The Catalan Numbers

Ishaan Patkar

1 Introduction

The Catalan numbers are one of the most ubiquitous sequences of numbers in mathematics, counting the solutions to many diverse problems. This makes them one of the richest integer sequences in mathematics.

In this paper, we will discuss some of the basic properties of the Catalan numbers. This will include a discussion of several objects counted by the Catalan numbers, known as "Catalan objects." There are hundreds of interesting Catalan objects, but in this paper we will focus on seven of the most fundamental. In order to show that these objects are in fact Catalan numbers, we will employ bijections:

Definition 1.1. For any two set A and B, a function $f : A \to B$ (from A to B) is bijective if it is:

- injective (one-to-one): for all $a, a' \in A$, if $f(a) = f(a')$, then $a = a'$
- surjective (onto): for all $b \in B$, there exists some $a \in A$ such that $f(a) = b$.

Alternatively, f is bijective if there exists some $g : B \to A$ such that for all $a \in A$, $g(f(a)) = a$ and for all $b \in B$, $f(g(b)) = b$.

It is left as an exercise as the reader to show that these two definitions of a bijection are equivalent. Throughout this text, we will use the terms function and mapping synonymously.

The crucial property of bijections is that if there exists a bijection between finite sets A and B, then A and B have the same number of elements. While there are many ways of showing this, bijections are one of the most methods, especially with Catalan objects where they arise naturally.

The first section details some fundamental properties of the Catalan numbers, and the second section discusses several Catalan objects and describes the bijections between them. While important properties on their own, the material on the Catalan generating function and formula will not be used in the second section. Much of the content of this text is taken from Richard P. Stanley's *Catalan Numbers* [1], which is the natural next step for an interested reader.

2 The Catalan Numbers

To begin our study of Catalan numbers, we must first define them. There are many equivalent definitions of the Catalan numbers, but ours involves lattice paths:

Definition 2.1. For a positive integer d, a lattice path with steps $S \subseteq \mathbb{Z}^d$ is a sequence $p_0, p_1, p_2, \ldots, p_n \in \mathbb{Z}^d$ such that for all integers $0 < i \leq k$, $p_i - p_{i-1} \in S$. We say that the path is *from* p_0 and *to* p_n and has length n. (Here, $\mathbb{Z}^d$ denotes the set of d-tuples of integers $(k_1, k_2, k_3, \ldots, k_d)$.)

One common type of lattice path is the path from $(0, 0)$ to (n, n) with steps of either $(0, 1)$ or $(1, 0)$. A Dyck path is a variation on this type of lattice path:

Definition 2.2. A Dyck path is a lattice path in $\mathbb{Z}^2$ with steps $S = \{(1, 0), (0, 1)\}$ from $(0, 0)$ to (n, n) such that for every point (x, y) on the path, $x \leq y$.

We often say that a Dyck path stays on or below the diagonal. See Figure 1 for some examples of Dyck paths. Some texts equivalently define Dyck paths as the set of lattice paths from $(0, 0)$ to $(2n, 0)$ with steps $(1, 0)$ and $(1, 1)$.

Dyck paths are how we will define the Catalan numbers:

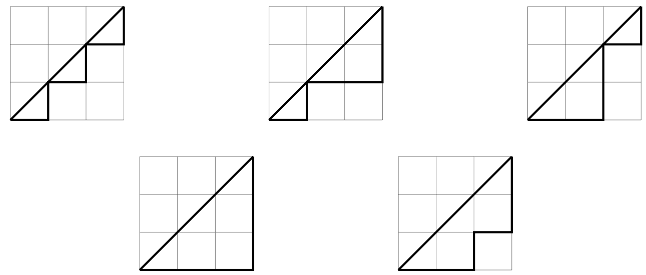


Figure 1: The five Dyck paths from $(0,0)$ to $(3,3)$. The bottom left corner of each grid is $(0,0)$ and the top right corner is $(3,3)$.

Definition 2.3. For all nonnegative integers n, the nth Catalan number, denoted C_n, is the number of Dyck paths from $(0,0)$ to (n,n).

We will continue to denote the nth Catalan number as C_n throughout this text.

This makes Dyck paths the first Catalan object we encounter. We can compute the first few Catalan numbers by listing Dyck paths as in Figure 1:

$$C_0 = 1, C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14, \ldots.$$

For larger n, this method becomes impractical. Fortunately, the Catalan numbers obey a simple recurrence relation.

Theorem 2.1. *For all positive integers n:*

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

Proof. We will prove this relation combinatorially with Dyck paths. For every positive integer n, consider the set of Dyck paths from $(0,0)$ to (n,n). For all integers $0 \le k \le n-1$, let S_k be the set of Dyck paths from $(0,0)$ to (n,n) with $(k+1,k+1)$ the first point on the diagonal. Every Dyck path must intersect the diagonal at some point after $(0,0)$, or else it cannot end at (n,n). Thus, every Dyck path is in one of these S_k. No Dyck path can be in S_i and S_j where $i \neq j$, so these S_k's partitions the set of Dyck paths from $(0,0)$ to (n,n). Thus, the number of Dyck paths is the sum of the elements in all of these S_k's. Now we will count the number of elements in S_k.

