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，那么妹妹摞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 $i \leq j$

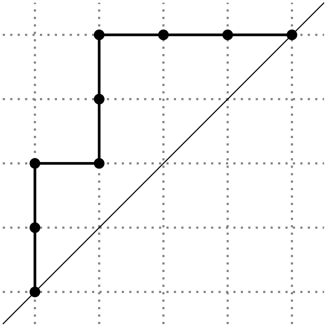
引例 (姐妹洗碗问题)

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, 那么妹妹摆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 $i \leq j$

上面例子可以叙述为如下的过程

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



- 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.

Catalan 数

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 $i \leq j$, 即在平面中位于 $y = x$ 以上.

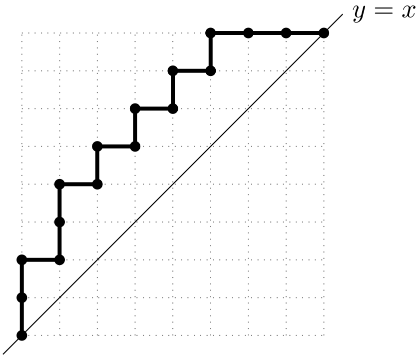


图: Dyck 路 $(n = 8)$

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.

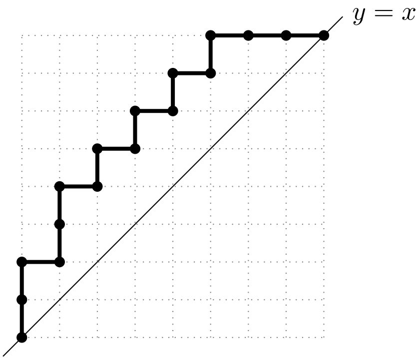


图: Dyck 路 $(n = 8)$

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.

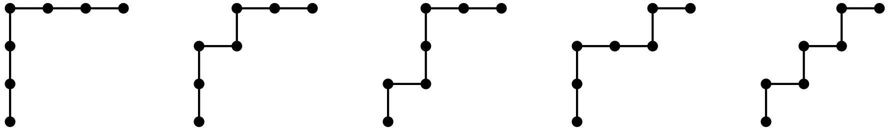


图: Dyck 路 $(n=3)$

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

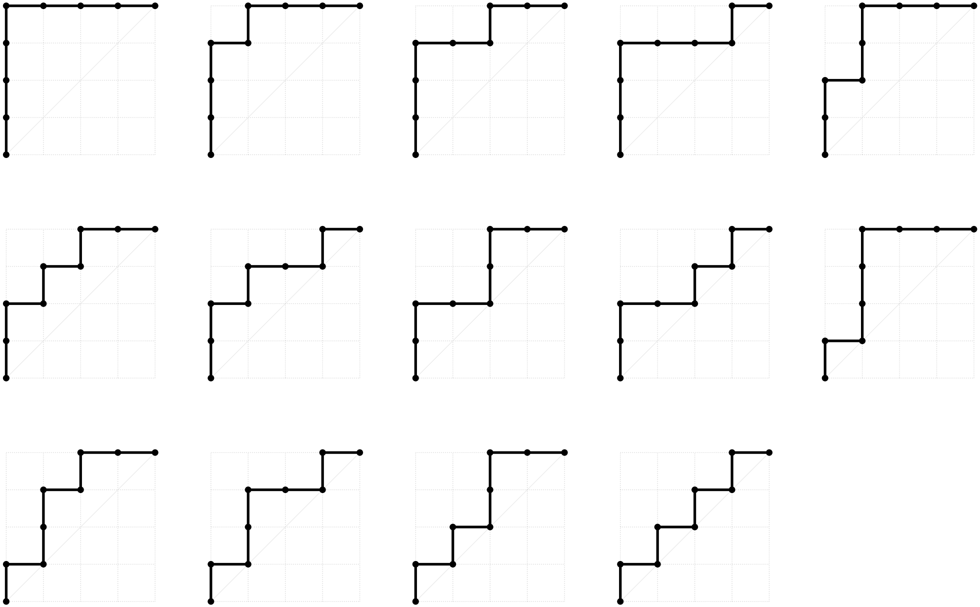
$$C_n = \#\mathcal{D}(n).$$

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).



