程序设计艺术与方法

第五讲　组合数学简介

合肥工业大学计算机与信息学院　2014-9

5.1 概述

5.2 补集转化

5.3 组合数学中的递推关系

5.4 母函数

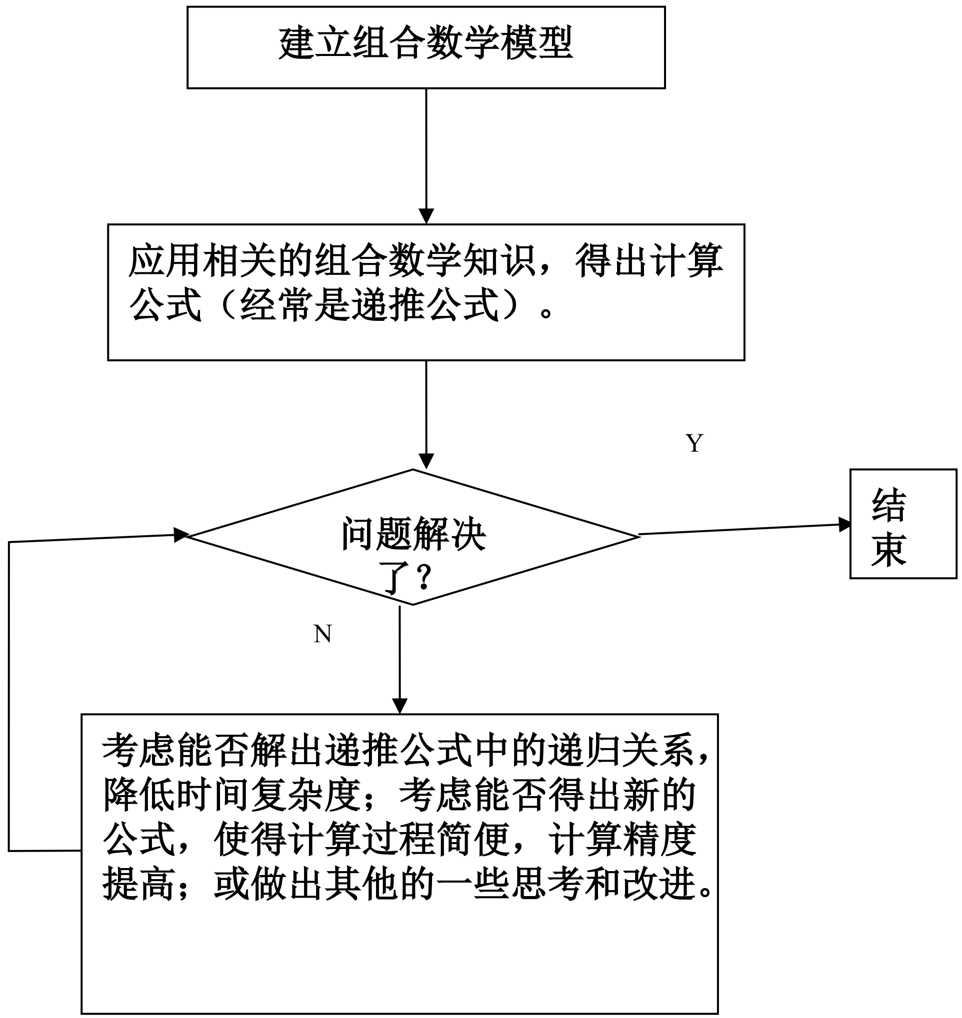
5.5 Pólya原理及应用

5.1 概述

◆组合数学是计算机出现以后迅速发展起来的一门数学分支

◆介绍一下组合数学在比赛中的一些应用

◆解组合数学题目的一般步骤



5.1 概述

5.2 补集转化

5.3 组合数学中的递推关系

5.4 母函数

5.5 Pólya原理及应用

5.2 补集转化

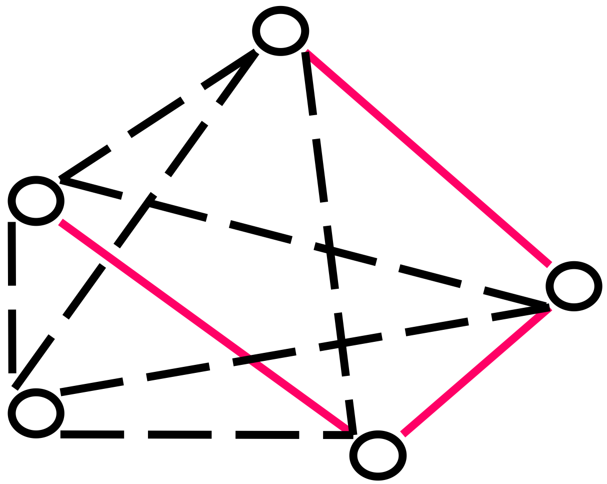
- 逆向思维方法，应用于组合数学问题中往往有着很好的效果

- 下面通过单色三角形问题的解题探讨补集转化思想

单色三角形问题

- 空间里有n个点，任意三点不共线。每两点之间都用红色或黑色线段（只有一条，非红即黑！）连接。如果一个三角形的三条边同色，则称这个三角形是单色三角形。对于给定的红色线段的列表，找出单色三角形的个数。
- 输入点数n、红色边数m以及这m条红色的边所连接的顶点标号，输出单色三角形个数R。3<=n<=1000，0<=m<=25万。

例如右图中有5个点，10条边，形成3个单色三角形



单色三角形问题

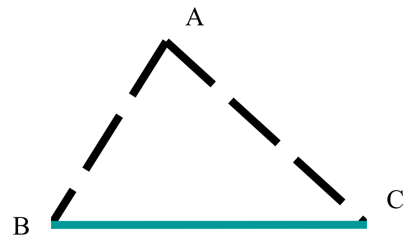
- 很自然想到了枚举所有的三角形（这是通过枚举三个顶点实现的），判断它的三条边是否同色，时间复杂度已经高达O(n^3)
- 单纯枚举不可以，组合计数是否可行？利用组合公式进行计算是非常高效的。
- 从总体上进行组合计数很难想到，尝试枚举每一个点，设法找到一个组合公式来计算以这个点为顶点的单色三角形的个数

单色三角形问题

- 对于枚举确定的点A，以A为一个顶点的单色三角形ABC不仅要满足边AB和边AC同色，而且边BC也要和AB、AC边同色，不可能仅仅通过枚举一个顶点A就可以确定单色三角形
- 枚举＋组合计数中组合公式的构造上我们遇到了障碍
- 障碍的本质是：从一个顶点A出发的两条同色的边AB、AC并不能确定一个单色三角形ABC，因为BC边有可能不同色

单色三角形问题

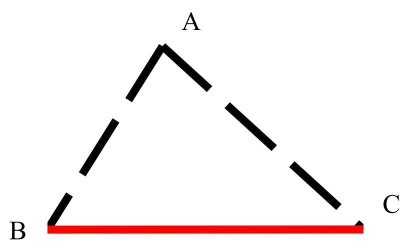
- 也就是说，无法在从同一个顶点出发的某两条边与所有的单色三角形之间建立一种确定的对应关系
- 换一个角度，从反面来看问题
- 三角形数S=C(n,3)
- 单色三角形数R加上非单色三角形数T就等于S，如果可以求出T，那么显然R = S － T
- 原问题就等价于：怎样高效地求出T



单色三角形问题

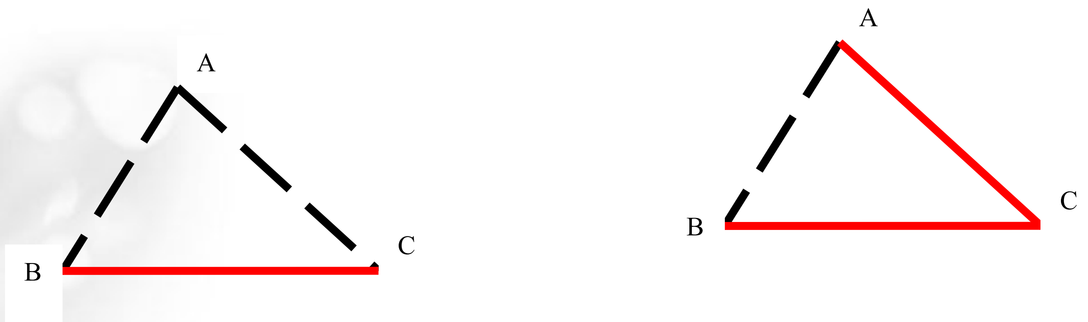
- 上面分析中夭折的枚举＋组合计数的算法的障碍是无法在“某两条边”与“单色三角形”之间建立确定的对应关系
- 那么有公共顶点的某两条边与非单色三角形之间是否有着确定的关系呢？