Let $p_0, p_1, p_2, \ldots, p_n \in \mathbb{Z}^2$ be a Dyck path in S_k. Note that if (a,b) is on the path, a steps of the form $(1,0)$ and b steps of the form $(0,1)$ have to be taken from $p_0 = (0,0)$ to reach (a,b), so $p_{a+b} = (a,b)$. So, as $(k+1,k+1)$ is on the path, have $p_{2k+2} = (k+1,k+1)$. For all $0 < i < 2k+2$, p_i must be strictly below the path or else we contradict the minimality of k. Thus, if $p_i = (a,b)$, $a < b$ and so $a \le b-1$. So, the path $p_1, p_2, p_3, \ldots, p_{2k+1}$ must stay on or below the line $y = x - 1$ at all times. Note that $p_1 = (0,1)$ and $p_{2k+1} = (k+1,k)$. Thus, this section of the path is just a Dyck path from $(0,0)$ to (k,k), only shifted by $(0,1)$ to the right. Likewise, consider the path $(k+1,k+1) = p_{2k+2}, p_{2k+3}, p_{2k+4}, \ldots, p_{2n} = (n,n)$. This path stays below the diagonal $y = x$, and so it is equivalent to a Dyck path from $(0,0)$ to $(n-k-1, n-k-1)$ shifted by $(k+1,k+1)$. We can thus form any Dyck path in S_k uniquely by combining a Dyck path from $(0,0)$ to (k,k) and a Dyck path from $(0,0)$ to $(n-k-1, n-k-1)$. Since there are C_k Dyck paths from $(0,0)$ to (k,k) and C_{n-k-1} Dyck paths from $(0,0)$ to $(n-k-1, n-k-1)$, there are thus $C_k C_{n-k-1}$ Dyck paths from $(0,0)$ to (n,n).

Summing over all k thus gives the number of Dyck paths from $(0,0)$ to (n,n). Thus:

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

□

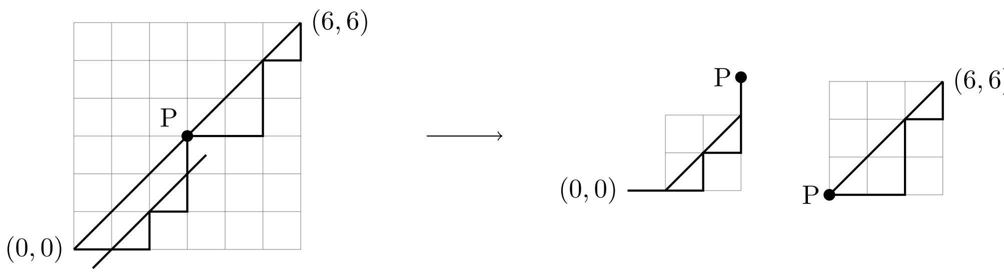


Figure 2: A Dyck path from $(0,0)$ to $(6,6)$ which intersects the diagonal for the first time at P and is broken down into two Dyck paths from $(0,0)$ to $(2,2)$ and $(0,0)$ to $(3,3)$.

See Figure 2 for an example of the deconstruction of a Dyck path into two smaller Dyck paths.

With the initial conditions of $C_0 = 0$, this recurrence relation uniquely defines the Catalan numbers. We can use it to compute larger values of the Catalan numbers:

$$C_5 = 42, C_6 = 132, C_7 = 429, C_8 = 1430, \ldots.$$

Recurrences are not uncommon in combinatorics and so we have techniques of dealing with recurrences in general. We might wonder if we can use these techniques to find a closed form for the Catalan numbers. Indeed we can:

Theorem 2.2. *Let*

$$C(x) = \sum_{n=0}^{\infty} C_n x^n$$

be the generating function for the Catalan numbers. Then

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

Proof. We can use the Catalan recurrence to expand out the generating function:

$$C(x) = \sum_{n=0}^{\infty} C_n x^n = \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} C_k C_{n-k-1} x^n + C_0.$$

We can actually put the sum on the right-hand side in terms of the Catalan generating function:

$$\begin{aligned}
C(x) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} C_k C_{n-k-1} x^n + C_0 \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} C_k C_{n-k} x^{n+1} + C_0 \\
&= \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} C_k C_{n-k} x^{n+1} + C_0 \\
&= x \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} (C_k x^k)(C_{n-k} x^{n-k}) + C_0 \\
&= x \sum_{k=0}^{\infty} (C_k x^k) \sum_{n=k}^{n} (C_{n-k} x^{n-k}) + C_0 \\
&= x \sum_{k=0}^{\infty} (C_k x^k) \sum_{n=0}^{n} (C_n x^n) + C_0 \\
&= x C(x)^2 + C_0
\end{aligned}$$

Rearranging gives

$$xC(x)^2 - C(x) + 1 = 0.$$

This is a quadratic, which we can solve. The quadratic formula gives:

$$C(x) = \frac{1 \pm \sqrt{1-4x}}{2x}.$$

Now we must determine the sign of the square root. We can expand the square root with the binomial theorem as follows

$$\sqrt{1-4x} = \sum_{n=0}^{\infty} \binom{1/2}{n} (-4x)^n = 1 - 2x - \cdots.$$

If the sign of the square root is positive, we have

$$C(x) = \frac{1}{2x}\left(1 + \sum_{n=0}^{\infty} \binom{1/2}{n} (-4x)^n\right) = \frac{1}{x} - 1 - \cdots$$

which is impossible, as $C(0) = C_0 = 1$, but the function on the right side is not defined at 0 because of the $\frac{1}{x}$. Thus, we have

$$C(x) = \frac{1 - \sqrt{1-4x}}{2x}.$$

□

Expanding this generating function gives us a closed form for the Catalan numbers:

Theorem 2.3. *For all nonnegative integers n,*

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

Proof. Expanding the square root in the Catalan generating function with the binomial theorem, we have

$$\begin{aligned} C(x) &= \frac{1 - \sqrt{1-4x}}{2x} \\ &= \frac{1}{2x}\left(1 - \sum_{n=0}^{\infty} \binom{1/2}{n} (-4x)^n\right) \\ &= -\frac{1}{2x} \sum_{n=1}^{\infty} \binom{1/2}{n} (-4x)^n \\ &= -\frac{1}{2x} \sum_{n=0}^{\infty} \binom{1/2}{n+1} (-4x)^{n+1} \\ &= \sum_{n=0}^{\infty} -\frac{1}{2} \binom{1/2}{n+1} (-4)^{n+1} x^n. \end{aligned}$$

