Catalan Numbers

Gauranga Kumar Baishya

IIT Madras, Tezpur University

gaurangabaishya15@gmail.com

June 25, 2021

Legal Strings

legal strings: definition

Let us call a 'legal string' of parentheses, the one with the property that, as we move along the string from left to right, we never will have seen more right parentheses than left.

$n = 3$

For $n = 3$, i.e. for 3 pairs of parentheses, it is easy to check that there are exactly 5 legal strings:

$$((()));\quad (()());\quad (())();\quad ()()();\quad ()(()) \tag{1}$$

counting legal strings

Legal Number ($= k$)

For any legal string, the integer, k is the smallest positive integer such that the first $2k$ characters of the legal string form a legal string! We say that a legal string of $2n$ parentheses is *primitive* if $k = n$. For instance, the first two strings of (1) are *primitive*. Both have $k = 3$.The value of LN ($= k$) associated with each strings in (1) are 3, 3, 2, 1, 1 respectively.

Let us ask a question now

How many legal strings of $2n$ parentheses will have a given legal number, say k ?

If ω be a legal string of $2n$ parentheses that have a given legal number, say k, then the first 2k characters of ω will not only be a legal string but will be a *primitive* string. But then again a question arise:

The Question

In how many ways can we choose the first $2k$ characters of ω, or in other words, how many *primitive* strings of length $2k$ are there?

The last $2n - 2k$ characters of ω, on the other hand will form an arbitrary legal string. Let $f(n)$ be the total number of legal strings of n pairs of parentheses for $n \geq 0$. $f(0) = 1$. Since it is arbitrary & legal, there are exactly $f(n-k)$ ways to choose the last $2n - 2k$ characters.

A lemma and its proof

Lemma

If $k \geq 1$ & $g(k)$ is the number of primitive legal strings of length $2k$, and $f(k)$, the number of all legal strings of $2k$ parentheses, then

$$g(k) = f(k-1) \tag{2}$$

Proof

Let ω be any legal string of $k-1$ pairs of parentheses. If we put a left parenthesis in the beginning and a right parenthesis in the end of ω then in the new legal string, $\omega^{'}$ we wont find that the parentheses that have been opened getting closed until the last parenthesis. Hence $\omega^{'}$ is a primitive legal string of length $2k$. Conversely, considering a primitive legal string, $\omega^{'}$ of length $2k$, then deleting the initial left and the terminal right parenthesis would result in an arbitrary legal string, ω of length $2k-2$.

The total number of legal strings counted recursively

Result

The total number of legal strings of $2n$ characters that have a given Legal Number, k is $g(k)f(n-k) = f(k-1)f(n-k)$, using (2). Since every legal string has an unique Legal Number, k, we have,

$$f(n) = \sum_k f(k-1)f(n-k) \qquad (n \neq 0; f(0) = 1) \tag{3}$$

We will use the above result (3) to find the generating function for the numbers counted by $f(n)$.

Cauchy product of two infinite power series

Theorem

Consider the following two power series $\sum_{i=0}^{\infty} a_i x^i$ *and* $\sum_{j=0}^{\infty} b_j x^j$. The Cauchy product of these two power series is (defined) as follows:

$$\left(\sum_{i=0}^{\infty} a_i x^i\right) \cdot \left(\sum_{j=0}^{\infty} b_j x^j\right) = \sum_{k=0}^{\infty} c_k x^k \quad \text{where} \quad c_k = \sum_{l=0}^{k} a_l b_{k-l}.$$

...continued

The result revisited

$$f(n) = \sum_k f(k-1)f(n-k) \qquad (n \neq 0; f(0) = 1) \tag{4}$$

Observation

So, let $F(x) = \sum_{k \geq 0} f(k)x^k$ be the opsgf of $\{f(n); n \geq 0\}$.It is now easy to observe that the RHS of (4) is the coefficient of x^n in the Cauchy product of the power series $F(x)$ and the power series $\sum_{k \geq 0} f(k-1)x^k$. The later series is nothing but $xF(x)$. Therefore, if we multiply both the sides of (4) by x^n and sum over all the terms and conditions that $n \neq 0; f(0) = 1$, the LHS becomes $F(x) - 1$ and the RHS becomes $xF(x)^2$.