非单色三角形的三条边，共有红黑两种颜色，也就是说，只能是两条边同色，另一条边异色。假设同色的两条边顶点为A，另外两个顶点为B和C，则从B点一定引出两条不同色的边BA和BC，同样，从C点引出两条不同色的边CA和CB



单色三角形问题

- 一个非单色三角形对应着两对“有公共顶点的异色边”
- 很明显要求的非单色三角形数T就等于所有“有公共顶点的异色边”的总对数Q的一半



单色三角形问题（例1）

- 总对数：每个顶点有n-1条边，根据输入的信息可以知道每个顶点i的红边数E[i]，那么其黑边数就是n-1-E[i]。枚举顶点A，则根据乘法原理，以A为公共顶点的异色边的对数就是E[i]*(n-1-E[i])。所以

$$Q=\sum_{i=1}^{n} E[i]*(n-1-E[i])$$

- 答案R=S-T=n*(n-1)*(n-2)/6-Q/2

单色三角形问题

直观的枚举算法无法实现

↓

考虑枚举＋组合计数

↓

正面分析，组合公式很难构造

↓

从反面思考（补集转化思想），得到突破

5.1 概述

5.2 补集转化

5.3 组合数学中的递推关系

5.4 母函数

5.5 Pólya原理及应用

5.3 组合数学中的递推关系

- 建立递推关系的关键在于寻找第n项与前面几项的关系式，以及初始项的值

- 下面介绍几种典型的递推关系

Fibonacci数列

- **有雌雄一对兔子，假定过两个月便可繁殖雌雄各一的一对小兔子。问过n个月后共有多少对兔子？**
- 解：设满x个月共有兔子F_x对，其中当月新生的兔子数目为N_x对。第x-1个月留下的兔子数目设为O_x对。则：

$$F_x=N_x+O_x$$

而 $O_x=F_{x-1}$，

$N_x=O_{x-1}=F_{x-2}$ (即第x-2个月的兔子到第x个月成熟了)

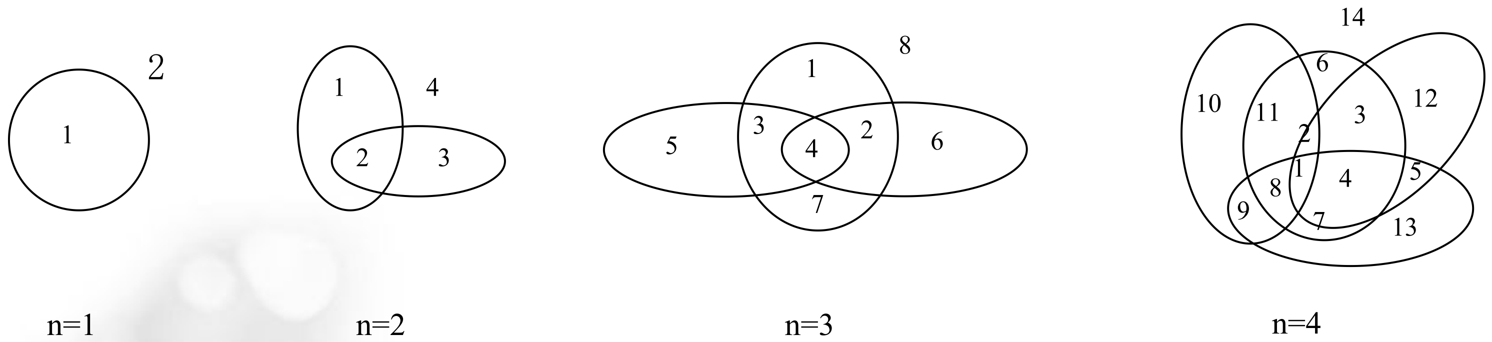
∴ $F_x=F_{x-1}+F_{x-2}$ 边界条件：$F_0=0$，$F_1=1$

平面分割问题

- 问题：
  - 设有n条封闭曲线画在平面上，而任何两条封闭曲线恰好相交于两点
  - 任何三条封闭曲线不相交于同一点
  - 问这些封闭曲线把平面分割成的区域个数

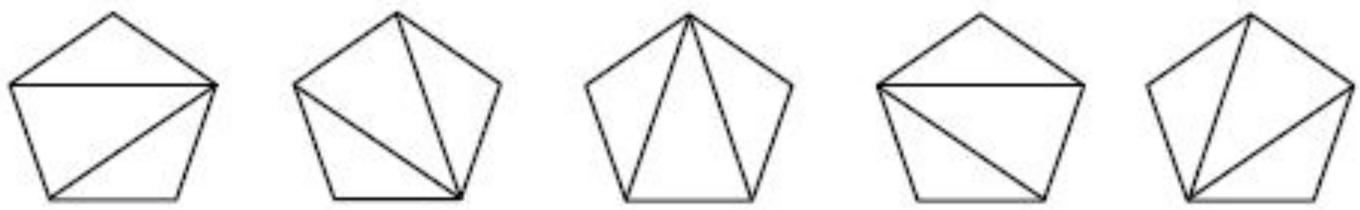
平面分割问题

- 设a_n为n条封闭曲线把平面分割成的区域个数。 由下图可以看出：$a_2-a_1=2$；$a_3-a_2=4$；$a_4-a_3=6$。可以看出$a_n-a_{n-1}=2(n-1)$
- 试着证明一下



Catalan数

- 一个凸n边形，通过不相交于n边形内部的对角线，把n边形拆分成若干三角形
- 不同的拆分数目用h_n表之，h_n即为Catalan 数
- 例如五边形有如下五种拆分方案，故$h_5=5$
- 求对于一个任意的凸n边形相应的h_n

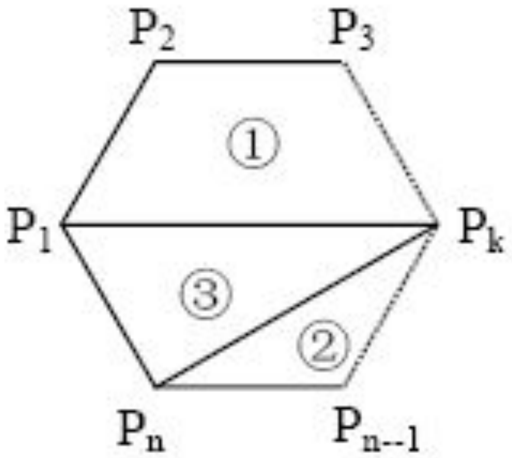


Catalan数（例2）

- 设C_n表示凸n边形的拆分方案总数。可以通过递推关系得到：

$$C_n = \sum_{i=2}^{n-1} C_i C_{n-i+1}$$

- 其中：$C_2=1$



5.1 概述

5.2 补集转化

5.3 组合数学中的递推关系

5.4 母函数

5.5 Pólya原理及应用

先看一个多项式乘法

$$(1+a_1x)(1+a_2x)\cdots(1+a_nx)$$
$$=1+(a_1+a_2+\cdots+a_n)x+(a_1a_2+a_1a_3+\cdots+a_{n-1}a_n)x^2$$
$$+\cdots+a_1a_2\cdots a_nx^n$$

可以看出：

x^2项的系数$a_1a_2+a_1a_3+\ldots+a_{n-1}a_n$中所有的项包括$n$个元素$a_1$，$a_2$，…，$a_n$中取两个组合的全体；

同理：x^3项系数包含了从n个元素a_1，a_2，…，a_n中取3个元素组合的全体；

以此类推。

特例

若令$a_1=a_2=\ldots=a_n=1$，在上式中$a_1a_2+a_1a_3+\ldots+a_{n-1}a_n$项系数中每一个组合有1个贡献，其他各项以此类推。故有：

$$(1+x)^n=1+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n$$

母函数的定义

- 对于序列a_0, a_1，a_2，…构造一函数：

$$G(x) = a_0 + a_1 x + a_2 x^2 + \cdots,$$

- 称函数G(x)是序列a_0, a_1，a_2，…的母函数

母函数的定义

$(1+x)^n$是序列C(n,0), C(n,1), ..., C(n,n)的母函数。

如若已知序列a_0, a_1， a_2， …则对应的母函数G(x)便可根据定义给出。

反之，如若已经求得序列的母函数G(x)，则该序列也随之确定。

序列a_0, a_1， a_2， …可记为$\{a_n\}$ 。

称重量

- 若有1克、2克、3克、4克的砝码各一枚，能称出哪几种重量？各有几种可能方案？

如何解决这个问题呢？考虑构造母函数。

如果用x的指数表示称出的重量，则：

1个1克的砝码可以用函数$1+x$表示，

1个2克的砝码可以用函数$1+x^2$表示，

1个3克的砝码可以用函数$1+x^3$表示，

1个4克的砝码可以用函数$1+x^4$表示，

称重量

几种砝码的组合可以称重的情况，可以用以上几个函数的乘积表示：

$$(1+x)(1+x^2)(1+x^3)(1+x^4)$$
$$=(1+x+x^2+x^3)(1+x^3+x^4+x^7)$$
$$=1+x+x^2+2x^3+2x^4+2x^5+2x^6+2x^7+x^8+x^9+x^{10}$$