By equating coefficients, we have

$$\begin{aligned} C_n &= -\frac{1}{2}\binom{1/2}{n+1}(-4)^{n+1} \\ &= -\frac{1}{2} \cdot \frac{\left(\frac{1}{2}\right) \cdot \left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \cdots \cdot \left(\frac{1}{2} - n\right)}{(n+1)!} (-1)^{n+1} 4^{n+1} \\ &= (-1)^n \cdot 2^{2n} \cdot \frac{\left(-\frac{1}{2}\right) \cdot \left(-\frac{3}{2}\right) \cdot \left(-\frac{5}{2}\right) \cdot \cdots \cdot \left(\frac{1-2n}{2}\right)}{(n+1)!} \end{aligned}$$

$$
\begin{aligned}
&= 2^n \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{(n+1)!} \\
&= 2^n \cdot \frac{n!}{n!} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{(n+1)!} \\
&= \frac{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)}{n!} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{(n+1)!} \\
&= \frac{(2n)!}{n!(n+1)!} \\
&= \frac{1}{n+1}\binom{2n}{n}.
\end{aligned}
$$

□

Note the equivalent forms:

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

There are many other ways to derive this closed form, including clever combinatorial proofs to count the Dyck paths.

3 Seven Catalan Objects

In this second part of the text, we will detail six more Catalan objects and prove they are counted by the Catalan numbers. We will do so by describing bijections between these new objects and objects which we already know are Catalan objects.

Before beginning, we note that the objects we discuss satisfies a stronger version of the Catalan recurrence that leads to naturally occurring relationships. Specifically, consider a Catalan object $A_0, A_1, A_2, \dots$ where for all nonnegative integers n, A_n has cardinality C_n. The objects we discuss satisfy the following relation: for all positive integers n, there exists a partition of A_n into the sets $S_0, S_1, S_2, \dots, S_{n-1}$ such that S_k is in bijection with $A_k \times A_{n-k-1}$. Notice that the Dyck paths satisfy this relation, which we showed while proving the Catalan recurrence.

If we have two Catalan objects, $A_0, A_1, A_2, \dots$ and $B_0, B_1, B_2, \dots$ that satisfies this relation, we can use it to inductively find a bijection between them. While the fact that A_n and B_n have the same cardinality implies that there exists a bijection between them, it may not be obvious how to construct this bijection. But the Catalan recurrence gives us a way to do so inductively.

Say for a positive integer n, we have bijections $f_k : A_k \to B_k$ for all integers $0 \le k < n$. By our conditions, there exists some partition of A_n into sets $S_0, S_1, S_2, \dots, S_{n-1}$ such that for all integers $0 \le k \le n-1$, S_k is in bijection with $A_k \times A_{n-k-1}$. Let $T_0, T_1, T_2, \dots, T_{n-1}$ be the analogous sequence for B. First, we can construct a bijection from $A_k \times A_{n-k-1}$ to $B_k \times B_{n-k-1}$ by combining the bijections f_k and f_{n-k-1}. Then, we can compose the bijection from S_k to $A_k \times A_{n-k-1}$, the bijection from $A_k \times A_{n-k-1}$ to $B_k \times B_{n-k-1}$, and the bijection from $B_k \times B_{n-k-1}$ to T_k to obtain a bijection from S_k to T_k. Finally, we can combine the bijections from S_k to T_k for all k to construct the bijection from A_n to B_n.

When we use this technique beginning with the bijection from A_0 to B_0 (there is only one, since both have $C_0 = 1$ element), we obtain a unique bijection from A_n to B_n. Such bijections stem wholly from the inherent relationship between these objects to the Catalan numbers and can thus be considered the natural relationship between these two Catalan objects. All of the bijections we describe in this section are these natural bijections, and, in this author's opinion, they have a great deal of elegance.

After Dyck paths, the second Catalan object we will study is the following:

Definition 3.1. For any nonnegative integer n, a *ballot sequence* is a sequence $a = (a_1, a_2, a_3, \dots, a_{2n})$ (or simply the empty set for $n = 0$) such that each term is either 1 or -1, there are exactly n 1's in the sequence

$$1-1-1- \qquad 1-11-- \qquad 11--1-$$

$$111--- \qquad 11-1--$$

Figure 3: The five ballot sequences with 6 terms. Here, the -1's are denoted $-$ for clarity.

and n -1's in the sequence, and, for all $0 \leq k \leq 2n$:

$$a_1 + a_2 + a_3 + \cdots + a_k = \sum_{i=1}^{k} a_i \geq 0.$$

This is known as a partial sum, and this condition is the same as saying that the partial sums are nonnegative.

For several examples of ballot sequences, see Figure 3.

The name "ballot sequences" come from the following question, which the ballot sequences answer:

Question. *A local election has $2n$ voters and two candidates, A and B. Assuming there are exactly n votes for candidate A and n votes for candidate B, how many sequences of ballots are there such that candidate A is always tied or in the lead?*

This is, of course, a ballot sequence: if we let 1 be a vote for A and -1 a vote for B, then the condition that A is tied or in the lead is equivalent to the partial sums being nonnegative.

This property of the ballot sequences should remind us of the Dyck paths. Indeed, this is how we will show the ballot sequences are a Catalan object:

Proposition 3.1. *The number of ballot sequences with length $2n$ is C_n.*

Proof. Take any Dyck path from $(0,0)$ to (n,n). We will construct a ballot sequence from this path. Scan the path from beginning to end. Every time we take a step in the form $(1,0)$, to the right, write down a 1. Every time we take a step in the form $(0,1)$ upward, write down a -1. Since there are $2n$ steps in the path, we have written down $2n$ integers, all of which are either 1 or -1. In fact, there are exactly n right steps and n up steps needed to go from $(0,0)$ to (n,n), so there will be exactly n 1's and n -1's. Now, at every point in the Dyck path, the number of up steps we have taken cannot exceed the number of right steps or we will be above the diagonal. Thus, the number of 1's is at least the number of -1's, so the partial sums of this sequence is nonnegative. Thus, this is a ballot sequence.

