Math 4707
The Catalan Nunbers

Dennis White

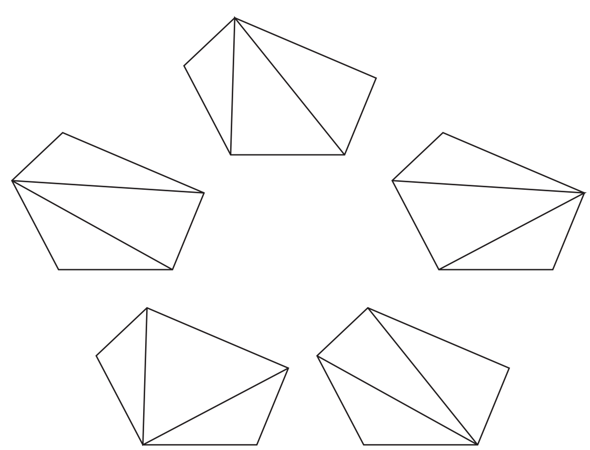
November 3, 2010

1 Introduction

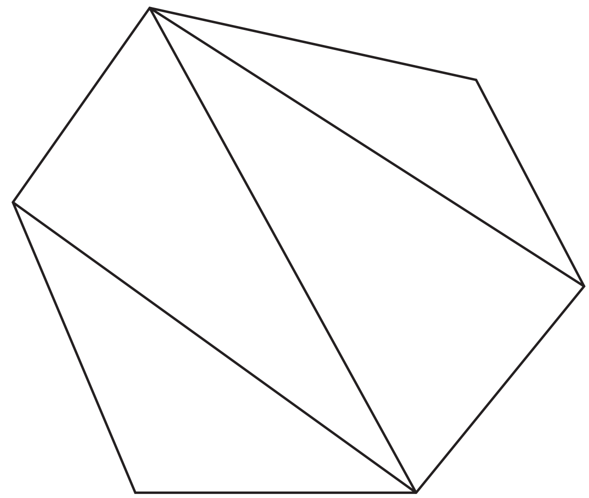
The Catalan numbers are a remarkable sequence of numbers that "solve" a number of seemingly unrelated counting problems. Richard Stanley's graduate textbook in combinatorics, *Enumerative Combinatorics*, contains a collection of 66 different sequences, all counted by the Catalan numbers. We will introduce four of these sequences here.

2 Some Counting Problems

How many ways are there to triangulate a polygon with n sides? That is, how many ways are there to draw non-intersecting diagonals so that the interior of the polygon is partitioned into triangles. For example, here are all the triangulations of the pentagon.



Here is one triangulation of a hexagon.



You should verify that there are 14 triangulations of the hexagon.

Next, how many ways are there for $2n$ friends seated around a round table to all shake hands, without crossing handshakes. For example, if $n = 3$ and the friends are A, B, C, D, E and F, seated around the table in that order, then there are five ways to do this:

- {AB}, {CD}, {EF}
- {AB}, {CF}, {DE}
- {AD}, {BC}, {EF}
- {AF}, {BC}, {DE}
- {AF}, {BE}, {CD}

Here is an example with $n = 4$ and the friends are A, B, C, D, E, F, G, H: {AD}, {BC}, {EF}, {GH}. You should verify that there are 14 ways to do this if $n = 4$.

Next, suppose Bill lives n blocks south and n blocks west from where he works. Suppose a railroad track runs diagonally from southwest to northeast, from just southeast of his home to just northeast of his workplace. We now declare that Bill must either walk north or east, and he cannot cross the railroad tracks. If $n = 3$, he has five choices of routes from home to work: NNNEEE, NNENEE, NNEENE, NENNEE and NENENE.

If $n = 4$, here is one possible route: NNEENENE. Again, verify that when $n = 4$, there are 14 different routes.

Note that the restriction that he not cross the railroad tracks implies that at any point in the N-E "word", the number of E's can be no more than the number of N's.

Finally, we count the number of possible outline structures with n headings at various levels. For example, when $n = 3$, we have these five possibilities:

I II III	I A II	I II A
I A B	I A 1	

When $n = 4$, here is an example:

I
  A
II
III

Once again, verify that when $n = 4$ there are 14 structures possible.

The remainder of these notes will solve three problems: show that all these sequences are the same, find a recursion for this sequence, and find an explicit formula for this sequence.

3 The Same Sequence

One way to show that the sequences are the same is to give bijections (one-to-one correspondences) between the objects in question. Another way is to show each sequence satisfies the same recurrence formula and initial condition. We will use a mix of these two methods, but you should be able to give bijections between any pair of these four objects, or prove the recurrence for each of the four examples.

