Counting

何俞均

常用组合恒等式

1. 对称恒等式：$\binom{n}{m} = \binom{n}{n-m}$
2. 吸收恒等式：$\binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1}$
   $m \times \binom{n}{m} = n \times \binom{n-1}{m-1}$
   $(n-m) \times \binom{n}{m} = n \times \binom{n-1}{m}$
3. 归纳恒等式：$\binom{n}{m} = \binom{n-1}{m} + \binom{n-1}{m-1}$

常用组合恒等式

1. $\binom{n}{r}\binom{r}{k} = \binom{n}{k}\binom{n-k}{r-k}$
2. $\sum\limits_{i=m}^{n}\binom{i}{m} = \binom{n+1}{m+1}$
3. 范德蒙德卷积：$\sum\limits_{i=0}^{k}\binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$

CF1264D2

给定一个长度为 n 的字符串，只有 (,),? 三种字符。
定义一个括号序列的权值为，任意删除字符之后剩下的形如 ((...()...)) 的括号序列的最大匹配数，例如 ()() 权值为 1,()(()) 权值为 2。
问把所有'?' 替换成左括号或者右括号的所有可行方案的权值和是多少。
$n \leq 10^6$

CF1264D2 Solution

1. 首先需要寻找一个方便计数的统计方式，可以发现每个括号序列，它的存在且仅存在一个分界点，分界线左边的'(' 与分界线右边的')' 数目相等，而在这里取到匹配数的最大值。
2. 然后考虑枚举分界点，解决计数的问题：假设目前分界点左边有 a 个'('，c_1 个'?'; 右边有 b 个')'，c_2 个'?'。
3. 枚举匹配数 i，可以得到答案的形式为 $\sum_{i=a}^{a+c_1} i\binom{c_1}{i-a}\binom{c_2}{i-b}$。
4. 将上式优化到 $O(1)$: $a\binom{c_1+c_2}{c_1+a-b} + c_1\binom{c_1+c_2-1}{c_1+a-b}$。

[洛谷 P7481] 梦现时刻

对所有 $1 \le a \le m, 1 \le b \le m$，求 $F(a,b) = \sum_{i=0}^{b} \binom{b}{i} \binom{n-i}{a}$。
$1 \le m \le 5000, 1 \le n \le 10^9$

[洛谷 P7481] 梦现时刻 Solution

1. 很难找到一个 $O(1)$ 或者卷积形式的 $F(a,b)$，考虑递推。
2. 首先将 $\binom{b}{i}$ 拆开，得到
   $F(a,b)=F(a,b-1)+\sum\binom{b-1}{i}\binom{n-i-1}{a}$。
3. 再考虑后面和式的变化：
   $\sum\binom{b-1}{i}\binom{n-i-1}{a}=F(a,b-1)-\sum\binom{b-1}{i}\binom{n-i-1}{a-1}$
4. 最后再整理 $\sum\binom{b-1}{i}\binom{n-i-1}{a-1}$，可以发现它等于
   $F(a-1,b)-F(a-1,b-1)$。
5. 得到递推式
   $F(a,b)=2F(a,b-1)+F(a-1,b-1)-F(a-1,b)$。

Lucas 定理

$$\binom{n}{m} \equiv \binom{\lfloor n/p \rfloor}{\lfloor m/p \rfloor}\binom{n \bmod p}{m \bmod p} \pmod p$$

证明?
推论:

$$\binom{n}{m} \bmod p \neq 0 \Leftrightarrow$$

n 在 p 进制下每一位都不小于 m 在 p 进制下的对应位

[CTSC2017] 吉夫特

给定一个长度为 n 的序列 $a_1, a_2 ..., a_n$，问有多少个长度 > 1 的不上升子序列满足 $\prod_{i=2}^{k} \binom{a_{b_{i-1}}}{a_{b_i}} \bmod 2 = 1$，其中 b 是这个不上升子序列的下标, $1 \leq b_1 < b_2 < ... < b_k \leq n$。
$n \leq 211985, a_i \leq 233333,\ a_i$ 互不相同

[CTSC2017] 吉夫特 Solution

1. 考虑 $\prod_{i=2}^{k}\binom{a_{b_{i-1}}}{a_{b_i}} \bmod 2 = 1$ 的条件，即子序列前一项是后一项的超集。
2. 利用上面的性质进行 dp。
3. 因为保证了 a_i 互不相同，故 dp 的复杂度是 $O(3^m)$ 的，m 是 a_i 的位数。

