Scan A3517
+
many
Shapiro
Catalan △

important paper/
add to several
seqs

f91

A108
~~A1118~~
A 2057
A 3517
A 3518
~~A39598~~
A 39598
A8315

Discrete Mathematics 14 (1976) 83–90.

A CATALAN TRIANGLE

L.W. SHAPIRO

Department of Mathematics, Howard University, Washington, D.C. 20001, USA

Received 6 August 1974
Revised 30 December 1974

We develop an arithmetic triangle similar to Pascal's triangle. The entries are interpreted in terms of numbers of pairs of nonintersecting paths in the first quadrant. The main applications are results about the Catalan numbers and various random walk problems.

1. Introduction

In this paper we consider lattice paths in the first quadrant and derive a triangle similar to Pascal's triangle that involves the Catalan numbers. We set up this Catalan triangle in Section 2. In Section 3 we examine the arithmetic properties of this triangle and in Section 4 we solve some random walk problems. In Section 5 two sequences derived from this triangle are discussed; one is a sequence used by Cayley [2] in a discussion of partitioning a polygon, the other arises in a paper of Fine [3] where he sets up an axiomatic theory of extrapolation.

We define the sequence $\{C_n\}_{n=1}^{\infty} = \{1, 2, 5, 14, 42, 132, 429, \ldots\}$ of Catalan numbers by

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

An extensive bibliography compiled by Gould [4] lists 243 references to the Catalan numbers including the famous problems of dividing a polygon into triangles, associating an n-product, and summing an even number of plus and minus ones so as to keep all partial sums nonnegative.

These results should be useful in the further study of arithmetic triangles, random walks, and the Catalan numbers themselves.

DM 14 84 76

2. Set up of the Catalan triangle

A *path* is a finite sequence of pairs $v_k = (a_k, b_k)$ of non-negative integers such that

(i) v_0 is $(0, 0)$;

(ii) if $v_k = (a_k, b_k)$, then $v_{k+1} = (1 + a_k, b_k)$ or $v_{k+1} = (a_k, 1 + b_k)$.

A path $(v_0, v_1, \ldots, v_n)$ is said to be of *length n*. The *distance* between $\{v_i\}_{i=0}^n = \{(a_i, b_i)\}_{i=0}^n$ and $\{w_i\}_{i=0}^n = \{(c_i, d_i)\}_{i=0}^n$ is $|a_n - c_n|$. We say two paths $\{v_i\}_{i=0}^n$ and $\{w_i\}_{i=0}^n$ *intersect* if $v_k = w_k$ for some $0 < k \leqslant n$.

We wish to find the number of pairs of non-intersecting paths. Observe that if two paths of length n have distance k, then this pair can be extended to four pairs of paths of length $n + 1$: one pair at distance $k + 1$, two pairs at distance k, and one pair at distance $k - 1$. Let B_{nk} denote the number of pairs of non-intersecting paths of length n and distance k. Tabulating the B_{nk} yields the triangle

n \ k	1	2	3	4	5	6	
1	1						...
2	2	1					...
3	5	4	1				...
4	14	14	6	1			...
5	42	48	27	8	1		...
6	132	165	110	44	10	1	...
	⋮	⋮	⋮	⋮	⋮	⋮	

Δ is A39598 also A8315 (read R to L)

~~1115~~

The recurrence relation derived from the observation above is $B_{nk} = B_{n-1,k-1} + 2B_{n-1,k} + B_{n-1,k+1}$ and the boundary conditions are $B_{n0} = 0 = B_{n,n+m}$, $m \geqslant 1$.

The first reason for calling this a Catalan triangle is that the numbers in the first column are indeed the Catalan numbers.

We now want a closed expression for B_{nk} and trial and error suggests $B_{nk} = (k/n)\binom{2n}{n-k}$.

Proposition 2.1. $B_{nk} = (k/n)\binom{2n}{n-k}$.

