Lecture 4
双重计数方法

中山大学 智能工程学院
金恺

Outline

- 双重记数方法（Double counting technique）
  - 组合计数三大常用方法：
    - 容斥原理
    - 一一映射（上周内容）
    - 双重记数

- **项链记数问题**：n个珠子的项链，珠子颜色可选1~k，方案数？

n=3
k=2

例1 求证 $1\binom{n}{1}+2\binom{n}{2}+\cdots+n\binom{n}{n}=n2^{n-1}$.

- **计数对象：**
  - 有多少二元对(a,B)，满足$B\subseteq\{1,\ldots,n\}$ 且 $a\in B$？
- **记数方法一：**
  - 先选a，有n种选法。
  - 再确定B。由于B要包含a，B的选择有2^{n-1}种。
  - 结果为右式
- **记数方法二：**
  - 先选B。设$|B|=i$。枚举i的每个可能性$1\le i\le n$。固定 i 后选法数为$\binom{n}{i}$。
  - 再确定a。选法数为i。
  - 结果为左式。

例2 求证 $\sum_{i\geq r}\binom{n}{i}\binom{i}{r}=\binom{n}{r}2^{n-r}$.

- **计数对象：**
  - 有多少二元对(A,B)，满足$|A|=r$ 且 $A\subseteq B\subseteq\{1,\ldots,n\}$？

难点在于，设计要计数的对象（构造性的）

- **记数方法一：**
  - 先选A。选法有$\binom{n}{r}$种。
  - 其次选B。当A确定后，B要包含A，因此B的选择有 2^{n-r}种。
- **记数方法二：**
  - 先选B。（假定$|B|=i$ 并枚举i的所有可能性：$r\leq i\leq n$。）
  - 当B确定后，再从B中选一个r元子集作为A。
- **对同一个集合的两种方法记数的结果必然相等！**

例3 Fibonacci序列的一个恒等式

- 求证 $\mathbf{F_0 + F_1 + \ldots + F_{n-2} = F_n - 1}$。 $(n \geq 2)$
  - Fibonacci序列的定义： $F_0=F_1=1$, $F_i= F_{i-1}+F_{i-2}$ $(i \geq 2)$。
- **计数对象：** 1* n 的区域有多少种铺砖方案满足：
  - (1) 砖块是1*1或者1*2的。
  - (2) 铺砖不重复不遗漏。
  - (3) 至少有一块1*2的转。换句话说，不允许全都用1*1的。
- **记数方法一：**
  - 根据动态规划，满足(1)和(2)的方案数为F_n。因此总方案数为右式。
- **记数方法二：**
  - 枚举**左起第一个**1*2砖块的位置。可以得到左式。 (板书解释)

例4 $(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^{n-r} b^r$ （此为下阶乘那次课的作业题）

- **计数对象：**
  - 假定有a个白球，b个黑球。编号各不相同（球有区分）。
  - 有多少种方案选出n个球排成一排？
- **方法一：** 易得，$(a+b)^n$。
- **方法二：**
  - 假定选出的球中有r个黑球，n-r个白球。
  - 先确定r，并且指定哪r个位置放黑球。
  - 在r个指定位置填入黑球，方案数为 b^r。
  - 在n-r个指定位置填入白球，方案数为a^{n-r}。
- **两种计算结果必然一样。**

例5 $x^{\overline{n}} = \sum_k \left[{n \atop k}\right] x^k, \qquad \text{integer } n \geq 0.$ (用归纳法证明过)

先考虑 x为正整数的情况（下一页将说明，对所有实数x，上式成立）

- **计数对象**：<u>找{1,…,n}的一个cycle partition并对其中的元素用颜色1~x着色，满足同一个cycle中的元素必须同色。</u> (不同cycles允许同色或异色）
- **第一种计数方案：**
  - 考虑cycles的个数。易得，方案数为右式。
- **第二种计数方案：**
  - 依次考虑元素1~n。 **//逐个添加的方式来生成cycle partition以及染色。**
  - 1有x种选择。
  - 2有x+1种选择（新开一个cycle，x种颜色可选，或者放到1后面）
  - 3有x+2种选择（新开一个cycle，x种颜色可选，或者放到1或2的后面）
  - 依此类推。 共有x的n次上阶乘 这么多种方案。