从上面的函数知道：可称出从1克到10克，系数便是方案数。

例如右端有$2x^5$ 项，即称出5克的方案有2：

5=3+2=4+1；同样，6=1+2+3=4+2；10=1+2+3+4。

故称出6克的方案有2，称出10克的方案有1

贴邮票

求用1分、2分、3分的邮票贴出不同数值的方案数。

- 因邮票允许重复，故母函数为：

$$G(x)=(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)\cdot(1+x^3+x^6+\cdots)$$

- 以展开后的x^4为例，其系数为4，即4拆分成1、2、3之和的拆分数为4；
- 即：4=1+1+1+1=1+1+2=1+3=2+2

整数拆分（例3）

所谓整数拆分即把整数分解成若干整数的和（相当于把n个无区别的球放到n个无标志的盒子，盒子允许空，也允许放多于一个球）。

$$G(x)=(1+x+x^2+\cdots)(1+x^2+x^4+\cdots)$$

$$\cdot(1+x^3+x^6+\cdots)\ \cdots$$

整数拆分（例3）

所谓整数拆分即把整数分解成若干整数的和（相当于把n个无区别的球放到n个无标志的盒子，盒子允许空，也允许放多于一个球）。

首先思考：如果让你手工计算，你是怎样处理的？

实际编程：让计算机按照自己的思路计算即可～

5.1 概述
5.2 补集转化
5.3 组合数学中的递推关系
5.4 母函数
5.5 Pólya原理及应用

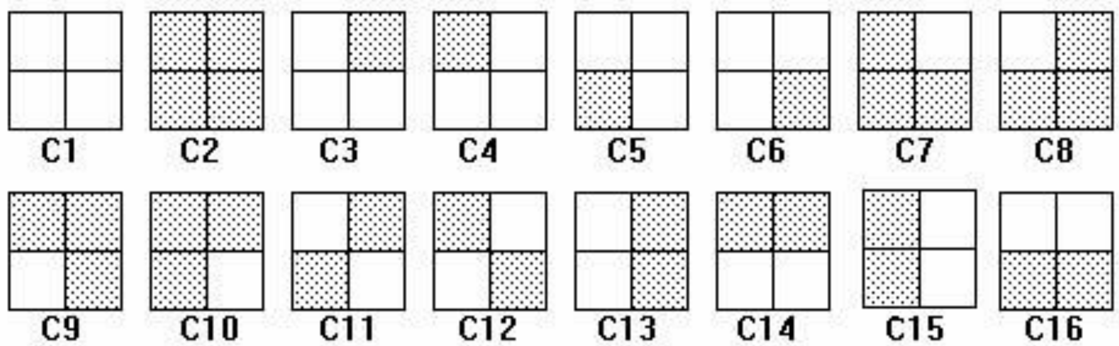
5.5 Pólya原理及应用

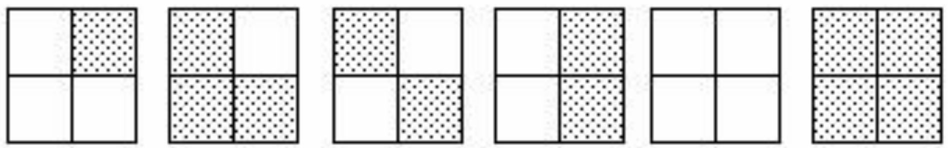
- Pólya原理是组合数学中，用来计算全部互异的组合状态的个数的一个十分高效、简便的工具

- 先看下面例题：

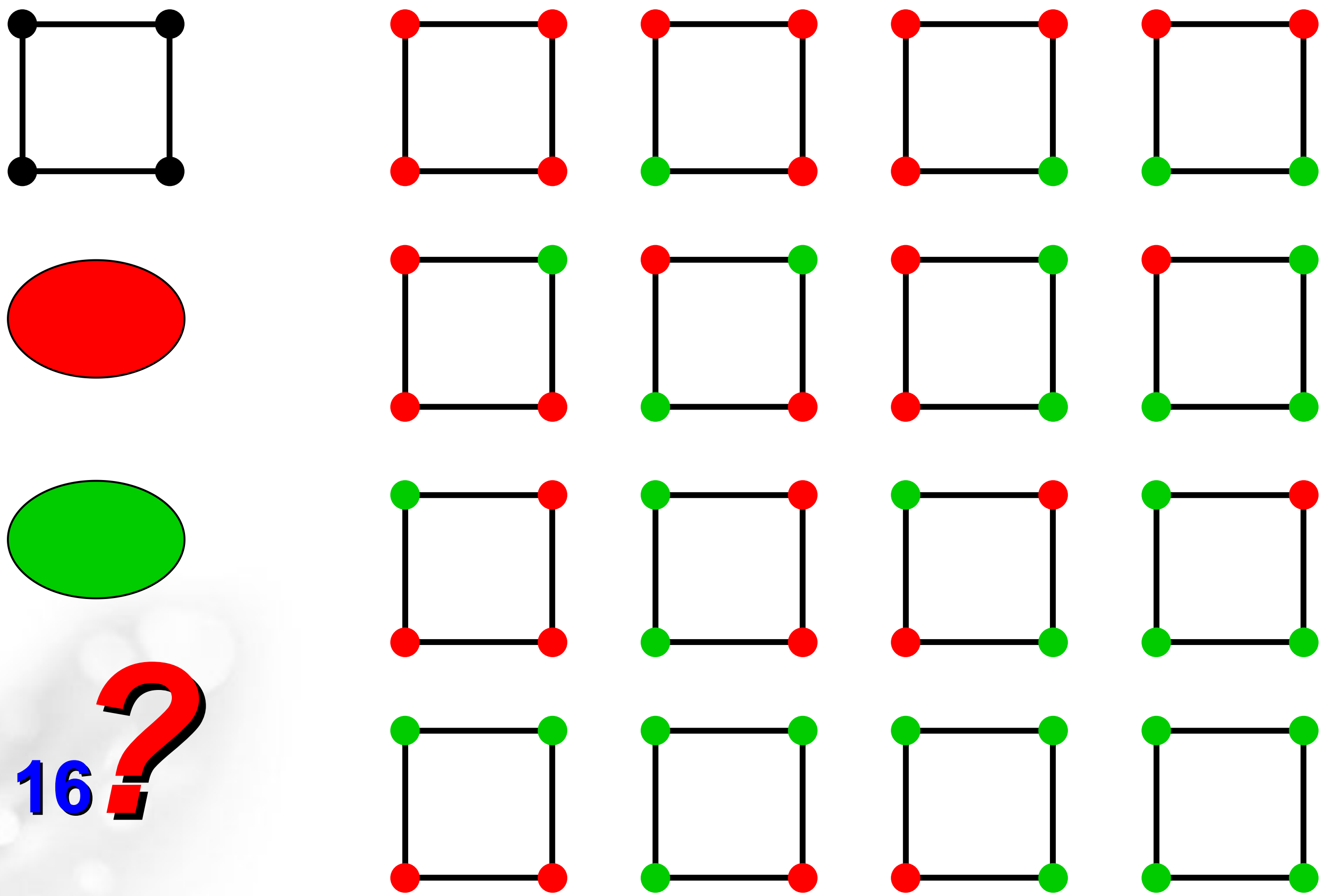
5.5 Pólya原理及应用

- 对2*2的方阵用黑白两种颜色涂色，问能得到多少种不同的图像？经过旋转使之吻合的两种方案，算是同一种方案
- 先把所有的涂色方案列举出来
- 旋转方法一共有4种：旋转0度、旋转90度、旋转180度和旋转270度

