The Catalan Numbers

by Arnav Kumar

1 Background Knowledge

Definition 1. *Given sets A and B. Additionally, the* ***set difference*** *$A\setminus B$ is the set of everything in A which is not in B.*

Exercise 1. **($\star$)** Given sets A, B, the set $A\cap B$ is the **intersection** between A and B and contains everything which is in both A and B. Show that A can be partitioned into $A\setminus B$ and $A\cap B$ so that everything in A is in exactly one of $A\setminus B$ and $A\cap B$, and nothing is contained either $A\setminus B$ or $A\cap B$ which is not in A.

Definition 2. *A* ***bijection*** *between two sets A and B is a map from set A to set B, $f: A\to B$, such that for any $b\in B$ there is a unique $a\in A$ satisfying $f(a)=b$. Any $f: A\to B$ is* ***bijective*** *if it is a bijection.*

Exercise 2. **($\star$)** Show that if $f: A\to B$ is a bijection, then it is invertible. Being invertible means that there is a function $g: B\to A$ such that for any $a\in A$ and any $b\in B$ we have $f(g(b))=b$ and $g(f(a))=a$.

Notation 1. *If S is a set with a finite number of elements in it, we use $|S|$ to notate the number of elements in S.*

Remark. This idea can be generalized to any set, but is not easy to define and thus omitted. If you are interested, it is called the **cardinality** of a set.

Proposition 1. *Let A and B be sets (not necessarily finite). If there is a $f: A\to B$ which is bijective, we have that $|A|=|B|$.*

Exercise 3. **($\star$)** Show that $|[0,1]|=|[0,7]|$ where $[a,b]$ is the set of all real numbers x satisfying $a\leq x\leq b$.

Exercise 4. **($\star$)** Show that $|(0,1)|=|(1,\infty)|$ where (a,b) is the set of all real numbers x satisfying $a<x<b$ and we have that every $x\in\mathbb{R}$ satisifes $x<\infty$. (Note here that ∞ is not a real number, but is a part of what is called the extended reals, $\bar{\mathbb{R}}$).

Definition 3. *For $n\in\mathbb{N}=\{0,1,2,\ldots\}$ we define* ***n factorial*** *to be $n!=\begin{cases}1 & \text{if } n=0\\ n\times(n-1)! & \text{otherwise}\end{cases}$.*

Exercise 5. **($\star\star$)** Show that the number of ways to order n distinct objects is $n!$.

Exercise 6. **($\star\star$)** The number of ways to choose k objects from a collection of n distinct objects where the selection order matters is called ***n* permute *k*** (sometimes denoted ${}_nP_k$). The **falling factorial** of n and k is $n^{\underline{k}}=n\times(n-1)\times\cdots\times(n-k+1)$ and is sometimes denoted $(n)_k$. Show that when n and k are non-negative integers with $k\leq n$, we have that $n^{\underline{k}}=\frac{n!}{(n-k)!}$ is the same as n permute k.

Notation 2. *We use $\binom{n}{k}$ (or sometimes ${}_nC_k$) to notate* ***n choose k****, the number of ways to choose k objects from a collection of n distinct objects where selection order does not matter.*

Exercise 7. **($\star\star$)** Show that $\binom{n}{k}=\frac{n^{\underline{k}}}{k!}=\frac{n!}{k!\,(n-k)!}$.

2 Motivation for the Catalan Numbers

Problem 1. We are playing a game on the Cartesian plane. Every move, if we were at the position (x, y) we may move to either $(x+1, y)$ or $(x, y+1)$. For $n \in \mathbb{N}$ arbitrary, how many distinct paths are there from $(0,0)$ to (n, n) given that we never move above the line $y = x$?

I choose to share the following solution because it is perhaps the easiest elementary solution which still provides a closed form solution, rather than a recurrence relation.

Solution. Let's represent a path as a finite sequence of moves. A move is one of U for up or R for right. Additionally, if S is a path, we define $S(x, y)$ to be the point we end up at as a result of following the path S starting at (x, y). By convention, if $k = 0$, we'll take $S(x, y) = (x, y)$. Let $\#_U(S)$ and $\#_R(S)$ be the number of Us and Rs in our path respectively.

