CATALAN NUMBERS:
What are they and What do they count?

Wong Kar Lyle

A Little History . . .

Catalan numbers were first studied by Johann Andreas von Segner in the eighteenth century but named after the Belgian mathematician Eugène Charles Catalan (1814-1894), then achieved a simple solution to the problem of enumerating the dissections of a polygon by non-intersecting diagonals into triangles in 1838. However, back in China, more that a century earlier, Ming An-Tu (1692-1763), a Mongolian mathematician had already established the Catalan numbers [5].

Introducing . . .

Many integer sequences arise in combinatorial contexts and here's an introduction to one such special integer counting sequence the Catalan numbers

$$\{C_n\}$$

1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, ...

which have the explicit formula:

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

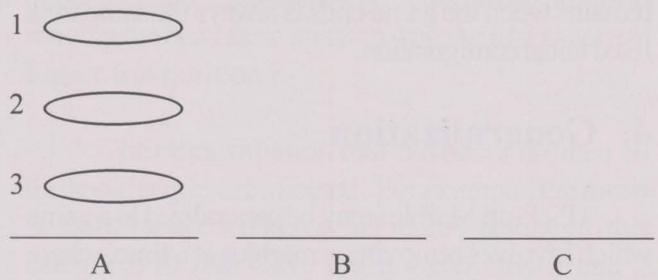
What is so special about the Catalan numbers? Well, besides occuring as sequences in 3 widely studied classes of combinatorial objects such as the partitioning of a polygon into triangles, the binary bracketing of $n + 1$ terms, and some integral lattice walks, they also have many manifestations in some unexpected places [7]. In fact, they are considered to be ubiquitious and Gould's biliography [2] reveals over 400 objects where they are used. [8] also provides a comprehensive introduction to this sequence.

Combinatorial fun . . .

Let's first begin with a game called Double-stacking.

Example 1

Consider an ordered sequence of elements (tokens) numbered 1, 2, ..., n piled up in stack A with token 1 on the top and token n at the bottom, while stacks B and C are initially empty. The only type of legal move in this game is one where only the top token of a stack can be moved to the top of the stack to its right. The object of the game then is to find the number of possible permutations of the tokens 1, 2, ..., n that can be created on stack C.



For n = 3, there are all in all only 5 permutations possible. Did you get them?

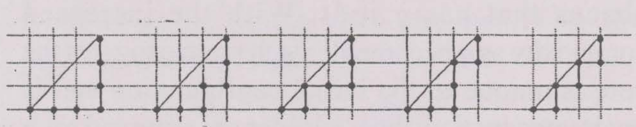
They are: (1 2 3) (3 2 1) (2 3 1) (1 3 2) (3 1 2)

Shown below is the unique sequence of moves that creates the permutation (2 3 1).

1																				2
2			2			2							3				3			3
3			3	1		3		1	3	2	1		2	1		2	1			1
A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C

Example 2

Let's try another puzzle. On a n–by–n integer grid, find the number of different possible integral walks allowing only n rightward moves and n upward moves form the origin (0,0) to the destination (n,n) which do not cross the diagonal $y = x$. Try for $n = 3$. The answer, you will find is also 5 and shown here are the solutions. Did you get them all this time?



Letting $n = 4, 5, 6, ...$ for both examples, we will get the same number of solutions corresponding to the Catalan number sequence.

Proof preliminaries...

Before we work out the formula for the Catalan numbers, let's discuss a late nineteenth century problem that stumped many mathematicians of that time:

> The Ballot Problem: An election is held in which there are two candidates, A and B. Suppose that A receives a votes, B receives b votes with $a > b$. What is the probability that, throughout the counting votes, A stays ahead of B?