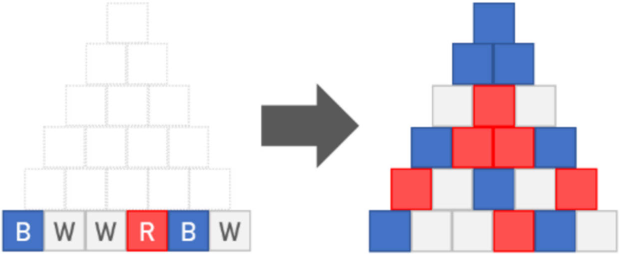
ARC117C Tricolor Pyramid

给定一个底部大小为 n 的二维金字塔，有三种颜色 `R,B,W`，你要将金字塔全部涂满颜色，规则如下：

1. 如果这个格子下面的两个格子颜色相同，那么这个格子的颜色就和下面的相同。
2. 如果这个格子下面的两个格子颜色不同，那么这个格子的颜色和下面两个都不同。

问塔顶的颜色。$n \le 400000$

ARC117C Tricolor Pyramid



ARC117C Tricolor Pyramid Solution

1. 考虑用一种其他的方式来描述颜色的变化：假设下面两块的颜色是 (x, y)，那么新的颜色为 $-(x+y) \bmod 3$。
2. 考虑每个底层方块对顶层的贡献。
3. 答案为 $\sum_{i=1}^{n}(-1)^{n-1}c_i\binom{n-1}{i-1}$，需要用 Lucas 定理进行计算。

组合意义

利用组合意义对问题进行分析有时可以获得比纯代数方法更简洁的式子以及性质。
常用组合意义：

- 格路计数
- 盒子与球
- 概率和偏序树
- ……

格路计数

基本思想是组合数 $\binom{n+m}{n}$ 可以看成从 $(0,0)$ 走到 (n,m)，每次只能向右或者向上的方案数。

- $\sum\limits_{i=m}^{n}\binom{i}{m}=\binom{n+1}{m+1}$
- 求从 $(0,0)$ 到 (n,m)，不碰到直线 $y=x+b$ 的方案数：$\binom{n+m}{m}-\binom{n+m}{m-b}$
- 格路计数在合法括号序列的计数中十分常用。

[JLOI2015] 骗我呢

有一个 $n \times m$ 的数组 $x_{i,j}(1 \le i \le n, 1 \le j \le m)$。
求满足

1. $0 \le x_{i,j} \le m$
2. $1 \le i \le n, 1 \le j < m, x_{i,j} < x_{i,j+1}$
3. $1 < i \le n, 1 \le j < m, x_{i,j} < x_{i-1,j+1}$

的 x 的个数。
$1 \le n, m \le 10^6$。

[JLOI2015] 骗我呢 Solution

1. 首先注意到 x 中每个元素只有 $m+1$ 种取值，只恰好比列数多 1。
2. 考虑利用上面的性质进行 dp，设 $f_{i,j}$ 表示目前在第 i 行且这一行缺少元素 j 的方案数，那么有转移 $f_{i,j}=\sum_{k=0}^{j+1} f_{i-1,k}=\sum_{k=0}^{j} f_{i-1,k}+f_{i-1,j+1}=f_{i,j-1}+f_{i-1,j+1}$。答案就是 $f_{n+1,m}$。
3. 把转移画在坐标轴上并将坐标拉直。可以发现答案为 $(0,0)$ 走到 $(n+m+1,n)$，且不能碰到 $l_1: y=x+1, l_2: y=x-m-2$ 的路径数。
4. 考虑如何解决上面的两条线问题：将每条路上依次碰到的直线写下来，并将相同的连续段所称一个，那么我们会得到类似于 1,12,121,2,21,212... 的许多路径。
5. 用总方案减去以 1,2 开头的方案数，会发现减去了多余的方案，需要通过对称并反复容斥解决。

[AGC018E]Sightseeing Plan

给定 $x_1, x_2 \dots x_6, y_1, y_2 \dots y_6$，求：

$$\sum_{x_a,x_b,x_c,y_a,y_b,y_c} \binom{x_b - x_a + y_b - y_a}{x_b - x_a} \binom{x_c - x_b + y_c - y_b}{x_c - x_b}$$

