# APPLICATIONS OF CATALAN NUMBERS

Xiaotong Jiang

Department of Mathematical Sciences, Sweet Briar College
Sweet Briar, VA 24595
jiang14@sbc.edu

July 12, 2012

## Abstract

The main purpose of this paper is to explore how Catalan numbers are applied in various cases, to derive the relationship for these applications with explicit formulas, and to produce an accessible write-up that can be understood by undergraduates and non-mathematicians alike. We will begin with a series of categorized applications that are counted by Catalan numbers, and then look at the basic description and recurrence relation of this sequence.

## Applications
### 1. Stacking Coins

The number of ways to stack coins on a bottom row that consists of $n$ consecutive coins in a plane, such that no coins are allowed to be put on the two sides of the bottom coins and every additional coin must be above two other coins, is the $n$th Catalan number.
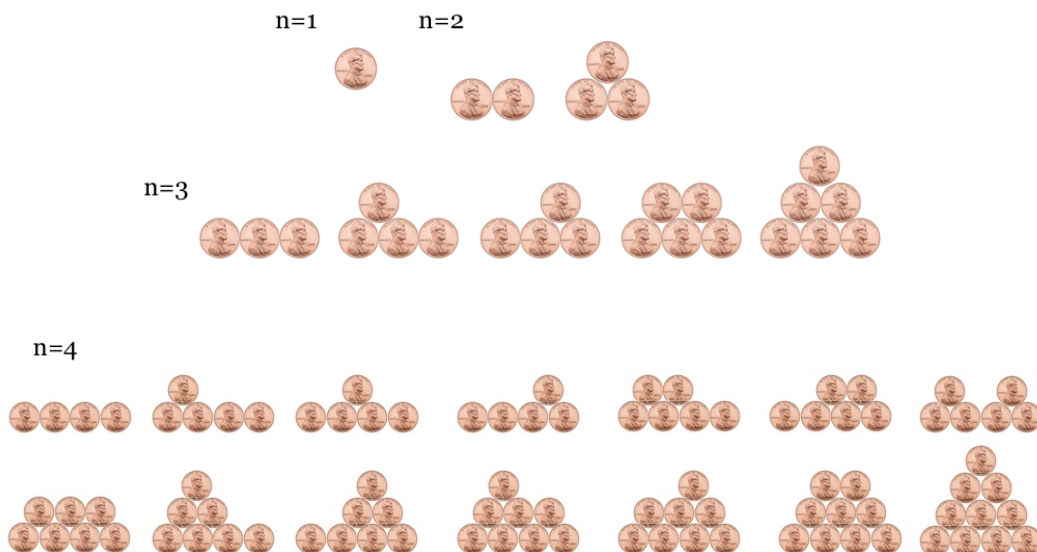
## 2. Balanced Parentheses

The number of ways to group a string of $n$ pairs of parentheses, such that each open parenthesis has a matching closed parenthesis, is the $n$th Catalan number. For example, (( )( )) is valid, but )( ))(( and ( ))( )( are not.

| $n = 0$ | Do Nothing. | 1way |
|---|---|---|
| $n = 1$ | ( ) | 1way |
| $n = 2$ | (( )), ( )( ) | 2 ways |
| $n = 3$ | ( )( )( ), ( ) (( )), (( )) ( ), ( ( )( ) ), ((( ))) | 5 ways |

Many other applications are equivalent to balanced parentheses, and here is one example. If we want to connect $2n$ dots lying on a horizontal line in the plane with $n$ nonintersecting arcs, the solution is also the Catalan sequence. Each arc connecting the two dots is equivalent to a pair of parentheses, with the left dot equivalent to an open parenthesis and the right dot equivalent to a closed parenthesis.



Compare the figure with the last row in the table above, and you will see the similarity.

## 3. Mountain Ranges

The number of ways to form mountain ranges on a line with $n$ upstrokes and $n$ downstrokes, such that each upstroke has a matching downstroke and the path does not go below the starting point, is the $n$th Catalan number. It is the same as the matching rule of the parenthesis grouping problem.

| n=0 | Do Nothing. | 1 way |
|---|---|---|
| n=1 | /\ | 1 way |
| n=2 | /\/\, / /\ \ | 2 ways |
| n=3 | /\/\/\, /\/ /\ \, / /\ \/\, / /\/\ \, / / /\ \ \ | 5 ways |

Note that a pair of strokes and a pair of parentheses are equivalent: upstrokes are equivalent to open parentheses, and downstrokes are equivalent to closed parentheses. The fact that one pair of parentheses are inside another pair corresponds to one pair of strokes being on top of another pair, thus forming the shapes of mountain ranges.

## 4. Polygon Triangulation

The number of ways to cut an $n+2$-sided convex polygon in a plane into triangles by connecting vertices with straight, non-intersecting lines is the $n$th Catalan number. This is the application in which Euler was interested.

