Лекция 2. Биномиальные коэффициенты

Рассмотрим обобщение задачи о биноме Ньютона: что получается после раскрытия скобок и приведения подобных в выражении

$$(x_1 + \cdots + x_k)^n = \underbrace{(x_1 + \cdots + x_k) \times \cdots \times (x_1 + \cdots + x_k)}_{n \text{ раз}}.$$

Иными словами, сколько будет слагаемых? Какой вид имеют эти слагаемые? Чему равны коэффициенты? Проще всего ответить на второй вопрос: слагаемое имеет вид $\alpha x_1^{i_1} \cdots x_k^{i_k}$, где $i_1 + \cdots + i_k = n$ и α — коэффициент (зависящий от $i_1, \ldots, i_k$).

Найти коэффициент α нетрудно: после раскрытия скобок в выражении выше, но до приведения подобных слагаемое вида $x_1^{i_1} \cdots x_k^{i_k}$ получается при выборе i_1 скобок из n, в каждой из которых мы берем x_1; при выборе i_2 скобок из оставшихся $(n - i_1)$, в каждой из которых мы берем x_2; при выборе i_3 скобок из оставшихся $(n - i_1 - i_2)$, в каждой из которых мы берем x_3, и так далее. В конце у нас останется $n - i_1 - i_2 - \cdots - i_{k-1} = i_k$ скобок, в каждой из которых мы берем x_k. Осуществить такую последовательность выборов можно

$$C_n^{i_1} \times C_{n-i_1}^{i_2} \times C_{n-i_1-i_2}^{i_3} \times \cdots \times C_{i_k}^{i_k}$$

способами. Что равно

$$\frac{n!}{(n-i_1)!\, i_1!} \times \frac{(n-i_1)!}{(n-i_1-i_2)!\, i_2!} \times \frac{(n-i1-i2)!}{(n-i_1-i_2-i_3)!\, i_3!} \times \cdots \times \frac{i_k!}{0!\, i_k!} = \frac{n!}{i_1!\, i_2! \cdots i_k!}.$$

Наиболее трудным является первый вопрос. Мы должны найти сколькими способами число n можно представить в виде суммы k неотрицательных целых чисел $i_1, \ldots, i_k$ (порядок важен!). Другими словами, нам нужно найти число целых неотрицательных решений уравнения $i_1 + \cdots + i_k = n$. Например, если $k = 4$, а $n = 3$, то число решений равно 20:

$$\begin{gathered} 0+0+0+3,\ 0+0+1+2,\ 0+1+0+2,\ 1+0+0+2,\ 0+0+2+1,\ 0+1+1+1,\ 1+0+1+1, \\ 0+2+0+1,\ 1+1+0+1,\ 2+0+0+1,\ 0+0+3+0,\ 0+1+2+0,\ 1+0+2+0,\ 0+2+1+0, \\ 1+1+1+0,\ 2+0+1+0,\ 0+3+0+0,\ 1+2+0+0,\ 2+1+0+0,\ 3+0+0+0. \end{gathered}$$

Это означает, что раскрытие скобок и приведение подобных в выражении $(x_1 + x_2 + x_3 + x_4)^3$ даст нам 20 слагаемых.

Общую формулу можно получить следующим образом. Рассмотрим $n + k - 1$ белых кружков

$$\circ \quad \circ \quad \circ \quad \circ \quad \circ \quad \circ$$

Выберем $k - 1$ из них и закрасим их в черный цвет

$$\circ \quad \bullet \quad \bullet \quad \circ \quad \circ \quad \bullet$$

У нас осталось n белых кружков, разбитых в групп: i_1 белых кружков слева от первого черного, i_2 белых кружков между первым черным и вторым, i_3 белых кружков между вторым черным и третьим, и так далее, i_k белых кружков справа от последнего черного (на рисунке $i_1 = 1$, $i_2 = 0$, $i_3 = 2$, $i_4 = 0$). Так как белых кружков в точности n, то $i_1 + \cdots + i_k = n$. Следовательно, имеется C_{n+k-1}^{k-1} целых неотрицательных решений этого уравнения.

Сформулируем полученный результат.