3518

2057 3517

Proof. We need only show that B_{nk} satisfies the recurrence relation and the boundary conditions.

$$B_{nk} = \frac{k}{n}\binom{2n}{n-k} = \frac{k-1}{n-1}\binom{2n-2}{n-k} + \frac{2k}{n-1}\binom{2n-2}{n-k-1} + \frac{k+1}{n-1}\binom{2n-2}{n-k-2}$$

$$= B_{n-1,k-1} + 2B_{n-1,k} + B_{n-1,k} \, .$$

The boundary conditions are obviously true.

A trivial corollary of this is that for integers $1 \leqslant k \leqslant n$ we always have that $(k/n)\binom{2n}{n-k}$ is an integer (see Birkhoff [1]).

To summarize we have $B_{nk} = (k/n)\binom{2n}{n-k}$ is the number of pairs of non-intersecting paths of length n and distance k. Thus

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{1}{n}\binom{2n}{n-1}$$

is both the number of pairs of non-intersecting paths of length n and distance 1 and also the number of pairs of paths that intersect for the first time (excluding $(0,0)$) at the $(n+1)$st step. Levine [5] presents an interesting alternate proof of this last fact.

3. Arithmetic properties

We want to evaluate $A_n = \Sigma_{k=1}^{n} B_{nk}$. To do this recall that each pair of non-intersecting paths extends to 4 pairs of paths of length n. Thus $A_n = 4A_{n-1} - C_{n-1}$, where C_{n-1} accounts for the paths of length $n-1$ that intersect at step n.

Proposition 3.1. $A_n = \frac{1}{2}\binom{2n}{n}$.

Proof. It suffices to show that this A_n satisfies the recurrence relation and to observe that $A_1 = 1$.

$$4A_{n-1} - C_{n-1} = 4\left(\frac{1}{2}\binom{2n-2}{n-1}\right) - \frac{1}{n}\binom{2n-2}{n-1} = \frac{2n-1}{n}\binom{2n-2}{n-1}$$

$$= \frac{1}{2}\binom{2n}{n} = A_n \, .$$

Though we won't require it, it is also true that $\Sigma_{m=0}^{n-k} B_{n,k+m} = \binom{2n-1}{n-k}$. The next proposition is analogous to the fact that $\binom{n+1}{k+1} = \Sigma_{j=0}^{n-k} \binom{k+j}{k}$.

Proposition 3.2. $B_{nk} = \Sigma_{j=1}^{n-k+1} C_j B_{n-j,k-1}$.

Proof. If two non-intersecting paths have distance k, then at some stage, say the $(n-j)$th, they must have distance $k-1$ for the last time. For the remaining j steps the two paths must never become closer than at the $(n-j+1)$st step. Thus the two paths may be continued in C_j ways and summing over j yields $B_{nk} = \Sigma_{j=1}^{n-k+1} C_j B_{n-j,k-1}$.

The same type of proof yields the more general result that

$$\sum_{m=0}^{r} B_{k+m,k} B_{l+r-m,l} = B_{k+l+r,k+l}.$$

Using Proposition 3.2 successively on the second, third, ... columns, yields the following version of the Catalan triangle:

$$\begin{array}{llll}
C_1 & & & \\
C_2 & C_1C_1 & & \\
C_3 & C_1C_2+C_2C_1 & \sum\limits_{i+j+k=3} C_iC_jC_k & \\
C_4 & C_1C_3+C_2C_2+C_3C_1 & \sum\limits_{i+j+k=4} C_iC_jC_k & \sum\limits_{i+j+k+l=4} C_iC_jC_kC_l \\
\vdots & \vdots & \vdots & \vdots
\end{array}$$

Here all indices run over positive integers. More formally we have:

Proposition 3.3. $B_{nk} = \Sigma_{i_1+i_2+\ldots+i_k=n} C_{i_1} C_{i_2} \cdots C_{i_k}$.

A converse relation between the B_{nk} and the C_i is given in the next proposition.

Proposition 3.4. $\Sigma_{k=1}^{\min(n,m)} B_{nk} B_{mk} = C_{n+m-1}$.

Proof. Consider all C_{n+m-1} pairs of paths which intersect for the first time after $n+m$ steps. At step n any such pair must have some distance k and the remaining m steps can be considered as a pair of non-intersecting paths of length m and also of distance k.