A convex polygon satisfies the following two properties: i) each interior angle is less than or equal to 180 degrees, and ii) each line segment connecting two of the vertices must remain inside the boundary of the polygon. As the term suggests, the vertices of a convex polygon point "outward" from the center of the polygon.
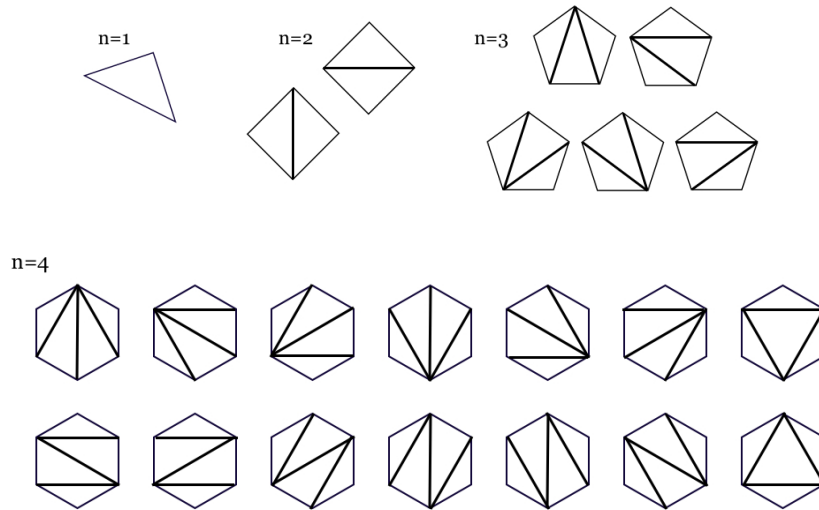


Note that a 2-sided polygon is set to be triangulated in exactly one way: do nothing, so it follows that $C_0 = 1$.

## 5. BALANCED TREES

The number of full binary trees with $n$ internal nodes is the $n$th Catalan number. An internal node is a node that connects to other nodes above it. A full binary tree is a rooted tree where each internal node has exactly two segments going up.



3
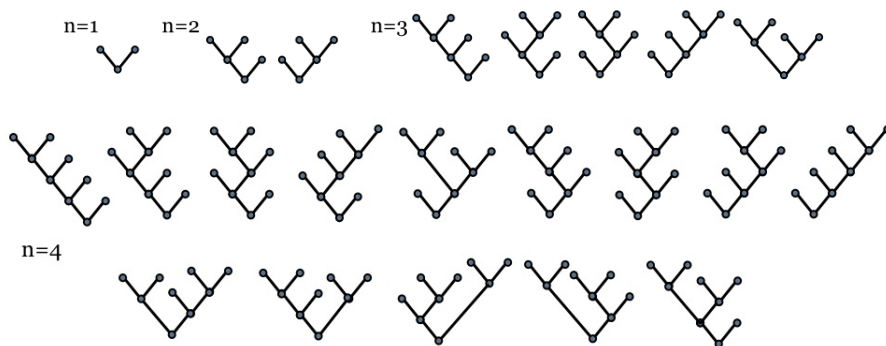
As for $n = 0$, there is one solution, which is a single node itself. In summary, a full binary tree with $n$ internal nodes has $2n + 1$ nodes, $2n$ branches and $n + 1$ leaves.

Other types of binary trees and plane trees contain Catalan numbers as well:

1. Binary trees with $n$ vertices [?].



2. Plane trees with $n + 1$ vertices [?].



## 6.Dyck Paths

Based on the Cartesian Coordinates system, a Dyck path is a walk from $(0, 0)$ to $(n, n)$ in an $n \times n$ lattice, such that

i) the path is only allowed to go in the positive x-axis or positive y-axis direction for each unit step,

ii) passing above the line $y = x$ is not allowed (see Figure 1).

The number of such paths in an $n \times n$ lattice is the $n$th Catalan number.
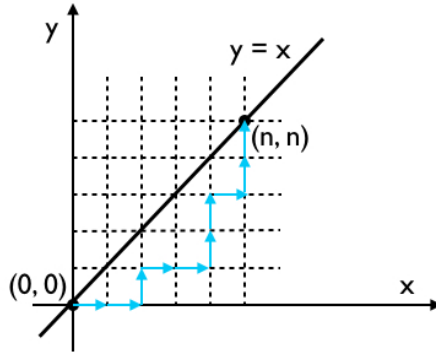


Figure 1: An example of Dyck path.

These kind of paths look a lot like mountain ranges if they are rotated counterclockwise about the origin until the diagonal is horizontal. Whatever value n is, the first step is always to the east and the last step is always to the north, because we cannot pass above the diagonal.

Other types of Dyck Paths turn out to follow the sequence of Catalan numbers as well:

1. Dyck Paths from $(0, 0)$ to $(2n + 2, 0)$ such that any maximal sequence of consecutive steps $(1, -1)$ ending on the x-axis has odd length [?].

