组合数学第六讲

授课时间: 2024 年 9 月 30 日　授课教师: 孙晓明　记录人: 赵夕然　葛景源

1 Catalan 数

例 1　有三个柱子, 从左到右分别记为 A, B, C, A 柱子上有 n 个盘子, 盘子从上到下依次进行编号, 为 $1, 2, \cdots n$, 现在要求将 A 柱子上的 n 个盘子全部移动到 C 柱子上, 移动过程中有如下规则:

1. 每次只能移动最上面一个盘子;
2. 盘子只能从 A 柱移动到 B 柱, 或者从 B 柱移动到 C 柱.

请问有多少种移动方法?

解　将此题的答案记作 C_n. 考虑 B 柱为一个栈, 把 A 柱子上的盘移动到 B 柱子上即为进栈, 把 B 柱子上的盘移动到 C 柱子上即为出栈, 所以问题转化为求栈的出栈序列数。考虑 B 柱首次为空时, 编号 1 的盘子一定在 C 柱的顶端, 假设此时 C 柱上有 k 个盘子。当所有盘子都移动到 C 柱上时, 1 号盘子将编号为 $2, \ldots, k$ 的盘子和编号为 $k+1, \ldots, n$ 的盘子分割成了独立的两部分, 而且每部分排列的生成规则跟原问题相同, 通过枚举所有的 k, 可得递推关系式

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

初始条件为 $C_0 = 1, C_1 = 1, C_2 = 2$. 设 $\{C_n\}$ 对应的生成函数为 $C(x)$. 由递推关系式可得

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

解得

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

因为 $C(0) = C_0 = 1$, 我们舍去上式中的正号（只有负号可以让 $C(x)$ 在 $x \to 0$ 时取到 1）. 对 $C(x)$ 中 $(1-4x)^{\frac{1}{2}}$ 应用二项式定理展开, 得

$$\begin{aligned}
C(x) &= \frac{1-(1-4x)^{\frac{1}{2}}}{2x} \\
&= \frac{1}{2x} \sum_{k \geq 1} \frac{1 \cdot 3 \cdots (2k-3)}{2^k \cdot k!} 2^{2k} x^k \\
&= \frac{1}{x} \sum_{k \geq 1} \frac{(2k-3)!!(2k-2)!!}{k!(2k-2)!!} 2^{k-1} x^k \\
&= \frac{1}{x} \sum_{k \geq 1} \frac{(2k-2)!}{2^{k-1}(k-1)!k!} 2^{k-1} x^k \\
&= \frac{1}{x} \sum_{k \geq 1} \binom{2k-2}{k-1} \frac{1}{k} x^k \\
&= \sum_{k \geq 1} \binom{2k-2}{k-1} \frac{1}{k} x^{k-1} \\
&= \sum_{k \geq 0} \binom{2k}{k} \frac{1}{k+1} x^k.
\end{aligned}$$

所以 $C_n = \frac{1}{n+1}\binom{2n}{n}$，此数被称作 Catalan 数.

定义 1 (Catalan 数). *Catalan 数指一个整数数列* $\{1,1,2,5,14,42,\ldots\}$，*其通项公式为*

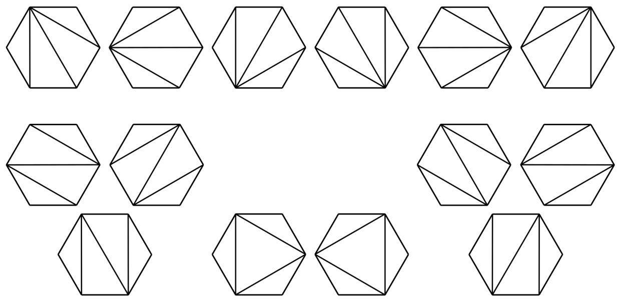
$$C_n = \frac{1}{n+1}\binom{2n}{n}, n \geq 0.$$

在本节的剩余部分，C_n 特指 Catalan 数，$C(x)$ 特指对应的生成函数.

例 2 “(” 和 “) ” 组成的序列称为合法序列, 如果对于任意的前缀序列, 左括号的数量都不小于右括号的数量. 例如, “(( ) ( ))” 和 “( )” 是合法序列, 而 “( )) (” 和 “(( ) ” 不是合法序列. 求长度为 $2n$ 的合法序列的个数.

解 同理, 可以把左括号看作进栈, 右括号看作出栈, 只有栈内有数据, 即有左括号才能出栈, 所以此问题与上例相同, 有 $C_n = \frac{1}{n+1}\binom{2n}{n}$ 种合法匹配.