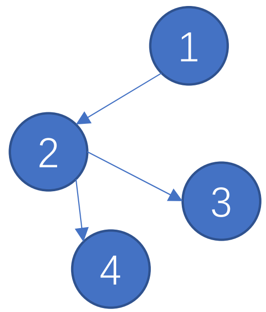
补充说明

- **观察**. 当一个多项式f(x) 对每一个正整数x，取值都为0时，
  那么只有一种可能：f(x)就是系数全为0的多项式。
- 反证法。 设 f(x)系数不全为0，比如$f(x) = 1.5x^4-(2.5x^3-2x+3)$。
  - 只要取 x 足够大，可以使得
    - ¼ $1.5x^4 > 2.5$ x^3
    - ¼ $1.5x^4 > 0$x^2
    - ¼ $1.5x^4 > -2x$
    - ¼ $1.5x^4 > 3$
    - 故 f(x)>0。 故假定不成立。
- **推论**：当一个多项式等式对所有正整数成立，则对所有实数成立。

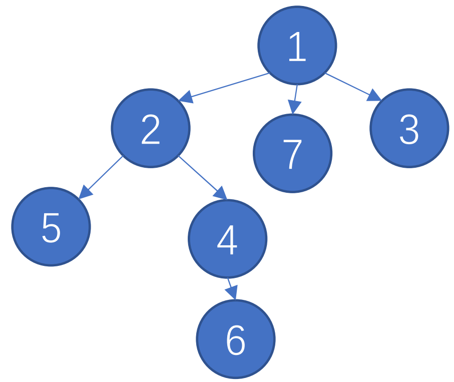
例6 Cayley公式的简单证明方法

- **计数对象：**
  - 序列($a_1,\ldots,a_{n-1}$)，其中$a_1,\ldots a_{n-1}$都是有向边且他们恰好能组成一棵**有根树**。

  - //特别留意：**计数的对象是“序列”，而不是“有根树”!**
  - **举例**：n=4时，
    - 一下序列都符合要求，且它们都组成右边这棵**有根树**。
    - (1,2),(2,3),(2,4)
    - (1,2),(2,4),(2,3)
    - (2,3),(1,2),(2,4)
    - (2,3),(2,4),(1,2)
    - (2,4),(1,2),(2,3)
    - (2,4),(2,3),(1,2)



何谓 **有根树**?
有向图,n 点，n-1条边
且恰有一个点入度为0
（叫根），其余点入度为1

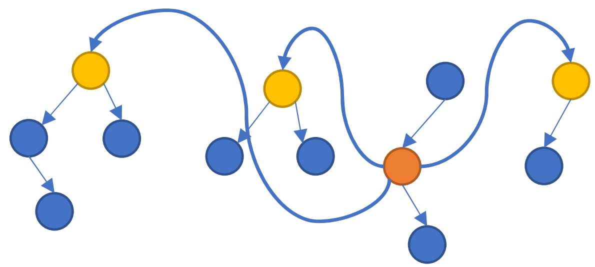


接下来会看到，这种序列有一种比较巧妙的计数方法。
因此，计数对象考虑序列，而不是直接考虑有根树。

- **第一种计数方法：**

符号S_n表示生成树个数。

  - 有根树个数 = $n\,S_n$ //找一个无根树，找一个根，得有根树。
  - 序列个数 = 有根树个数* $(n\text{-}1)! = S_n\, n!$ // 从一个有根树获得(n-1)!个序列
- **第二种数树方法：**
  - **思路**：逐一选择 a_1、a_2、…、$a_{n\text{-}1}$，来得到序列。
    // 任何时刻确保（已选择的边构成的有向图中）每个点**入度≤1**，即已选的边组成**有根森林**。则，最终确定完$a_{n\text{-}1}$后，选中的n-1条边恰组成有根树。
  - **重要观察**：假设当前森林有k个根时，下一步的连法个数为 **$n(k\text{-}1)$**。
    - Why? 每个点都可以连出去，到另一个树的根。
  - 故，答案为 $$\prod_{k=2}^{n} n(k-1) = n^{n-1}(n-1)! = n^{n-2} n!.$$

