组合数学 — Catalan 数

张 彪

天津师范大学

zhang@tjnu.edu.cn

引例 (姐妹洗碗问题)

姐姐和妹妹一起洗 4 个碗，姐姐洗好的碗一个一个往上摆，妹妹再从最上面一个一个地拿走放入碗柜摆成一摞，姐姐一边洗，妹妹一边拿，那么妹妹摆好的碗一共有多少种不同的方法?

图: 姐妹洗碗问题

引例 (姐妹洗碗问题)

姐姐和妹妹一起洗 4 个碗, 姐姐洗好的碗一个一个往上摆, 妹妹再从最上面一个一个地拿走放入碗柜摆成一摞, 姐姐一边洗, 妹妹一边拿, 那么妹妹摆好的碗一共有多少种不同的方法?

图: 姐妹洗碗问题

例如

- 姐姐洗 2 个碗,
- 妹妹摆 1 个碗,
- 姐姐再洗 2 个碗,
- 妹妹再摆 3 个碗.

怎么去描述数学的语言这种取法?

引例 (姐妹洗碗问题)

姐姐和妹妹一起洗 4 个碗, 姐姐洗好的碗一个一个往上摆, 妹妹再从最上面一个一个地拿走放入碗柜摆成一摞, 姐姐一边洗, 妹妹一边拿, 那么妹妹摆好的碗一共有多少种不同的方法?

图: 姐妹洗碗问题

例如

- 姐姐洗 2 个碗,
- 妹妹摆 1 个碗,
- 姐姐再洗 2 个碗,
- 妹妹再摆 3 个碗.

怎么去描述数学的语言这种取法?

- 令 j 为姐姐洗完的碗的个数，i 为妹妹摆碗的个数
- 条件为妹妹摆碗的个数不能超过姐姐洗完的碗的个数，即 $i \le j$

上面例子可以叙述为如下的过程

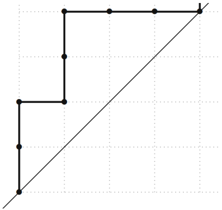
$$(0,0) \to (0,2) \to (1,2) \to (1,4) \to (4,4)$$

引例 (姐妹洗碗问题)

姐姐和妹妹一起洗 4 个互不相同的碗, 姐姐洗好的碗一个一个往上摆, 妹妹再从最上面一个一个地拿走放入碗柜摆成一摞, 姐姐一边洗, 妹妹一边拿, 那么妹妹摞好的碗一共有多少种不同的方法?

令 j 为姐姐洗完的碗的个数，i 为妹妹摞碗的个数，条件转化为 $i \leq j$

- 纵坐标表示姐姐洗完的碗的个数，横坐标表示妹妹摞碗的个数,
- 要求摆法的方案数实际上是求从坐标 (0，0) 到坐标 (4,4) 的所有满足条件的路径数.



定义 (Dyck 路)

一个长度为 $2n$ 的 *Dyck* 路是一个从 $(0,0)$ 到 (n,n) 的格路，它

- 含有 n 个水平步骤 “*E*” $(i,j) \to (i+1,j)$
  和 n 个垂直步骤 “*N*” $(i,j) \to (i,j+1)$,
- 且路径上所有整数格点满足 $i \le j$, 即在平面中位于 $y = x$ 以上.

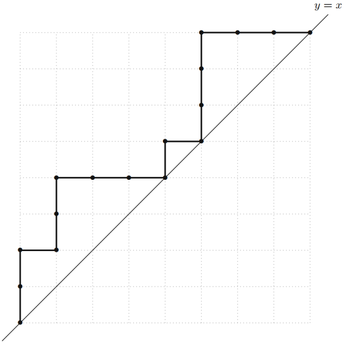


图: Dyck 路 $(n = 8)$

定义 (Dyck 路)

一个长度为 $2n$ 的 *Dyck* 路是一个从 $(0,0)$ 到 (n,n) 的格路，它

- 含有 n 个水平步骤 "*E*" $(i,j) \to (i+1,j)$
  和 n 个垂直步骤 "*N*" $(i,j) \to (i,j+1)$,
- 且路径上所有整数格点满足 $i \le j$, 即在平面中位于 $y = x$ 以上.

