Scan

for

A6196

Kenny

(again)

6196

Information Processing Letters 25 (1987) 227–232
North-Holland

17 June 1987

A NOTE ON THE NUMBER OF LEFTIST TREES

Rainer KEMP

Fachbereich Informatik (20), Johann Wolfgang Goethe-Universität, Postfach 11 19 32, D-6000 Frankfurt am Main, Fed. Rep. Germany

Communicated by R. Wilhelm
Received 31 October 1986
Revised 16 December 1986

In this paper we show that the number t(n) of all leftist trees with n leaves is asymptotically given by $t(n) \sim \alpha c^n n^{-3/2}$, where $\alpha = 0.250\,363\,429\ldots$ and $c = 2.749\,487\,902\ldots$ are constants. This solves an open problem due to Knuth (1973, on p. 159).

Keywords: Data structures, discrete mathematics, combinatorial problems

1. Introduction

Let $T = (I, L, r)$ be an *extended binary tree* [3, p. 399] with I the set of internal nodes, L the set of leaves, and root $r \in I$. For any two nodes $u, v \in I \cup L$, the *distance* d(u, v) from u to v is defined by the number of nodes appearing in the shortest path from u to v in T. The *left branch length* LBL(T) of T is defined by $\mathrm{LBL}(T) := d(r, a)$, where $a \in L$ is the leftmost leaf of T. The subtree of T with root $x \in I \cup L$ is denoted by $T_x = (I_x, L_x, x)$.

The tree is said to be a *leftist tree* if

$$(\forall x \in I \cup L)\left(\min\{d(x, v) \mid v \in L_x\} = \mathrm{LBL}(T_x)\right);$$

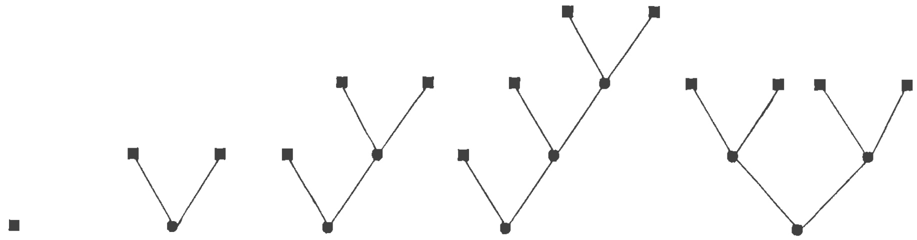
the tree T is called a *leftist tree of type* $\lambda \in \mathbb{N}$ if T is a leftist tree with the property $\mathrm{LBL}(T) = \lambda$. Leftist trees play an important part in particular sorting and merging algorithms based on selection; they involve an efficient way to represent priority queues as linked binary trees [4, pp. 150–152,159]. The leftist trees with less than five leaves are shown in Fig. 1.

Henceforth, let t(n) ($t_\lambda(n)$, $\lambda \in \mathbb{N}$) be the number of all leftist trees (of type λ) with n leaves. Obviously, we have

$$t(n) = \sum_{\lambda \geq 1} t_\lambda(n). \tag{1}$$

It is well known (see [4, p. 620]) that $t(n + m) \geq t(n)t(m)$, a fact that implies the existence of $\lim_{n \to \infty} t(n)^{1/n}$. In this note we derive a recursion formula for t(n) (Corollary 2.2) and show that $t(n) \sim \alpha c^n n^{-3/2}$, where

$$\alpha = 0.250\,363\,429\ldots \quad \text{and} \quad c = 2.749\,487\,902\ldots$$



n=1 n=2 n=3 n=4

Fig. 1. All leftist trees with less than five leaves.

Table 1
The exact and asymptotic number of leftist trees with n leaves

n	t(n) (exact)	t(n) (asymptotic)	n	t(n) (exact)	t(n) (asymptotic)
1	1	0.688	30	2.323 015E+10	2.293 207E+10
2	1	0.692	40	3.713 340E+14	3.677 504E+14
3	1	1.001	50	6.547 511E+18	4.496 902E+18
4	2	1.789	60	1.228 176E+23	1.220 260E+23
5	4	3.519	70	2.404 127E+27	2.390 839E+27
6	8	7.360	80	4.854 969E+31	4.831 484E+31
7	17	16.058	90	1.004 018E+36	9.997 003E+35
8	38	36.137	100	2.115 613E+40	2.107 423E+40
9	87	83.267	150	1.055 184E+62	1.052 459E+62
10	203	195.474	200	6.283 911E+83	6.271 733E+83
11	482	465.856	250	4.123 702E+105	4.117 304E+105
12	1 160	1 124.141	300	2.877 358E+127	2.873 634E+127
13	2 822	2 741.129	350	2.094 518E+149	2.092 193E+149
14	6 929	6 743.793	400	1.572 631E+171	1.571 102E+171
15	17 149	16 719.034	450	1.209 043E+193	1.207 998E+193
16	42 736	41 727.264	500	9.470 195E+214	9.462 814E+214
17	107 144	104 755.855	550	7.530 607E+236	7.525 263E+236
18	270 060	264 359.388	600	6.063 345E+258	6.059 394E+258
19	683 940	670 231.540	650	4.933 308E+280	4.930 336E+280
20	1 739 511	1 706 326.354	700	4.049 798E+302	4.047 529E+302