Catalan 数

递推关系的建立

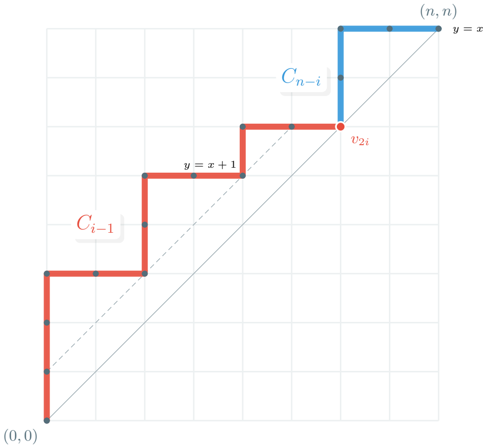


图: Dyck 路径分解 (First Return)

设 Dyck 路: $P = v_0 v_1 \cdots v_{2n}$,

其中 $v_0 = (0,0)$, $v_{2n} = (n,n)$.

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,

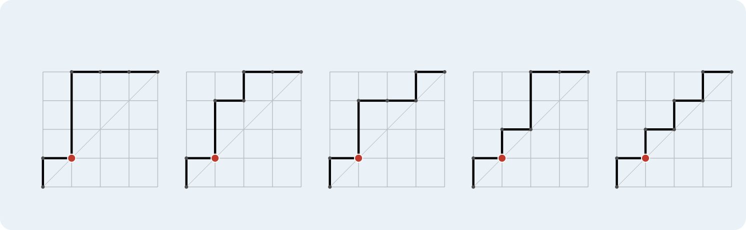
1. $v_0 = (0,0) \to v_{2i} = (i,i)$.
   此时第一步一定向上, 最后一步一定向右,
   故只需考虑 $(0,1) \to (i-1,i)$ 的格路条数,
   计数为 C_{i-1}.
2. $v_{2i} = (i,i) \to v_{2n} = (n,n)$.
   计数为 C_{n-i}. (为什么 ?)

于是有

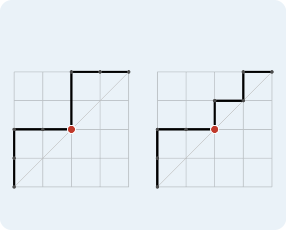
$$C_n = \sum_{i=1}^{n} C_{i-1} C_{n-i}. \tag{1}$$

Dyck 路径分类：根据首次返回点 $v_{2i} = (i, i)$

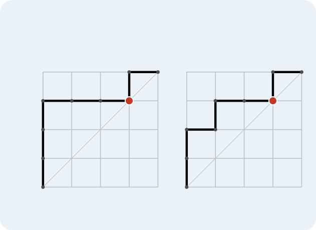
Group 1: $v_2 = (1, 1) \implies C_0 \times C_3 = 5$



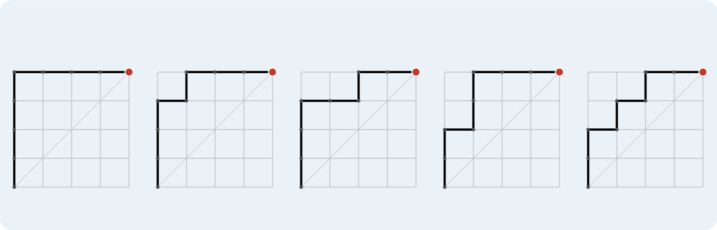
Group 2: $v_4 = (2, 2) \implies 2$



Group 3: $v_6 = (3, 3) \implies 2$



Group 4: $v_8 = (4, 4) \implies C_3 \times C_0 = 5$



Catalan 数生成函数与通项公式

规定 $C_0 = 1$. 令生成函数

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

由递推关系求和可得方程

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

解得

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

利用广义二项式定理展开:

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

若取 "+" 号，当 $x \to 0$ 时 $C(x) \to \infty$，舍去.