We can now construct a mapping from the Dyck paths from $(0,0)$ to (n,n) to the ballot sequences of length $2n$ that maps every Dyck path to this constructed ballot sequence. In order to show that this mapping is bijective, we will construct an inverse map; that is, a map that reverses this process.

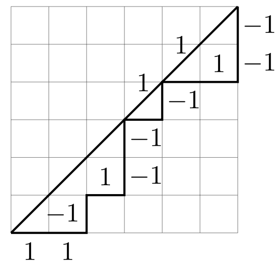
Take any ballot sequence with length $2n$. Scan the sequence from left to right. Starting from $(0,0)$, take a right step whenever there is a 1 in the sequence, and take an up step whenever there is a -1 in the sequence. Since there are exactly n 1's and n -1's in the ballot sequence, the lattice path will end at (n,n). Since the partial sums are nonnegative, at every point on the path, there will be at least as many right steps as there are up steps, meaning that the path will be on or below the diagonal. Thus, this is a Dyck path.

It is left as an exercise to the reader to verify this is an inverse of the previous mapping. □

See Figure 4 for a diagram of this bijection.

Ballot sequences are a fairly simple Catalan object, and not all the Catalan objects we study will be as straightforward:

Definition 3.2. A *plane tree* P on a set V of vertices is either $P = v$ if $V = \{v\}$ or a tuple $P = (v, P_1, P_2, P_3 \ldots, P_n)$ for a positive integer n. In this case, $P_1, P_2, P_3, \ldots, P_n$ are each respectively plane trees on the sets $V_1, V_2, V_3, \ldots, V_n$, which partition $V \backslash \{v\}$. We call v the root of the plane tree, and P_i the ith subtree of v. If P_i has root v_i then we say v *and* v_i *have an edge between them*, v_i *is a child of* v, and v *is the parent of* v_i.



(a) A Dyck path from (0, 0) to (6, 6) labeled with 1's on the right steps and −1's on the left steps.

$11 - 1 - -1 - 11 - -$

(b) The ballot sequence of length 12 constructed from the Dyck path on the left.

Figure 4: An illustration of the bijection between the Dyck paths from (0, 0) to (6, 6) and the ballot sequences of length 12.

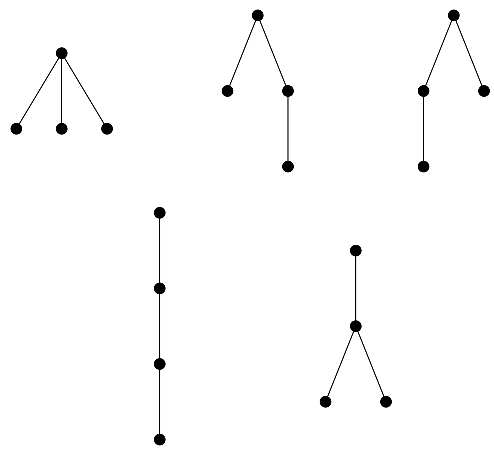


Figure 5: The five plane trees with 4 vertices.

This is a recursive definition, which amounts to saying that a plane tree has a root vertex which has some (possibly zero) number of children, each of which are roots of other plane trees. Notice that the order of the children is important: a plane tree (v, P_1, P_2) is not equal to the plane tree (v, P_2, P_1) when $P_1 \neq P_2$.

We can also describe plane trees visually with dots and lines, as with graphs and other types of trees. The vertices are represented by dots, and two vertices with an edge between (a parent-child relationship) are connected by a line, known unsurprisingly as edges. The root is always listed at the top of the tree. The ith child of a vertex is the ith child from the left in the diagram. See Figure 5 examples of these plane trees in this visual representation.

Plane trees are often defined as a type of graph-theoretic tree. For the purposes of this paper, however, this recursive definition satisfies our needs.

Plane trees are, of course, another Catalan object, and we will define a bijection from plane trees to ballot sequences. In order to describe this bijection, however, we first need the notion of a preorder:

Definition 3.3. Given a plane tree P, the *depth first order* or *preorder* $\text{ord}(P)$ is defined as follows. If $P = v$ is a single vertex, then $\text{ord}(P) = v$. Otherwise, we have $P = (v, P_1, P_2, \ldots, P_n)$, and we define:

$$\text{ord}(P) = v, \text{ord}(P_1), v, \text{ord}(P_2), v, \ldots, v, \text{ord}(P_n), v \text{ (concatenation)}.$$

By concatenation, we mean that if $\alpha = a_1, a_2, a_3, \ldots, a_m$ and $\beta = b_1, b_2, b_3, \ldots, b_n$, then the concatenation α, β is

$$a_1, a_2, a_3, \ldots, a_m, b_1, b_2, b_3, \ldots, b_n.$$

We can think of preorder as the order in which an ant would encounter each vertex of the plane tree if the ant starts from the root and walks counterclockwise around outside the plane tree. If the plane tree

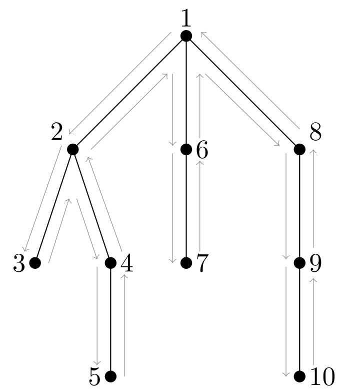


Figure 6: An illustration of the preorder of the plane tree. The ant starts at the root vertex 1 and follows the gray arrows around the outside of the tree. The order in which they see the vertices of the tree is $1, 2, 3, 2, 4, 5, 4, 2, 1, 6, 7, 6, 1, 8, 9, 10, 9, 8, 1$, which is the preorder. We can also use the recursive definition to find this preorder.

is $P = (v, P_1, P_2, P_3, \ldots, P_n)$, then the ant must first travel from v to the root of P_1, then walk around P_1 to come back to v, then walk around P_2 to come back to v, and so on, until it finishes walking around P_n and comes back to v for the final time. Notice that this is consistent with the recursive definition of the preorder. See Figure 6 for an example of this walk and a preorder. We often refer to the preorder as a traversal, imagining ourselves being this ant and walking along this tree.