n=1  n=2  n=3  n=4
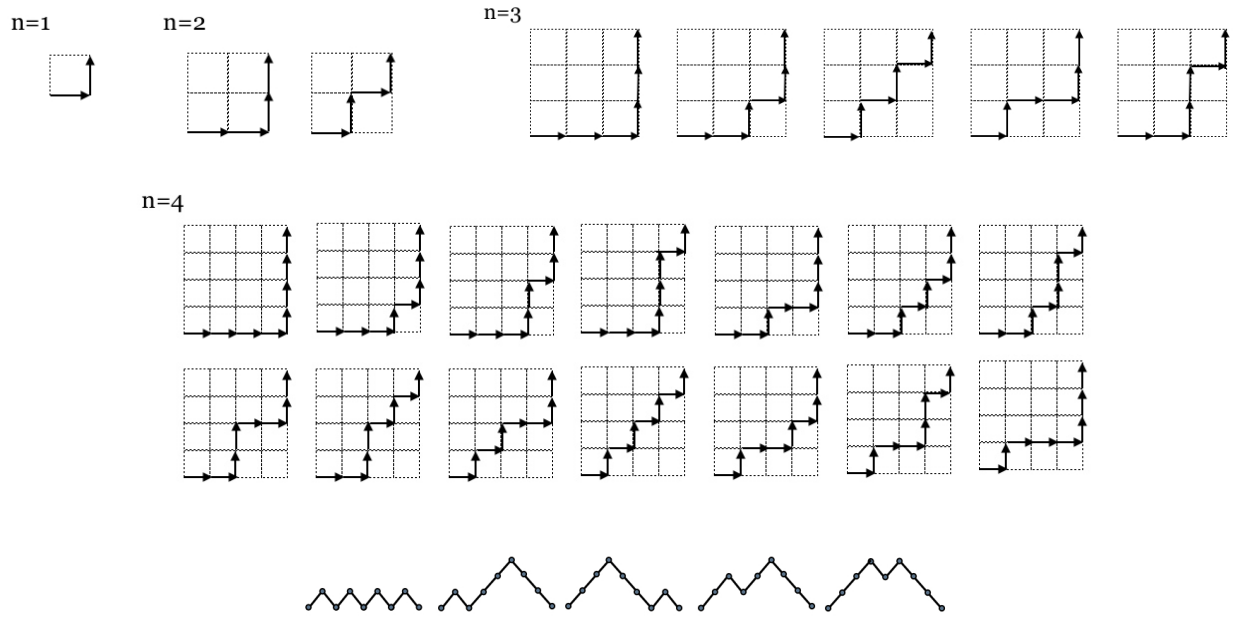
2. Dyck Paths from (0, 0) to (2n + 2, 0) with no peaks at height two [**?**].



## 7. Permutations

A permutation of $\{1, 2, ..., n\}$ is an rearrangement of the $n$ numbers. For example, the permutations of $1, 2, 3$ includes 6 arrangements: $(1, 2, 3), (1, 3, 2), (2, 1, 3, ), (2, 3, 1), (3, 1, 2), (3, 2, 1)$. A 123-avoiding permutation is a permutation that avoids an increasing subsequence of 3 terms (the 3 terms do not have to be consecutive). Therefore, we should avoid $(1, 2, 3)$ for $n = 3$. Taking $n = 4$ as another example, $(4, 3, 1, 2)$ is valid, but $(4, 1, 2, 3)$ is not valid because of the subsequence 123, and neither is $(2, 3, 1, 4)$ because of 234. The number of permutations of $\{1, 2, ..., n\}$ that avoid 123 is the the $n$th Catalan number.

| $n = 0$ | Do Nothing. | 1 way |
|---|---|---|
| $n = 1$ | $(1)$ | 1 way |
| $n = 2$ | $(1, 2), (2, 1)$ | 2 ways |
| $n = 3$ | $(1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)$ | 5 ways |
| $n = 4$ | $(1, 4, 3, 2), (2, 1, 4, 3), (2, 4, 1, 3), (2, 4, 3, 1), (3, 1, 4, 2), (3, 2, 1, 4), (3, 2, 4, 1)$ $(3, 4, 1, 2), (3, 4, 2, 1), (4, 1, 3, 2), (4, 2, 1, 3), (4, 2, 3, 1), (4, 3, 1, 2), (4, 3, 2, 1)$ | 14 ways |

Note that a 123-avoiding permutation only avoids an increasing subsequence of *three* terms, regardless of the value of $n$. Therefore, $(1, 2), (2, 1)$ are valid even though they have increasing subsequences as well.

5

Similarly, there are 321-avoiding permutations of $[n]$, which avoid a decreasing subsequence of 3 terms. Taking $n = 3$ as an example, we have $(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2)$. Moreover, a permutation of $[n]$ is called 132-avoiding "if it does not have three entries $a < b < c$ so that $a$ is the leftmost of them and $b$ is the rightmost of them" [?].

Catalan numbers also count shuffles of the permutation $1, 2, \cdots, n$ with itself, i.e., permutations of the multiset $\{1^2, 2^2, \cdots, n^2\}$ which are a union of two disjoint subsequences $1, 2, \cdots, n$ [?]. These shuffles have no weakly decreasing subsequence of length three.

A shuffle of $1, 2, \cdots, n$ and itself $1, 2, \cdots, n$ is obtained by intermixing the letters in each string of numbers, while the letters in each string stay in the original order. For example, a shuffle of 123 and 456 could be: 124536. No weakly decreasing sequence of length three means that the subsequence either is strictly increasing or has at most two equal entries not followed or preceded by a decrease. For example, 12133 is valid (because it has only two equal entries instead of three with no decrease), but 112332 is not (because there is a decrease, 2, after 33).

