THE ENUMERATION OF ROOTED TREES BY TOTAL HEIGHT

JOHN RIORDAN and N. J. A. SLOANE

(Received 22 July 1968)

1. Introduction

The *height* (as in [3] and [4]) of a point in a rooted tree is the length of the path (that is, the number of lines in the path) from it to the root; the *total height* of a rooted tree is the sum of the heights of its points. The latter arises naturally in studies of random neural networks made by one of us (N.J.A.S.), where the enumeration of greatest interest is that of trees with all points distinctly labeled.

Wiite J_{ph} for the number of rooted trees with p labeled points and total height h, and

$$J(x, y) = \sum_{p=1}^{\infty} \sum_{h=0}^{\infty} J_{ph} \frac{x^p}{p!} y^h = \sum_{p=1}^{\infty} \frac{x^p}{p!} J_p(y)$$

for the enumerator (= enumerating generating function) of such trees by number of points and by total height. It will be shown that

$$J(x, y) = x \exp J(xy, y) \tag{1}$$

which is an analogue of George Pólya's familiar formula, [7],

$$R(x) = x \exp R(x),$$

with $R(x)$ the enumerator of rooted trees by number of labeled points; indeed $R(x) = J(x, 1)$. Also if

$$W_p = \sum hJ_{ph} = J'_p(1)$$

with the prime denoting a derivative (W_p/R_p is the mean total height of all rooted trees with p labeled points and also the mean height) the enumerator $W(x) = \sum W_p x^p/p!$ is found to be

$$W(x) = [xR'(x)]^2 \tag{2}$$

with $R'(x)$ the derivative of $R(x)$. From (2), it follows that

$$\begin{aligned} W_n &= n! \sum_{k=0}^{n-2} \frac{n^k}{k!}, \qquad n = 2, 3, \ldots \\ &\sim n^n \sqrt{\pi n/2}. \end{aligned} \tag{3}$$

https://doi.org/10.1017/S1446788700007527 Published online by Cambridge University Press

Since $R_n = n^{n-1}$, it follows that

$$\frac{W_n}{R_n} \sim n\sqrt{\pi n/2}. \tag{4}$$

For completeness, it may be noted that the enumerator $H(x, y, z)$ by number of points, by total height, and by number of points labeled, that is, the generating function having, as coefficient of $x^p y^h z^k/k!$, H_{phk} the number of rooted trees with p points, k of which are labeled (distinct labels), and with total height h, is given by

$$\begin{aligned} H(x, y, z) = x \exp[H(xy, y, z)+H(x^2y^2, y^2)/2+ \dots \\ +H(x^k y^k, y^k)/k+ \dots] \end{aligned} \tag{5}$$

with $H(x, y, 1) \equiv H(x, y)$.

For oriented trees, the results corresponding to (1) and (5) are

$$L(x, y) = x \exp 2L(xy, y) \tag{6}$$

$$\begin{aligned} K(x, y, z) = x \exp 2[K(xy, y, z)+K(x^2y^2, y^2)/2+ \dots \\ +K(x^k y^k, y^k)/k+ \dots]. \end{aligned} \tag{7}$$

2. Labeled rooted trees by total height

The simplest argument leading to (1) is but a slight modification of Pólya's [7] derivation of the formula for rooted trees (cf [5], p. 127 et seq). Write $J_n(y; m)$ for the enumerator by total height with n labeled points and m lines at the root. Then first

$$J_n(y; 1) = ny^{n-1}J_{n-1}(y); \tag{8}$$

for there are n possible labels for the root, and for each of these the height of any of the $n-1$ points of the tree connected to the single line at the root is increased by unity. Next

$$J_n(y; 2) = \frac{n}{2!}\sum_{k=1}^{n-2}\binom{n-1}{k} y^k J_k(y)\cdot y^{n-1-k}J_{n-1-k}(y). \tag{9}$$

The factor $2!$ on the right accounts for the permutations of the two lines at the root; the binomial coefficient in the sum accounts for the assignment of labels to the two trees connected to the two lines, and, as above, each point of a tree connected to a line at the root has its height increased by one. The symbolic version of (9) is

$$2!J_n(y; 2) = ny^{n-1}(J+J)^{n-1}, \qquad J^k \equiv J_k(y), \tag{9a}$$

and of course $J_0(y) = 0$. It is clear that

(10) $$m!J_n(y;m) = ny^{n-1}(J+\ldots+J)^{n-1}, \qquad J^k \equiv J_k(y),$$

with m terms in the symbolic sum on the right, and since

$$J_n(y) = \sum_{m=1}^{\infty} J_n(y;m),$$

it follows at once that

(1) $$J(x,y) = x \exp J(xy, y).$$

Rewriting (1) as

(1a) $$\begin{aligned} x^{-1}J(x,y) &= \sum_{n=1}^{\infty} \frac{x^{n-1}}{n!} J_n(y) \\ &= \exp\left[xyJ_1(y) + \frac{x^2y^2}{2!} J_2(y) + \ldots\right] \end{aligned}$$