are constants (Theorem 2.3). Our method is similar to the one used by Otter [5] in order to derive the asymptotic behaviour of rooted trees of given order.

The first few values of t(n) are displayed in Table 1; the exact (asymptotic) values were computed by means of Corollary 2.2 (Theorem 2.3) presented in the next section.

2. The main result

First, let us derive an expression for the generating function

$$H(z) = \sum_{n \geqslant 1} t(n) z^n$$

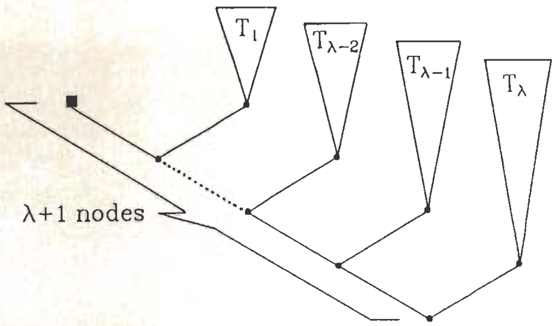


Fig. 2. The structure of leftist trees of type $\lambda+1$.

of the numbers t(n) of leftist trees with n leaves. For this purpose, we introduce the generating functions

$$T_\lambda(z) = \sum_{n\geqslant 1} t_\lambda(n) z^n, \quad \lambda \in \mathbb{N},$$

of the numbers $t_\lambda(n)$ of leftist trees with n leaves and left branch length λ. The following lemma presents a functional relation satisfied by H(z).

2.1. Lemma. *We have*

$$H(z) = z + \tfrac{1}{2}H^2(z) + \tfrac{1}{2}\sum_{\lambda\geqslant 1} T_\lambda^2(z).$$

where

$$T_1(z) = z, \qquad T_2(z) = z\,H(z), \qquad T_{\lambda+1}(z) = T_\lambda(z)\Big[H(z) - \sum_{1\leqslant i<\lambda} T_i(z)\Big], \quad \lambda > 1.$$

Proof. Obviously, the one-node tree is a leftist tree of type 1; thus $T_1(z) = z$. A leftist tree of type 2 has the structure drawn in Fig. 2 with $\lambda = 1$, where T_1 is an arbitrary leftist tree; therefore, $T_2(z) = z\,H(z)$.

A leftist tree of type $\lambda+1$, $\lambda > 1$, has the structure given in Fig. 2, where T_i is a leftist tree of type greater than or equal to i, $1 \leqslant i \leqslant \lambda$; hence,

$$T_{\lambda+1}(z) = z \prod_{1\leqslant i\leqslant\lambda} \sum_{r\geqslant i} T_r(z)$$

and therefore

$$T_{\lambda+1}(z) = T_\lambda(z)\sum_{r\geqslant\lambda} T_r(z) = T_\lambda(z)\Big[H(z) - \sum_{1\leqslant r<\lambda} T_r(z)\Big], \tag{2}$$

because, by (1),

$$H(z) = \sum_{i\geqslant 1} T_i(z).$$

Using this relation again together with $T_1(z) = z$ and $T_2(z) = zH(z)$, we further find

$$\begin{aligned} H(z) &= T_1(z) + T_2(z) + \sum_{\lambda \geqslant 2} T_{\lambda+1}(z) \\ &= z + zH(z) + \sum_{\lambda \geqslant 2} T_\lambda(z)H(z) - \sum_{\lambda \geqslant 2} T_\lambda(z) \sum_{1 \leqslant r < \lambda} T_r(z) \\ &= z + zH(z) + H(z)[H(z) - T_1(z)] - \frac{1}{2}\left[\left(\sum_{\lambda \geqslant 1} T_\lambda(z)\right)^2 - \sum_{\lambda \geqslant 1} T_\lambda^2(z)\right] \\ &= z + \tfrac{1}{2}H^2(z) + \tfrac{1}{2}\sum_{r \geqslant 1} T_r^2(z). \end{aligned}$$

This completes the proof of our lemma. □

Passing to the coefficients of $T_j(z)$, $j \in \mathbb{N}$, relation (2) immediately leads to the following recursive definition of $t(n)$.