Back to the application: let us see an example of of $n = 3$, where we want to know the number of shuffles of permutations of 1, 2, 3 with 1, 2, 3. It turns out that there are 5 distinct shuffles:

$$112233 \quad 112323 \quad 121233 \quad 121323 \quad 123123.$$

Note that these are distinct shuffles. In fact, for each single shuffle, some of the entries may come from either string of $1, 2, 3$, although the sequence appears the same. For instance, the first shuffle 112233 has multiplicity 8: there are 8 ways to obtain this shuffle from the original words.

$$112233 \quad 112233 \quad 112233 \quad 112233 \quad 112233 \quad 112233 \quad 112233 \quad 112233$$

In addition, the Catalan sequence counts the number of permutations of $a_1 a_2 a_3 \cdots a_{2n}$, formed from integers $1, 1, 2, 2, 3, 3, \ldots n, n$ such that

i) these integers 1, 2, 3, ... , n are in increasing order when they *first* occur, and

ii) there is no form like $\alpha \beta \alpha \beta \cdots$, where the integers $\alpha, \beta, \cdots$ do not have to be consecutive.

For example, **1212** and **122313** are not valid. Here are the solutions where $n = 3$:

$$112233, 112332, 122331, 122133, 123321.$$

## 8. Young Diagrams

In combinatorics, a partition of a positive integer $n$ is an expression rewriting $n$ as a sum of positives integers, where the order of the summands does not matter. (The sum would be a composition if order mattered.) Taking $n = 4$ as an example, there are 5 ways to partition 4: 4, 3+1, 2+2, 2+1+1, 1+1+1+1. Remember, the ordering of the integers does not matter, i.e. 1+1+2, 1+2+1, 2+1+1 are equivalent partitions. Partitions can be visualized in explicit graphs, and the most commonly used one is called a Young diagram. Again, take n = 4 and n = 5 as two examples for better understanding. Young diagrams of partitions of 4 and 5 are shown in Figure 2 and Figure 3, respectively.
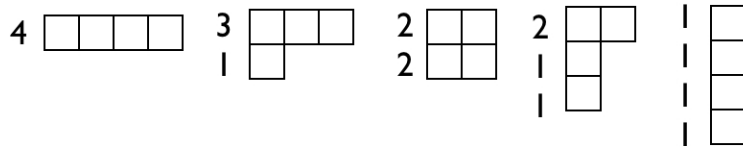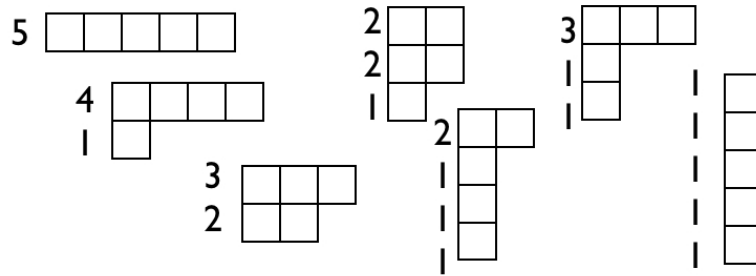
Figure 2: Partition of 4



Figure 3: Partition of 5

Young diagrams that fit in the shape $(n - 1, n - 2, \cdots, 1)$ are counted by the Catalan sequence[?].

The shape $(n - 1, n - 2, \cdots, 1)$ is a Young diagram that looks like a upside down staircase. Its top stair consists of $n - 1$ blocks, next stair $n - 2$ blocks, and so on, until the last stair has only 1 block. By "fit," we mean that we try to find Young diagrams that could be a part of the shape $(n - 1, n - 2, \cdots, 1)$ or the entire shape.

In this example image below, the last figure is the shape $(2, 1)$ where $n = 3$. Each of the other four figures, including the empty set, could be a part of the shape. Adding the original shape, we have five solutions for the shape $(2, 1)$ in total.
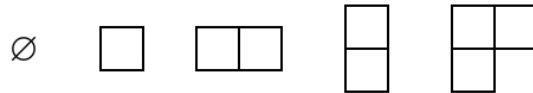


## 9. POSETS

A Partially Ordered Set $P$, or Poset $P$ for short, is a set together with a binary relation denoted by $\leq$, satisfying the following three axioms, where x and y are arbitrary objects of the set:

  i) For all $x \in P, x \leq x$. (*reflexivity*)

  ii) If $x \leq y$, and $y \leq x$, then $x = y$. (*antisymmetry*)

  iii) If $x \leq y$, and $y \leq z$, then $x \leq z$. (*transitivity*)[?]

(a) Hasse diagram of a 3-element set.

(b) Hasse diagram of Young diagrams.