其中 $x_a \in [x_1, x_2], y_a \in [y_1, y_2], x_b \in [x_3, x_4], y_b \in [y_3, y_4], x_c \in [x_5, x_6], y_c \in [y_5, y_6]$。
$x_1 \le x_2 \le \dots \le x_6 \le 10^6, y_1 \le y_2 \le \dots \le y_6 \le 10^6$

[AGC018E]Sightseeing Plan Solution

1. 三个矩形的形式比较复杂，考虑将问题化简，按照这样的顺序考虑问题：点到线段 → 点到矩形 → 矩形到矩形 → 三个矩形
2. 发现点到矩形实际上就是若干个点到线段，拼起来之后转化为一个点到四个点。
3. 矩形到矩形就是若干个一个点到四个点，就是四个点到一个矩形，就是四个点到四个点。
4. 三个矩形的话可以 4×4 地枚举代表第一个和第三个矩形的点，然后再枚举进入第二个矩形边界的点，但是这样有个问题，因为每种相同的路径可能会多次贡献。
5. 考虑进入第二个矩形的路线长度 len，它会被贡献 len 次，但是我们可以把 len 拆成 $x_2 + y_2 - x_1 - y_1$ 的形式来解决这个问题。

Raney 引理

假设当前有一个长度为 n 的整数序列 $x_1, x_2, \cdots x_n$ 满足 $\sum_{i=1}^{n} x_i = 1$ 那么在这个序列的 n 个循环移位中恰好有一个满足 $\forall i, \sum_{j=1}^{i} x_j > 0$。

可以由此给出卡特兰数 $C_n = \frac{\binom{2n}{n}}{n+1}$ 的组合意义证明。

[清华集训 2016] 你的生命已如风中残烛

有 n 个特殊牌和 $m-n$ 张普通牌，有一个长度为 $m+1$ 的序列，最后一个元素固定为 0，其他元素编号为 $1\sim m$，每张特殊牌有一个特殊的权值 w_i 表示获得了其之后可以摸 w_i 张牌，保证 $\sum w=m$，初始您获得第 1 张牌，求所有 $m!$ 种方案中使得您摸完牌的方案数。

$m-n\geq 4, n\leq 40, w_i\leq 10^5$

[清华集训 2016] 你的生命已如风中残烛 Solution

1. 考虑更加具体化计数的条件：普通牌看作 -1，特殊牌 w_i 减一，那么最后就是要求这 m 个元素重排且任意前缀和非负的排列数。
2. 类似求卡特兰数的方法，我们在序列最后丢一个-1，那就是求前 $m!$ 种方案中前 m 个前缀和大于等于 0 的方案数。
3. 可以发现上面的问题可以转化为每个后缀和都小于 0，然后总和为 -1，和 Raney 引理的形式本质上是一样的。
4. 那么这 $m+1$ 个元素的排列方案数为 $m!$，因为任意一个-1 都可能是我们添加的，所以还要除掉 $m+1-n$。

盒子与球

另一类比较重要的组合意义转化方式。
应用：

1. 插板
2. 斯特林数
3. $k^m = \sum_{i=0}^{k} \left\{ {m \atop i} \right\} i! \binom{k}{i}$

概率

注意到一个很简单的事情：有一个排列，这个排列某一个给定位置为 1 的概率是 $\frac{1}{n}$; 这个排列中某 k 个位置按照一定的大小关系排布的概率为 $\frac{1}{k!}$。
类似这样的东西在把期望拆成概率的过程中经常用到。

偏序树

这个名字可能在网上找不到，因为是某个学长编的。
其问题如下：给定一个 n 个点的外向树，每个点上有权值，这些权值构成一个排列，问符合外向树偏序关系的赋权方法有多少种。
答案：$n!\prod\frac{1}{sz_i}$

ARC114E

1. 有一块 $h \times w$ 的巧克力，那么横向有 $h-1$ 条切割线，纵向有 $w-1$ 条。
2. 现在这块巧克力上有两个特殊坐标 (h_1, w_1),(h_2, w_2)，现在进行如下操作。
3. 从目前巧克力中的横纵切割线中随机等概率选择一条，并把巧克力沿着这个切割点一分为二。
4. 如果特殊点仍在一块巧克力中，用它们处在的巧克力继续操作，否则结束操作。
5. 求切割操作次数的期望。$1 \leq h, w \leq 10^5$