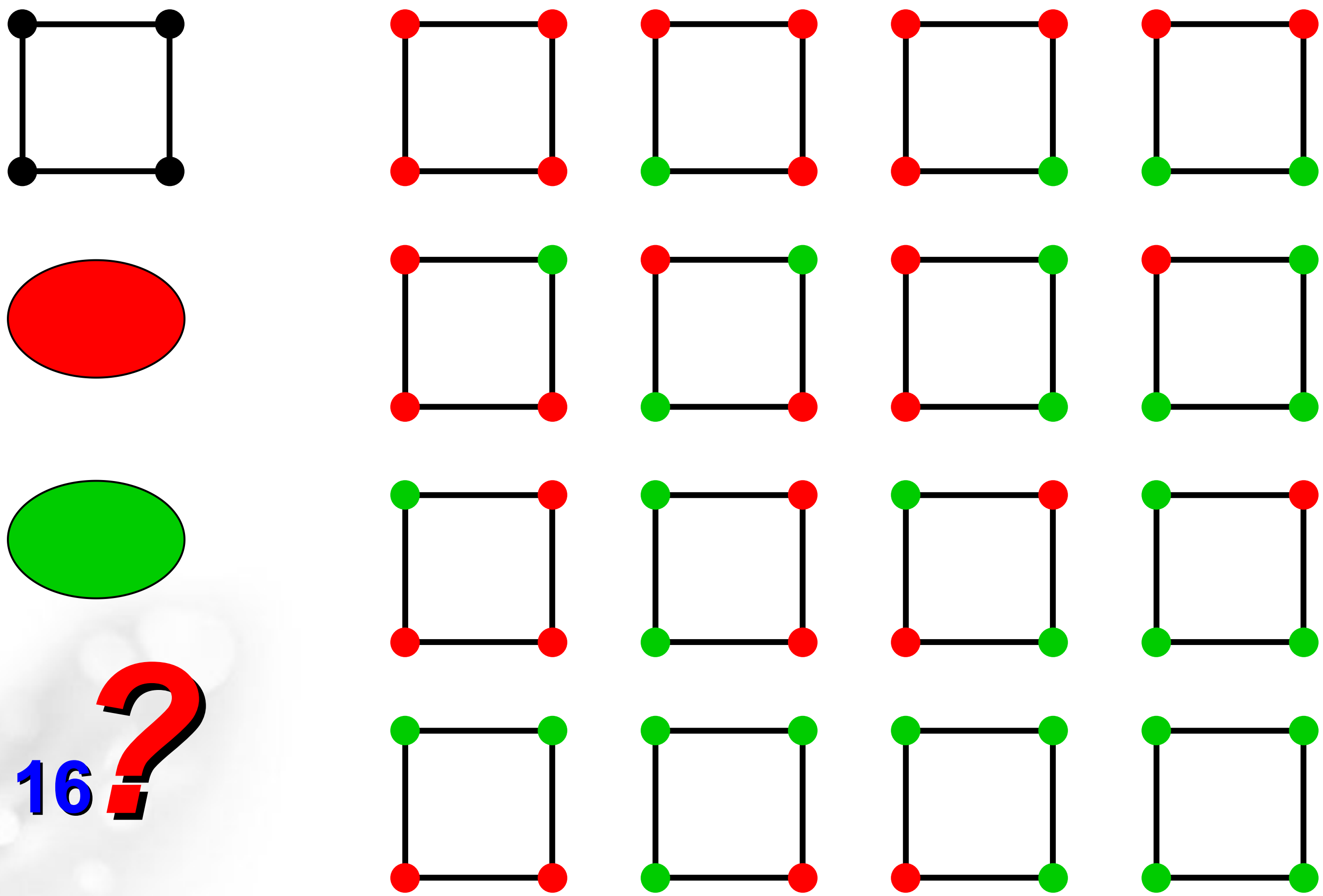




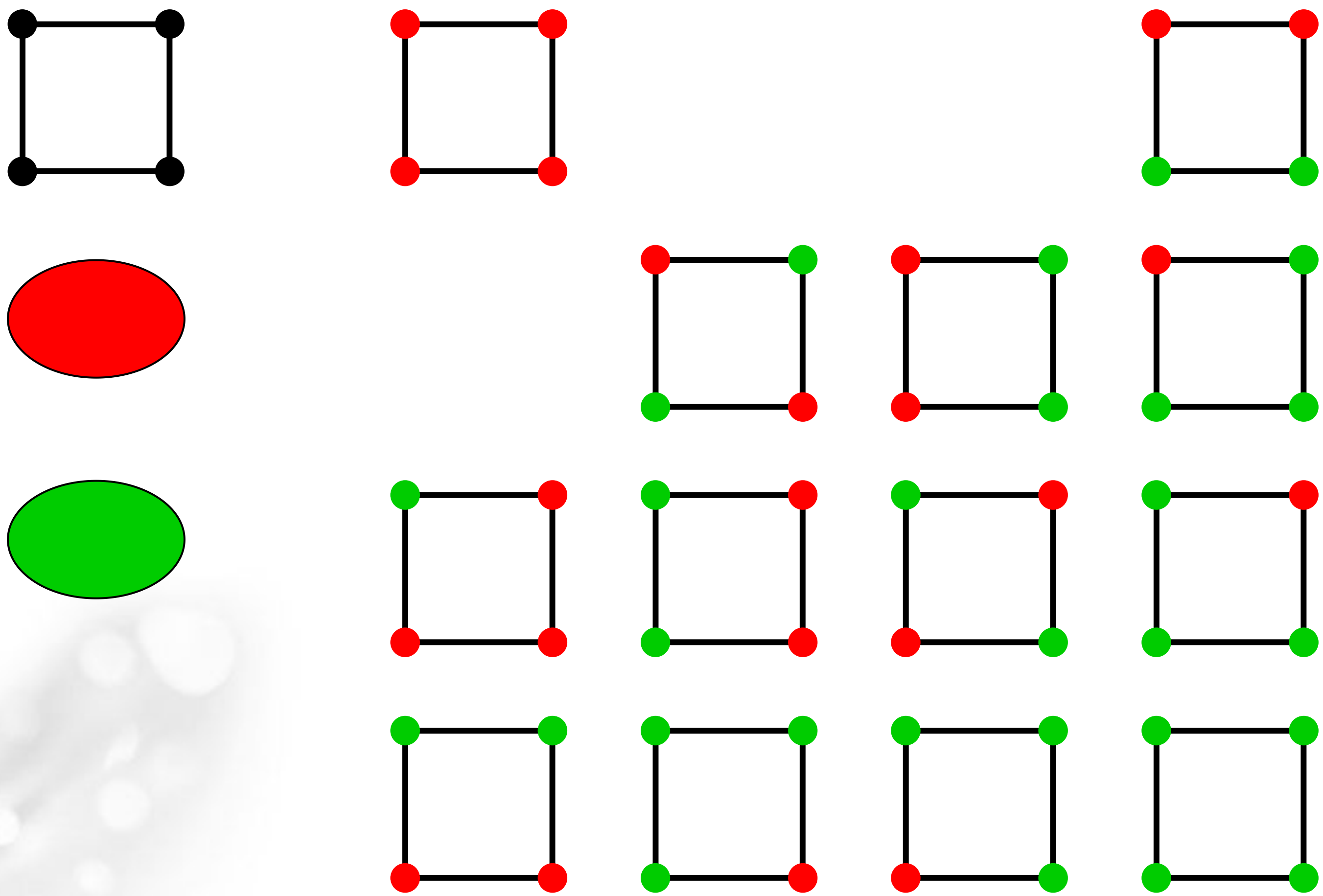
问题



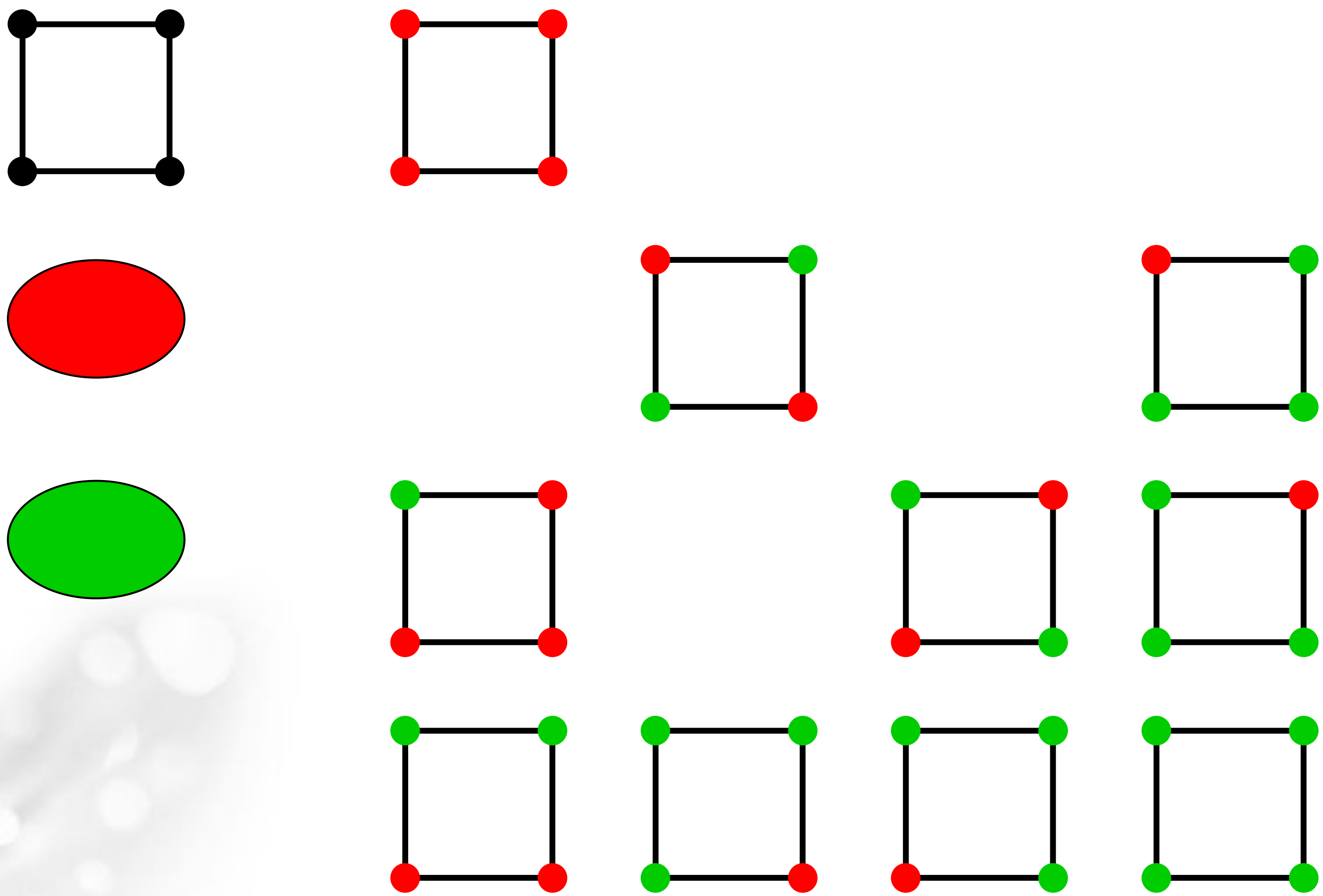
问题



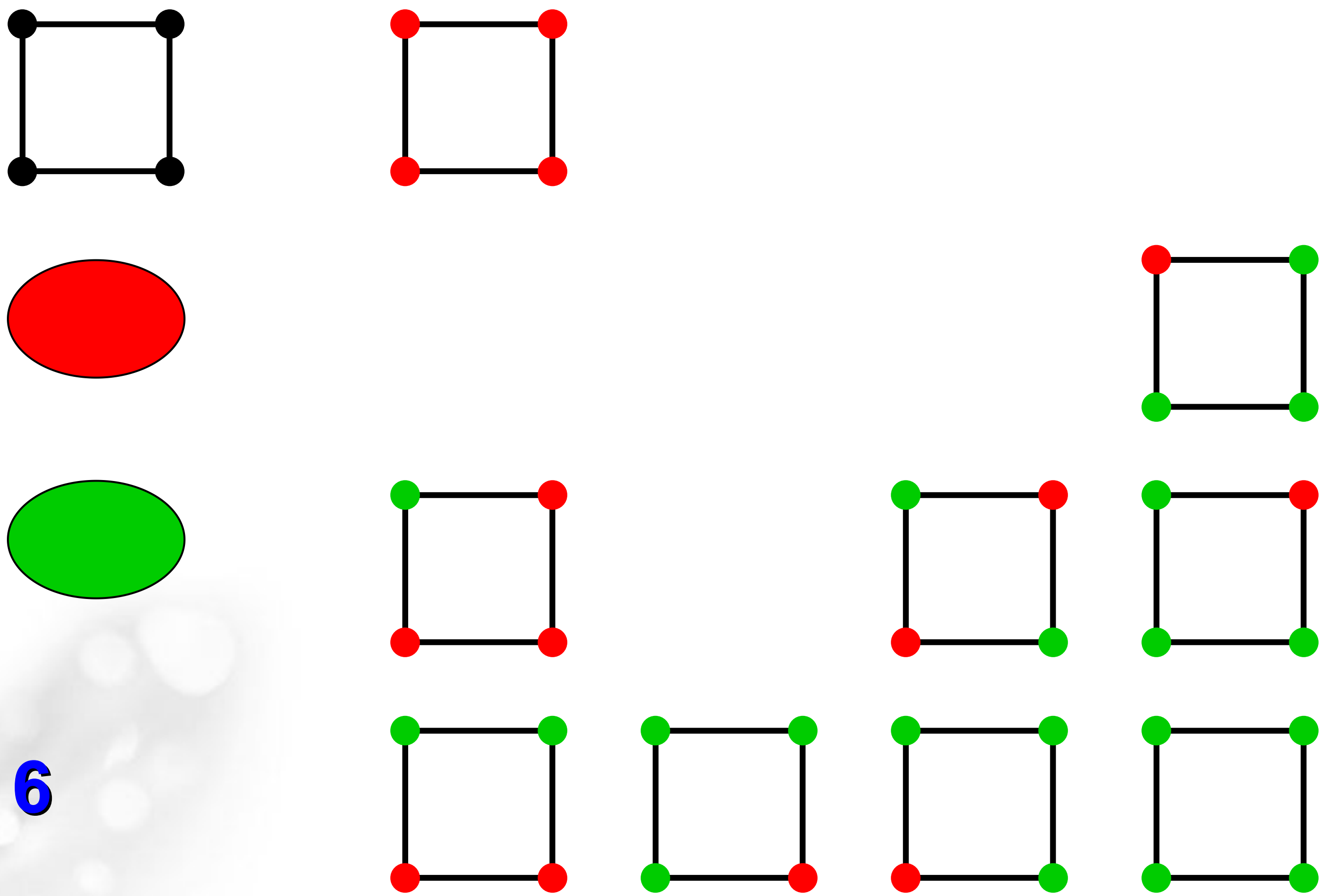
问题



问题

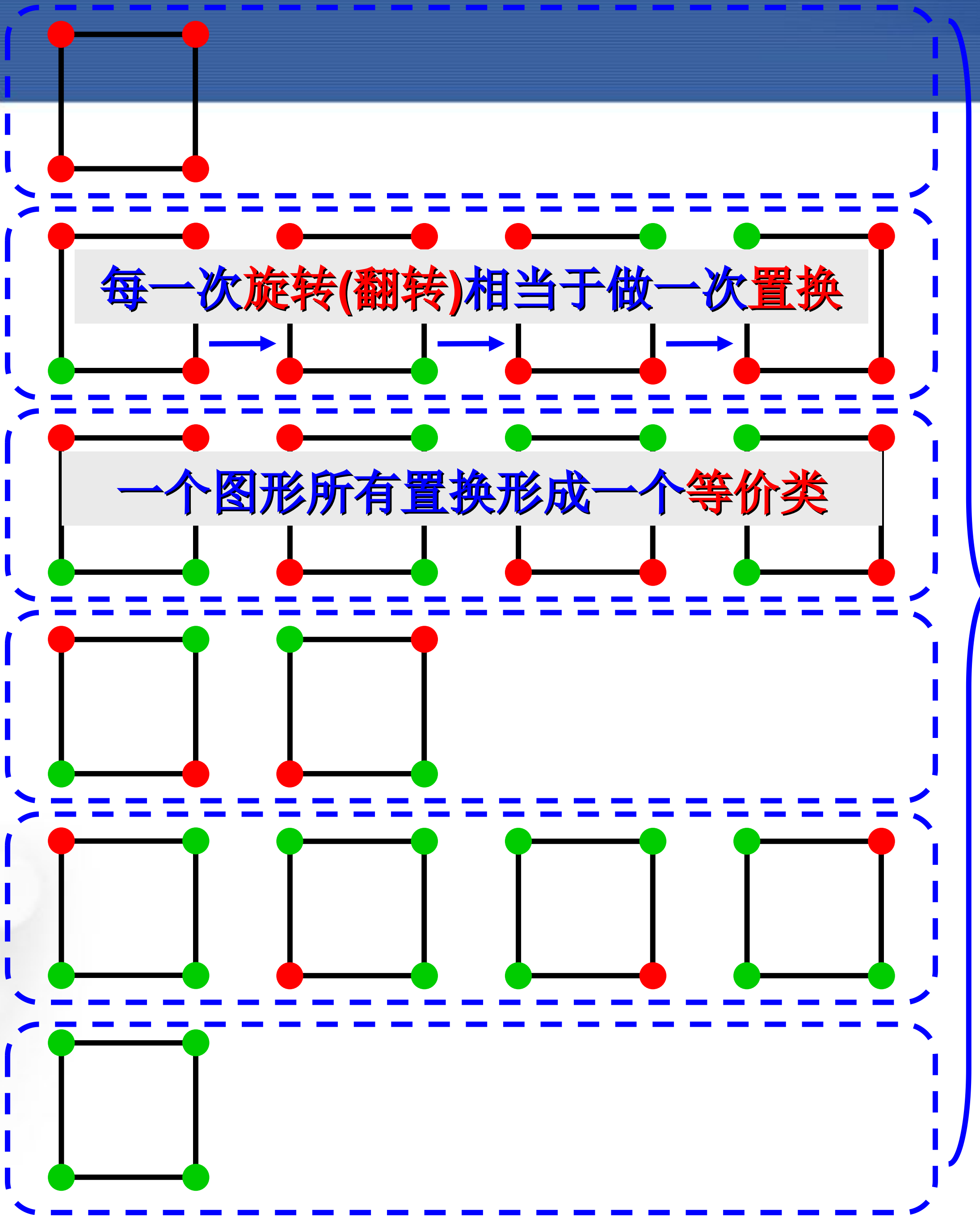


问题



6

问题



计数问题↓求等价类个数问题

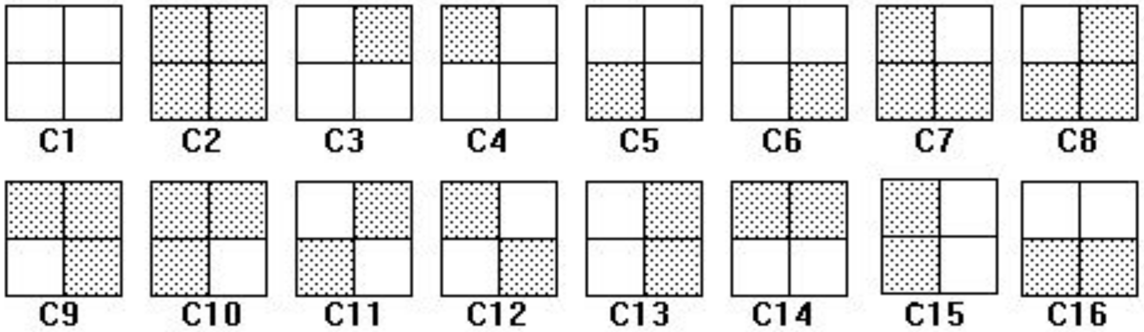
5.5 Pólya原理及应用

- 群：给定一个集合$G=\{a,b,c,\ldots\}$和集合G上的二元运算，并满足：
  - (a) 封闭性：$\forall a,b\in G, \exists c\in G, a * b=c$。
  - (b) 结合律：$\forall a,b,c\in G, (a * b) * c=a * (b * c)$。
  - (c) 单位元：$\exists e\in G, \forall a\in G, a * e=e * a=a$。
  - (d) 逆元：$\forall a\in G, \exists b\in G, a * b=b * a=e$，记$b=a^{-1}$。

  则称集合G在运算$*$之下是一个群，简称G是群。一般$a * b$简写为ab。
- 置换：n个元素$1,2,\ldots,n$之间的一个置换 $\begin{pmatrix} 1 & 2 & \cdots & n \\ a_1 & a_2 & \cdots & a_n \end{pmatrix}$

  表示1被1到n中的某个数a_1取代，2被1到n中的某个数a_2取代，直到n被1到n中的某个数a_n取代，且$a_1,a_2,\ldots,a_n$互不相同。

5.5 Pólya原理及应用



- 本例中有4个置换
  - 转0° a1= $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \end{pmatrix}$
