组合数学

张昕渊

October 6, 2023

上半节大纲

- 容斥原理、莫比乌斯反演与min-max容斥

容斥原理

Theorem (容斥原理)

对于集合$A_1, A_2, \ldots, A_n$,

$$|\bigcup_{i=1}^{n} A_i| = \sum_{S \subseteq [n]} (-1)^{|S|-1} |\bigcap_{i \in S} A_i|.$$

容斥原理

Theorem (容斥原理等价形式)

对于集合$B_1, B_2, \dots, B_n \subseteq U$，

$$|\bigcap_{i=1}^{n} \overline{B_i}| = \sum_{S \subseteq [n]} (-1)^{|S|} |\bigcap_{i \in S} B_i|,$$

当S为空集时，$\bigcap_{i \in S} B_i = U$（全集）。

- 如何理解且应用容斥原理：B_i为一系列的“坏事件”。计算所有坏事件都不发生困难，但计算某些坏事件发生要来得简单的时候就可以考虑容斥原理！
- 本质是对一些难以处理的条件的“放松”。

容斥原理与莫比乌斯反演

- 莫比乌斯反演实际上可以看作是容斥原理的一个示例：
- 令$f(n) = \sum_{d|n} g(d)$。已知$f(1), f(2), \dots, f(n)$，求$g(n)$。
- 令$G_1, G_2, \dots, G_n$为不交集合满足$|G_i| = g(i)$，$F_i = \cup_{d|i} G_i$。
- 则$G_n = F_n \setminus \bigcup_{p|n} F_{n/p}$，此时全集为$F_n$，坏事件为元素落在某个$F_{n/p}$中。
- 假设n有$p_1, p_2, \dots, p_m$这m个素因子，由容斥原理：$|G_n| = \sum_{S\subseteq[m]} (-1)^{|S|} |F_{n/p_S}| = \sum_{d|n} \mu(d) f(n/d)$。

错排

- 有多少个排列$p = [p_1, p_2, \ldots, p_n]$满足对于所有的$i$都有$p_i \neq i$。
- 令B_i为$p_i = i$的坏事件，则k个坏事件的交为k个位置确定之后的排列方案数，为$(n-k)!$种。因此，

$$
\begin{aligned}
|\bigcap_{i=1}^{n} \overline{B_i}| &= \sum_{S \subseteq [n]} (-1)^{|S|} |\bigcap_{i \in S} B_i| \\
&= \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)! \\
&= n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.
\end{aligned}
$$

不定方程非负整数解计数

- 求$x_1, x_2, \ldots, x_n$满足$0 \le x_i < C$且$\sum_{i=1}^n x_i = m$的整数解个数。
- 考虑容斥原理：$x_i < C$这个条件难以处理，我们令B_i为这个条件被违反的坏事件。
- $\bigcap_{S\in[n]} B_i$这个集合指的是所有$i \in S$有$x_i \ge C$，且$\sum_{i=1}^n x_i = m$的非负整数解个数。
- 这等价于求解$\sum_{i=1}^n x_i = m - kC$的非负整数解个数。
  - 插板法解决：考虑$m - kC$个球，$n - 1$个板子将这一些球分成n块，每一块内球的个数对应着x_i的取值，因此解的个数为$\binom{m-kC+n-1}{n-1}$。
  - $\sum_{i=1}^n x_i \le m$呢？只需引入一个冗余变量即可。

简单复形体积问题

- 求下列n维几何体的体积：

$$S = \{0 < x_i < 1 | \sum_{i=1}^{n} x_i \leq m\}.$$

- $x_i < 1$不好处理，我们考虑坏事件B_i为$x_i > 1$，容斥后只需求$S = \{0 < x_i | \sum_{i=1}^{n} x_i \leq m - C\}$的体积，为$(m - C)/n!$。

带标号的DAG计数

- 求n个带标号顶点的DAG个数。
- 令f_n为答案，A_i为i号顶点入度为0的事件，则$\bigcap_{i\in S} A_i$为$|S|$个顶点入度均为0，满足该种条件的DAG方案数为$2^{|S|(n-|S|)} f_{n-|S|}$。因此

$$f_n = |\bigcup_{i\in S} A_i| = \sum_{S\subseteq[n]} (-1)^{|S|-1} |\bigcup_{i\in S} A_i|$$
$$= \sum_{k=1}^{n} \binom{n}{k} (-1)^{k-1} 2^{k(n-k)} f_{n-k}$$