例 3 求凸 $n+2$ 边形用其 $n-1$ 条对角线分割为互不重叠的三角形的方法种数，下图为 $n=4$ 时的情形，即凸六边形全部 14 种三角剖分。



解 假设总共有 a_n 种三角剖分，我们可以将凸 $n+2$ 边形的顶点标记为 $v_1, v_2, \ldots, v_{n+2}$. 我们考虑 v_{n+1} 和 v_{n+2} 之间的边所在的三角形，这个三角形的第三个顶点可以是 $v_1, v_2, \ldots, v_n$ 中的任意一个. 当第三个顶点是 v_k 时，此问题变成了对多边形 $v_{n+2}, v_1, v_2, \ldots, v_k$ 和多边形 $v_k, v_{k+1}, \ldots, v_{n+1}$ 分别进行剖分，二两者相互独立，于是再次有递推关系式

$$a_n = \sum_{k=1}^{n} a_{k-1}a_{n-k}.$$

$a_n = C_n$.

例 4 求具有 n 个叶子节点的满二叉树的个数。

解 假设总共有 a_n 个满二叉树。假设根节点的左子树有 k 个叶子节点，则右子树有 $n-k$ 个叶子节点. 而对左子树和右子树的构造亦是相互独立的，于是再次有递推式 $a_n = \sum_{k=1}^{n-1} a_k \cdot a_{n-k}$ 由此可得 $a_n = C_{n-1}, n \geq 3$.

例 5 有 $2n$ 个人去商店购买一个价值为 5 元的物品，其中 n 个人各只有一张 5 元的钞票，另外 n 个人各只有一张 10 元的钞票, 且商店内最开始没有任何零钱，请问这 $2n$ 个人有多少种购买顺序，可以避免商店出现无法找钱的情况?

解 我们可以将拥有 5 元钞票的人视作左括号，拥有 10 元钞票的人视作右括号，此时该问题与括号匹配问题完全相同。在这里我们通过另一种方法来计算 Catalan 数.

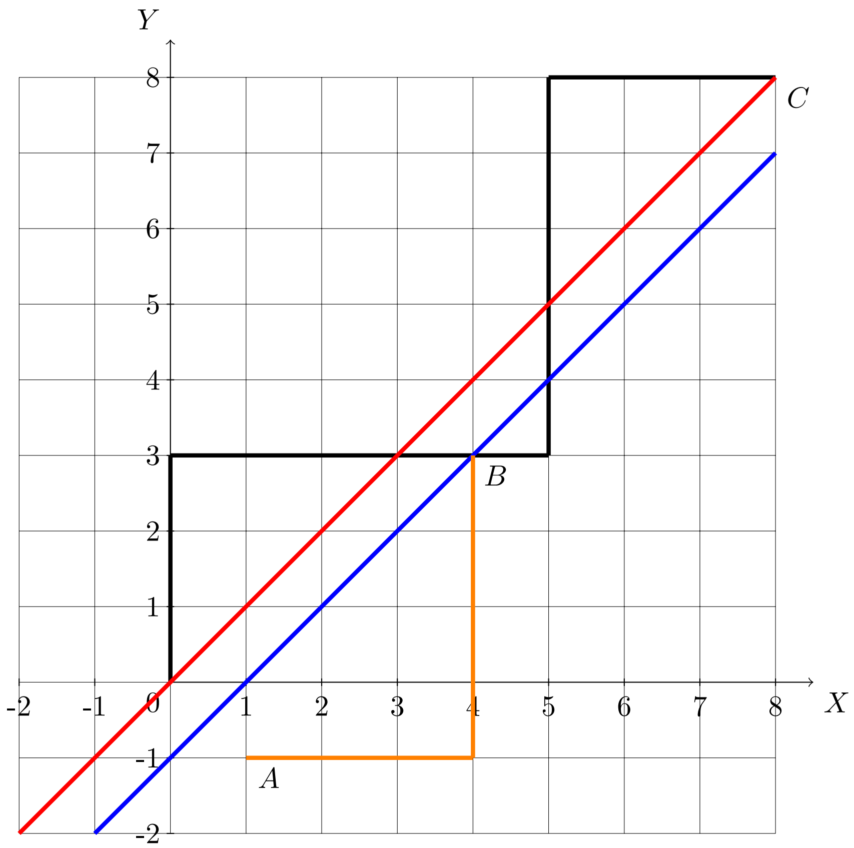
我们把问题转化为路径问题：平面上一蚂蚁从原点出发，每次只能向上或向右移动一个单位长度 (向上代表 5 元，向右代表 10 元)。现规定蚂蚁的轨迹必须在直线 $y=x$ 上方（允许接触），问蚂蚁到达 C 点 (n,n) 有多少种路径?

如果不考虑限制条件，要到达 C，一共需要向上和向右各走 n 次，确定在哪 n 步向右走即可唯一确定一条到达 C 的路径，因此路径数目有 $N_1=\binom{2n}{n}$ 种。我们只需求出其中不合法的路径数即可.