The preorder allows us to describe elegant bijection between plane trees and ballot sequences.

Proposition 3.2. *The number of plane trees with $n + 1$ vertices is C_n.*

Proof. Take a ballot sequence with $2n$ terms, for any nonnegative integer n. We will construct a plane tree from this sequence using the following algorithm.

Begin with a plane tree with only a root vertex. For every 1 in the ballot sequence, construct a child of the vertex we are currently on such that it is the rightmost child. Then, let us traverse down to this new vertex.

Now, for every -1 in the ballot sequence, let us traverse up to the parent of the vertex we are currently on. This is possible since for each -1 in the ballot sequence, the current vertex will have a parent. We can show this through a clever counting argument: let d be the number of steps from the root to the current vertex. For each 1 in the ballot sequence, d is incremented by 1 as we travel one step away from the root. For each -1, d is decremented by 1 as we travel one step closer to the root. Since d starts at 0, d at every point is equal to the partial sum of the ballot sequence at that point. Thus, if -1 is the next term in the ballot sequence d must be positive or else we have a negative partial sum. That means $d > 0$ which means it is not the root vertex, so it must have a parent.

This algorithm will construct a plane tree, of course. The tree will also have $n + 1$ vertices, since we create n vertices for each of the n 1's in the ballot sequence in addition to the root, which we start with.

To show this is a bijection, we will describe an inverse of this mapping. Given a plane tree P with $n + 1$ vertices, we can construct a ballot sequence as follows: traverse the plane tree in preorder. Every time we step from a parent to a child, write down a 1. Every time we step up from a child to a parent, write down a -1.

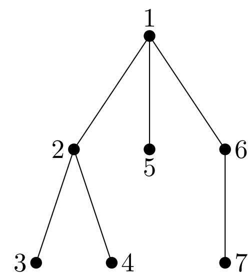
Now we show that this is a ballot sequence of length $2n$. Since each vertex except the root has a unique edge that leads to its parent and there are $n + 1$ vertices, there must be n edges. In the preorder, we traverse each edge exactly twice: once going down the edge, and once going up the edge. So, the sequence will have exactly n 1's and n -1's. Next, in order to show that the partial sums of this sequence are nonnegative, consider the distance between the current vertex and the root. We start at the root, which has distance 0, and with an empty sequence, which has partial sum 0. Whenever we go down an edge, we increase the distance by 1 and write down a 1, and whenever we go up an edge, we decrease the distance by 1 and write down a -1. So, the distance and the partial sum starts at the same number and changes identically, meaning that they are equal. Since the distance is always nonnegative, the partial sums are all nonnegative as well.

It is again left as an exercise to the reader to verify that these two mappings are reverse each other. □

See Figure 7 for a diagram of this bijection.

$$11-1--1-11--$$

(a) A ballot sequece of length 12.



(b) The plane tree of 7 vertices constructed from the ballot sequence on the left. The vertices are labeled the order of their creation.

Figure 7: An illustration of the bijection between ballot sequences of length 12 and plane trees of 7 vertices.

A related type of tree is the binary tree, which is our next Catalan object:

Definition 3.4. A *binary tree* T on a set of vertices V is either $T =$ if $V =$ or $T = (v, T_1, T_2)$ where $v \in V$ and T_1 and T_2 are binary trees on disjoint partitions of $V\backslash\{v\}$ (not necessarily nonempty). We call v the root. If T_1 is nonempty and has root v_1, then we say v_1 is the left child of v, v is the parent of v_1, and there exists an edge between v and v_2. Likewise, if T_2 is nonempty and has root v_2, we say v_2 is the right child of v, v is the parent of v_2, and there exists an edge between v and v_2. Likewise, we call T_1 the left subtree of v and T_2 the right subtree of v.

Like plane trees, binary trees have many equivalent definitions, though for our purposes this is all we need. Do note that binary trees are not a type of plane tree, since if a vertex in a binary tree has exactly one child, it can either be the left child or the right child. However, there is only one way for a vertex in a plane tree to have one child.

Binary trees have a similar visual dot-line representation as plane trees. Dots represent vertices and edges between them represent parent-child relationships. The root is also always the topmost vertex. Unsurprisingly, left children are written to left of the parent and right children are written to the right of the parent. See Figure 8 some examples of binary trees.

Despite their differences, binary trees and plane trees do share several similar properties. For instance, binary trees also have preorders, which are defined in the same way:

Definition 3.5. Let T be a binary tree, and define $\text{ord}(T)$ to be the preorder of T. If $T =$, let $\text{ord}(T) =$. If $T = (v, T_1, T_2)$, then define

$$\text{ord}(T) = v, \text{ord}(T_1), v, \text{ord}(T_2), v \ (\textit{concatention})$$

where if one of these is , we concatenate nothing.

The preorder of a binary tree is likewise the order in which an ant visits the vertices of the binary tree, if it starts from the root and walks counterclockwise around the outside of the tree. We will see the preorder of binary trees briefly later on.

Binary trees, are, of course, another Catalan object. We will describe a bijection between the plane trees and the binary trees, though it is a more involved process than the bijections we have previously seen.