例题：逆序对个数

- 求长度为n且逆序对个数为K的排列个数，模10^9+7。
- $n, K \leq 10^5$。

例题：逆序对个数

- 题目等价于求$x_1, x_2, \ldots, x_{n-1}, x_n$满足下述条件的方案数。
  - $0 \le x_i < i$;
  - $\sum_{i=1}^{n} x_i = K$
- 考虑容斥原理，令B_i为$x_i < i$被违反的坏事件，则对于任意的$S \subseteq [n]$，

$$|\bigcap_{i \in S} B_i| = \binom{K - \sum_{i \in S} i + |S| - 1}{|S| - 1},$$

  因此，我们只需统计$f_{i,j}$使得$|S| = i$且$\sum_{k \in S} k = j$的方案数即可。
- $f_{i,j} = f_{i,j-i} + f_{i-1,j-i} - f_{i-1,j-(n+1)}$，且第一维为$O(\sqrt{K})$量级。
- 整体复杂度为$O(K\sqrt{K})$。

例题：LEQ and NEQ (easy version)

- 求满足下列条件的序列$A_1, A_2, \ldots, A_n$个数，模10^9+7。
  - $A_i \leq X_i$;
  - $A_i \neq A_{i+1}$.
- $n \leq 5000, X_i \leq 10^9$。

例题：LEQ and NEQ(easy version)

- 非常自然的考虑对条件$A_i \neq A_{i+1}$进行容斥：令B_i为事件$A_i = A_{i+1}$，则$|\bigcap_{i \in S} B_i|$为一系列区间最小值的乘积。
- 令$dp_{i,k}$为A_i和A_{i+1}不在同一个区间，坏事件交的个数奇偶性为k的总代价。
- 考虑$i+1$到$i+\ell$被合并至一起，则$dp_{i,k}$可以转移到$dp_{i+\ell, k+(\ell-1) \bmod 2}$。
- 答案为$dp_{n,0} - dp_{n,1}$。
- 本题可以加速至线性时间。

例题：Perfect matching

- 给一个$2n$个点的树T，求T的补图中完美匹配的个数，答案对998244353取模。
- $n \leq 2000$。

例题：Perfect matching

- 由容斥原理（令B_i为第i条边在树上的坏事件），问题转化为求树上大小为K的匹配个数，这可以在$O(n^2)$的时间内通过经典的树dp求出。

例题：ABC string

- 求有多少个由A, B, C构成的序列s，使得A, B, C的出现次数分别为a, b, c，且序列中不出现连续子串ABC，BCA或者CAB。
- $a, b, c \leq 10^6$。

例题：ABC string

- 考虑容斥原理，问题转化成如下：
  - 选择k个位置$x_1, x_2, \ldots, x_k$使得这$s[x_i : x_i + 2]$恰好为ABC, BCA或者CAB。此时方案数乘以$(-1)^k$贡献给总答案。求总答案为多少
- 对于每个位置x_i，我们将x_i, x_{i+1}, x_{i+2}相连。则最终$[n]$可以划分为$I_1, I_2, \ldots, I_\ell$这ℓ个区间的并。
- 固定某种特定的区间划分，我们首先计算所有可以导出这种区间划分的位置选择加权和，权重为$(-1)^k$。
- 每个区间单独处理，令L_i为第i个区间长度，则第L_i个区间权重和为1如果$L_i \mod 3 = 1$，权重和为0如果$L_i \mod 3 = 2$，否则为-1。
- 问题转化成如下：
  - $[n]$划分成若干段，每一段都是$ABC/BCA/CAB$的若干次重复加上$A/B/C$或者不加上任何元素。每一种划分对答案的贡献是此时填ABC满足条件的方案数乘以$(-1)^w$，其中w是不加$A/B/C$的区间数。
- 将ABC打包看成D，固定D的个数，对于每一种$ABCD$的序列，划分总贡献为$(-2)^D$ 或者$-3 * (-2)^{D-1}$，取决于最后一个位置是否为D。