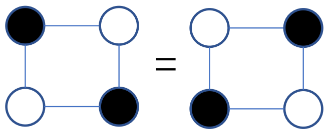
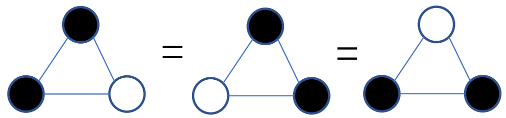


- 因此，$S_n\, n! = n^{n\text{-}2}\, n!$。化简得 $S_n = n^{n\text{-}2}$。

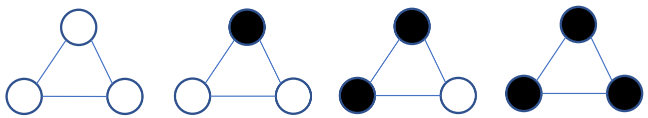
——映射方法证明cayley公式更复杂，但是能给出一些推论（更深刻）

例7 项链计数问题（较难）

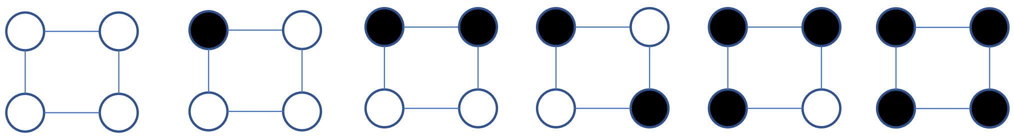
- 长度为n，珠子颜色数为$1\sim k$，不同项链个数记作$T_{n,k}$。
  - //循环相同的视作同一条项链。

  - 例如$T_{3,2}=4$。

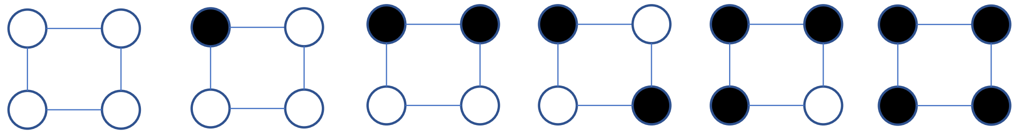


  - 例如$T_{4,2}=6$。



- 求 $\boldsymbol{T_{n,k}}$的通项公式。

显然，每个项链可以用一个字符串来描述。长度为n、字符集为{0,⋯,k-1}代表颜色。



0000 0001 0011 0101 0111 1111

注意：0101=1010 描述的是同一个项链。0001=0010=0100=1000是同一条项链。也就是说，一条项链表示为串，可能有多种表示方式。

- 构造多重集合**B**：（列成表格）
  - 1、$\mathbf{T_{n,k}}$条项链对应的串表示，构成**B**的第一列。
  - 2、每个串的n个cyclic-shifts 构成**B**的一行。
- 例（n=4, k=2）
  - **0000** 0000 0000 0000
  - **0001** 1000 0100 0010
  - **0011** 1001 1100 0110
  - **0101** 1010 0101 1010
  - **0111** 1011 1101 1110
  - **1111** 1111 1111 1111

显然|**B**|= $n\cdot T_{n,k}$。

另一方面，可以按下述方式计算|**B**|：

$$\sum_{(a_0\dots a_{n-1})\in\{0,\dots,k-1\}^n} \text{串}a_0 \dots a_{n-1}\text{在}\mathbf{B}\text{中出现的次数}$$

子问题:

串$a_0 \dots a_{n-1}$在**B**中出现的次数？？

0000	0000	0000	0000
0001	1000	0100	0010
0011	1001	1100	0110
0101	1010	0101	1010
0111	1011	1101	1110
1111	1111	1111	1111

- 根据**B**的定义，$a_0 \dots a_{n-1}$肯定属于**B**的**恰好**一行中
- 但是可能在该行出现多次。具体出现的次数=几?

$$\boxed{\sum_{i=0}^{n-1}[a_0a_1\dots a_{n-1} = a_i\dots a_{n-1}a_0\dots a_{i-1}]}$$

why?