故取 "−" 号 (使 $C(0) = 1$):

$$\begin{aligned} C(x) &= \frac{1 - \sqrt{1-4x}}{2x} \\ &= \frac{1}{2x}\left[1 - \sum_{n\geq 0} \binom{\frac{1}{2}}{n} (-4x)^n\right] \\ &= \dots \quad \text{(整理中间项)} \\ &= \sum_{n\geq 0} \frac{1}{n+1} \binom{2n}{n} x^n. \end{aligned}$$

提取 x^n 项系数, 得通项公式:

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

1. 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 的个数不少于 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

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

1. 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 的个数不少于 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

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

长度为 $2n$ 的 Ballot 序列 $w = w_1 \cdots w_{2n}$ 与 Dyck 路径 P 之间存在如下双射:

- $w_i = 1 \to P$ 的第 i 步为垂直步骤 “N”;
- $w_i = 2 \to P$ 的第 i 步为水平步骤 “E”.

2. 括号的匹配

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

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.

2. 括号的匹配

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

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.

为给出 Catalan 数下面递推关系的组合意义,

$$C_n = \sum_{k=1}^{n} C_{k-1} C_{n-k}, \qquad C_0 = 1,$$

考虑由 $n+1$ 个 x 通过插入 n 对合法括号所形成的完全括号化表达式.

任一非平凡表达式可唯一分解为 $(A)(B)$, 其中 A 包含 k 个 x (对应 C_{k-1} 种括号化方式), B 包含 $n-k+1$ 个 x (对应 C_{n-k} 种方式), $k=1,\ldots,n$. 通常省略最外层括号, 故总数为 C_n.

3. 完全二叉树

定义

- 树是一种无回路连通图. 若指定某个节点为根, 则称为根树.
- 度为 0 的节点称为叶子; 度为 2 的节点称为内节点.
- 若每个内节点的度数最多为 2, 则称为二叉树.
- 所有内节点度均为 2 的二叉树称为完全二叉树 (full binary tree).

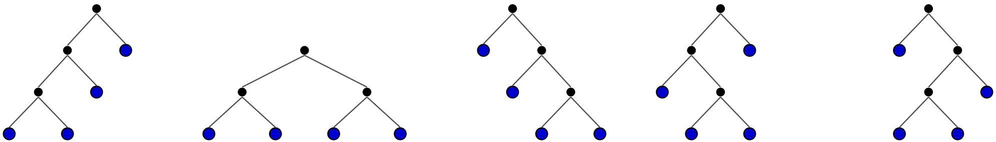


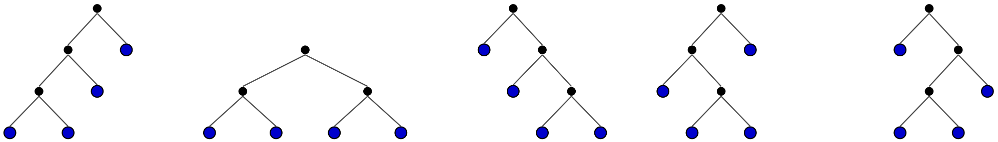
图: 4 个叶子的完全二叉树的 5 种结构

完全二叉树与括号表达式的对应

设一棵有 n 个叶子的完全二叉树的叶子依次标为 $a_1, a_2, \ldots, a_n$.

每个内节点与其左右子树分别对应两部分，若两个子树的叶子表达式分别为 x_1, x_2，则该内节点表示：$(x_1 x_2)$. 于是整棵树对应一个 n 元乘积 $a_1 a_2 \cdots a_n$ 的一种括号化方式.

结论： 含 n 个叶子的完全二叉树与 n 个元素乘积的所有合法括号化之间存在一一对应.



$(((a_1a_2)a_3)a_4)$ $(a_1a_2)(a_3a_4)$ $a_1(a_2(a_3a_4))$ $(a_1(a_2a_3))a_4$ $a_1((a_2a_3)a_4)$

4. 二叉树

- 二叉树：顶点度小于等于 2 的树.

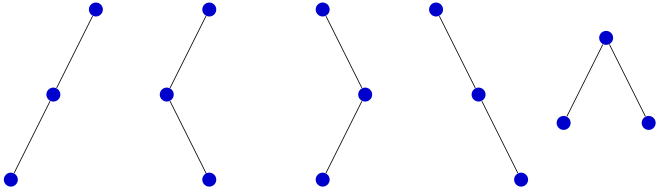


图: 三节点二叉树的 5 种结构

$n+1$ 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 $w = xx$, 则 B_w 只包含一个顶点 (根);
- 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, 对 3 个顶点的的二叉树, 其对应的括号为

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

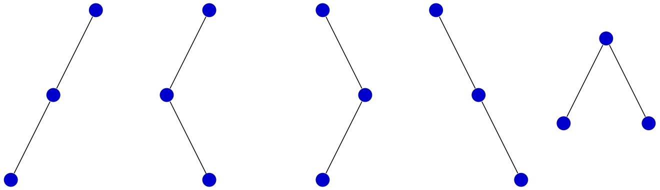


图: 三节点二叉树的 5 种结构

5. 凸多边形的三角剖分

- 将 $n+2$ 边形添加对角线, 使其被切割为 n 个三角形.

