REFERENCES

1. N. G. DeBruijn & B. J. M. Morselt. "A Note on Plane Trees." *J. Comb. Th.* 2 (1967):27-34.
2. R. Donaghey. Article to appear in *J. Comb. Th.*
3. Martin Gardner. "Mathematical Games." *Scientific American.*
4. D. Knuth. *Fundamental Algorithms.* Vol. I, 2nd ed. Reading: Addison-Wesley, 1973. Pp. 234-239.
5. N. O. Sloane. *A Handbook of Integer Sequences.* New York: Academic Press, 1975.

*****

PROFILE NUMBERS

ARNOLD L. ROSENBERG

IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598

ABSTRACT

We describe a family of numbers that arises in the study of balanced search trees and that enjoys several properties similar to those of the binomial coefficients.

1. INTRODUCTION

In the course of a recent investigation [4] concerning balanced search trees [2, Section 6.2.3], the following combinatorial problem arose. We encountered in the investigation a family $\{T_L\}$ of $(2L + 1)$-level binary trees, $L = 1, 2, \ldots$; the problem was to determine, as a function of L and $l \in \{0, 1, \ldots, 2L\}$, the number of nonleaf nodes at level l of the $(2L + 1)$-level tree T_L. (By convention, the root of T_L is at level 0, the root's two sons are at level 1, and so on.) The numbers solving this problem, which we call *profile numbers* since, fixing L, the numbers yield the *profile* of the tree T_L [3], that is, the number of nodes at each level of T_L, enjoy a number of features that are strikingly similar to properties of binomial coefficients. Foremost among these similarities are the generating recurrences and summation formulas of the two families of numbers. Let us denote by $P(n,k)$, $n \geq 1$ and $k \geq 0$, the number of nonleaf nodes at level k of the tree T_n, conventionally letting $P(n,k) = 0$ for all $k > 2n$; and let us denote by $C(n,k)$, $n \geq 1$ and $k \geq 0$, the binomial coefficient, conventionally letting $C(n,k) = 0$ for $k > n$. The well-known generating recurrence

$$C(n + 1,k + 1) = C(n,k + 1) + C(n,k), \quad k \geq 0$$

for the binomial coefficients is quite similar to the generating recurrence

$$P(n + 1,k + 1) = P(n,k) + 2P(n,k - 1), \quad k > 0 \tag{1}$$

for profile numbers. Further, the simple closed-form solution of the well-known summation

$$\sum_{0 \leq k < n} C(n,k) = 2^n - 1$$

for binomial coefficients corresponds to the equally simple solution of the summation

$$(2) \qquad \sum_{0 \le k < 2n} P(n,k) = 3^n - 1$$

for our new family of numbers. Further examples of relations between these two families of numbers will manifest themselves in the course of the development. As an aid to the reader, we close this introductory section with a portion of the triangle of numbers defined by the recurrence (1) with the boundary conditions

$$(3) \qquad \begin{array}{ll} P(n,0) = 1 & \text{for all } n \ge 1 \\ P(1,1) = 1 & \\ P(n,1) = 2 & \text{for all } n > 1 \\ P(1,k) = 0 & \text{for all } k > 1 \end{array}$$

(4)

$k \backslash n$	1	2	3	4	5	6	7	8	9	10
0	1	1	1	1	1	1	1	1	1	1
1	1	2	2	2	2	2	2	2	2	2
2		3	4	4	4	4	4	4	4	4
3		2	7	8	8	8	8	8	8	8
4			8	15	16	16	16	16	16	16
5			4	22	31	32	32	32	32	32
6				20	52	63	64	64	64	64
7				8	64	114	127	128	128	128
8					48	168	240	255	256	256
9					16	176	396	494	511	512
10						112	512	876	1004	1023

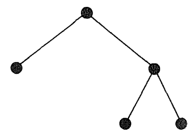
2. THE NUMBERS $P(n,k)$

Our goal in this section is to derive the basic properties of the profile numbers. By describing the tree-oriented origins of the numbers, we verify in Subsection A that they do indeed obey recurrence (1) with boundary conditions (3). We then proceed in Subsection B to solve recurrence (1), obtaining an explicit expression for $P(n,k)$ in terms of exponentials and binomial coefficients. In Subsection C, we derive the generating recurrences for individual rows and columns of the triangular array (4). These recurrences permit us in Subsection D to derive the summation formula (2) for profile numbers. Finally, in Subsection E, we use the summation formula to determine the so-called *internal path length* of the trees $\{T_L\}$, which determination was one of the motivations for studying the profile numbers. Our investigation will then have gone full circle.

In what follows, we shall refer often to binomial coefficients. These references will be very much facilitated by the convention $C(n,i) = 0$ whenever $i < 0$ or $i > n$, which should always be understood.