ARC114E Solution

1. 考虑将每个切割线被切到的概率算出来，最后加和。
2. 怎么算这个概率？发现组合方法推式子会十分复杂，考虑能不能直接算概率。
3. 不管切割方法是否合法，切割的排列有 $(h-1+w-1)!$ 种。也就是说我们需要算，在这些排列中，我们枚举的切割是合法的的概率。
4. 发现要满足的条件就是枚举的切割要在所有能 ban 掉它的切割之前，把这些线的个数 x 算出来，贡献即为 $\frac{1}{x+1}$。

CCPC Final 2022. M

有 n 个数 $a_1, a_2 \ldots a_n$ 和 $n-1$ 个加减操作符 $op_1, op_2, \ldots, op_{n-1}$，它们被排列成 $a_1\ op_1 a_2\ op_2 \ldots op_{n-1}\ a_n$ 的形式。
一次擦除操作是选择一个运算符，然后把两边的数按照这个运算符的运算变成一个数，可以发现进行 $n-1$ 次操作就可以把所有数变成一个。
问所有进行 $n-1$ 次擦除操作的方案的结果之和是多少。
$2 \le n \le 2 \times 10^5, 0 \le a_i \le 10^9$

CCPC Final 2022. M Solution

1. 考虑每个数的贡献，也就是它最后被算到的时候是正的方案数减去负的的方案数乘上它自己的大小。
2. 考虑哪些符号会对第 i 个数起作用：把下标 $[1...i-1]$ 的运算符拿出来，优先级是后缀最小值的运算符，这个概率是 $\frac{1}{i-j}$，此时贡献为 sgn_j，否则为 1。最后这个数的贡献也就是 $(\frac{sgn_j}{i-j}+\frac{i-j-1}{i-j})$。
3. 那么答案的式子就是 $a_1+\sum_{i=2}^{n}\frac{a_i sgn_{i-1}}{(i-1)!}\prod_{j=1}^{i-1}(sgn_j+i-j-1)$。考虑如何快速求。
4. 考虑采用离散对数将后面的连乘变为加法，答案最后可以用卷积优化，变成 $\sum coe_i \ln i$ 的形式，复杂度 $O(n \times \text{O(离散对数)} + n\log n)$

CCPC Final 2022. M Solution

考虑使用如下方法计算 x 的离散对数：
做带余除法，$p = qx + r$，同时 $p = (q+1)x + (r-x)$。
那么 $\log x \equiv \log(-r) - \log q \equiv \log(x-r) - \log(q+1) \pmod{\varphi(p)}$
线性筛 +bsgs 预处理出 $\leq \sqrt{p}$ 内所有数的离散对数，那么只需要考虑 $x > \sqrt{p}$。
此时 $q < \sqrt{p}$，每次选用 $\min\{r, x-r\}$ 可以将问题规模缩小一半。
这样的话可以做到 $O(\sqrt{\pi(\sqrt{p})p})$ 预处理 $+O(\log n)$ 询问。

基本容斥

最简单的容斥形式:

$$ans = |U \bigcap \overline{A}_1 \bigcap \overline{A}_2 \cdots \bigcap \overline{A}_n| = \sum_{S \subseteq A} (-1)^{|S|} |U \bigcap S_1 \bigcap S_2 \cdots \bigcap S_{|S|}|$$

如果只和违反限制的个数 i 有关，就是 $\sum_{i=0}^{n} (-1)^i \binom{n}{i} f(i)$。
简单的例子:

1. 错排
2. $\left\{ \begin{matrix} m \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^m$

CCPC2021 Weihai. M

求有多少长度为 n，恰有 m 个 1 的 01 串，满足最长的 1 连续段恰好为 k。
$0 \le n, m, k \le 10^5$

CCPC2021 Weihai. M Solution

1. 考虑求 $\le k$ 的方案数，最后答案就是 $\le k$ 的减去 $\le k-1$ 的结果。
2. 把每个 0 看作隔板，那么就相当于求有 m 个 1 组成 $n-m+1$ 个 1 连续段，连续段长度可以为 0，每段长度 $\le k$ 的方案数。
3. 容斥一下可以得到答案为 $\sum_{i=0}^{n-m+1}(-1)^i\binom{n-m+1}{i}\binom{m-i(k+1)+n-m}{n-m}$。