  - 转90° a2= $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 1 & 2 & 6 & 3 & 4 & 5 & 10 & 7 & 8 & 9 & 12 & 11 & 16 & 13 & 14 & 15 \end{pmatrix}$
  - 转180° a3= $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 1 & 2 & 5 & 6 & 3 & 4 & 9 & 10 & 7 & 8 & 11 & 12 & 15 & 16 & 13 & 14 \end{pmatrix}$
  - 转270° a4= $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 1 & 2 & 4 & 5 & 6 & 3 & 8 & 9 & 10 & 7 & 12 & 11 & 14 & 15 & 16 & 13 \end{pmatrix}$

置换群

- 置换群的元素是置换，运算是置换的连接
- 例如

$$\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 1 & 2 & 4\end{pmatrix}\begin{pmatrix}1 & 2 & 3 & 4\\ 4 & 3 & 2 & 1\end{pmatrix}$$

$$=\begin{pmatrix}1 & 2 & 3 & 4\\ 3 & 1 & 2 & 4\end{pmatrix}\begin{pmatrix}3 & 1 & 2 & 4\\ 2 & 4 & 3 & 1\end{pmatrix}$$

$$=\begin{pmatrix}1 & 2 & 3 & 4\\ 2 & 4 & 3 & 1\end{pmatrix}$$

- 设 σ 为 A 上的一个置换，若A 中存在 k 个元素a_1, a_2, ..., a_k，使 $\sigma(a_1) = a_2$, $\sigma(a_2) = a_3$,..., $\sigma(a_k) = a_1$, 其余 $\sigma(a_i) = a_i$, 称 σ 为长度为 k 的循环，简称 ***k*-循环**，记为 $(a_1a_2...a_k)$。
- 性质