Теорема 1. *Раскрывая скобки и приводя подобные в выражении $(x_1 + \cdots + x_k)^n$, мы получаем сумму из C_{n+k-1}^{k-1} слагаемых, причем каждое слагаемое имеет вид*

$$\frac{n!}{i_1! \times \cdots \times i_k!}\, x_1^{i_1} \cdots x_k^{i_k},$$

где $i_1 + \cdots + i_k = n$.

Замечание. Естественно сразу пытаться найти число целых неотрицательных решений уравнения $a_1x_1 + \cdots + a_kx_k = n$, но рассуждения, работающие в случае $a_1 = \cdots = a_k = 1$, в общем случае ответа не дают. Решение такой задачи оказывается существенно более сложным. Мы рассмотрим этот вопрос далее.

Замечание. Полученная формула описывает ситуацию *выбора с повторениями*: сколькими способами можно составить набор из n шаров k различных цветов, если шаров каждого цвета неограниченное количество?

Числа Каталана

Существуют десятки определений чисел Каталана. Вот одно из них. Рассмотрим выражение $2+3\times 4-5$. Понятно, что для того, чтобы произвести вычисления в этом выражении, мы должны определить порядок выполнения операций. Сколькими способами это можно сделать? В нашем случае — пятью:

$$\begin{gathered}((2+3)\times 4)-5=15\\ (2+3)\times(4-5)=-5\\ (2+(3\times 4))-5=9\\ 2+((3\times 4)-5)=9\\ 2+(3\times(4-5))=-1\end{gathered}$$

Число Каталана C_{n-1} — это количество способов вычислить выражение $a_1*\cdots*a_n$, где звездочка обозначает операцию, а $a_1,\ldots,a_n$ — некоторые числа. (Мы нашли, что $C_3=5$).

Оказывается, что C_n — это количество „правильных “ скобочных структур из $2n$ скобок (n открывающих и n закрывающих). Скобочная структура называется *правильной*, если в любом левом отрезке структуры количество закрывающих скобок не превышает количества открывающих. Например, структура $((()())())$ — правильная, а структура $(()())())(()$ — нет (в левом отрезке из 9-ти скобок открывающих — 4, а закрывающих — 5).

Правильных скобочных структур из шести скобок — пять. Вот они:

$$((())),\quad (()()),\quad (())(),\quad ()(()),\quad ()()().$$

Докажем сформулированное выше утверждение. Рассмотрим выражение, котором n переменных связаны $(n-1)$-м знаком операций. В этом выражении расставлены скобки (их $2n-2$, включая и самые внешние), которые указывают на порядок выполнения операций. Например, так: $(((a_1*a_2)*a_3)*a_4)$. Такому выражению мы поставим в соответствие скобочную структуру по правилу: мы оставляем открывающие скобки, вместо знаков операций ставим закрывающие скобки, а остальные символы удаляем. Вот как это выглядит для $n=3$:

$$\begin{array}{rcl}(((a_1*a_2)*a_3)*a_4) & \Rightarrow & ((()))\\ ((a_1*(a_2*a_3))*a_4) & \Rightarrow & (()())\\ ((a_1*a_2)*(a_3*a_4)) & \Rightarrow & (())()\\ (a_1*((a_2*a_3)*a_4)) & \Rightarrow & ()(())\\ (a_1*(a_2*(a_3*a_4))) & \Rightarrow & ()()()\end{array}$$

Восстановление выражения по скобочной структуре производится так. Вместо закрывающих скобок ставим знаки операций. Если в полученной строке есть два знака $*$ подряд, то рассмотрим первое появление фрагмента $(**$. Это означает, что в выражении в этом месте стоит фрагмент $(a*a)$. Удалим фрагмент $(*$ и запомним, что в соответствующем месте должен стоять не операнд a, но фрагмент $(a*a)$, продолжим преобразование.

Пример. Преобразуем скобочную структуру $((()())())$.

$$((()())())\Rightarrow (((*(**(**\Rightarrow (((*.*(**$$

Здесь точка указывает, что в этом месте в выражении будет стоять $(a*a)$. Продолжаем.

$$(((*.*(**\Rightarrow ((..*(**$$

Здесь две точки указывают, что в выражении в этом месте будет стоять $(a*(a*a))$. Продолжаем.

$$((..*(**\Rightarrow ((..*.*\Rightarrow (...*$$

Здесь три точки указывают, что в выражении в э том месте будет стоять фрагмент $((a*(a*a))*(a*a))$. Окончательно получаем $(((a*(a*a))*(a*a))*a)$. Обратное преобразование

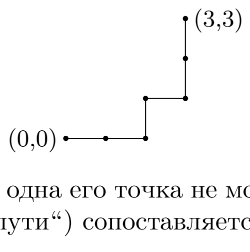
$$(((a*(a*a))*(a*a))*a)\Rightarrow ((()())())$$

дает нам исходную скобочную структуру.