Min-Max 容斥

$$ans = \max_{i=1}^{n} a_i = \sum_{T \subseteq A} (-1)^{|T|+1} \min T$$

证明是基于两个非常简单的式子的:

$$\min\{\max\{a, b\}, c\} = \max\{\min\{a, c\}, \min\{b, c\}\}$$

$$\max\{a, b\} = a + b - \min\{a, b\}$$

51nod 1355 斐波那契的最小公倍数

给定 $a_1, a_2 ... a_n$，求 $\mathrm{lcm}(fib_{a_1}, fib_{a_2} ..., fib_{a_n})$，对 $10^9 + 7$ 取模。
$n \le 5 \times 10^4, a_i \le 10^6$

51nod 1355 斐波那契的最小公倍数 Solution

1. 对每个质因数，lcm 相当于取 max，因此可以考虑 Min-Max 容斥，变成 gcd 求答案。
2. 由斐波那契数列的性质, 可以枚举 gcd, 答案变为 $\prod_{d\geq 1} fib(d)^{\sum_{T\subset S}[\gcd(T)=d](-1)^{|T|+1}}$
3. 令 $a_d = \sum_{T\subset S}[\gcd(T) = d](-1)^{|T|+1}$, $b_d = \sum_{T\subset S}[d|\gcd(T)](-1)^{|T|+1}$，莫比乌斯反演可得 $a_i = \sum_{i|d}\mu(\frac{d}{i})b_d$。
4. 那么最后的答案就是 $\prod_{d\geq 1} f(d)^{\sum_{d|i}\mu(\frac{i}{d})b_i}$。

二项式反演

$$g_n = \sum_{i=0}^{n} \binom{n}{i} f_i \Leftrightarrow f_n = \sum_{i=0}^{n} (-1)^{n-i} \binom{n}{i} g_i$$

f_i 为恰好 i 个匹配上的方案数，g_i 为至多 i 个匹配上的方案数。

$$g_k = \sum_{i=k}^{n} \binom{i}{k} f_i \Leftrightarrow f_k = \sum_{i=k}^{n} (-1)^{i-k} \binom{i}{k} g_i$$

f_i 为恰好 i 个匹配上的方案数，g_i 为至少 i 个匹配上的方案数。

[CTS2019] 随机立方体

有一个 $n \times m \times l$ 的立方体，立方体中每个格子上都有一个数，如果某个格子上的数比三维坐标至少有一维相同的其他格子上的数都要大的话，我们就称它是极大的。
现在将 $1 \sim n \times m \times l$ 这 $n \times m \times l$ 个数等概率随机填入 $n \times m \times l$ 个格子（即任意数字出现在任意格子上的概率均相等），使得每个数恰出现一次，求恰有 k 个极大的数的概率。
$1 \leq n, m, l \leq 5 \times 10^6, 1 \leq k \leq 100, 1 \leq T \leq 10$

[CTS2019] 随机立方体 Solution

1. 恰好的条件不好处理，考虑转化为至少然后反演回来，现在的问题是钦定有 k 个极大值之后要怎么做。
2. 把所有格子分为被最大值影响到的和没有影响到的，那么主要问题是被影响到的部分不好计数。
3. 假设放置 i 个数被影响到的位置为 v_i，注意我们钦定的的 i 个值是有大小关系的，用概率算出的方案数为 $(v_i)!\prod_{k=1}^{i}\frac{1}{v_k}$。
4. 需要注意的是所有 $\frac{1}{v_i}$ 需要线性时间内求出。

单位根反演

$$\forall k, [k|n] = \frac{1}{k}\sum_{i=0}^{k-1}\omega_k^{in}$$

常常用来求特定倍数的系数之和。
例如我们要求 $\sum_{i=0}^{[\frac{n}{k}]}[x^{ik}]f(x)$，可以用上面的公式把整除替换成 $\frac{1}{k}\sum_{j=0}^{k-1} f(\omega_k^j)$

LOJ6247 九个太阳

给定 n, K，满足 K 是二的幂，求 $\sum_{0\le i\le n, K|i} \binom{n}{i}$。

$n \le 10^{15}, K \le 2^{20}$

LOJ6247 九个太阳 Solution

1. 这题就相当于求 $(1+x)^n$ 被 K 整除次项的系数，把上一页的公式带进去就可以得到 $\frac{1}{K}\sum_{j=0}^{K-1}(1+\omega_K^j)^n$。
2. 因为 K 是二的幂，所以 ω 在 mod 998244353 意义下可以取 $3^{\frac{mod-1}{K}}$。