Corollary 3.5. $\Sigma_{k=1}^{n} (B_{nk})^2 = C_{2n-1}$ *and* $\Sigma_{k=1}^{n} B_{nk} B_{n+1,k} = C_{2n}$. *Thus*

$$\sum_{k=1}^{n} \Big(\sum_{i_1+i_2+\ldots+i_k} C_{i_1} C_{i_2} \ldots C_{i_k} \Big)^2 = C_{2n-1} \, .$$

4. Some random walk problems

We can use the results in the preceding sections to do some random walk problems. Assume for instance that a cop and a robber both start at (0, 0) and both walk in the first quadrant. At each lattice point they each flip a fair coin to determine if they should then proceed north or east to the next lattice point. If both the cop and robber move simultaneously, what is the chance that they will meet again after leaving (0,0)?

Since both the cop and the robber have available 2^n equal likely paths of length n, there are 4^n pairs of paths of length n, $2(\frac{1}{2}\binom{2n}{n})$ of which do not meet. The extra factor of 2 accounts for the fact that we have labeled the two paths. However

$$\lim_{n\to\infty} \frac{\binom{2n}{n}}{4^n} = \lim_{n\to\infty} \frac{\sqrt{4\pi n}(2n\,\mathrm{e}^{-1})^{2n}}{(\sqrt{2\pi n}(n\,\mathrm{e}^{-1})^n)^2 \cdot 4^n} = \lim_{n\to\infty} \frac{1}{\sqrt{\pi n}} = 0 \, .$$

We have used Stirling's formula here. Thus the unfortunate robber has probability 1 of meeting the cop again.

What if we change the problem to where they start at different locations but both initial locations are n steps from (0, 0)? Let us call their initial locations I_c and I_r and the probability of not meeting after starting from I_c and I_r we will call p_2. Let p_1 be the probability of starting at (0, 0) and going by non-intersecting paths to I_c and I_r. Obviously $p_1 > 0$. Then $p_1 p_2 \leqslant$ chance of two paths starting at (0, 0) and not intersecting, so $p_1 p_2 = 0$ and thus $p_2 = 0$.

Similar reasoning says that if we allow the robber a few steps lead but now ask that the cop (now equipped with a bloodhound) cross the robber's trail, then we obtain the same results.

We can pass to the classical theory of random walks as follows. Let a particle start at the origin and move one step east or west according as the cop moves east or north. Simultaneously the particle moves one step south or north depending on the move of the robber east or north. A return to the origin of this particle is equivalent to the cop and the robber meeting at a point (k, k). Thus by the classical theory of random walks the cop and robber must meet infinitely often at points on the diagonal $\{(k, k)\}$.

5. Two sequences

The second column of the Catalan triangle yields the sequence

$$\left\{\frac{2}{n}\binom{2n}{n-2}\right\}_{n=2}^{\infty} = \{1, 4, 14, 48, 165, 572, 2002, \ldots\} .$$

These numbers appear in a paper of Cayley's [2] where he discusses dissecting polygons into smaller polygons by adding non-intersecting diagonals. Simplifying Cayley's computations leads to

$$(1 + 2x + 5x^2 + 14x^3 + 42x^4 + \ldots)^2 = 1 + 4x + 14x^2 + 48x^3 + 165x^4 + \ldots .$$

The proof of this is immediate using Proposition 3.2. Similarly $(1 + 2x + 5x^2 + 14x^3 + 42x^4 + \ldots)^n$ is the generating function of the nth column of the Catalan triangle. Thus we have given a geometric interpretation to these numbers of Cayley as well as their closed form.

Fine [3], in a paper where he sets up an abstract theory of extrapolation, develops the number sequence $\{0, 1, 2, 6, 18, 57, 186, 622, 2120, 7338, \ldots\}$ which he computes recursively. Consider the set, T, of real numbers $x_1 < x_2 < x_3 < \ldots < x_n$. A *similarity relation* $\sim$ on T is a reflexive, symmetric relation where if $x_a < x_b < x_c$ and $x_a \sim x_c$, then $x_a \sim x_b$ and $x_b \sim x_c$. These requirements all seem reasonable; if a is similar to b, then b should be similar to a, certainly every element should be similar to itself, and if two items are similar, then every element between them should be similar to both.