因此，

$$nT_{n,k} = |\mathbf{B}| = \sum_{(a_0\dots a_{n-1})\in\{0,\dots,k-1\}^n} a_0 \dots a_{n-1}\text{在}\mathbf{B}\text{中出现的次数}$$

$$= \sum_{(a_0\dots a_{n-1})\in\{0,\dots,k-1\}^n} \sum_{i=0}^{n-1}[a_0a_1\dots a_{n-1} = a_i\dots a_{n-1}a_0\dots a_{i-1}]$$

$$nT_{n,k} = \sum_{(a_0a_1\ldots a_{n-1})} \sum_{i=0}^{n-1} [a_0a_1\ldots a_{n-1} = a_i\ldots a_{n-1}a_0\ldots a_{i-1}]$$

$$= \sum_{i=0}^{n-1} \sum_{(a_0a_1\ldots a_{n-1})} [a_0a_1\ldots a_{n-1} = a_i\ldots a_{n-1}a_0\ldots a_{i-1}]$$

$a_0..a_{n-1}$=$a_i\ldots a_{n-1}a_0\ldots a_{i-1}$

⇔ a以gcd(i,n)循环

板书举例: n=6, i=4 。

严格证明留为作业 ☺

$$= \sum_{i=0}^{n-1} \{\text{有多少串}(a_0a_1 \ldots a_n)\text{以}\gcd(i,n)\text{为循环节}\}$$

$$= \sum_{i=0}^{n-1} k^{\gcd(i,n)} = \sum_{d|n} k^d \cdot |\{0 \le i < n \mid \gcd(i,n) = d\}|$$

$$|\{0 \le i < n \mid \gcd(i,n) = d\}|$$
$$= \left|\left\{0 \le i < n \mid \gcd\left(\frac{i}{d}, \frac{n}{d}\right) = 1\right\}\right|$$
$$= \left|\left\{0 \le j < \frac{n}{d} \mid \gcd\left(j, \frac{n}{d}\right) = 1\right\}\right|$$
$$= \varphi\left(\frac{n}{d}\right)$$

$$= \sum_{d|n} k^d \cdot \varphi\left(\frac{n}{d}\right)$$

$$T_{n,k} = \frac{1}{n}\sum_{d|n} k^d \varphi\left(\frac{n}{d}\right) = \frac{1}{n}\sum_{d|n} k^{\frac{n}{d}} \varphi(d)$$

板书证明 (trivial)

例8 费马小定理

- **费马小定理**：
  - 已知 (a, p)互素且p为质数，则$\boldsymbol{a^p - a = 0 \pmod p}$。
- **计数**：有多少个字符串满足：
  - 长度为p、字符为$1\sim a$、且包含了至少两种不一样的字符。
- 1. 答案为a^p-a。（所有方案，扣除掉只含一个字符的。）
- 2. 这些字符串每p个循环同构的为一组
  - 当$\boldsymbol{s}=s_1s_2\ldots s_p$ 满足要求，则$\boldsymbol{s}$的所有cyclic-shift都满足要求，构成一组。
    - （注意：**由于p为素数且s有2种以上字符，cyclic-shift会各不相同**）
  - 因此答案为 gp，其中g为组数（即循环同构下本质不同的方案数）。
- 故，a^p-a=gp。因此p 整除 a^p-a。

双重计数方法证明一些不等式

- **例9** 平面图存在度数<=5的点。
- **例10** 无三角的图的边数的最大值
- **例11** LYM不等式
  - EKR定理 （见本ppt中**更难的思考题一**） 等等。