  - $(a_1a_2...a_k) = (a_2...a_ka_1) = (a_ka_1a_2...a_{k-1})$
  - 不相交循环的乘积可交换
  - 若$P = (1,2,...,n)$，则$P^n = (1)\ (2)...\ (n)$
  - 任何一个置换可分解为若干个循环的乘积

循环

- $C(P_i)$：表示 P_i 中循环的个数
- 循环的阶：循环中元素的个数
- $C_k(P_i)$：表示 P_i 中 k 阶循环的个数
- 例

$$P_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 1 & 5 & 2 \end{pmatrix} = (1\,4\,5\,2\,3)$$

$$P_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 1 & 4 \end{pmatrix} = (1\,5\,4)(2)(3) = (2)(1\,5\,4)(3)$$

- 于是， $C(P_2) = 3, C_1(P_2) = 2, C_2(P_2) = 0, C_3(P_2) = 1$

k不动置换类

- 设 G 是 $N=\{1,2,...,n\}$ 上的置换群，若 $P_i\in G$ 将元素$k\in N$映射到其自身，称元素 k 为在置换 P_i 作用下的**不动元**。
- 设 G 是 $N=\{1, 2, ..., n\}$ 上的置换群，G中使k为不动元的置换的全体为 **k不动置换类，用Z_k表示**。
- **例** $N=\{1, 2, 3, 4\}$，$G=\{P_1, P_2, P_3, P_4\}$

  $P_1=(1)(2)(3)(4)$，$P_2=(12)(3)(4)$

  $P_3=(1)(2)(34)$，$P_4=(12)(34)$, 求k不动置换类
- 解： $Z_1=\{P_1, P_3\}=Z_2$

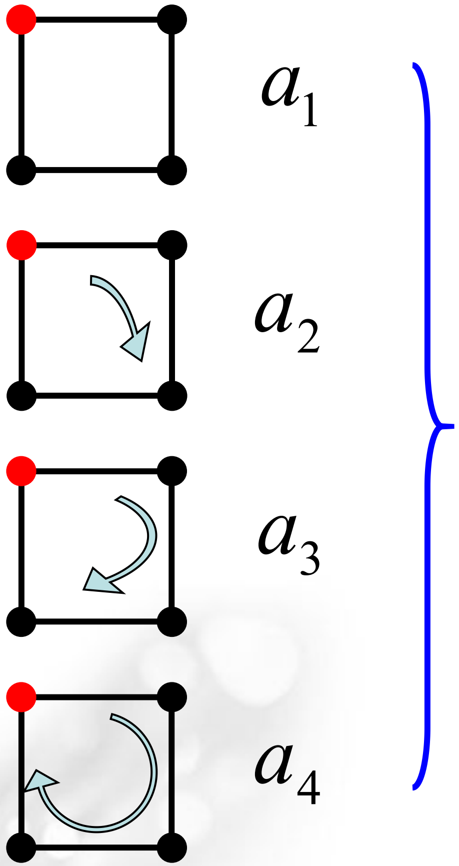
  $Z_3=\{P_1, P_2\}=Z_4$

等价类

- 定义：设G为N上的置换群，对于数k,若存在置换$P_i \in G$,将k映射为j,则称k与j等价，记作$k \sim j$，N中与k等价的元素的全体称为k的**等价类**，记作E_k。