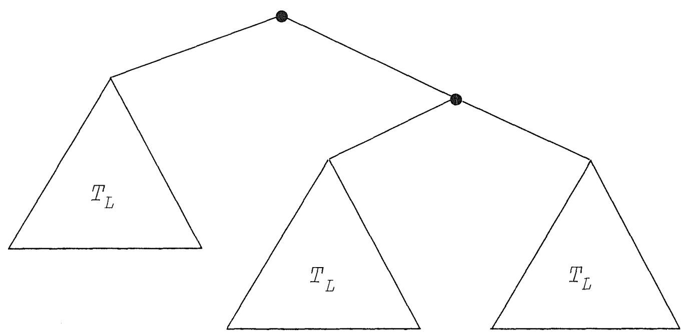
A. The Family $\{T_L\}$ of Trees: The trees T_L are specified recursively as follows. T_1 is the 3-leaf binary tree

T_1:



and, for each $L \geq 1$, the tree T_{L+1} is obtained by appending a copy of the tree T_L to each of the three leaves of T_1, as in

T_{L+1}:



The fact that $P(n,k)$ denotes, when $k \in \{0, \ldots, 2n\}$, the number of nonleaf nodes at level k of the tree T_n renders obvious the validity of recurrence (1) and boundary conditions (3) in addition to verifying the reasonableness of the convention

$$P(n,k) = 0 \quad \text{whenever} \quad n \geq 1 \quad \text{and} \quad k > 2n.$$

B. The Solution of Recurrence (1):

Theorem 1: For all $n \geq 1$ and all $k \geq 0$,

$$P(n,k) = 2^{k-n} \sum_{0 \leq i < 2n-k} C(n,i).$$

The theorem asserts, in particular, that $P(n,k) = 2^k$ for all $k < n$, and $P(n,k) = 0$ for all $k > 2n$.

Proof: We proceed by induction on n. The case $n = 1$ being validated by the boundary conditions (3), we assume for induction that the theorem holds for all $n < m$, and we consider an arbitrary number $P(m,k)$.

If $k \in \{0, 1\}$, then the boundary conditions (3) assure us that

$$P(m,k) = 2^k = 2^{k-n} \cdot 2^n = 2^{k-n} \sum_{0 \leq i < 2n-k} C(n,i),$$

which agrees with the theorem's assertion.

If $k > 1$, then recurrence (1) and the inductive hypothesis yield

$$\begin{aligned} P(m,k) &= P(m-1,k-1) + 2P(m-1,k-2) \\ &= 2^{k-m} \sum_{0 \leq i < 2m-k-1} C(m-1,i) + 2^{k-m} \sum_{0 \leq j < 2m-k} C(m-1,j) \end{aligned}$$

$$= 2^{k-m}C(m,0) + 2^{k-m}\sum_{0<i<2m-k} C(m-1,i) + C(m-1,i-1)$$

$$= 2^{k-m}\sum_{0\le i<2m-k} C(m,i),$$

which agrees with the theorem's assertion.

Since k was arbitrary, the induction is extended, and the theorem is proved. □

C. The Triangle of Profile Numbers: Yet more of the relation between profile numbers and binomial coefficients is discernible in the recurrences that generate individual rows and columns of the triangle (4).

Theorem 2: For all $n \ge 1$ and all $k \ge 0$,

(a) $P(n,k+1) = 2P(n,k) - 2^{k-n+1}C(n,k-n+1);$

(b) $P(n+1,k) = P(n,k) + 2^{k-n-1}\{C(n,k-n) + C(n+1,k-n)\}.$

Proof: Recurrence (1) translates to the three recurrences

(5) $$P(n,k) = P(n-1,k-1) + 2P(n-1,k-2).$$

(6) $$P(n,k+1) = P(n-1,k) + 2P(n-1,k-1).$$

(7) $$P(n+1,k) = P(n,k-1) + 2P(n,k-2).$$

Combining (5) and (6) leads, via Theorem 1, to the chain of equalities

$$\begin{aligned}
P(n,k+1) &- 2P(n,k)\\
&= P(n-1,k) - 4P(n-1,k-2)\\
&= 2^{k-n+1}\left\{\sum_{0\le i<2n-k-2} C(n-1,i) - \sum_{0\le i<2n-k} C(n-1,i)\right\}\\
&= -2^{k-n+1}\{C(n-1,2n-k-2) + C(n-1,2n-k-1)\}\\
&= -2^{k-n+1}C(n,k-n+1),
\end{aligned}$$

whence part (a) of the theorem.