We will first give a bijection between outlines and blockwalks. Our bijection is constructed recursively. Suppose we have a bijection between blockwalks and outlines for all values of $0 \leq k < n$. Suppose B is a blockwalk with $2n$ steps. It starts along the diagonal, ends on the diagonal and may revisit the diagonal at various points in between. These visits to the diagonal will occur exactly when the number of N's equals the number of E's. Between these visits, B does not touch the diagonal. In fact, we partition B into sections which lie between these visits. Furthermore, if we remove the initial N and final E of each of these sections, we will have a smaller blockwalk. Let's list these smaller blockwalks, B_1, B_2, ..., B_m. There will then be m major headings in our outline. Now each of these smaller blockwalks B_i will correspond, recursively, to the minor headings under each major heading. This construction is clearly reversible.

Here is an example. Suppose the blockwalk is NENNNENEENEENNEE. Then the segments between visits to the diagonal are NE, NNNENEENEE, and NNEE. Removing the initial N and final E on each gives $\emptyset$, NNENEENE, and NE. The outline structure then has three major headings, I, II and III. Working recursively, the first is simply I, the second has these subheadings: II, II-A, II-A-1, II-A-2, and II-B, and the third is III, III-A. Notice that there are 8 headings, and the blockwalk sequence had 16 steps.

Next, we give a bijection between blockwalks and handshakes. As before, we work recursively. We decompose B just as described above. Suppose the length of blockwalk B_i is $2k_i$, where $0 \leq k_i < n$. Now number the $2n$ people around the table by 1, 2, ..., $2n$. Then 1 will shake hands with $2(k_1+1)$, $2(k_1+1)+1$ will shake hands with $2(k_1+k_2+2)$, $2(k_1+k_2+2)+1$ will shake hands with $2(k_1+k_2+k_3+3)$, and so on. In between these handshakes, the handshakes are assigned recursively.

In the example above, $k_1 = 0$, $k_2 = 4$ and $k_3 = 1$. Therefore, 1 shakes with 2, 3 with 12 and 13 with 16. Then, working recursively, 4 shakes with 9, 5 with 6, 7 with 8, 10 with 11, and 14 with 15.

Again, the construction reverses naturally.

4 The Recursion

We now know that the same number counts handshakes, blockwalks and outlines. Let's let C_n denote the number of handshakes involving $2n$ people (or blockwalks with $2n$ steps or outlines with n headings). We concentrate on the handshake model. Again, assume the people are numbered from 1 to $2n$ around the table, clockwise. We can decompose the handshakes according to whom person 1 shakes with. If person 1 shakes with person 2, then the remaining $2n-2$ people can shake in C_{n-1} ways. If person 1 shakes with person 4, then persons 2 and 3 can shake in $C_1 = 1$ ways, and the remaining $2n-4$ people can shake in C_{n-2} ways. If person 1 shakes with person $2k+2$, then the $2k$ people to person 1's left can shake in C_k ways and the $2(n-k-1)$ people to person 1's right can shake in C_{n-k-1} ways, for $k = 0, 1, \ldots, n-1$ (as long as we define $C_0 = 1$).

By the way, why does person 1 always shake with an even numbered person?

Theorem 1.

$$C_n = \sum_{k=0}^{n-1} C_k C_{n-1-k}$$

Starting with $C_0 = 1$, this theorem allows us to compute C_n. For example, $C_1 = 1$, $C_2 = 2$, $C_3 = 5$, $C_4 = 14$, and $C_5 = 42$. You should verify these numbers.

But what about the triangulations of the polygon? One confusion in this problem is the relationship between the parameter of the Catalan number and the number of sides of the polygon. Rather than associate the number of sides with the Catalan number, it is better to associate with the number of triangles.

Theorem 2. *The number of triangulations of a polygon with $n+2$ sides into n triangles is C_n.*

Proof. We prove this be showing that triagulations satisfy the Catalan recursion. Since the initial values are the same as the Catalan numbers, the sequences must be the same.

As with the handshakes, we must divide a triangulation into two pieces. This is accomplished as follows. Fix a side of the polygon, calling it the *base*. The base must be one side of a triangle in any triangulation. Call that triangle the *base triangle*. The base triangle divides the triangulation into two pieces. Clockwise from the base will be a polygon defined by the polygon sides and one non-base side of the base triangle. Suppose this polygon has $k+2$ sides (and therefore k triangles in its triangulation). Counterclockwise from the base will be a polygon defined by the polygon sides and the other non-base side of the base triangle. This polygon must have $n-k+1$ sides and $n-k-1$ triangles. Since k can vary from 0 to $n-1$ (where, in the extreme two cases, the base triangle uses two polygon edges), the recurrence follows. □

5 Block Walking and the Explicit Formula

Before we proceed to the explicit formula, let's look at the general problem of "block walking" Suppose there are n north-south blocks and m east-west blocks in a rectangular street system. How many paths are there from the southwest corner to the northeast corner, where the path proceeds either north or east? Notice that we've dropped the railroad track condition. The answer is simply $\binom{m+n}{n}$, because any sequence of n N's and m E's would constitute such a blockwalk. And we know that the number of "words" using the two letters N and E, with n N's and m E's is this binomial coefficient.