LOJ6358 前夕

求 $\{0,1,...,2^n-1\}$ 的所有数按位与二进制位数为 4 的倍数的子集数量。定义空集按位与为 0。

$n \leq 10^7$

LOJ6358 前夕 Solution

1. 发现恰好为 k 的数目并不好算，考虑钦定 k 个位置然后反演回来。
2. 那么有 $f(k)=\binom{n}{k}(2^{2^{n-k}}-1)=\sum_{i=k}^{n}\binom{i}{k}g(i), g(k)=\sum_{i=k}^{n}(-1)^{i-k}\binom{i}{k}f(i)$
3. 把单位根带进上式可求得答案为 $\sum_{i=0}^{n}f(i)\frac{1}{4}\sum_{j=0}^{3}(\omega_4^j-1)^i$。

莫比乌斯反演

约数形式：

$$g(n)=\sum_{d|n}f(d)\Leftrightarrow f(n)\quad=\sum_{d|n}\mu(\frac{n}{d})g(d)$$

倍数形式：

$$g(n)=\sum_{n|d}f(d)\Leftrightarrow f(n)\quad=\sum_{n|d}\mu(\frac{d}{n})g(d)$$

在计数里面常常用来容斥与约数有关的计数问题，例如我要求某条件恰好某个数的个数，如果这个数的倍数比较好求的话可以先求倍数再反演回来。

常用的生成函数形式

1. $[x^m](1+x)^n = \binom{n}{m}$
2. $\frac{1}{1-x} = 1 + x + x^2 + \dots$
3. $\frac{1}{(1-x)^n} = \sum_{i\geq 0} \binom{n+i-1}{n-1} x^i$

常用的生成函数形式

1. $e^x = \sum_{i \geq 0} \frac{x^i}{i!}$
2. $\frac{e^x + e^{-x}}{2} = \sum_{i \bmod 2 = 0} \frac{x^i}{i!}$
3. $\frac{e^x - e^{-x}}{2} = \sum_{i \bmod 2 \neq 0} \frac{x^i}{i!}$

$\frac{1}{1-F}$,exp,ln 的意义

1. $\frac{1}{1-F} = 1 + F + F^2 + \ldots$ 表示大小为 i 的物品系数为 $[x^i]F$ 的完全背包。
2. $\exp F = \sum_i \frac{F^i}{i!}$，如果把 $F(x)$ 看成单个元素的 EGF，那么 $\exp$ 就描述了这些元素构成的有标号的集合。
例如 n 个点的带标号无向连通图生成函数是 $F(x)$，那么 $\exp F(x)$ 就是 n 个点带标号的无向图的 EGF。
3. $\ln F$ 可以理解为 $\exp$ 的逆运算，将有标号集合容斥回单点。

CCPC2022 Guilin. D

抽卡，每次有 p 的概率抽中，需要中 n 次，假设一共抽了 x 次。
求 $E(x^0), E(x^1), E(x^2) ... E(x^m)$。
$1 \le n \le 10^5, 0 \le m \le 10^5$

CCPC2022 Guilin. D Solution

由 $E((x+y)^k)=\sum_i \binom{k}{i}E(x^i)E(y^{k-i})$，可以发现，如果我们可以得出 $n=1$ 的 $0,1,\dots,k$ 次方答案，那么把这个期望的 EGF 的 n 次方即为最后的答案。

$n=1$ 时，设 f_k 为期望的 k 次方，有（$q=1-p$）

$$f_k=\sum_{i\geq 1} i^k q^{i-1}$$

$$qf_k=\sum_{i\geq 1} i^k q^i$$

CCPC2022 Guilin. D Solution

$$
\begin{aligned}
(1-q)f_k =& 1 + q\sum_{i\geq 1}((i+1)^k - i^k)\,q^{i-1} \\
=& 1 + q\sum_{i\geq 1}(\sum_{j=0}^{k-1}\binom{k}{j}i^j)\,q^{i-1} \\
=& 1 + q\sum_{j=0}^{k-1}\binom{k}{j}\sum_{i\geq 0} i^j q^{i-1} \\
f_k =& 1 + q\sum_{j=0}^{k}\binom{k}{j}f_j
\end{aligned}
$$

令 $F(x)=\sum_{i=0} f_i x^i$，有 $F(x)=1+qe^x F(x)$，得出 $F(x)=\frac{1}{1-qe^x}$，最后求 $F^n(x)$ 即可。