min-max容斥

- 给定集合$U=\{a_1,a_2,\ldots,a_n\}$，

$$\min(U)=\sum_{\emptyset\neq S\subseteq U}(-1)^{|S|-1}\max(S),$$

  $\min,\max$交换也成立。
- 不妨假设$a_1\leq a_2\leq\ldots\leq a_n$
- 当$S=\{a_1\}$时，$\max(S)=a_1$。
- 当$S\neq\{1\},\emptyset$，$\max(S)=\max(S\oplus\{1\})$，其他会相互抵消。
- 推论：当$a_1,a_2,\ldots,a_n$为随机变量时，

$$E[\min(U)]=\sum_{\emptyset\neq S\subseteq U}(-1)^{|S|-1}E[\max(S)].$$

min-max容斥

- 给定集合$U = \{a_1, a_2, \ldots, a_n\}$,

$$\mathrm{kth} - \min(U) = \sum_{\emptyset \neq S \subseteq U} (-1)^{|S|-k} \binom{|S|-1}{k-1} \max(S),$$

- 同样的有期望版本。

例题：按位或

- 给定正整数n以及变量$x = 0$，每一秒从$[0, 2^n - 1]$中以概率分布p独立选取正整数a，将x更新至$x|a$。当$x = 2^n - 1$时过程停止，求过程终止的期望时间。
- $n \leq 20$。

例题：按位或

- 每一位单独考虑，令X_i为第i位被置为1所需时间，则答案为$E[\max(X_0, X_1, \ldots, X_{n-1})]$。
- 由min-max容斥：

$$E[\max(X_0, X_1, \ldots, X_{n-1})] = \sum_{\emptyset \neq S \subseteq [n]} (-1)^{|S|-1} E[\min(X_S)],$$

- $E[\min(X_S)]$：S中某一位被置为1的期望时间。令q_S为抽取数字a包含S中某一位的概率，则$E[\min(X_S)] = 1/q_S$。
- 对于每一个S，计算q_S的问题等价于计算高维前缀和，可以在$O(n2^n)$的时间内求解。

一维带碰撞壁随机游走模型

Problem

求满足以下条件的长度为n的序列$a = [a_1, a_2, \ldots, a_n]$个数：

- $a_1 = 0,\ a_n = A$;
- $|a_{i+1} - a_i| = 1$;
- $a_i > -C$。

- 对于这种问题，我们一般考虑“André反射原理”。
- 我们将不合法的折线序列与从$-2C$出发的折线序列一一对应起来：即我们考虑所有满足下列条件的序列b
  - $b_1 = 2C,\ b_n = A$;
  - $|b_{i+1} - b_i| = 1$。
- 非法的a与b的一一对应关系：从开始到第一个达到$-C$的点的这段区间上下翻转。

Figure: Andre's reflection principle

一维带碰撞壁随机游走模型

Problem

求满足以下条件的长度为n的序列$a = [a_1, a_2, \dots, a_n]$个数：

- $a_1 = 0$, $a_n = A$;
- $|a_{i+1} - a_i| = 1$;
- $a_i > -C$。

- 总方案数：0到A的长度为n的折线方案数：$\binom{n}{\frac{n+A}{2}}$;
- 坏的方案数：$-2C$到A的长度为n的折线方案数：$\binom{n}{\frac{n+A+2C}{2}}$。

一维带碰撞壁随机游走模型

Problem

求满足以下条件的长度为n的序列$a = [a_1, a_2, \ldots, a_n]$个数：

- $a_1 = 0$, $a_n = A$;
- $|a_{i+1} - a_i| = 1$;
- $-C < a_i < C$。

- 如果有两侧的限制将会如何：不满足条件的折线可能会来回穿过$y = -C$与$y = C$。
- 令A_i为折线存在碰壁$y = C$, $y = -C$, $y = C$, $y = -C$...的长度为i的子序列；B_i为折线存在碰壁$y = -C$，$y = C$，$y = -C$，...的长度为i的子序列。
- 合法方案数为总方案数减去$\sum_{i=1}^{+\infty}(-1)^{i-1}(|A_i| + |B_i|)$（容斥原理中的算两次技巧）。
- $|A_i|$方案应该如何计算：类似于反射原理，我们可以将原点按$y = C$,$y = -C$, $y = C$,...依次翻转n次，则A_i中的折线与起始点$2kC$，终点为A的折线方案数一一对应。