- **例** 求上例中等价类。
- 解：$E_1=E_2=\{1,2\}, E_3=E_4=\{3,4\}$

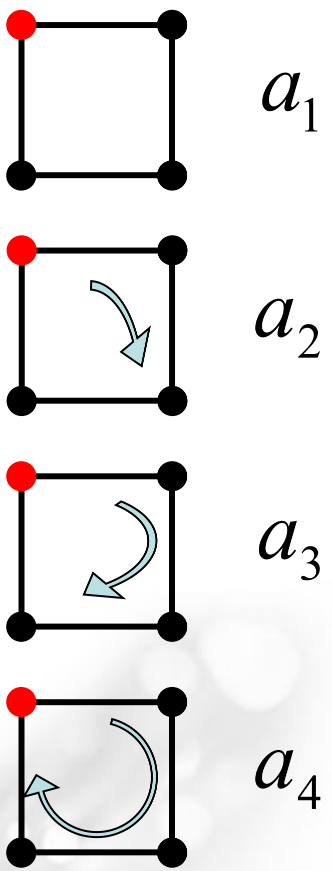
置换群

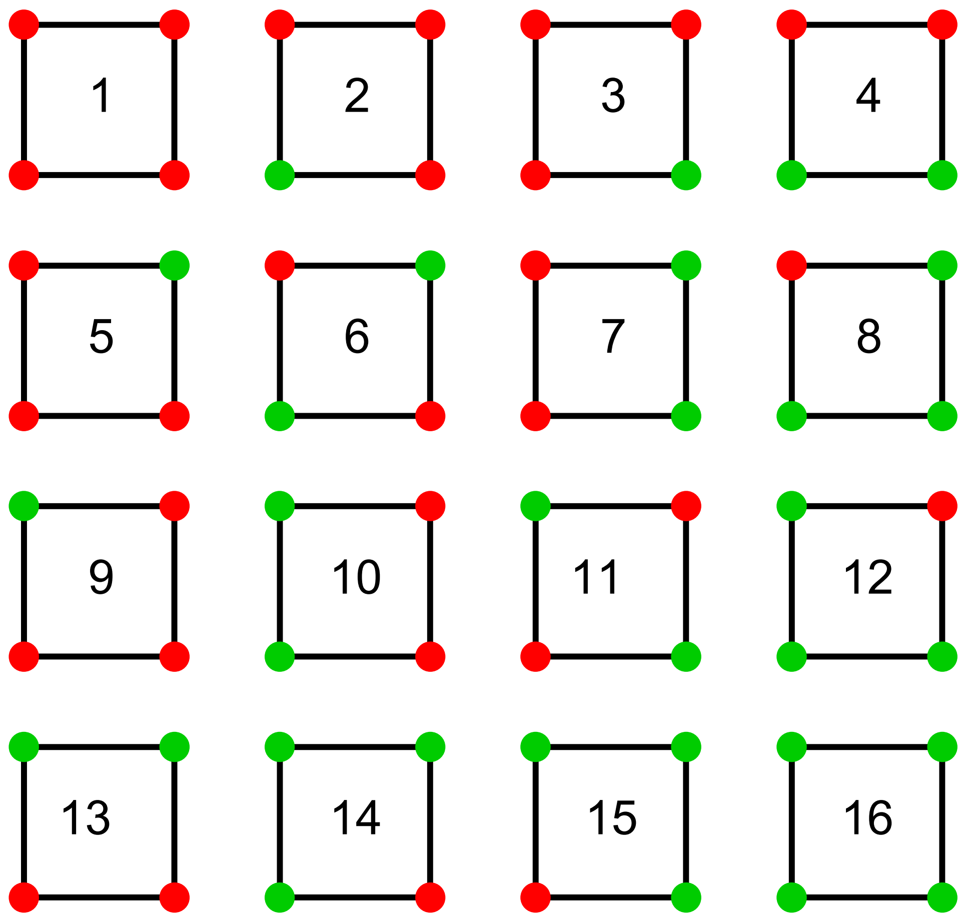


$$G = \{a_1, a_2, a_3, a_4\}$$

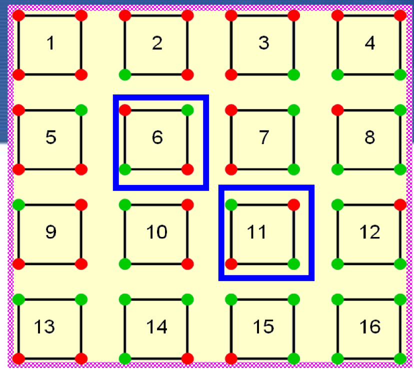
G为置换群

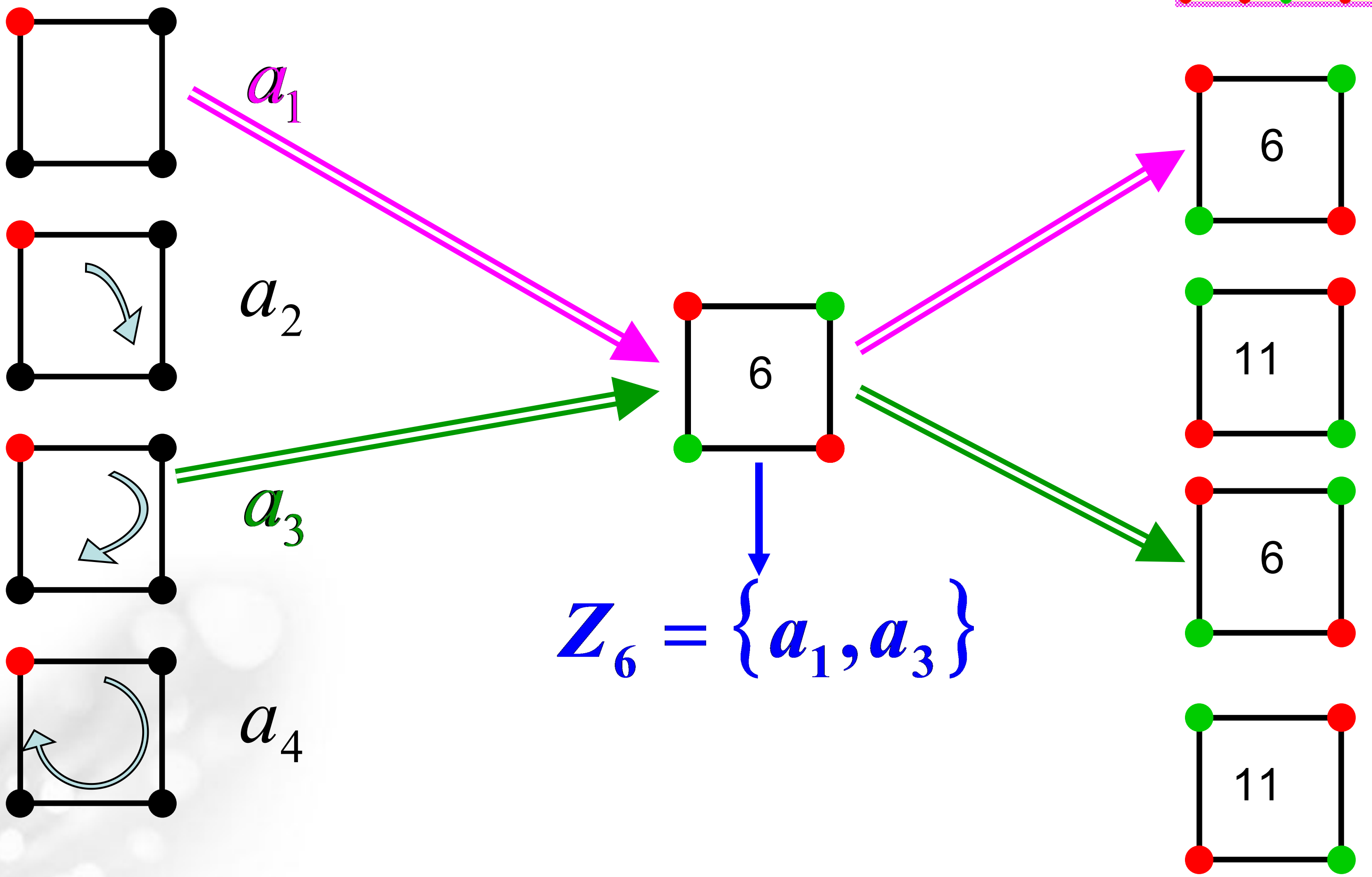
k不动置换类





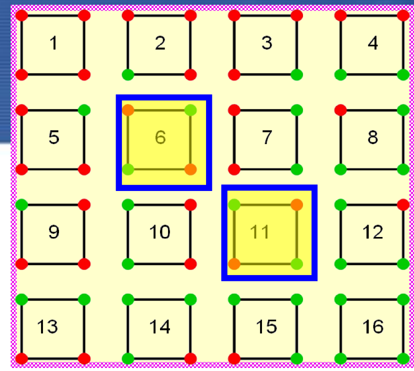
k不动置换类

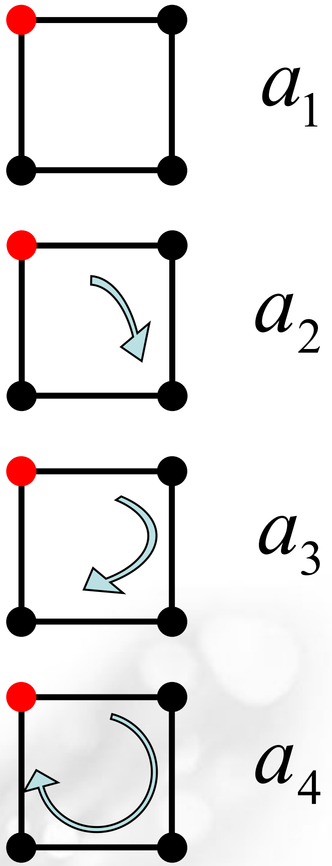


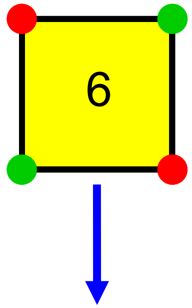


$$Z_6 = \{a_1, a_3\}$$

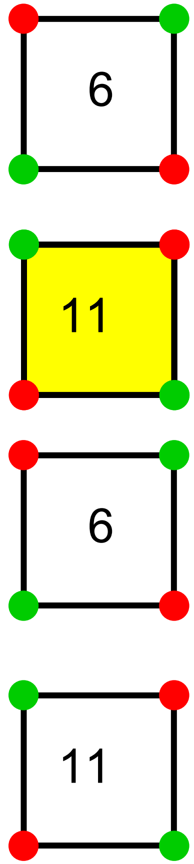
k等价类







$$E_6=\{6,11\}$$



- 公式： $|E_k| \cdot |Z_k| = |G| \quad k=1\ldots n$
  - Z_k (k不动置换类)：设G是$1\ldots n$的置换群。若k是$1\ldots n$中某个元素，G中使k保持不变的置换的全体，记以Z_k，叫做G中使k保持不动的置换类，简称k不动置换类
  - E_k(等价类)：设G是$1\ldots n$的置换群。若k是$1\ldots n$中某个元素，k在G作用下的轨迹，记作E_k

