王振宇), Some Properties of Ordered Trees, *Acta Mathematica Scientia*, **2** (1982), 81—83.

[2] Knuth, D. E., *The Art of Computer Programming*, Vol. 1, *Fundemental Algorithms*, Addison-Wesley, New York, 1973.