Если же в строке нет двух знаков $*$ подряд, то строка имеет вид $(*(*(*\cdots$ и ей отвечает выражение $(a*(a*(a*(a*\cdots$.

Правильной скобочной структуре из $2n$ открывающих и $2n$ закрывающих скобок мы поставим в соответствие путь в квадрате $[0,n]\times[0,n]$. Путь начинается в точке (0,0) и заканчивается в точке (n,n). Открывающей скобке мы сопоставляем горизонтальный отрезок длины 1, а закрывающей — вертикальный.

$(()())\quad\Longrightarrow$



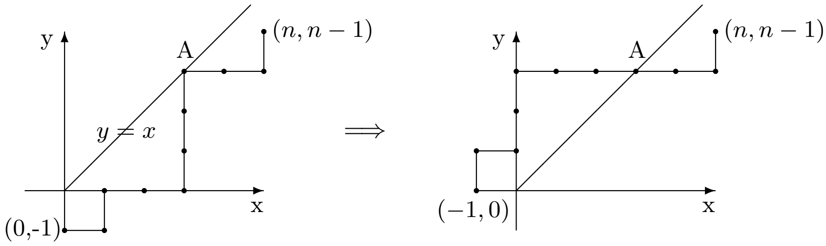
Если путь сопоставлен правильной структуре, то ни одна его точка не может лежать выше главной диагонали квадрата. Обратно, такому пути („правильному пути“) сопоставляется правильная скобочная структура.

Геометрическое представление правильных скобочных структур позволяет найти выражение для чисел Каталана.

Сместим правильный путь на 1 клетку вниз. Теперь правильный путь начинается в точке $(0,-1)$, заканчивается в точке $(n,n-1)$ и не имеет общих точек с прямой $y=x$ — биссектрисой первого квадранта. Нам нужно найти количество правильных путей. Для этого мы найдем количество *неправильных*, и из общего числа путей вычтем количество неправильных.

Мы рассматриваем пути из точки $(0,-1)$ в точку $(n,n-1)$. Длина такого пути равна $2n$ и он содержит n вертикальных сегментов и n горизонтальных. Количество всех таких путей равно числу способов выбрать n вертикальных сегментов из общего числа $2n$ сегментов, т.е. равно C_{2n}^{n}.

Рассмотрим неправильный путь и его первую точку на прямой $y=x$ (точка A). Отрезок пути от точки $(0,-1)$ до точки A заменим симметричным относительно прямой $y=x$. Мы получим путь длины $2n$, идущий из точки $(-1,0)$ в точку $(n,n-1)$.



Такой путь обязательно пересекает прямую $y=x$.

Обратно, пусть нам дан путь длины $2n$ из точки $(-1,0)$ в точку $(n,n-1)$ и пусть A — первая точка этого пути, лежащая на прямой $y=x$. Заменив участок пути от точки $(-1,0)$ до точки A на симметричный относительно прямой $y=x$, мы получим неправильный путь из точки $(0,-1)$ в точку $(n,n-1)$. Следовательно, неправильных путей из точки $(0,-1)$ в точку $(n,n-1)$ столько же, сколько путей из точки $(-1,0)$ в точку $(n,n-1)$. Такой путь длины содержит $n+1$ горизонтальных и $n-1$ вертикальных участков. Поэтому, количество таких путей равно C_{2n}^{n-1}. Значит, количество правильных путей (т.е. число Каталана C_n) равно

$$C_n = C_{2n}^{n} - C_{2n}^{n-1} = \frac{(2n)!}{n!\,n!} - \frac{(2n)!}{(n+1)!\,(n-1)!} = \frac{(2n)!}{n!\,(n+1)!} = \frac{1}{2n+1}\,C_{2n+1}^{n}.$$

В частности,

$$C_4 = \frac{1}{9}\,C_9^4 = 14.$$