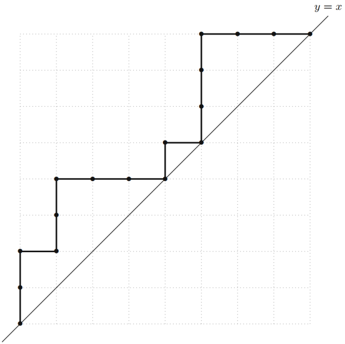


图: Dyck 路 $(n = 8)$

我们可以画出

- 格路的图 或
- 遵循路径的步骤列表写成一个多重集合 $\{N, E\}$ 的排列.

例如，左边的 Dyck 路所对应的多重集合 $\{N, E\}$ 的排列如下

$NNENNEEENENNNEEE$

如图所示.

图: Dyck 路 ($n=3$)

我们用 $\mathcal{D}(n)$ 表示长度为 $2n$ 的 Dyck 路的集合，且定义 *Catalan* 数如下

$$C_n = \#\mathcal{D}(n).$$

我们将证明,

定理

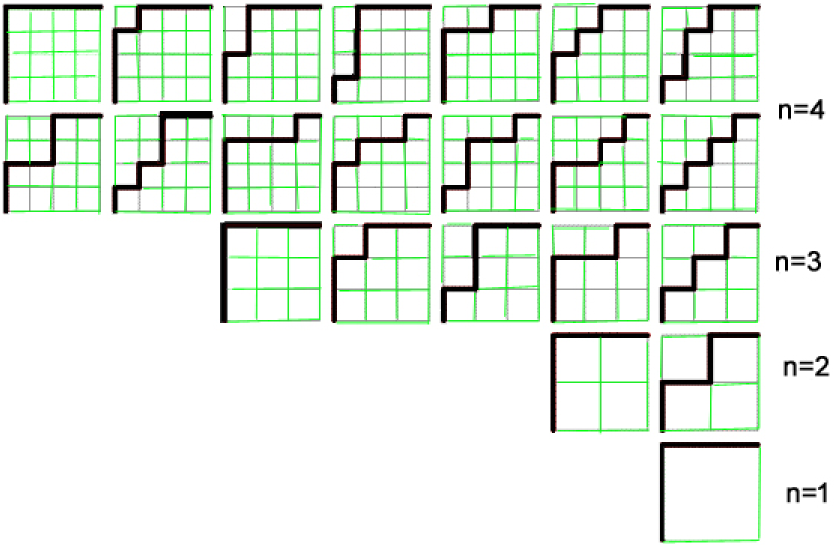
第 n 项 *Catalan* 数的表达式为

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

历史上，清代数学家明安图 (1692-1763) 在其《割圆密率捷法》最早用到“卡塔兰数”，远远早于比利时的数学家卡塔兰 (1814–1894).

有中国学者建议将此数命名为“明安图数”或“明安图–卡塔兰数”．

经过计数, 可得 $C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14.$



递推关系的建立

以半 n 长 Dyck 路为例. 设满足条件的 Dyck 路的条数即 Catalan 数为 C_n. 设满足条件的一条 Dyck 路: $P = v_0v_1\cdots v_{2n}$, 其中 $v_0 = (0,0)$, $v_{2n} = (n,n)$.

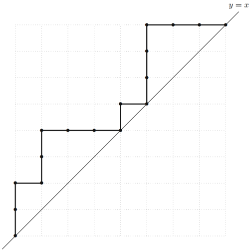


图: Dyck 路 $(n = 8, i = 4)$

设第一个与 $y = x$ 相交的顶点为 $v_{2i} = (i,i)$. 将 Dyck 路 P 按照 v_{2i} 分成两段,

1. $v_0 = (0,0) \to v_{2i} = (i,i)$. 此时第一步一定向上, 最后一步一定向右, 故只需考虑 $(0,1) \to (i-1,i)$ 的格路条数, 计数为 C_{i-1}.
2. $v_{2i} = (i,i) \to v_{2n} = (n,n)$. 计数为 C_{n-i}.