Now for the explicit formula. The approach is through a device called the "reflection principle". Instead of counting the blockwalks directly, we will count the number of blockwalks which do cross the railroad tracks, and then subtract from the total number of blockwalks.

Let B be a blockwalk which crosses the railroad track, and let B_i denote the letter (N or E) in the ith step of the walk. Any blockwalk which crosses the tracks must, at some point, have the number of E's exceed the number of N's. Let m be the first point at which this happens, that is B_m is E. Let A denote the first m steps of B, $(B_1, B_2, \ldots, B_m)$. Then A has exactly one more E than N. It follows that m is odd, say $m = 2k+1$, and A has k N's and $k+1$ E's. Now construct a new blockwalk B' by changing all N's to E's and vice versa in A, but leaving the remaining steps in B, $B_{m+1}, \ldots, B_{2n}$, unchanged. Then B' will still have $2n$ steps, but $n+1$ of them will be N's and $n-1$ of them will be E's. This is because k N's became E's and $k+1$ E's became N's.

This describes a method of transforming a blockwalk B with n N's and n E's which crosses the railroad tracks into B' which has $n+1$ N's and $n-1$ E's. But this process is reversible!

Start with a blockwalk B' with $n+1$ N's and $n-1$ E's. Since the total number of N's is greater than the total number of E's, at some point along this blockwalk, the number of N's must first exceed the number of E's. Let A' denote the portion of the blockwalk up to and including this point. Now change all N's to E's and E's to N's in A', but leave the rest of B' unchanged, to form a new blockwalk B. Then B will have n N's, n E's, and will have the number of E's first exceed the number of N's at the same point.

Therefore, the "bad" blockwalks are in one-to-one correspondence with all blockwalks in a $(n-1)\times(n+1)$ grid. Subtracting the bad from all blockwalks in the $n \times n$ grid gives the explicit formula.

Theorem 3.

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

The Catalan number has other forms

Corollary 1.

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

and

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

Exercise 1. Compute C_6 and C_7 using both the recursion and the explicit formula.

Exercise 2. Suppose 40 people are seated around a table. Suppose person 1 shakes hands with person 8 and person 14 shakes with person 29. How many possible handshakes are there (without crossing hands) which include these two? How many of these will have person 32 shaking with person 11?

Exercise 3. How many ways are there to triangulate a 24-sided polygon if the triangle 1-8-18 is one of the triangles?

Exercise 4. How many ways are there to triangulate a 24-sided polygon if edges 1-8, 1-15, and 16-22 are used?

Exercise 5. How many blockwalks on a 14 by 14 grid are there which do not cross the railroad tracks, but which visit the diagonal after 6 steps and after 20 steps (and perhaps elsewhere)?

Exercise 6. How many blockwalks on a 14 by 14 grid are there which do not cross the railroad tracks, but which visit the diagonal after 6 steps and after 20 steps, and only at these two points and at the start and finish?

Exercise 7. Suppose Bill lives 4 blocks south and 6 blocks west of work. Suppose the railroad track runs diagonally from 1/2 block south of work to 2 and 1/2 blocks east of home. How many blockwalks does Bill have which do not cross the tracks?

Exercise 8. Suppose Bill lives 9 blocks south and 9 blocks west of work. Suppose there is a lake which prevents him from using the block which is 4 blocks east of home and between 6 and 7 blocks north of home. How many paths (no railroad) does he have from home to work? How many paths are there which do not cross the railroad track?

Exercise 9. A certain town has a rectangular street system. It is 15 blocks in the north-south direction and 15 blocks in the east-west direction. Bill lives 9 blocks north of the southern boundary, and on the western boundary of town. His workplace is 11 blocks east of the western boundary, and on the northern boundary of town. Mary lives 6 blocks east of the western boundary, and on the southern boundary of town. She works 11 blocks north of the southern boundary, and on the eastern boundary of town. How many path pairs (B, M), where B is Bill's path and M is Mary's path are there? How many of these pairs do not touch or cross?

Exercise 10. Suppose the numbers from 1 to $2n$ are arranged in a $2 \times n$ rectangular array in such a way that each row is increasing and each column is increasing. For example, if $n = 6$, then

$$\begin{array}{cccccc} 1 & 2 & 4 & 7 & 9 & 10 \\ 3 & 5 & 6 & 8 & 11 & 12 \end{array}$$

is such an array. Show that the number of ways to do this is C_n. You may either construct a bijection with one of the four objects in these notes, or show it satisfies the Catalan recurrence.

Exercise 11. Show that the number of n-tuples, $(x_1, x_2, \ldots, x_n)$, such that x_i is an integer with $x_i < i$ and $x_1 \leq x_2 \leq \cdots \leq x_n$, is given by C_n. For example, when $n = 6$, $(0, 0, 2, 2, 2, 3)$ is such an n-tuple. You may either construct a bijection with one of the four objects in these notes, or show it satisfies the Catalan recurrence.