Part (b) follows by direct calculation from recurrence (7) and Theorem 1:

$$\begin{aligned}
P(n+1,k) &- P(n,k)\\
&= P(n,k-1) + 2P(n,k-2) - P(n,k)\\
&= 2^{k-n-1}\left\{\sum_{0\le i<2n-k+1} C(n,i) + \sum_{0\le i<2n-k+2} C(n,i) - 2\sum_{0\le i<2n-k} C(n,i)\right\}\\
&= 2^{k-n-1}\{2C(n,2n-k) + C(n,2n-k+1)\}\\
&= 2^{k-n-1}\{C(n,k-n) + C(n+1,k-n)\}.\ \square
\end{aligned}$$

D. The Summation Formula (2): Theorem 2(a) permits easy verification of the summation formula for profile numbers.

Theorem 3: For all $n > 0$, $\sum_{0\le k<2n} P(n,k) = 3^n - 1.$

Proof: Theorems 1 and 2(a) justify the individual equalities in the following chain.

$$\sum_{0 \le k < 2n} P(n,k) = 1 + \sum_{0 \le k < 2n} P(n,k+1)$$

$$= 1 + 2\sum_{0 \le k < 2n} (P(n,k) - 2^{k-n}C(n,k-n+1)).$$

Thus, we have

$$\sum_{0 \le k < 2n} P(n,k) = \sum_{0 \le j \le 2n-1} 2^{n-j}C(n,j) - 1$$

$$= 2^n \cdot (3/2)^n - 1$$

$$= 3^n - 1. \square$$

E. The Internal Path Lengths of the Trees $\{T_L\}$: Theorems 2(a) and 3 greatly facilitate the determination of the *internal path length* [1, Section 2.3.4.5] $I(L)$ of the tree T_L of Section 2A, which is given by

$$I(L) = \sum_{0 \le k < 2L} kP(L,k).$$

Theorem 4: For all $L > 0$, $I(L) = \frac{5}{3}L3^L - 2(3^L - 1)$.

Proof: $$I(L) = \sum_{1 \le k < 2L+1} kP(L,k) = \sum_{0 \le k < 2L} (k+1)P(L,k+1)$$

$$= A + B + C + D$$

where

$$A = 2\sum_{0 \le k < 2L} kP(L,k) = 2I(L);$$

$$B = 2\sum_{0 \le k < 2L} P(L,k) = 2(3^L - 1);$$

$$C = -\sum_{0 \le k < 2L} 2^{k-L+1}C(L,k-L+1) = -3^L;$$

$$D = -\sum_{0 \le k < 2L} k2^{k-L+1}C(L,k-L+1)$$

$$= -\sum_{0 \le j \le L} (j+L-1)2^jC(L,j) = -5L3^{L-1} + 3^L.$$

Combining terms yields the theorem. $\square$

We close by remarking that the quantity $I(L)$ can be determined just as easily from the recurrence

$$I(L) = 3I(L-1) + \frac{5}{3}3^L - 4,$$

which is derived easily from the form of the trees $\{T_L\}$.

REFERENCES

1. D. E. Knuth. *The Art of Computer Programming I: Fundamental Algorithms*. Reading, Mass.: Addison-Wesley, 1968.
2. D. E. Knuth. *The Art of Computer Programming III: Sorting and Searching*. Reading, Mass.: Addison-Wesley, 1973.
3. R. E. Miller, N. Pippenger, A. L. Rosenberg, & L. Snyder, "Optimal 2,3-Trees." *SIAM J. Comput.* 8 (1979):42-59.
4. A. L. Rosenberg & L. Snyder. "Minimal-Comparison 2,3 Trees." *SIAM J. Comput.* 7 (1978):465-480.

*****

A STUDY OF THE MAXIMAL VALUES IN PASCAL'S QUADRINOMIAL TRIANGLE

CLAUDIA SMITH and VERNER E. HOGGATT, JR.
San Jose State University, San Jose, CA 95112

1. INTRODUCTION

In this paper we search for the generating function of the maximal values in Pascal's quadrinomial triangle. We challenge the reader to find this function as well as a general formula for obtaining all generating functions of the $(H - L)/k$ sequences obtained from partition sums in Pascal's quadrinomial triangle.

Generalized Pascal triangles arise from the multinomial coefficients obtained by the expansion of

$$(1 + x + x^2 + \cdots + x^{j-1})^n,\ j \geq 2,\ n \geq 0,$$

where n denotes the row in each triangle. For $j = 4$, the quadrinomial coefficients produce the following triangle:

1									
1	1	1	1						
1	2	3	4	3	2	1			
1	3	6	10	12	12	10	6	3	1

The partition sums are defined by

$$S(n,j,k,r) = \sum_{i=0}^{M} \left[\!\left[{n \atop r + ik} \right]\!\right]_j;\ 0 \leq r \leq k - 1,$$

where

$$M = \left[\frac{(j - 1)n - r}{k}\right];$$

the brackets [ ] denote the greatest integer function. To clarify, we give a numerical example. Consider $S(3,4,5,1)$. This denotes the partition sums in the third row of the quadrinomial triangle, in which every fifth element is added, beginning with the first column. Thus,

$$S(3,4,5,1) = 3 + 10 = 13.$$