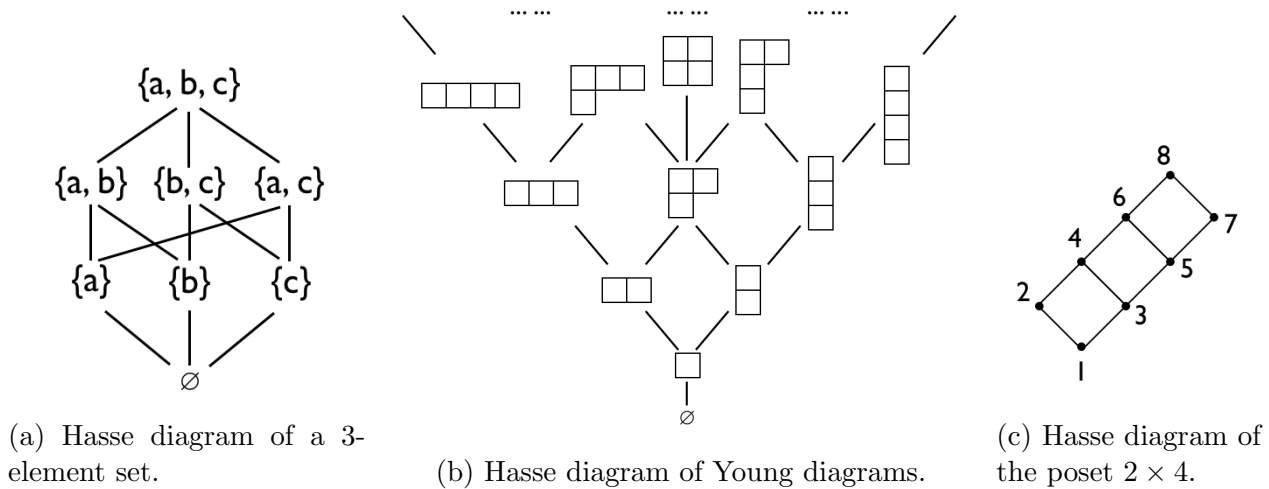(c) Hasse diagram of the poset $2 \times 4$.

Figure 4: Examples of Hasse diagrams to illustrate posets.

A Hasse diagram is used to represent a finite poset. Each element in the poset is a vertex in the Hasse diagram. The transitive relation in the poset is represented by lines going up from one vertex to another in the Hasse diagram. The lines may cross each other but cannot touch other vertices before they reach the endpoint. The line segments and labeled vertices in the diagram illustrate the partial order of a set. See Figure 4 for several examples of Hasse diagrams. As you can see, the element in the poset is any object; it could be a set, a diagram or a number.

How do Hasse diagrams embody the 3 axioms of posets? In Figure 4a,

i) Each element in the diagram reflects itself, i.e., the set $\{a\}$ is less or equal to itself $\{a\}$. This shows reflexivity.
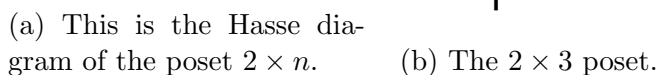
ii) Since set $\{a\}$ is less or equal to $\{a\}$ and $\{a\}$ is less or equal to $\{a\}$, then $\{a\} = \{a\}$. This symmetry does not work between, for example, $\{a\}$ and $\{b\}$, because $\{a\}$ and $\{b\}$ are two different sets. This is the idea of antisymmetry.

iii) Since the empty set is less or equal to set $\{a\}$ and set $\{a\}$ is less or equal to set $\{a, b\}$, the empty set is less or equal to $\{a, b\}$. In the diagram, the three sets are connected with 2 line segments. This shows that if we can follow the lines going from the bottom element up to the top one, then all the elements on the way obey transitivity. Likewise, we can tell that $\{b\}$ is less or equal to $\{a, b, c\}$ according to the diagram.

Linear extensions of the poset $2 \times n$ follow the Catalan sequence [?]. (See Figure 5a)

If $n = 3$, then the poset $2 \times 3$ looks like Figure 5b. A linear extension is obtained by rewriting the Hasse diagram on a line, reading the diagram from the bottom up. How do we know where to put the elements? Starting from the bottom 1, we can write each element in the Hasse diagram only after the elements with which it connects from the bottom are already written. Hence, there is more than one linear extension of a poset. In the example of $n = 3$, there are 5 linear extensions. The starting and ending point will never change, whereas the points in between vary.

$$123456, 123546, 132456, 132546, 135246$$

8

(a) This is the Hasse diagram of the poset $2 \times n$.

(b) The $2 \times 3$ poset.

Figure 5

## 10. PASCAL'S TRIANGLE

Catalan numbers lie in Pascal's Triangle.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
\end{array}
$$

```
                        1
                     1     1
                  1     2     1
               1     3     3     1
            1     4     6     4     1
         1     5    10    10     5     1
       1    6    15    20    15    6    1
     1    7   21   35   35   21   7   1
   1   8   28   56   70   56  28   8   1
  1  9  36  84 126 126  84 36  9  1
 1 10 45 120 210 252 210 120 45 10 1
  ⋮   ⋮   ⋮   ⋮   ⋮   ⋮   ⋮   ⋮   ⋮   ⋮
```

If we take the difference of numbers in the middle column on even rows (starting with row 0) and the adjacent column, we will find the Catalan sequence.