In 1887, the French mathematician Desire Andr (1840-1917) produced a solution which later became known as the "Reflection Principle". Observe first that a sequence of the counting of votes corresponds to a lattice path from (0,0) to (a,b) which remains below the diagonal $y=x$ at all times. Each vote for A is interpreted as a horizontal move while a vote for B is indicated by a vertical move. Since at any point (x,y), $y \leq x$ thus satisfying the criteria for the ballot problem. The task remains to count the number of paths required (called good paths) which is really just the total number of paths from (0,0) to (a,b) minus the total number of bad paths whereby a bad path is one which crosses the diagonal.

Now, to get the closed form for the Catalan numbers, we consider the previous example 2. We want to count the number of good paths (not crossing the diagonal $y = x$) from O (0,0) to N (n,n). This is equivalent to counting the total number of paths from O (0,0) to N (n,n) minus the total number of bad paths. It is known (for example, see [4]) that the number of paths from (x_1, y_1) to (x_2, y_2) is

$$\binom{x_2 - x_1 + y_2 - y_1}{y_2 - y_1}.$$

And so the total number of paths from O (0,0) to N (n,n) is

$$\binom{2n}{n}.$$

Let's consider a bad path which crosses the main diagonal $y=x$ and first intersects the line $y=x-1$ at the point Q. We then reflect the sub-path OQ (denoted by P_1) about the line $y = x - 1$ to get a new path $O'N$, comprising of the new reflected sub-path P'_1 and the sub-path QN (denoted by P_2). Clearly, each bad path has a one-to-one correspondence with the new path $P'_1 P_2$. Since each new path starts from $O'(-1,1)$ and ends at N (n,n), the number of bad paths is

$$\binom{n+1+n-1}{n-1}.$$

Hence, the number of good paths is

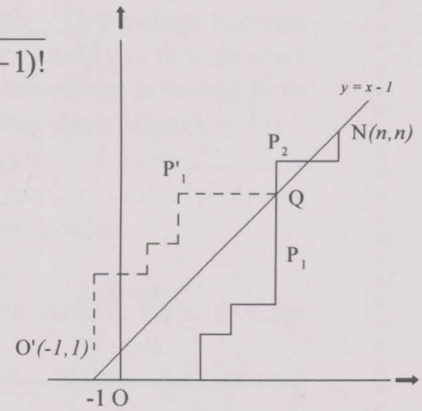
$$=\binom{2n}{n}-\binom{n+1+n-1}{n+1}$$

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

$$=\frac{(2n)!}{n!n!}-\frac{(2n)!}{(n+1)!(n-1)!}$$

$$=\frac{(2n)!}{n!(n-1)!}\left(\frac{1}{n(n+1)}\right)$$

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



Also, by considering the ratio of C_{n+1} / C_n we can easily obtain a recurrence relation to write C_n in terms of n and C_{n-1}. Try it yourself.

Generalizations of Catalan numebrs . . .

It is not surprising that the Catalan numbers can also be generalized. We know that the Catalan numbers count those integral lattice walks as defined previously. What if we decided to move the diagonal line? And by how many units? Upwards or downwards? Is there then a neat formula to count the number of new paths corresponding to a different diagonal line? In [3], Hilton and Pederson provide the answers and give us the generalized formulae

$${}_pC_k=\frac{1}{k}\binom{pk}{k-1}$$

which, as suspected, counts the number of p-good paths from (0, -1) to $(k,(p-1)k-1)$. A path is p-good if it does not cross the line $y = (p-1)x$.

Besides the generalizations of the lattice walks, ${}_pC_k$ is also

1. the number of ways of dividing a convex polygon into k disjoint $(p+1)$-gons by non-intersecting diagonals.

2. the number of p-ary trees with k source-nodes.

3. the number of ways of associating k applications of a given p-ary oparator.

Let's then see some pictorial examples of objects counted by the generalized Catalan number ${}_pC_k$.