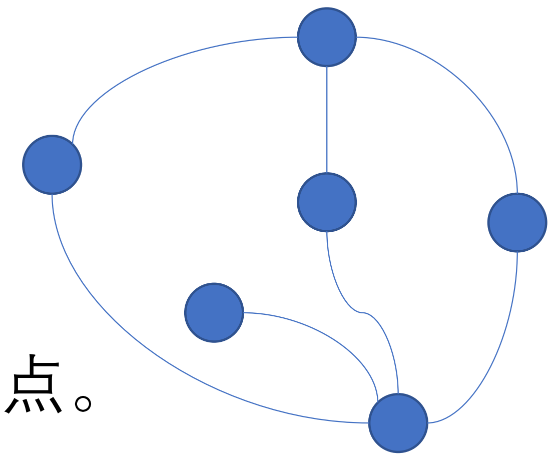
例9 平面连通图中若无重边，求证则存在一个度数 ≤ 5的点。

反证法。假定每个点度数 ≥ 6。

- 不妨设图中有**F**个面（包括无限大那个面）、**E**条边、**V**个点。
- **观察：** **$3F \le 2E$**（double counting）。下页证明。
- **观察：** **$6V \le 2E$** (double counting)。下页证明。
- 所以$F \le \frac{2}{3}E$。$V \le \frac{1}{3}E$。从而**$F+V \le E$**。
- 由欧拉公式，**$F+V=E+2$**
- 矛盾！
- 因此不可能每个点度数≥6。也就是存在度数≤ 5的点。

归纳法。初始F=1, E=V-1。
每加一条边，V不变，E++,F++。

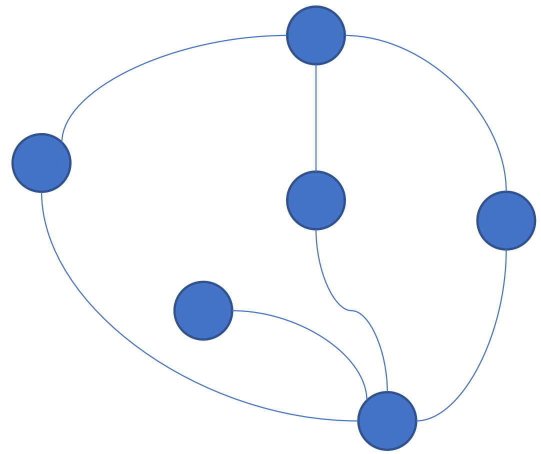
V=6,E=7,F=3.



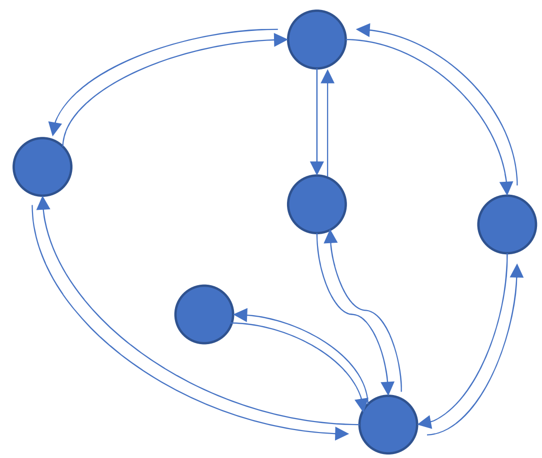
实际上，存在一个度数 ≤ 4的点（证明难很多）。

证明$3F \leq 2E$、$6V \leq 2E$。

- **观察**$\sum_{面f} e(f) = 2E$，其中e(f) 表示 f 的边数。
- 注意$e(f) \geq 3$ （因为无重边）。
- 故$2E = \sum_{面f} e(f) \geq 3F$。因此$3F \leq 2E$ 。



- **观察**$\sum_{点v} d(v) = 2E$，其中d(v)表示 v的度数。
- 注意d(v) ≥ 6 （因为假定每个点度数 ≥ 6）
- 故$2E = \sum_{点v} d(v) \geq 6V$ 。因此$6V \leq 2E$ 。