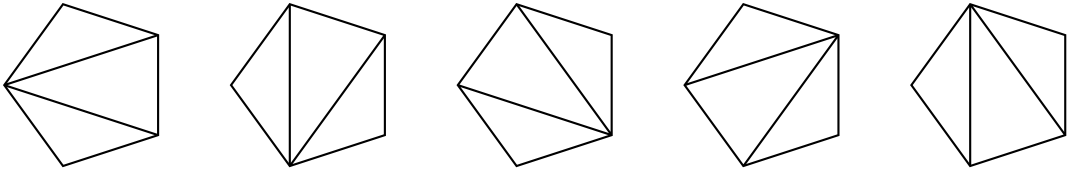
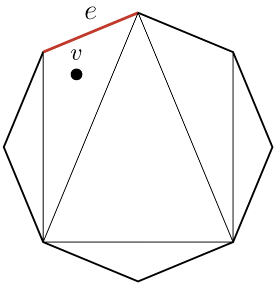


图: 凸 5 边形切割

从三角剖分到二叉树

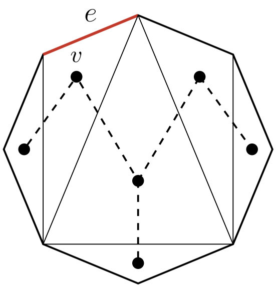
- 固定边 e，令包含 e 的三角形对应根顶点 v (取三角形内心).



(a) 取内心构造对偶图

从三角剖分到二叉树

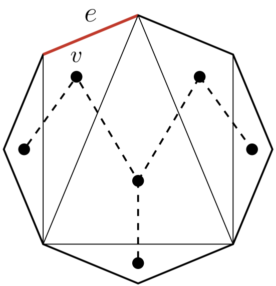
- 固定边 e，令包含 e 的三角形对应根顶点 v (取三角形内心).
- 在所有小三角形中取内心，连接相邻点构造对偶图 (虚线).



(a) 取内心构造对偶图

从三角剖分到二叉树

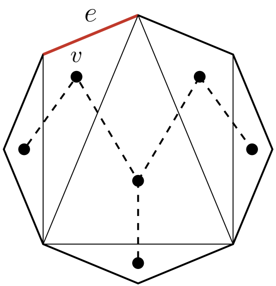
- 固定边 e，令包含 e 的三角形对应根顶点 v (取三角形内心).
- 在所有小三角形中取内心，连接相邻点构造对偶图 (虚线).
- 从 v 出发，按逆时针顺序:
  - 第一条边对应 左子树
  - 第二条边对应 右子树



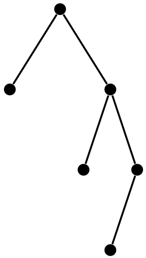
(a) 取内心构造对偶图

从三角剖分到二叉树

- 固定边 e，令包含 e 的三角形对应根顶点 v (取三角形内心).
- 在所有小三角形中取内心，连接相邻点构造对偶图 (虚线).
- 从 v 出发，按逆时针顺序:
  - 第一条边对应 左子树
  - 第二条边对应 右子树
- 得到如图 (b) 所示的二叉树.



(a) 取内心构造对偶图



(b) 生成二叉树

课程结束寄语

- 组合数学源于朴素直观的问题，却蕴含深刻隽永的思想，是一片值得长期耕耘的学术沃土
- 鼓励大家以组合数学为题开展毕业论文研究，借此系统梳理知识脉络，深化理论认知
- 欢迎有兴趣的同学在研究生阶段继续深造，在组合数学及相关前沿领域探索求真
- 期待各位在未来的教学与科研实践中，能够自觉融入并灵活运用组合学的智慧与方法

愿组合思维，成为你开启数学世界奥秘的一把钥匙