Proposition 3.3. *The number of binary trees with n vertices is C_n.*

Proof. Take any plane tree P with $n+1$ vertices. Remove all edges between a parent and a child that is not a first child. Then remove the root vertex. For all vertices v, if v has a child remaining, declare that child

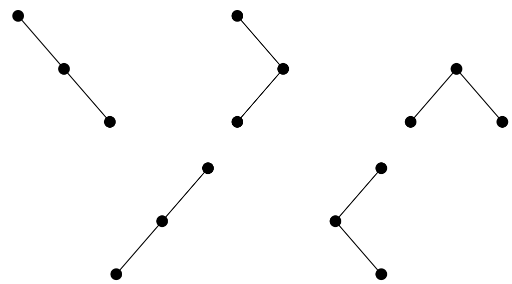


Figure 8: The five binary trees with 3 vertices.

to be the left child of v. Now, add an edge between every vertex and the vertex that was its next sibling, if it exists, and declare that to be its right child. That is, v is a vertex that was originally the ith child of another vertex w, then if w has an $(i+1)$th child u, we we add an edge between v and u and let u be the right child of v. If w does not have an $(i+1)$th child, then we let the right child of v be empty. Finally, declare the root of this new tree to be the first child of the original root of the plane tree; note this vertex has no parent. Every other vertex either has a parent remaining or is a child of the original root, in which case it is now a child of one of its siblings. Since each vertex now has a (possibly empty) left child and right child, we have constructed a binary tree. This binary tree also has n vertices, since we have removed only one vertex.

Thus, we have established a mapping from the set of plane trees with $n+1$ vertices to the set of binary trees with n vertices. Now, we will go the other way and outline the inverse mapping.

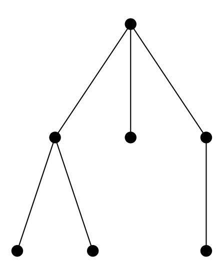
Take any binary tree T with n vertices. First, construct a new binary tree with left subtree T and no right child. Now remove all edges between parents and their right child, if it exists. Now, for any vertices v in the tree, we declare the children of v as follows. First, if v has no left children, we declare v to have no children. If not, let the first child of v be the left child of v. Now for every $i > 1$, we declare the ith child of v to be the right child of the $i-1$the child of v, if the right child exists.

Every vertex v in the tree will thus have a parent, except for the root: if v was originally a left child, it will have a parent. If v was originally a right child, then it will have some ancestor that is the parent of a left child, as the new root is such an ancestor. (By ancestor we mean a vertex w for which there is a sequence of vertices starting at w and ending at v with every vertex a child of the previous.) Then v will become the child of the closest such ancestor.

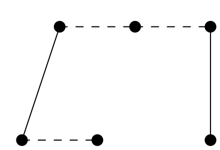
Since each vertex has a parent, there exists a root vertex, and there is an ordering among the children of every vertex, we have constructed a plane tree.

It is likewise left as an exercise to the reader to verify that these mappings reverse each other. ☐

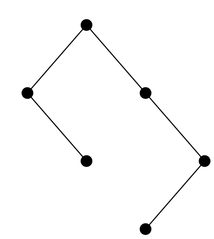
See Figure 9 for an illustration of this process:



(a) The original plane tree.



(b) The plane tree after removing the root and all edges except those between a parent and the first child. The dashed lines connect vertices to their next sibling. The top left vertex is the new root.



(c) The binary tree constructed after converting the dashed edges on the left into edges that connect parents to right children. The first children become left children.

Figure 9: An illustration of the bijection the plane tree with 7 vertices and the binary trees with 6 vertices.

This elegant bijection due to De Brujin and Morselt [2], and like all the bijections we have presented

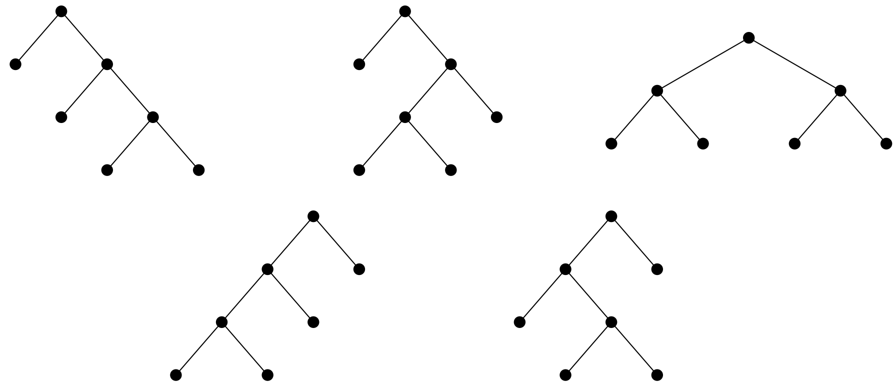


Figure 10: The five complete binary trees with 4 leaves and 3 internal vertices.

here, can be constructed from the Catalan recursion. Notice that since binary trees and complete binary trees are in bijection, there are C_n complete binary trees with $n+1$ leaves.

The next Catalan object we consider is closely related to the binary tree:

Definition 3.6. A *complete binary tree* is a binary tree where every vertex either exactly two children or no children. (No children means both children are the empty set.) We call vertices with no children *leaves* and vertices with exactly 2 children *internal vertices*.

See Figure 10 for some examples of complete binary trees.

Consider a complete binary tree with n internal vertices. Notice that each vertex has exactly one parent, except for the root. So, since there are n internal vertices, there are $2n$ children in the tree, which accounts for all the vertices except for the root. Thus, there are $2n+1$ vertices in the tree, which also means there are $n+1$ leaves.

This fact helps us find a simple bijection between binary trees with n vertices and complete binary trees with $n+1$ vertices:

Proposition 3.4. *The number of complete binary trees with $n+1$ leaves is C_n.*

Proof. Take a binary tree with n vertices. For every vertex in this tree with less than 2 children, add all of the missing children. These new vertices will have no children, and all of the old vertices will have exactly 2 children, so this new binary tree is complete. Specifically, the original vertices are internal, and the new vertices are leaves. Thus, this tree has n internal vertices and $n+1$ leaves.

Now, let us construct an inverse map. Take a complete binary tree with $n+1$ leaves. Remove all of the leaves from this tree. We are left with a binary tree with the n internal vertices.