2.2. Corollary. *The number* $t_\lambda(n)$ *of leftist trees of type* $\lambda \in \mathbb{N}$ *with* n *leaves is given by*

$$t_\lambda(1) = \delta_{\lambda,1}, \qquad t_1(n) = \delta_{n,1}, \qquad t_\lambda(n) = \sum_{1 \leqslant i < \lambda} t_{\lambda-1}(i) \sum_{\lambda - 1 \leqslant r \leqslant \lfloor \ell d(2(n-i)) \rfloor} t_r(n-i)$$

for $n \geqslant 2$, $1 \leqslant \lambda \leqslant \lfloor \ell d(2n) \rfloor$. *The number of leftist trees with* n *leaves is defined by means of* $t_\lambda(n)$ *according to* (1).

Note that the equation for $H(z)$ presented in Lemma 2.1 implies the 'symmetric' equation

$$2t(n) = \sum_{1 \leqslant i < n} t(i)t(n-i) + \sum_{\lambda \geqslant 1} \sum_{1 \leqslant i < n} t_\lambda(i) t_\lambda(n-i). \tag{3}$$

Now for the asymptotics. We essentially apply the method presented in [2, pp. 209–213]. For this purpose, we define the complex-valued function $F(z, y)$ by

$$F(z, y) = z + \tfrac{1}{2}y^2 - y + \tfrac{1}{2}\sum_{\lambda \geqslant 1} f_\lambda^2(z, y), \tag{4}$$

where

$$\begin{aligned} &f_1(z, y) = z, \qquad f_2(z, y) = zy, \\ &f_{\lambda+1}(z, y) = f_\lambda(z, y)\left[y - \sum_{1 \leqslant i < \lambda} f_i(z, y)\right], \quad \lambda > 1, \end{aligned} \tag{5}$$

and consider the equation $F(z, y) = 0$. By (1), we have $t_\lambda(n) \leqslant t(n)$ for all $\lambda, n \in \mathbb{N}$. Thus, by (3), analogous observations as in [2, pp. 209–213] show that $y = H(z)$ is the unique analytic solution of $F(z, y) = 0$ which has a singularity at $z = \eta \geqslant \frac{1}{4}$; the constants η and $H(\eta)$ are determined by the equations $F(\eta, H(\eta)) = 0$ and $F_y(\eta, H(\eta)) = 0$, where F_y stands for the partial derivative of F with respect to y. Thus, by Lemma 2.1, we have to solve the following system of nonlinear equations:

$$F(z, y) = z + \tfrac{1}{2}y^2 - y + \tfrac{1}{2}\sum_{\lambda \geqslant 1} f_\lambda^2(z, y) = 0, \tag{6a}$$

$$F_y(z, y) = y - 1 + \sum_{\lambda \geqslant 1} f_\lambda(z, y) h_\lambda(z, y) = 0, \tag{6b}$$

where $f_\lambda(z, y)$ is given by (5) and $h_\lambda(z, y) := (\partial/\partial y)f_\lambda(z, y)$ by

$$h_1(z, y) = 0, \qquad h_2(z, y) = z,$$

$$h_{\lambda+1}(z, y) = h_\lambda(z, y)\left[y - \sum_{1 \leqslant i < \lambda} f_i(z, y)\right] + f_\lambda(z, y)\left[1 - \sum_{1 \leqslant i < \lambda} h_i(z, y)\right], \quad \lambda > 1. \tag{7}$$

The solution of the system (6a), (6b) is the fixed point (z, y) of the function $A := \lim_{m \to \infty} A_m$ with

$$A_m(z, y) := \left(y - \tfrac{1}{2}y^2 - \tfrac{1}{2}\sum_{1 \leqslant \lambda \leqslant m} f_\lambda^2(z, y), \quad 1 - \sum_{1 \leqslant \lambda \leqslant m} f_\lambda(z, y)h_\lambda(z, y)\right), \quad m \geqslant 0,$$

which can be computed by the iteration formula $(z_{n+1}, y_{n+1}) = A_m(z_n, y_n)$ for well-chosen initial values (z_0, y_0), for example $(z_0, y_0) = (0, 0)$. We successively find

m	fixed point (z, y) of $A_m(z, y)$
0	(0.5, 1)
1	(0.414 213 562 373 095..., 1) = ($\sqrt{2} - 1$, 1)
2	(0.370 810 235 208 032..., 0.879 120 700 951 101...)
3	(0.363 856 547 050 700..., 0.825 686 491 896 708...)
4	(0.363 704 167 949 654..., 0.823 302 164 490 486...)
5	(0.363 704 091 587 335..., 0.823 299 737 553 244...)
$\geqslant 6$	(0.363 704 091 587 316..., 0.823 299 737 552 035...)