Let T be the set of all paths S such that $S(0,0) = (n, n)$. Additionally, define $A \subseteq T$ to be all of those paths which have a point on the path above $y = x$.

Note that the value we are searching for is $|T \setminus A| = |T| - |A|$ since T and $A \subseteq T$ are finite. This is because we want the number of paths from $(0,0)$ to (n, n) which are never above $y = x$.

Further, we know that for any $S \in T$, we have $\#_R(S) = \#_U(S) = n$. Thus we see that $|T| = \binom{2n}{n}$ which is the number of ways to choose which of the $2n$ moves is U (the rest have to be Rs).

For any $S \in A$, let's suppose that move l was the first move that brought us above $y = x$. Since it is the first, we know that after move l, we are on the line $y = x + 1$. We split up S into S_0 and S_1 which are the parts of S before reaching $y = x + 1$ for the first time, and after reaching $y = x + 1$ for the first time. We see that $\#_R(S_0) = 1 + \#_U(S_0)$. Additionally, $\#_R(S_0) + \#_R(S_1) = \#_R(S) = \#_U(S) = \#_U(S_0) + \#_U(S_1)$ which means that $\#_R(S_1) + 1 = \#_U(S_1)$.

Now take $\widetilde{S}_1$ to be the path S_1 but with every R now a U and vice versa. We see that $\#_R(\widetilde{S}_1) = \#_U(S_1) = 1 + \#_R(S_1) = \#_U(\widetilde{S}_1)$. Thus, we see that the path S' (which is the combination of the paths S_0 and $\widetilde{S}_1$) satisfies $\#_R(S') = 2 + \#_U(S')$. This means that $S'(0,0) = (n-1, n+1)$ since path S' has the same length as S. Furthermore, every path from $(0,0)$ to $(n-1, n+1)$ must pass through $y = x + 1$, so can demonstrate a bijection between paths from $(0,0)$ to $(n-1, n+1)$ and those in A (if it's not obvious to you, see if you can figure out what the bijection is).

This means $|A| = \binom{2n}{n+1}$ which is the number of ways to choose which of the $2n$ terms in the sequence are going to be Us.

Overall, $|T \setminus A| = \binom{2n}{n} - \binom{2n}{n+1} = \binom{2n}{n}\left(1 - \frac{n}{n+1}\right) = \frac{1}{n+1}\binom{2n}{n}$. Q.E.F.

3 The Catalan Numbers

Definition 4. *For $n \geq 0$, we define the n^{th}* ***Catalan number*** *to be* $C_n := \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{n!\,(n+1)!}$.

Remark. Note that C_n is the solution to the previous problem. By finding bijections between the set we are counting in the first problem and sets we are counting in others, we can find alternate, equivalent characterizations of the Catalan numbers.

Exercise 8. ($\star$) A well formed bracket expressions is an example of a **Dyck word**. Show that the number of well-formed bracket expressions with n bracket pairs is C_n.

Exercise 9. ($\star\star$) Based on the interpretation of Catalan numbers as in the previous exercise, show that we can find an recursive formula for C_n of the form $C_n = C_0 C_{n-1} + C_1 C_{n-2} + \cdots + C_{n-1} C_0 = \sum_{i=1}^{n} C_{i-1} C_{n-i}$.

Exercise 10. ($\star\star$) Show that the number of ways to triangulate a convex $(n+2)$-gon (split it into n triangles) is C_n.

Exercise 11. ($\star$) Show C_n is the number of ways to draw n non-crossing diagonals between the $2n$ verticies of a convex $2n$-gon such that every vertex is only part of 1 diagonal.

Exercise 12. ($\star\star\star$) Again, we are playing the game on the Cartesian plane introduced earlier. We define the **exceedence** of a path from $(0,0)$ to (n, n) as the number of times we move U given that we were on or above the line $y = x$ initially. For any $k \in \{0, \ldots, n\}$ show that the number of distinct paths from $(0,0)$ to (n, n) with exceedence k is C_n.