It is again left as an exercise to the reader to verify that these mappings invert one another. □

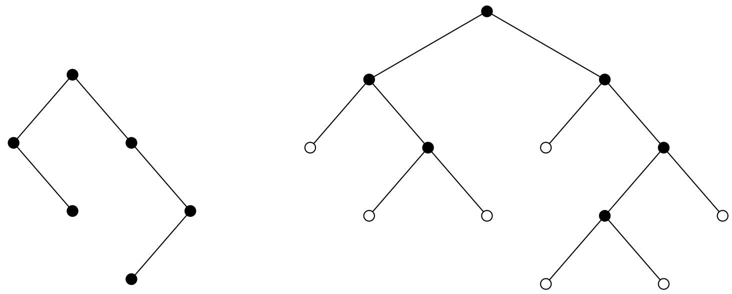
See Figure 11 for an illustration of this bijection.

This simple relationship between binary trees and complete binary trees are one reason why some authors define all of their binary trees to be complete binary trees. In this text, we treat the complete binary trees as a separate Catalan object because it has an elegant relationship with our penultimate Catalan object, parenthesizations:

Definition 3.7. For any nonnegative integer n, a *parenthesization* α of symbols $(a_1, a_2, a_3, \ldots, a_{n+1})$ is $\alpha = a_1$ if $n = 0$. If $n > 0$, then $\alpha = (\beta, \gamma)$, where β is a parenthesization of $(a_1, a_2, a_3, \ldots, a_k)$ and γ is a parenthesization of $(a_{k+1}, a_{k+2}, a_{k+3}, \ldots, a_{n+1})$ for some integer $1 \le k \le n+1$.

A parenthesization is essentially a way of bracketing the symbols $(a_1, a_2, a_3, \ldots, a_{n+1})$ in that order. We often omit the commas in a parenthesization. See Figure 12 for some examples of parenthesizations. Notice that a parenthesization of $n+1$ symbols uses n pairs of parentheses, as each pair of parentheses pairs two terms into one term, reducing the number of terms by 1. Since we end up with 1 terms, we must have $(n+1) - 1 = n$ pairs of parentheses.

Parenthesizations are closely related to complete binary trees.



(a) A binary tree with 6 vertices.

(b) A complete binary tree with 7 leaves constructed from the binary tree on the left. The constructed vertices (and leaves) are denoted with white dots.

Figure 11: A diagram of the bijection from binary trees with 6 vertices to complete binary trees with 7 leaves.

$(a(b(cd)))$ $(a((bc)d))$ $((ab)(cd))$

$(((ab)c)d)$ $((a(bc))d)$

Figure 12: The five parenthesizations of a, b, c, and d.

Proposition 3.5. *The number of parenthesizations of $n+1$ symbols is C_n.*

Proof. Let us take any complete binary tree with $n+1$ leaves, and any list of symbols $(a_1, a_2, a_3, \ldots, a_{n+1})$. Now, consider the preorder of this tree. Label the leaves with the symbols $a_1, a_2, a_3, \ldots, a_{n+1}$ in order of their first appearance in the preorder. That is, label the ith vertex to appear in the preorder with a_i. Now, label each invernal vertex recursively as follows: if the vertex has left child α and right child β, where α and β are some parenthesized expressions, label the vertex (α, β).

Now, suppose for contradiction, we cannot label some vertices recursively in this way. Then let us take a vertex v with the maximum distance from the root for which this is the case. Vertex v must be internal as all the leaves have labels. Now consider the two children of v: they have a greater distance from the root, and so are labeled. Let the left child have label α and the right child have label β; we can then label v (α, β), which contradicts the fact that v cannot be labeled. Thus, all vertices in the tree can be labeled.

Now, take the parenthesized expression we obtain from the label of the root. This will have all $n+1$ symbols. In fact, it will also have them in the order $a_1, a_2, a_3, \ldots, a_{n+1}$ since the leaves of the tree were labeled in order of appearance in the preorder, which is from left to right. Thus, we have constructed a parenthesization.

Now for the inverse map. Take any parenthesization α of $a_1, a_2, a_3, \ldots, a_{n+1}$. We will construct a complete binary tree with n vertices from α. Start with a root, and label it α. Now for every vertex in the tree with label (β, γ), where β and γ are parenthesized expressions, construct a left child with label β and right child with label γ.

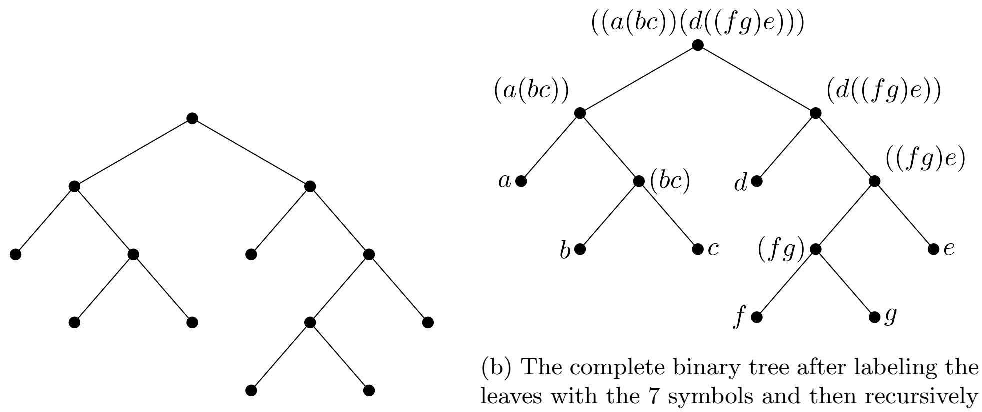
Notice that this is a complete binary tree: either a vertex can be written as (β, γ) in which case it has exactly two children and is an internal vertex, or it cannot in which case it has no children and is a leaf. So we have one internal vertex for every pair of parentheses, of which there are n, so we have n internal vertices and $n+1$ leaves.

It is once again left as an exercise to the reader to show that these two mappings invert each other. □

See Figure 13 for an illustration of this bijection.

Our final Catalan object is the set of triangulations of a convex polygon:

Definition 3.8. A triangulation of a convex polygon is a set of lines connecting vertices of the polygon such that no two lines intersect within the polygon, and the lines divide the polygon into triangles.