We can now ask various questions. How many such similar relations are there? What if the relation must also be transitive? What if we require each element to be similar to some other element than itself? What if we require this last property together with transitivity? The respective answers are the Catalan numbers, $\{2^n\}$, Fine's sequence, and the Fibonacci numbers.

The reason for requiring each element to be similar to some other element is to avoid extrapolating with no similar data. Another way that Fine's sequence shows up is by adding diagonals in the Catalan triangle as follows.

1	1				
2	2	1			
6	5	4	1		
18	14	14	6	1	
57	42	48	27	8	1

This of course is analogous to a procedure for developing the Fibonacci numbers.

$$D_n = \sum_{j=1}^{[n+1/2]} B_{n+1-j,j}\,.$$

If we call the nth number of Fine's sequence D_n, then we have $2D_n + D_{n-1} = C_{n+1}$. Solving this in terms of C_n yields

$$D_n = -\sum_{j=1}^{n} (-2)^{-j} C_{n+2-j}\,.$$

This gives as a corollary that

$$2^n \;\Big|\; \sum_{j=1}^{n} C_{n+2-j}\,(-2)^{n-j}\,.$$

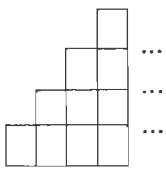
The proof that Fine's sequence is identical to this Fibonacci–Catalan sequence is complicated and not included here. The details are available from the author.

We conclude with a list of related open questions.

(1) When does $B_{nk} = B_{n,k+l}$ other than when $n = 2k(k+1)$?

(2) Do other lattices yield similar results?

(3) If we used the following lattice what results would be obtained?



(4) What happens if we change to 3-space? In particular must the cop and robber meet again with probability 1?

(5) Is there a theory of arithmetic triangles where a simple function of the generating function of the first column yields the generating function of the nth column?

(6) What happens if biased coins are used in the cop and robber problem?

(7) When examining similarity relations what happens if we require that each element be similar to at least 2 (or k) other elements?

(8) Is there a simple interpretation of D_n in terms of pairs of paths?

(9) Would we expect three paths in the northeast quadrant to have a common intersection? (If we translate as in Section 4 to the classical

theory of random walks, we find that the chance of their meeting again on the diagonal (k, k) is about 0.239. Dr. J. Komlos (unpublished) has proved that the probability of the 3 paths meeting infinitely often is 1. He has also proved that for 4 or more paths the chance of a common intersection is less than one.)

We can ask similar questions about k paths in n dimensions with various boundaries.

One interesting way many of these problems can be rephrased is in terms of dyadic expansions. For instance, is it true that any three dyadic expansions will, for some n, have the same number of 1's among the first n digits, where each digit is chosen randomly? This problem is equivalent to (9).

Many of the above results can be rephrased in terms of matrix multiplication. We conclude by giving one such example.

$$\begin{bmatrix} B_{11} & 0 & 0 & \cdots \\ B_{21} & B_{22} & 0 & \cdots \\ B_{31} & B_{32} & B_{33} & \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} B_{11} & B_{21} & B_{31} \\ 0 & B_{22} & B_{32} \\ 0 & 0 & B_{33} \\ \vdots & \vdots & \vdots \end{bmatrix} = \begin{bmatrix} C_1 & C_2 & C_3 & \cdots \\ C_2 & C_3 & C_4 & \cdots \\ C_3 & C_4 & C_5 & \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

References

[1] G. Birkhoff, Problem 3674 (proposed), Am. Math. Monthly 4 (1934).
[2] A. Cayley, On the partitions of a polygon, Proc. London Math. Soc. 22 (1891) 237–262.
[3] T. Fine, Extrapolation when little is known, Information and Control 16 (1970) 331–360.
[4] H. Gould, Research bibliography of two special number sequences, Mathematica Monongaliae, Number 12 (1971).
[5] J. Levine, Note on the number of pairs of nonintersecting routes, Scripta Math. 24 (1959) 335–338.