对于每一条不合法的路径，我们将其对应到另一条路径. 首先不合法的路径一定与直线 $y=x$ 接触并越过，考虑其第一次越界的位置，其必定在直线 $y=x-1$ 上，设之为 B，我们将该路径在 B 点之前的部分关于直线 $y=x-1$ 翻折，而保持 B 点之后的部分不变. 对应后的新路径是从 $(1,-1)$ 到 (n,n) 且碰到 $y=x-1$ 的路径. 显然不同的原路径，其生成的新路径也不同；反之对于每一条新路径，其对应的原路径也不同，形成一一对应的关系，从而二者的路径数目相同，设为 N_2.

由于 $(1,-1)$ 和 (n,n) 分别位于 $y=x-1$ 的两边，因此它们之间的路径一定会碰到直线 $y=x-1$，从而原来的限制条件此时天然成立，故有 $N_2=\binom{2n}{n-1}$。最终求得 Catalan 数的形式

$$\begin{aligned}C_n&=N_1-N_2\\&=\binom{2n}{n}-\binom{2n}{n-1}\\&=\frac{1}{n+1}\binom{2n}{n}.\end{aligned}$$



例 6 有一个赌徒有一块钱，每次赌博有 1/2 的几率输一元，有 1/2 的几率赢一元，他会一直玩直到把钱输光。试求

(i) 赌徒输光所有钱的概率是多少?

(ii) 赌徒期望会在赌场里面赌多少次?

解 该问题可等价于一维随机游走问题，动点 P 从坐标轴 1 点出发，掷硬币以 $\frac{1}{2}$ 概率向左走一格，以 $\frac{1}{2}$ 概率向右走一格，每次掷硬币相互独立，P 第一次走到 0 即表示赌徒输光钱.

设 p_{2i+1}，$i=0,1,2,\ldots$ 表示第 $2i+1$ 步踏入原点的概率（动点只有可能在奇数步回到原点）. 考虑动点 P 在 $2k+1$ 步后到达原点，最后一步一定是从点 1 到原点（有 $\frac{1}{2}$ 的概率），之前的 $2k$ 步要有 k 步向右走，有 k 步向左走，但是不能走到原点. 在这里，向右走可视为进栈，向左走可视为出栈，所以前 $2k$ 步有 C_k 种走法. 所以有：

$$
\begin{aligned}
p_{2k+1} &= \frac{C_k}{2^{2k+1}} \\
\sum_{k\geq 0} p_{2k+1} &= \frac{1}{2}\sum_{k\geq 0} C_k\left(\frac{1}{4}\right)^k = \frac{1}{2}C\left(\frac{1}{4}\right).
\end{aligned}
$$

因为 $C(x)=\frac{1-(1-4x)^{\frac{1}{2}}}{2x}$，所以 $\sum_{k\geq 0} p_{2k+1}=\frac{1}{2}C\left(\frac{1}{4}\right)=1$，即赌徒输光所有钱的概率是 1. 问题 2 相当于计算游走次数 T 的期望：

$$
\begin{aligned}
\mathbb{E}(T) &= \sum_{k\geq 0}(2k+1)p_{2k+1} \\
&= 1+\sum_{k\geq 0} 2kp_{2k+1} \\
&= 1+\sum_{k\geq 0} kC_k\left(\frac{1}{4}\right)^k.
\end{aligned}
$$

变换 Catalan 数的生成函数，

$$
\begin{aligned}
C(x) &= \frac{1-(1-4x)^{\frac{1}{2}}}{2x} = \sum_{k\geq 0} C_k x^k \\
xC'(x) &= \sum_{k\geq 0} C_k k x^k.
\end{aligned}
$$

从而

$$
\begin{aligned}
\mathbb{E}(T) &= 1+\left(\frac{1}{(1-4x)^{\frac{1}{2}}}-\frac{1-(1-4x)^{\frac{1}{2}}}{2x}\right)\Bigg|_{x=\frac{1}{4}} \\
&= +\infty.
\end{aligned}
$$

2 选做题

考虑汉诺塔问题的变种。如下图所示，现在有从左到右编号为 I、II、III 和 IV 的 4 根柱子，初始状态 I 号柱上有 n 个圆盘，从上到下编号依次为 $1, 2, \ldots, n$，其他三个柱子均为空。目标是把所有圆盘通过若干步移动转移到 IV 号柱上。只允许移动某个柱子最上面的盘子，到其右边相邻的柱子上（即 I 号柱到 II 号柱、II 号柱到 III 号柱、III 号柱到 IV 号柱）。

1. 完成目标时，IV 号柱上的盘子能形成多少种排列?
2. 扩展问题到有多个柱子的情况。至少需要多少柱子，才能使完成目标时，最后一根柱子上的盘子能形成所有可能的排列?