and noting the equation of definition for the Bell multivariable polynomials $Y_n(y_1, y_2, \ldots, y_n)$ (cf. [5], equation (45) of chapter 2), namely

$$\exp uY(y_1, y_2, \ldots) = \exp\left[uy_1 + \frac{u^2y_2}{2!} + \ldots\right]$$

it follows that

$$J_{n+1}(y)/(n+1) = y^n Y_n(J_1(y), \ldots, J_n(y))$$

or

(11) $$J_{n+1}(y) = (n+1)y^n Y_n(J_1(y), \ldots, J_n(y)).$$

The first few instances of (11) are, omitting functional arguments,

$$\begin{aligned} J_1 &= Y_0 = 1 \\ J_2 &= 2yY_1(J_1) = 2y \\ J_3 &= 3y^2Y_2(J_1, J_2) = 3y^2(J_2+J_1^2) = 3y^2(1+2y) \\ J_4 &= 4y^3(J_3+3J_2J_1+J_1^3) = 4y^3(1+6y+3y^2+6y^3). \end{aligned}$$

Next, (partial) differentiation of (1) with respect to x, results in, using the suffix notation for partial derivatives,

(12) $$xJ_x(x,y) = J(x,y)[1+xJ_x(xy,y)]$$

which entails the recurrence

(13) $$nJ_n(y) = J_n(y) + \sum_{k=0}^{n-1} \binom{n}{k} ky^k J_k(y) J_{n-k}(y).$$

https://doi.org/10.1017/S1446788700007527 Published online by Cambridge University Press

Similarly

$$(14) \qquad J_y(x, y) = J(x, y) \left\{ \sum_{k=1} k y^{-1} \frac{(xy)^k}{k!} J_k(y) + \sum_{k=1} \frac{(xy)^k}{k!} J_k'(y) \right\}$$

so that

$$\begin{aligned} W(x) &= J_y(x, y)|_{y=1} \\ &= R(x)[xR'(x) + W(x)] \\ &= R(x)[1-R(x)]^{-1} xR'(x) \\ &= [xR'(x)]^2 \end{aligned}$$

which is (2). The identity $xR'(x) = R(x)[1-R(x)]^{-1}$ used in finding the last line follows from differentiating $R(x) = x \exp R(x)$.

Since

$$xR'(x) = \sum_{n=1}^{\infty} n^n \frac{x^n}{n!}$$

it follows from (2) that

$$(15) \qquad W_n = \sum_{k=1}^{n-1} \binom{n}{k} (n-k)^{n-k} k^k.$$

From the Cauchy formula (equation (1c) of [6]), associated with Abel's generalization of the binomial formula, (15) is readily reduced to the first equation of (3); the second of (3) follows from Stirling's formula and [8].

It is worth noting that the first of (3) implies relations of W_n with other graph entities. First, from equation (10) of [1], the number of connected graphs with n labeled points and n labeled lines, $C_{nn} = T_{n,n,1}$, is given by

$$C_{nn} = \tfrac{1}{2} n!(n-1)! \sum_{k=0}^{n-2} \frac{n^k}{k!}.$$

Hence

$$(n-1)!\, W_n = 2C_{nn}.$$

Next, the number of connected graphs with n labeled points is given in [2] by

$$C_n = (n-1)! \sum_{k=0}^{n-1} \frac{n^k}{k!} = n^{-1} W_n + n^{n-1}.$$

Hence

$$W_n = n(C_n - R_n).$$

3. More general enumeration

The results given in (5), (6), and (7) are straightforward consequences of the use of George Pólya's fundamental theorem in enumerative combi-

https://doi.org/10.1017/S1446788700007527 Published online by Cambridge University Press

natorial analysis. The essential modification of the procedure in [5] for enumeration of rooted trees without regard to total height, is that used above: a tree added to a line at the root occasions an increase by unity of the heights of each of its points.

References

[1] T. L. Austin, R. E. Fagen, W. F. Penney, and John Riordan, 'The number of components in random linear graphs', *Ann. Math. Statist.* 30 (1959), 747–754.

[2] Leo Katz, 'Probability of indecomposability of a random mapping function', *Ann. Math. Statist.* 26 (1955), 512–517.

[3] John Riordan, 'The enumeration of trees by height and diameter', *IBM Jnl. Res. Dev.* 4 (1960), 473–478.

[4] A. Rényi and G. Szekeres, 'On the height of trees', *J. Aust. Math. Soc.* 7 (1967), 497–507.

[5] John Riordan, *An Introduction to Combinatorial Analysis*, (Wiley, New York, 1958).

[6] H. Salié, 'Über Abel's Verallgemeinerung der binomischen Formel', *Ber. Verh. Säch. Akad. Wiss. Leipzig Math. — Nat. Kl.* 98, No. 4 (1951), 19–22.

[7] G. Pólya, 'Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen', *Acta Math.* 68 (1937), 145–253.

[8] G. N. Watson, 'Theorems stated by Ramanujan (V): Approximations connected with e^x, *Proc. London Math. Soc.* (2), 29 (1929), 293–308.

Rockefeller University
and
Cornell University