This is not a coincidence but certainty. Each entry in Pascal's Triangle is in the form of $\binom{k}{r}$, where $k$ is the row number starting at 0 from top to bottom, and $r$ is the entry number starting at 0 from left to right. For example, the first three rows can be expressed as:

$$\binom{0}{0}$$

$$\binom{1}{0} \quad \binom{1}{1}$$

$$\binom{2}{0} \quad \binom{2}{1} \quad \binom{2}{2}$$

9

Since all of the blue numbers are in even rows $0, 2, 4, ...$, we use $2n$ to represent their row number. Moreover, since the blue column is in the middle, the entry of each number is always half of the its row number. Therefore, the numbers in the middle column can be expressed as $\binom{2n}{n}$. The numbers in the pink column are in the same row as those in the middle column and they are right next to the middle column, so they can be expressed as $\binom{2n}{n+1}$. It is now obvious that the differences of numbers in the middle column on odd rows and their adjacent column are Catalan numbers. This is true because

$$\binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1} \cdot \binom{2n}{n} = C_n, n = 0, 1, 2, ...$$

where the latter equation is given below.

## Definitions and Proofs

The $n$th Catalan number is defined as

$$C_n = \frac{1}{n+1} \cdot \binom{2n}{n} = \frac{(2n)!}{n!(n+1)!}, \quad n = 0, 1, 2, ...$$

The binomial coefficient, $\binom{n}{r}$, pronounced as n choose $r$, represents the number of possible combinations of $r$ objects from a collection of $n$ objects:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

Therefore,

$$\binom{2n}{n} = \frac{(2n)!}{n!(2n-n)!} = \frac{(2n)!}{n!n!} = \frac{2n \cdot (2n-1) \cdot (2n-2) \cdots (n+1)}{n!}.$$

Example: $\binom{11}{4} = \frac{11!}{4!7!} = \frac{11!}{4!(11-4)!} == \frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{4 \times 3 \times 2 \times 1 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{11 \times 10 \times 9 \times 8}{4 \times 3 \times 2 \times 1} = \frac{11 \times 10 \times 9 \times 8}{4 \times 3 \times 2 \times 1}$.

Catalan numbers could be described in various but equivalent ways, including

$$C_n = \binom{2n}{n} - \binom{2n}{n+1}, \quad n = 0, 1, 2, \cdots.$$

Note that $\binom{2n}{n}, \binom{2n}{n+1}$ are natural numbers and $\binom{2n}{n} > \binom{2n}{n+1}$. Therefore, $C_n$ is the difference between two positive, natural numbers which are entries in Pascal's Triangle.

Catalan numbers grow rapidly.

| n | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $C_n$ | 1 | 1 | 2 | 5 | 14 | 42 | 132 | 429 | 1430 | 4862 | 16796 | 58786 |

## Recurrence Definition and Bijection

We have seen various kinds of applications of Catalan numbers so far: "Stacking Coins," "Balanced Parentheses," "Mountain Ranges," "Polygon Triangulation," "Binary Trees," "Binary Paths," "Permutations,"' "Young Diagrams" and "Posets." In fact, all the sequences are equivalent, and we will show that there is a common formula that counts them all.

$$C_0 = 1, C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i} \text{ for } n \geq 1$$

In the application Balanced Parentheses, it is already known that for each open parenthesis, there is a closed parenthesis. Now, let's try to find a pattern in these paired parentheses with example $n = 3$:

$$((( \ ))) - (( \ )( \ )) - (( \ ))( \ ) - ( \ )(( \ )) - ( \ )( \ )( \ ).$$

The "pattern" is that we can always separate them in two collections. For example, we can separate the set $(( \ ))( \ )$ into: $(( \ ))$ and $( \ )$. We name them collection A and collection B, either of which is able to contain zero pairs of parentheses. Similarly, $((( \ )))$ could be separated into $((( \ )))$ and nothing. For $(( \ )( \ ))$ , we treat it as a whole and put it in collection A, so B is, again, empty.

What about $( \ )( \ )( \ )$? At first glance, we see three pairs of parentheses, but we have only two collections. We could choose to put the first two pairs of parentheses in collection A and the last pair should be in B, or put two in B and only one in A. This is exactly the same and we do not want to count them twice, thus there is a need for a regulation in order to avoid repetition. Since $n$ is no less than 1 in the recurrence definition mentioned above, it is certain that there is at least a pair of parentheses and we will fix it in collection A. Thus, the simplest form where $n = 1$ is:

$$( \ \ )$$

$$A \quad B,$$

and this is our base form. For values of n that are greater than 1, we simply "add" more pairs of parentheses "inside" the fixed black parentheses to collection A, and place the rest in collection B. In this way, both A and B are able to contain up to $n-1$ pairs of parentheses (the base parentheses in the base form does not count as one of them). If collection A contains k pairs, then it is not hard to find that there are $n - (k + 1)$ pairs in collection B.

What is the purpose of separating the parentheses into 2 collections? It is because we want to count the combinations of parentheses systematically, that is, If A has 0 pairs, then B has $n - 1$ pairs; if A 1 pair, then B $n - 2$ pairs; if A 2 pairs, then B $n - 3$ pairs, etc.