Hence we have,

$$F(x) - 1 = xF(x)^2 \tag{5}$$

Solving for $F(x)$

The equation revisited

Solving for quadratic equation in $F(x)$, we get,

$$F(x) = \frac{1 \pm \sqrt{1-4x}}{2x} \tag{6}$$

Now, if we choose the '+' signed solution, then $\lim_{x\to 0} F(x) = \infty \neq 1$. But if we choose the '-' sign, then using L'Hospital's rule one can easily show that $\lim_{x\to 0} F(x) = \lim_{x\to 0} \frac{1-\sqrt{1-4x}}{2x} = 1 = f(0)$.

F(x) is one of the most celebrated generating functions in combinatorics.

Finding the explicit formula for the numbers $f(n)$

Newton's Generalised Binomial Theorem

Before finding the explicit formula for the numbers, $f(n)$, we need to get ourselves acquainted with the Newton's Generalised Binomial Theorem:

Let us define for an arbitrary number r, $\binom{r}{k} = \frac{r(r-1)\cdots(r-k+1)}{k!}$. If x and y are real numbers with $|x| > |y|$, and r is any complex number, then we have, $(x+y)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^{r-k} y^k$.

...continued

Calculating $\binom{\frac{1}{2}}{k}$

$$\binom{\frac{1}{2}}{k} = \frac{\frac{1}{2}(\frac{1}{2}-1)\cdots(\frac{1}{2}-k+1)}{k!} = \frac{(-1)^{k-1}}{2^k}\ \frac{1\times2\times3\times4\times\cdots\times(2k-3)\times(2k-2)}{2\times4\times6\times\cdots\times(2k-2)\times(k!)} = \frac{(-1)^{k-1}}{k\times2^{2k-1}} \times \frac{(2k-2)!}{(k-1)!^2}$$

Calculating $\sqrt{1-4x}$

$$\sqrt{1-4x} = \ (1-4x)^{\frac{1}{2}} = 1 + \sum_{k=1}^{\infty} \binom{\frac{1}{2}}{k}(-4x)^k$$

simplification

$$\frac{(-1)^{k-1}}{k\times2^{2k-1}}\ \binom{2k-2}{k-1}\ (-4x)^k = \frac{(-1)^{2k-1}2^{2k}}{k\times2^{2k-1}}\ \binom{2k-2}{k-1}\ x^k = \text{-}\ \frac{2}{k}\ \binom{2k-2}{k-1}\ x^k$$

$F(x)$

Simplification of $F(x)$

$\frac{1-\sqrt{1-4x}}{2x} = \sum_{k=1}^{\infty} \frac{1}{k}\binom{2k-2}{k-1}x^{k-1} = \sum_{k=0}^{\infty} \frac{1}{k+1}\binom{2k}{k}x^{k} = \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n}x^{n}$

$F(x)$

Hence,

$$F(x) = \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n}x^n = 1 + x + 2x^2 + 5x^3 + 14x^4 + 42x^5 + 132x^6 + \cdots \tag{7}$$

$f(n)$

The Catalan Numbers

$f(n) = \frac{1}{n+1}\binom{2n}{n}$ and these numbers are called the Catalan numbers. It can be verified in (1) that for $n = 3$ pairs of parentheses, there were exactly 5 legal strings. 5 is the coefficient of x^3 in $F(x)$.

Other combinatorial objects counted by the catalan numbers

There are more than 100 combinatorial objects that are counted by the catalan numbers apart from counting the number of legal strings. Another such combinatorial object are the dyck paths.

Dyck paths

A Dyck Path is a series of equal length steps that form a staircase walk from $(0, 0)$ to (n, n) that will lie strictly below, or touching, the diagonal of the $n \times n$ square. We say it then, that it is a "a dyck path of order n".

A brief history of the Catalan numbers

1. Euler, Goldbach & Segner
2. The French school, 1838-1843
3. The British school, 1857-1891
4. The ballot problem
5. Later years

References

- Introductory Combinatorics, fifth edition, Brualdi, Richard A.
- Catalan numbers, Stanley, Richard P. (2015). Cambridge University Press, ISBN 978-1-107-42774-7.
- Enumerative Combinatorics Vol 1, Vol 2, Stanley, Richard P.
- Generating functionology, Herbert Wilf