于是有

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

设 C_n 为 Catalan 数, 规定 $C_0 = 1$. 令

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

对 (1) 式两边同乘以 x^n 并关于 $n \geq 0$ 求和得

$$A(x) - 1 = xA(x)^2,$$

解得

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

利用二项式定理

$$\sqrt{1-4x} = (1-4x)^{\frac{1}{2}} = \sum_{n\geq 0} \binom{\frac{1}{2}}{n} (-4x)^n .$$

因此,

$$
\begin{aligned}
A(x) &= \frac{1-\sqrt{1-4x}}{2x} \quad \text{(舍去正号，为什么？)} \\
&= \frac{1}{2x}\left(1-\sum_{n\geq 0}\binom{\frac{1}{2}}{n}(-4x)^n\right) \\
&= -\frac{1}{2}\sum_{n\geq 0}\binom{\frac{1}{2}}{n+1}(-4)^{n+1}x^n \\
&= \sum_{n\geq 0}\frac{1}{n+1}\binom{2n}{n}x^n. \quad \text{(计算过程请自行补充完整)}
\end{aligned}
$$

提取 x^n 项系数, 得

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

建立 $n+1$ 个 x 组成的括号字符串

一个 $n+1$ 个 x 组成的括号字符串由插入 n 个左括号和 n 个右括号组成.

$n=3$ 时,

$$x(x(xx)) \quad x((xx)x) \quad (xx)(xx) \quad (x(xx))x \quad ((xx)x)x$$

注

对于 (((xx)x)((xx)(xx))) 这种形式的元素, 我们通常忽略最左和最右的括号.

$2n$ 长 Ballot 序列

设 $w = w_1 \cdots w_{2n}$ 是由 n 个 1 和 n 个 2 组成的序列, 对任意 $i = 1, 2, \cdots, 2n$, 要求前 i 个字 $w_1 \cdots w_i$ 中 1 的个数大于或等于 2 的个数. 满足上述条件的序列称为 $2n$ 长 Ballot 序列.

如 $n = 3$ 时, Ballot 序列如下

$$111222 \quad 112122 \quad 112212 \quad 121212 \quad 121122$$

n 对圆括号合法排列

n 对圆括号排在一起, 从左往右看, 左括号的个数大于等于右括号的个数, 则称为合法.

$n=3$ 时,

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

注

Catalan 数任意两个组合解释之间都可建立双射. 这里 Ballot 序列与 n 对圆括号合法排列之间的双射, 只需将 “1” 与 “(”, “2” 与 “)” 对应起来即可.

凸 $n+2$ 边形切割

- 将 $n+2$ 边形添加对角线, 使其被切割为 n 个三角形.

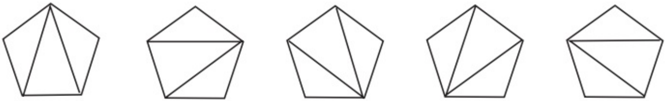


图: 凸 $n+2$ 边形切割 $(n=3)$

n 个顶点的二叉树

- 树：连通且无圈.
- 二叉树：顶点度小于等于 2 的树.

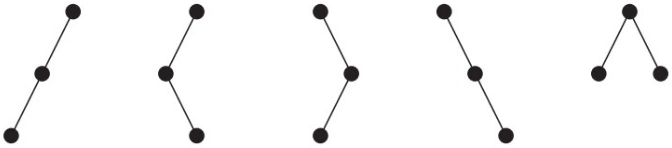


图: n 个顶点的二叉树 $(n = 3)$

$n+1$ 个 x 组成的括号字符串与 n-二叉树间的双射

- 设 $n+1$ 个 x 组成的括号字符串为 w, 定义二叉树 B_w 的递推关系满足: 如果 $n=0$, 则 $B_w=\emptyset$; 否则, 从 w 最外层括号开始, 如果 $w=st$, 则 B_w 有一个根顶点 v、左子树 B_s 及右子树 B_t. 例如, 如果 $w=xx$, 则 B_w 只包含一个顶点（根）. 对下图中的二叉树, 其对应的括号为

$$((xx)x)x \quad (x(xx))x \quad x((xx)x) \quad x(x(xx)) \quad (xx)(xx)$$

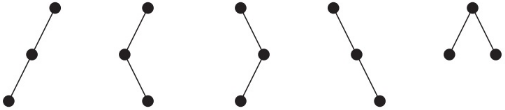


图: n 个顶点的二叉树 (n=3)

建立 n-二叉树与凸 $n+2$ 边形切割间双射

- 固定多边形的边 e, 在 T 的每个三角形内部放置一个顶点, 让根顶点对应于以 e 为边的三角形. 连接相邻三角形中的点, 如图 (a). 从顶点 v 出发, 确定边 f_1 及 f_2, 沿着边 f_1 逆时针旋转定义第一条边为左侧边 f_{11}, 第二条边为右侧边 f_{12}. 以此类推, 即可得到二叉树如 (b).