(a) A complete binary tree with 7 vertices.

(b) The complete binary tree after labeling the leaves with the 7 symbols and then recursively labeling the internal vertices.

Figure 13: An illustration of the bijection between binary trees and parenthesizations of a, b, c, d, e, f, and g. The label of the root of (b) is the parenthesization constructed from the binary tree.

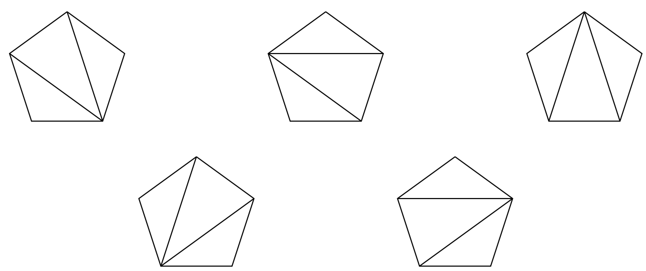


Figure 14: The five triangulations of a pentagon.

See Figure 14 for some examples of triangulations.

In a triangulated n-gon, all of the triangles in the polygon will have their vertices among the vertices of the polygon. This means that the sum of all of the angles of the triangles in the n-gon is equal to the sum of the angles of the n-gon. If we have m triangles, this sum will be $180m^\circ = 180(n-2)^\circ$. Thus, $m = n-2$, so all triangulations of an n-gon form $n-2$ triangles. In order to produce these $n-2$ regions, we will need $n-3$ lines, one for each additional region after the n-gon itself.

These facts help us prove the following:

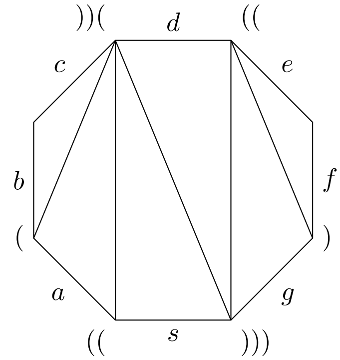
Proposition 3.6. *The number of triangulations of an* $(n+2)$*-gon is* C_n*.*

Proof. Take any parenthesization of the $n+1$ symbols $a_1, a_2, a_3, \ldots, a_{n+1}$, and fix a side s of the $(n+2)$-gon. In counterclockwise order, label the edges of the $(n+2)$-gon with the $n+1$ symbols, beginning by letting a_1 be the edge directly counterclockwise to e and a_{n+1} be the edge directly clockwise to e. Label the vertices $0, 1, 2, \ldots, n+2$ in counterclockwise order where 0 is the vertex directly counterclockwise to e and $n+2$ the vertex directly clockwise to e. Now, we draw a line between vertices i and j with $i < j$ if there exists a matching pair of parentheses that opens directly before a_i and closes directly after a_j.

Next, we verify this is a legal triangulation. If one parenthesis opens before another one, then the second parenthesis will close before the second pair. This means that the line formed by the second pair of parentheses will terminate before the line formed by the first pair. Thus, the parentheses will not cross. We also need to show that the lines divide the $(n+2)$-gon into n triangles. In order to do this, we will show that we have constructed exactly $n-1$ lines. First, notice that no two pairs of parentheses can open and close in the same places, so all the lines are distinct. Next, notice that there will be a matching pair of parentheses before a_1 and after a_{n+1}, which will coincide with the fixed edge e. This is the only line that will coincide in such a way: if another line coincides with a side of the polygon a_i, that would mean that there is a pair of parentheses around a_i such as (a_i), which is impossible. Thus, we have constructed exactly $n-1$ legal lines, meaning we have divided this $(n+2)$-gon into n triangles.

$((a(bc))(d((ef)g)))$

(a) A parenthesization of a, b, c, d, e, f, and g.



(b) The triangulation of an octagon constructed from the parenthesization on the left. The vertex numbers are omitted here for clarity and are instead replaced by the parentheses between the symbols.

Figure 15: An illustration of the bijection between parenthesizations of a, b, c, d, e, f, and g and triangulations of an octagon.

For the inverse map, consider a triangulation of an $(n + 2)$-gon with the same fixed side s. Let $a_1, a_2, a_3, \ldots, a_{n+1}$ be the symbols we wish to parenthesize, and in that order. Label the edges of the $(n + 2)$-gon $a_1, a_2, a_3, \ldots, a_{n+1}$ with a_{n+1} and the vertices $0, 1, 2, \ldots, n + 1$ as before. Now, if there is a line between vertices i and j where $i < j$, put a matching pair of parentheses with the open parenthesis directly before a_i and the close parenthesis directly after a_j. Then put a pair of parentheses around the whole expression.

This is a valid parenthesization as there are n pairs in total, no pairs of parentheses overlap, no parentheses open and close in the same place, and there are no parentheses around a single symbol. □

See Figure 15 for an example of this bijection.

4 Conclusion

It is the hope of the author that, after reading this article, the reader now has a foundation for further study of the Catalan numbers, if they so wish. Indeed, there is much to the Catalan numbers beyond this paper: for instance, even among the seven objects discussed here, only seven of the $\binom{7}{2} = 21$ naturally occurring bijections have been described. There are many more Catalan objects to be studied, and if the reader is still curious beyond the contents of this paper, they are suggested to refer to Richard P. Stanley's *Catalan Numbers* [1] or his *Catalan Addendum*[3] available on his website.

References

[1] Richard P. Stanley. *Catalan Numbers.* Cambridge University Press, 32 Avenue of the Americas, New York, NY, 2015.

[2] N. G. de Brujin and B. J. M. Morselt. A note on plane trees. *Journal of Combinatorial Theory*, 2(1):27–34, 1967.

[3] Richard P. Stanley. Catalan addendum. http://www-math.mit.edu/ rstan/ec/catadd.pdf, 2013. Accessed 2021-07-11.