将每条无向边拆成两条相反的有向边。
上述两个绿色等式都是对这些有向边做double counting。

例10 无三角的图边数不能超过 $n^2/4$

- 设无向图G=(V,E)不含三角形，|V|=n。求证$|E| \le \frac{n^2}{4}$。
- 由于不含三角形。对于任意一条边(x,y)，d(x)+d(y) ≤ n.
- 因此$\sum_{(x,y)\in E}(d(x)+d(y)) \le \mathrm{n}|\mathrm{E}|$。
- $\sum_{(x,y)\in E}(d(x)+d(y)) = \sum_{x\in \mathrm{V}} \mathrm{d}^2(x)$ (由double counting)
- $\sum_{x\in \mathrm{V}} \mathrm{d}^2(x) \ge n\left(\frac{\sum_{x\in \mathrm{V}} d(x)}{n}\right)^2 = n^{-1}(2|E|)^2 = \frac{4|E|^2}{n}$。
- 因此$\frac{4|E|^2}{n} \le n|E|$。因此$|E| \le \frac{n^2}{4}$。

Double counting technique 总结

- 对同一个**集合**，用两种不同的 方法计数
  - 常用的方法：交换求和次序。
- 难点大多在于：定义一个好用的“待计数的**集合**”。

- Double counting technique可应用于
  - 证明等式、简化等式。
  - 证明一些不等式。
  - 计算一些染色数（比如项链计数）

练习题 LYM不等式 （备用例题）

- 假定**A**包含$\{1,\ldots,n\}$的若干个“互不包含的”子集$S_1,\ldots,S_m$，并且其中大小为k的子集有a_k个，则 $\Sigma_k\ a_k/\binom{n}{k} \le 1$。
- 证明:
  - 首先，两边同时乘以$n!$ 转证：$\Sigma_k a_k k!\,(n-k)! \le n!$。
  - 令$s_i=|S_i|$。转证 $\Sigma_i s_i!\,(n-s_i)! \le n!$。
  - 考虑$\{1,\ldots,n\}$的排列。
    - 若这个排列的前s_i项的集合刚好为S_i，它就归为S_i“管辖”。
    - 由于互不包含，可知一个排列排列最多被一个集合管辖。
    - 因此，恰好有 $\Sigma_i s_i!\,(n-s_i)!$ 个（不同的）排列被管辖。
  - 另一方面，排列的总数为$n!$。因此，$\Sigma_i s_i!\,(n-s_i)! \le n!$。

练习题

Grid https://cses.fi/problemset/task/2210

- **题意**： n*n的格子黑白染色。求本质不同的方案数（旋转后90,180,270后相同的染色方案认为是同一种）。
- 以n=4为例子。
  - 假设C_1~C_T是本质不同的着色。 把它们每一个 都转4次。列成表格。
  - C表示一种着色。f_C表示这个着色C 在 表中出现几次。（1，2，4）。
- $f_C=4$的 有2^4个。
- $f_C=2$的 有2^8-2^4个。
- $f_C=1$的 有$2^{16}-2^8$个。
- 因此， $4T = 4\ 2^4 + 2(2^8-2^4) + 1(2^{16}-2^8)$。 By double counting。
- 故， $T = 2^4 + 2^7 - 2^3 + 2^{14} - 2^6$。

拓展阅读1：二次互反律

- [Proofs of quadratic reciprocity - Wikipedia](#)
  这里Eisenstein的证明是double counting。精妙。

Law of quadratic reciprocity — Let p and q be distinct odd prime numbers, and define the Legendre symbol as:

$$\left(\frac{q}{p}\right) = \begin{cases} 1 & \text{if } n^2 \equiv q \bmod p \text{ for some integer } n \\ -1 & \text{otherwise} \end{cases}$$

Then:

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

拓展阅读2： Burnside引理

Burnside引理（主要也是用了一个double counting）

In the following, let G be a finite group that acts on a set X. For each g in G let X^g denote the set of elements in X that are fixed by g (also said to be left invariant by g), i.e. $X^g = \{ x \in X \mid g.x = x \}$. Burnside's lemma asserts the following formula for the number of orbits, denoted $|X/G|$:[2]

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

这是一个更一般的结论。 旋转、翻转意义下本质相同，仍然可以解决。
（可以解决许多计数的问题）

看懂Burnside引理的证明，大约需要1-2天（需要比较多基础引理）