Add up all of the situations, and we get the total number:

$$C_n = C_0 C_{n-1} + C_1 C_{n-2} + C_2 C_{n-3} + \cdots + C_{n+2} C_1 + C_{n-1} C_0.$$

This formula is the recursive relation that we are looking for. Plugging in actual numbers may help you understand this great formula.

$$C_1 = C_0 C_0$$

11

| Number of Pairs Contained in A | Number of Pairs Contained in B | Illustration | Number of Solutions for Each Situation |
|:---:|:---:|:---:|:---:|
| 0 | $n-1$ | $( )(\cdots)$ <br> $A \quad B$ | $C_0 C_{n-1}$ |
| 1 | $n-2$ | $(( ))(\cdots)$ <br> $A \quad B$ | $C_1 C_{n-2}$ |
| 2 | $n-3$ | $((( )))(\cdots)$ <br> $A \quad B$ | $C_2 C_{n-3}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $n-1$ | 0 | $((\cdots))$ <br> $A \quad B$ | $C_{n-1} C_0$ |

$$C_2 = C_0 C_1 + C_1 C_0$$

$$C_3 = C_0 C_2 + C_1 C_1 + C_2 C_0$$

By slightly changing the base form into

$$/A\backslash B,$$

we can generate the recursive relation formula in similar examples, such as Mountain Ranges and Binary Paths.

### Another Approach

The method introduced above may not be so powerful for those applications with figures. Hence, we need another way of thinking to approach problems like Polygon Triangulation, in order to obtain the recursive definition.

Here, we use the case of a 6 - sided polygon where $n = 4$.

We will start with drawing the first triangle based on the horizontal side at the top of the hexagon. This horizontal side is going to be part of the triangle. Since there are four vertices left, besides the two vertices that the horizontal line connects, we can draw four different cases. See the figure below.
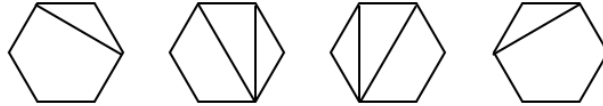


Each triangle divides the hexagon into two polygons, on the left and right of the original triangle. Our next step is trying to triangulate these two separated polygons. Recall that a polygon with $k > 3$ sides can be triangulated into $C_{k-2}$ ways. In the first case (the first hexagon in the figure above), there is a pentagon on the left that has $C_3$ ways of triangulation, and nothing on the right of the selected triangle, which has $C_0$ triangulations. Thus, the first case has $C_3 \cdot C_0$

solutions in total. In Case 2, we have a quadrangle and a triangle, so there are $C_2 \cdot C_1$ solutions. Similar methods can be applied to Case 3 and 4.
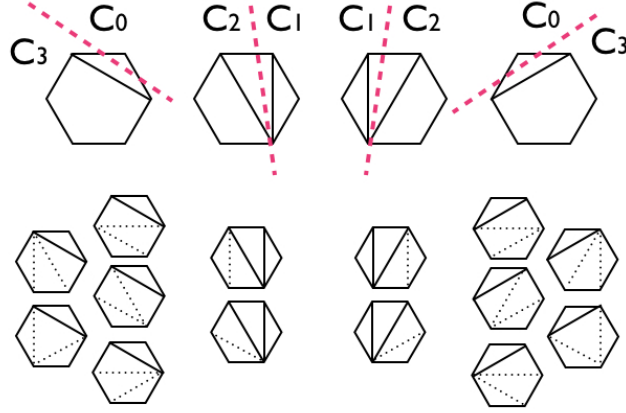


Adding up the number of solutions in each case, we get the total number of ways to triangulate a 6 - sided polygon, $C_4 = C_3C_0 + C_2C_1 + C_1C_2 + C_0C_3$.

Generally, an $n + 2$- sided polygon will have $n$ different first triangles after the initial step of triangulation. On the left and right sides of each of those triangles, there are an $n + 1$- sided polygon and nothing, respectively ($C_{n-1}C_0$), an $n$- sided polygon and a triangle ($C_{n-2}C_1$), an $n-1$ - sided polygon and a quadrangle ($C_{n-3}C_2$), or an $n - 2$ sided polygon and a 5 - sided polygon ($C_{n-4}C_3$), and so on.

Take the sum of these, and the total number of ways to triangulate an $n + 2$ - sided polygon is

$$C_n = C_{n-1}C_0 + C_{n-2}C_1 + C_{n-3}C_2 + \cdots + C_1C_{n+2} + C_0C_{n-1}.$$

## BIJECTIONS

A bijection is a one-to-one correspondence of two sets (or a one-to-one and onto function). In a more understandable way, we can always pair every element in one set with exactly one element in the other set. Hence, there are no unpaired elements in either set and the total numbers of elements in both sets are the same.

To have an exact pairing between $X$ and $Y$ (where $Y$ need not be different from $X$), four properties must hold:

i). each element of $X$ must be paired with at least one element of $Y$,

ii). no element of $X$ may be paired with more than one element of $Y$,

iii). each element of $Y$ must be paired with at least one element of $X$, and

iiii). no element of $Y$ may be paired with more than one element of $X$ [?].

Property 1 and 2 guarantee that the bijection is a function on domain $X$. Functions satisfying property 3 are called "onto." Functions satisfying property 4 are called "one-to-one."