Burnside引理

由组合方案的序号组成的集合

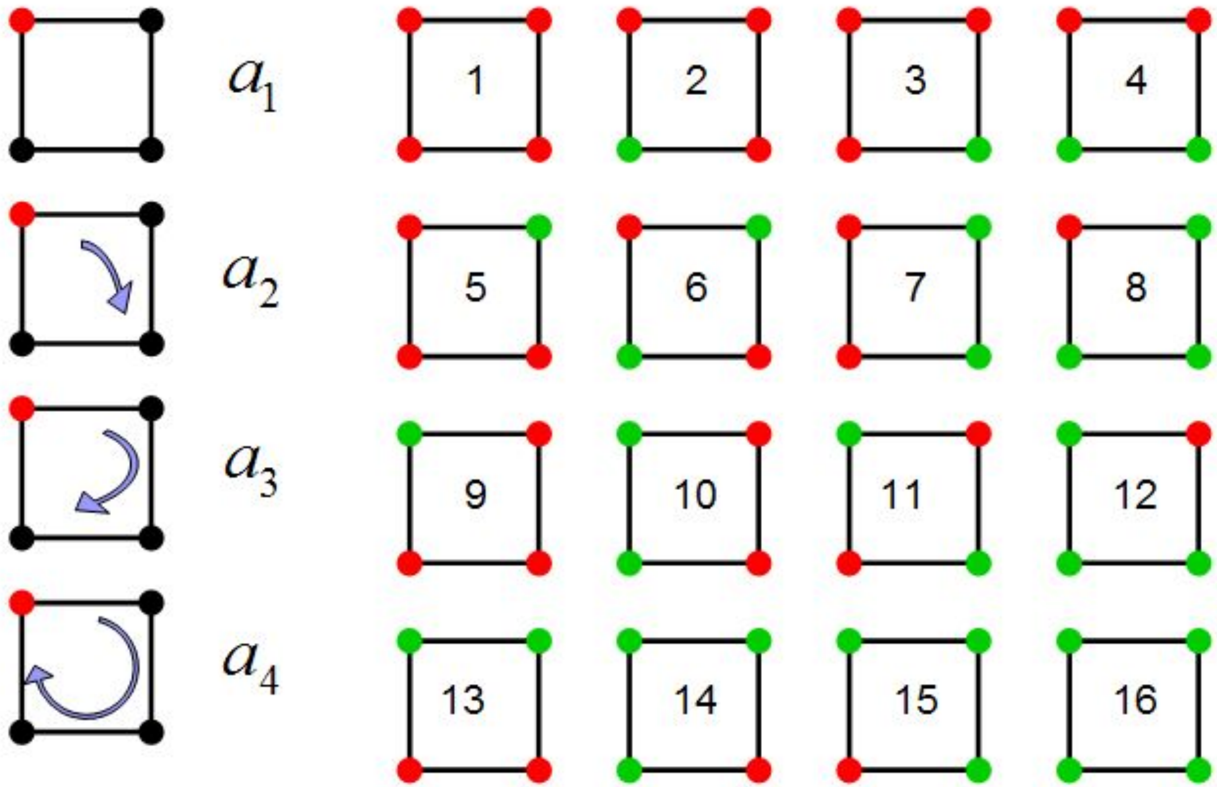
组合方案的序号，并不是真正的自然数

- 设G是$N=\{1,2,\ldots,n\}$上的置换群，G在N上可引出不同的等价类，其不同的等价类的个数为：

$$l=\frac{1}{|G|}\left[c_1(a_1)+\cdots c_1(a_i)+\cdots c_1(a_g)\right]$$

置换a_i作用后不变的方案个数，也就是置换中1阶循环的个数

- **例** 对一个正方形的4个顶点用两种颜色进行着色，问能得到多少种不同的方案？



- 解：

$$a_1=(1)(2)(3)(4)\cdots(15)(16)$$

$$a_2=(1)(16)(6\ 11)(2\ 9\ 3\ 5)(4\ 10\ 13\ 7)(8\ 12\ 14\ 15)$$

$$a_3=(1)(6)(11)(16)(2\ 5)(3\ 9)(4\ 13)(7\ 10)(8\ 14)(12\ 15)$$

$$a_4=(1)(16)(2\ 3\ 5\ 9)(4\ 7\ 13\ 10)(6\ 11)(8\ 15\ 14\ 12)$$

$$l=\frac{1}{|G|}\left[c_1(a_1)+c_1(a_2)+c_1(a_3)+c_1(a_4)\right]$$

$$=\frac{16+2+4+2}{4}=6$$

Burnside引理-小结

- 存在的问题：置换是作用在所有方案上的，如果颜色数过多，方案数随之剧增，这种情况下Burnside引理则有些无能为力！

那该怎么办？

Pólya定理

- 设$\overline{G}$是n个对象的一个置换群，用m种颜色对这n个对象进行着色，则不同的染色方案数为：

$$l=\frac{1}{\left|\overline{G}\right|}\left[m^{c(\overline{a}_1)}+\cdots m^{c(\overline{a}_i)}+\cdots m^{c(\overline{a}_g)}\right]$$

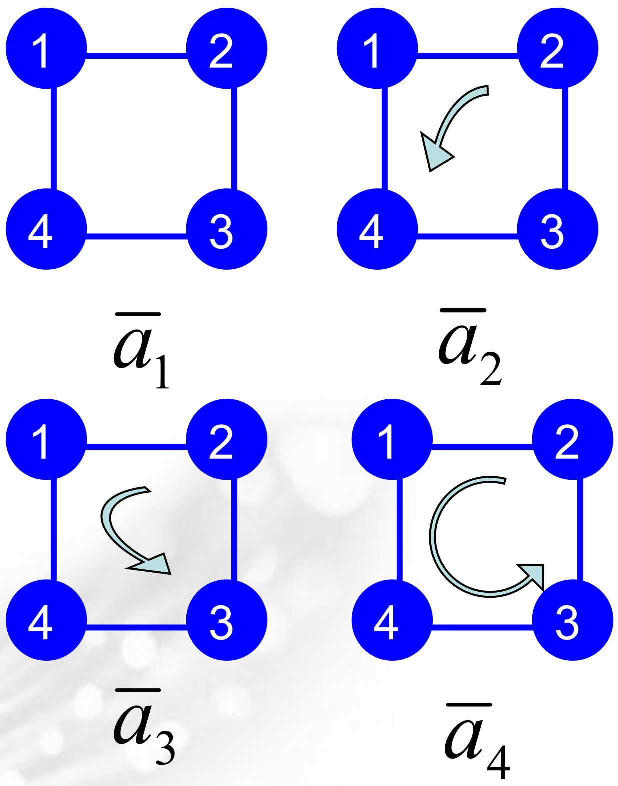
置换$\overline{a}_i$的循环节数，如：

$\overline{a}_i=(1)(2)(3\ 4)$, 则$c(\overline{a}_i)=3$

Pólya

Pólya定理-举例

- **例** 对一个正方形的4个顶点用两种颜色进行着色，问能得到多少种不同的方案？
- 解：



$$\bar{a}_1=(1)(2)(3)(4)\quad c(\bar{a}_1)=4$$

$$\bar{a}_2=(1\ 4\ 3\ 2)\quad c(\bar{a}_2)=1$$

$$\bar{a}_3=(1\ 3)(2\ 4)\quad c(\bar{a}_3)=2$$

$$\bar{a}_4=(1\ 2\ 3\ 4)\quad c(\bar{a}_4)=1$$

$$l=\frac{1}{4}\left[2^4+2^1+2^2+2^1\right]=6$$