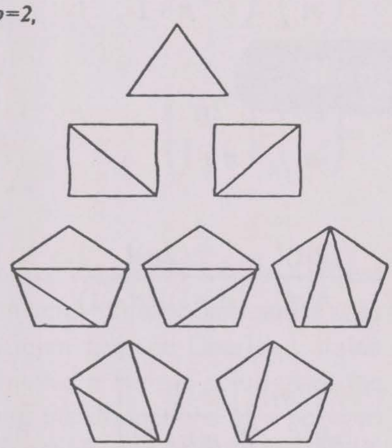
Dividing a convex polygon into k disjoint $(p + 1)$-gons. p-ary trees with k source-nodes

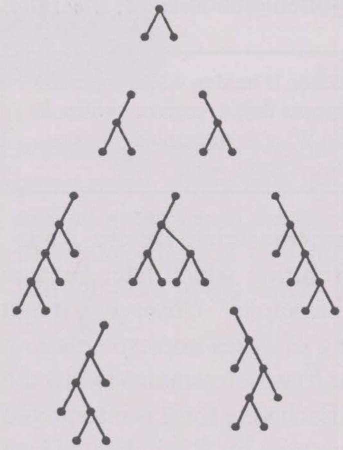
For $p=2$,

$k=1$

$k=2$

$k=3$



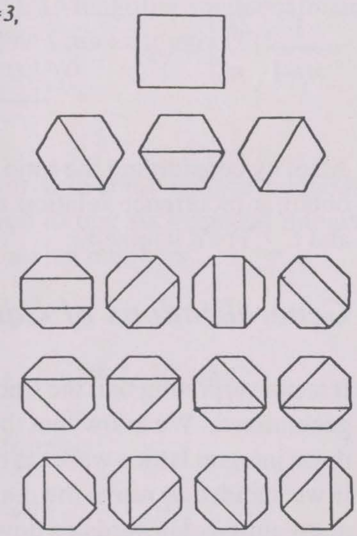


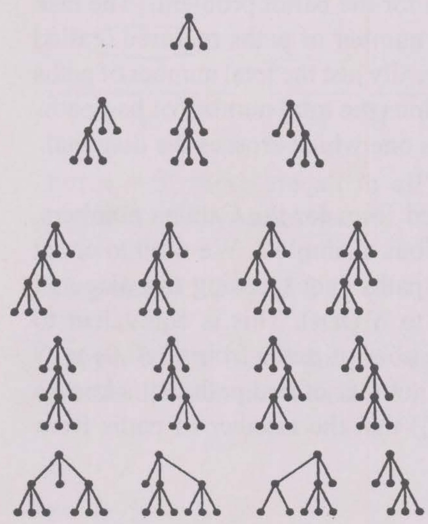
For $p=3$,

$k=1$

$k=2$

$k=3$





References

[1] Conway, J.H. and Guy, R.K. *The Book of Numbers.* (New York, Springer-Verlag, 1996) 96-106.

[2] Gould, Henry W. *Research bibiliography of two special number sequences.* (6th ed. Math. Monongaliae. 12, Morgantown, WV).

[3] Hilton, P. and Pederson, J. *Catalan Numbers, Their Generalization, and Their Uses.* Mathematical Intelligencer 13 (1991) 64-75.

[4] Koh, K.M. and Tan, B.P. *Counting-Its Principles and Techniques (3)* Mathemiatcal Medley 23 (Mar 1996) 9-14.

[5] Lou, Jian-Jin *Catalan Numbers in the History of Mathematics in China* Proceedings of the Spring School and the International Conference on Combinatorics (1992).

[6] Stanley, Richard P. *Problems on Catalan and related numbers* Eumerative Combinatorics 2 91994).

[7] Gardner, Martin *Mathematical Games, Catalan numbers: an integer that materializes in unexpected places.* Sci. Ameri. 234 (June 1976) 120-125.

[8] http://www.astro.virginia.edu/~eww6n/math/Catalan Number.html.

Acknowledgement

This article was submitted as part of the author's

Mathematics honours thesis

under the kind supervision of

Dr. Lam Tao Kai.

Miss Wong Kar Lyle

is a

Mathematics Honours year

(1997/98) student

at the

National University of Singapore