[CTS2019] 珍珠

有 n 个在范围 $[1, D]$ 内的整数均匀随机变量。求至少能选出 m 个瓶子，使得存在一种方案，选择一些变量，并把选出来的每一个变量放到一个瓶子中，满足每个瓶子都恰好装两个值相同的变量的概率。

$0 \le m \le 10^9, 1 \le n \le 10^9, 1 \le D \le 10^5$

[CTS2019] 珍珠 Solution

首先考虑什么情况可以满足题目条件：奇数个数不超过 $n-2m$ 个。

枚举奇数个数，可以发现我们答案就是要求 $\sum_{i=0}^{\min(n-2m,D)} f_i$，其中 f_i 为出现次数恰好为 i 的方案数。

那么有 $f_i=\binom{D}{i}\sum_{j=0}^{n}\binom{n}{j}[x^j](\frac{e^x-e^{-x}}{2})^i[x^{n-j}](\frac{e^x+e^{-x}}{2})^{D-i}=\binom{D}{i}[x^n](\frac{e^x-e^{-x}}{2})^i(\frac{e^x+e^{-x}}{2})^{D-i}$

发现这个式子不是很好往下推，用 g_i 表示至少 i 个奇数反演回来，$g_i=\binom{D}{i}[x^n](\frac{e^x-e^{-x}}{2})^i e^{x(D-i)}$。

二项式定理展开，$g_i=\binom{D}{i}\frac{1}{2^i}[x^n]\sum_{j=0}^{i}\binom{i}{j}(-1)^{i-j}[x^n]e^{(D-2i+2j)x}=\frac{D!}{2^i(D-i)!}\sum_{j=0}^{i}\frac{(-1)^j(D-2j)^n}{j!}\times\frac{1}{(i-j)!}$，可以通过卷积解决。

CF891E Lust

给定长度为 n 的序列，进行 k 次操作，每次等概率选一个数减一，并将答案加上序列中其他数的乘积。求操作完成后答案的期望，对 10^9+7 取模。
$n \le 5000, k, a_i \le 10^9$

CF891E Lust Solution

1. 首先发现答案实际上只和初末状态有关，即 $\prod_i a_i - \prod_i a'_i$。
2. 可以构造 EGF 来描述每个数的变化：$F_a(x) = \sum_{i\geq 0} \frac{a-i}{i!} x^i$，那么答案就是 $\frac{1}{n^k}[x^k]\prod_i F_{a_i}(x) = \frac{1}{n^k}[x^k]e^{nx}\prod_i(a_i - x)$。
3. $O(n^2)$ 求出 $\prod_i(a_i - x)$ 做个卷积就行了。

和二元选择有关的计数有时可以转化为基环树边的定向问题。

CCPC Final 2022. A

一个无重边无自环的图包含 n 个点 $2n-2$ 条边，问它可以表示成多少对 (T_1, T_2) 的并。
其中 T_1 是以 1 号点为根的有根树，且满足除根之外，每个点的父亲的编号小于他。T_2 是以 n 号点为根的有根树，且满足除根之外，每个点的父亲的编号大于他。
$n \le 5 \times 10^5$

CCPC Final 2022. A Solution

1. 直接计数不方便，考虑转化一下题目的意义：给每个点确定一个父亲，[1, n − 1] 种每个点找一个更大的父亲，[2, n] 找一个更小的父亲。
2. 那么一条边 (x, y) 有两个选择：成为 x 的父边或 y 的父边。
3. 构造一个两边均为 $n-1$ 个点的二分图，那么这个图构成基环树森林，需要给边定向满足每个点入度为 1。
4. 答案为二的连通块次方。

平方的组合意义：在所有选的点中有顺序地选出两个点。

2nd Ucup Stage 3 A

定义两个长度相等的字符串 A, B 的距离为 $A_i \neq B_i$ 的 i 的数目，称字符串 Q 为字符串 P 的一个 Almost Prefix 当且仅当 $|Q| \leq |P|$ 且 $P_{[1...|Q|]}$ 与 Q 的距离至多为 1。
给定字符串 S,T，求

$$\sum_{S=P_1+P_2+...+P_n} n^2$$

其中每个 P_i 均为 T 的 Almost Prefix，'$+$' 表示字符串拼接。
$|S|, |T| \leq 10^6$

2nd Ucup Stage 3 A. Solution

1. 这是一个对 S 串的划分问题，考虑序列 dp。
2. 序列 dp 中 Almost Prefix 的限制如何解决。
3. 平方项怎么处理。

The End

Thanks