Thus,

$$(\eta, H(\eta)) = (0.363\,704\,091\,587\,316\ldots, 0.823\,299\,737\,552\,035\ldots).$$

Since, by (4),

$$F_{yy}(z, y) := \frac{\partial^2}{\partial y^2}F(z, y) = 1 + \sum_{\lambda \geqslant 1} h_\lambda^2(z, y) + \sum_{\lambda \geqslant 1} f_\lambda(z, y)g_\lambda(z, y),$$

where $f_\lambda(z, y)$ and $h_\lambda(z, y)$ are given by (5) and (7), respectively, and $g_\lambda(z, y) := (\partial/\partial y)h_\lambda(z, y)$ by

$$g_1(z, y) = 0, \qquad g_2(z, y) = 0,$$

$$g_{\lambda+1}(z, y) = g_\lambda(z, y)\left[y - \sum_{1 \leqslant i < \lambda} f_i(z, y)\right] + 2h_\lambda(z, y)\left[1 - \sum_{1 \leqslant i < \lambda} h_i(z, y)\right] - f_\lambda(z, y)\sum_{1 \leqslant i < \lambda} g_i(z, y), \quad \lambda > 1,$$

we find

$$F_{yy}(\eta, H(\eta)) = 1.492\,174\,401\,506\,950\ldots \neq 0. \tag{8}$$

Thus, by the theorem presented in [2, p. 211], the generating function H(z) has an expansion around η of the form

$$H(z) = H(\eta) + \sum_{k \geqslant 1} b_k(\eta - z)^{k/2}.$$

Working out the proof of this theorem (see also [1]), it is easily verified that $b_1 = -2\sqrt{\pi/\eta}\,a_1$ and

$$a_1 = \left(\eta\frac{F_z(\eta, H(\eta))}{2\pi F_{yy}(\eta, H(\eta))}\right)^{1/2}, \tag{9}$$

where F_z stands for the partial derivative of F with respect to z. By (4), we find

$$F_z(z, y) = 1 + \sum_{\lambda \geqslant 1} f_\lambda(z, y)\varphi_\lambda(z, y),$$

where $\varphi_\lambda(z, y) := (\partial/\partial z)f_\lambda(z, y)$ is recursively given by

$$\varphi_1(z, y) = 1, \qquad \varphi_2(z, y) = y,$$

$$\varphi_{\lambda+1}(z, y) = \varphi_\lambda(z, y)\left[y - \sum_{1 \leqslant i < \lambda} f_i(z, y)\right] - f_\lambda(z, y) \sum_{1 \leqslant i < \lambda} \varphi_i(z, y), \quad \lambda > 1.$$

Therefore, we obtain

$$F_z(\eta, H(\eta)) = 1.615\,820\,264\,188\,719\ldots$$

and finally, by (9),

$$a_1 = 0.250\,363\,429\,392\,523\ldots\ .$$

Hence, by the result stated in [2, p. 212], which is implied by the Darboux–Pòlya theorem [1], we have proved the following result.

2.3. Theorem. *The number of leftist trees* t(n) *satisfies*

$$t(n) = \alpha c^n n^{-3/2} + O(c^n n^{-5/2}),$$

where

$$\alpha := \alpha_1 = 0.250\,363\,429\ldots \quad \textit{and} \quad c := \eta^{-1} = 2.749\,487\,902\ldots$$

are constants.

References

[1] E.A. Bender, Asymptotic methods in enumeration, SIAM Rev. 16 (4) (1974) 485–515.

[2] F. Harary and E. Palmer, Graphical Enumeration (Academic Press, New York, 1973).

[3] D.E. Knuth, The Art of Computer Programming, Vol. 1 (Addison-Wesley, Reading, MA, 1973).

[4] D.E. Knuth, The Art of Computer Programming, Vol. 3 (Addison-Wesley, Reading, MA, 1973).

[5] R. Otter, The number of trees, Ann. Math. 49 (1948) 583–599.