13

Here is an example of a bijection. Let $X, Y$ be the two sets, and $f : X \to Y$. If $X = \{A, B, C, D\}$ and $Y = \{a, b, c, d\}$, one possible bijective function is:

$$f(A) = b,$$
$$f(B) = c,$$
$$f(C) = d,$$
$$f(D) = a.$$

The followings are not bijections:

$$f(A) = b, \qquad f(A) = b,$$
$$f(B) = b, \qquad f(A) = c,$$
$$f(C) = d, \qquad f(C) = d,$$
$$f(D) = a. \qquad f(D) = a.$$

The formula for the recursive definition is essential because it connects any two applications of Catalan numbers and shows their bijection. Try it if you want to analyze and find bijections among other examples with the same idea and method.

A Few Hints:

* Binary Trees: The base is the one full binary tree at the bottom. Similar to the first method for' "Balanced Parentheses," collection A contains the "baby" trees branching out from the "left" node of the base tree, while collection B contains the "baby" trees branching out from the "right" node of the base tree. All of the trees in both collections have a total number of $n - 1$.

* Binary Paths: Besides rotating counterclockwise, there is another way to look at the bijection between binary paths and parentheses. Starting with the origin $(0, 0)$, the unit path to the east is equivalent to an open parenthesis, so the unit path to the north is equivalent to a closed parenthesis. The reason we could do this kind of bijection is that there are definitely the same numbers of steps to the east and steps to the north, because the grid has the same length and width.

* Stacking Coins: If you outline a border that is tangent to the coins for each coin stack, it will be obvious that it looks like mountain ranges. For instance, Figure 6 is equivalent to the second mountain range when $n = 3$ as it is illustrated in the section Mountain Ranges.



Figure 6

## History

The first person in Europe who discovered Catalan numbers was Leonhard Euler. In 1751, Euler discussed the number of ways to cut a polygon with lines into triangles without any of the lines intersecting in his letter to Christian Goldbach, a German mathematician.

It was a French and Belgian mathematician, Eugéne Charles Catalan, who described this number sequence in a well-defined formula, and introduced this subject to solve parentheses expressions.

Before Euler, a Mongolian mathematician Minggatu was the first person in the East who established and applied what were later to be known as Catalan numbers. In the 1730s, he brought forward this sequence of numbers and continued using it when he was trying to express series expansions of $\sin(ma)$, where $m = 2, 3, 4, 5, 10, 100, 1000,$ and $10000$. This topic was included in his book, *Ge Yuan Mi Lu Jie Fa* (The Quick Method for Obtaining the Precise Ratio of Division of a Circle).

## Summary

So far we have seen a certain number of applications of Catalan numbers and how they are related to each other. In fact, Catalan numbers arise in over 600 examples.

We have been convinced that, however the examples vary, applications of Catalan numbers are related to each other in an equivalent way. "In a sequence of $2n$ items with $n$ +'s and $n$ −'s, if there are no more −'s than +'s anywhere in the sequence (in other words, if the partial sums of this sequence are always nonnegative), then the number of ways of counting these items is the $n$th Catalan number." Such a sequence is called a ballot sequence.

Think about the parentheses. If there are more closed parentheses than open parentheses somewhere in the sequence, then it will not make sense.

Any problem that follows this rule can be solved by Catalan numbers. Here is one example. A class of 400 college students is voting for their class president. 200 students vote for A, and 200 students vote for B. If in the voting process B always trails or equals A, then can you tell me the number ways of the sequence in which the votes could appear?

# References

[1] Stanley, Richard P. *Enumerative Combinatorics*, vol.2. Cambridge University Press. New York/Cambridge. 1999.

[2] Stanley, Richard P. *Catalan Addendum*. MIT Mathematics. Version of 22 October 2011. Retrieved from Richard Stanley's home page

http://www-math.mit.edu/ rstan/ec/.

[3] Wikipedia. *Bijection*. Retrieved from http://en.wikipedia.org/wiki/Bijection.

[4] Davis, Tom. *Catalan Numbers*. Mathematical Circles Topics. November 24, 2010. Retrieved from http://www.geometer.org/mathcircles/catalan.pdf.

[5] Stanley, Richard P. *Enumerative Combinatorics*, vol.1. Second printing, Cambridge University Press. New York/Cambridge. 1997.

[6] Campbell, Douglas M. *The Computation of Catalan Numbers*. Mathematics Magazine, Vol. 57, No.4 (Sep., 1984), pp. 195 - 208.

[7] Choo, Koo-Guan. *Catalan Numbers*. Mathematics Department of University of Sydney, AU. Retrieved from

http://www.maths.usyd.edu.au/u/kooc/catalan.html.

[8] Britz, Thomas. Cameron, Peter. *Partially ordered sets*. For Combinatorics Study Group talks by Professor Rafael Sorkin. Syracuse University. 2001. Retrieved from

http://www.maths.qmul.ac.uk/ pjc/csgnotes/posets.pdf.

[9] Bona, Miklos. *A self-dual poset on objects counted by the Catalan numbers*. School of Mathematics, Institute of Advanced Study. November 10, 1998.

[10] Dowling, Thomas A. *Catalan Numbers*. Department of Mathematics of Ohio State University. Retrieved from

http://www.mhhe.com/math/advmath/rosen/r5/instructor/applications/ch07.pdf.