组合数学

whzzt

2019 年 7 月 12 日

安徽师范大学附属中学

前言

讲课内容可能和动态规划等专题有较多重合，求轻喷。

Atcoder Grand Contest 012 D Colorful Balls

有 n 个有颜色和重量的球，你每次可以交换同色的重量和不超过 X 的球，或是异色的重量和不超过 Y 的球，求有多少种可能的排列。

有 n 个有颜色和重量的球，你每次可以交换同色的重量和不超过 X 的球，或是异色的重量和不超过 Y 的球，求有多少种可能的排列。

$n \leq 2 \times 10^5$

Solution

将可以交换的球连边，每个联通块内部的球可以任意排列。

Solution

将可以交换的球连边，每个联通块内部的球可以任意排列。

可以发现只有一个联通块，计算大小即可。

求满足对于所有的 i 均有 $|a_i - i| \neq k$ 的排列 a 数目。

$n, k \leq 2000$。

Solution

考虑计算至少 i 个位置不满足条件的方案数，随后我们就可以容斥求出答案了。

Solution

考虑计算至少 i 个位置不满足条件的方案数，随后我们就可以容斥求出答案了。

我们将不满足条件的数和位置连边，那么不满足条件的位置个数不少于 k 就是匹配了 k 个位置。

Solution

考虑计算至少 i 个位置不满足条件的方案数，随后我们就可以容斥求出答案了。

我们将不满足条件的数和位置连边，那么不满足条件的位置个数不少于 k 就是匹配了 k 个位置。

于是直接 DP 即可。

三维偏序

给三个排列 a, b, c，求有多少对 (x, y) 满足 $a_x < a_y, b_x < b_y, c_x < c_y$。

三维偏序

给三个排列 a, b, c，求有多少对 (x, y) 满足 $a_x < a_y, b_x < b_y, c_x < c_y$。

$n \leq 5 \times 10^6$

Solution

直接统计三维偏序是 $O(n \log^2 n)$ 的，难以通过。考虑优化，我们只确定其中的两维，进行二维数点。那么容易发现，一对 (x, y) 要么被统计一次，要么被统计三次。于是减去总方案数即可。

BZOJ 4036 **按位或**

一开始你有一个数字 0，每秒从 $[0, 2^n)$ 中以 p_i 的概率或上 i，问数字变为 $2^n - 1$ 的期望时间。

BZOJ 4036 **按位或**

一开始你有一个数字 0，每秒从 $[0, 2^n)$ 中以 p_i 的概率或上 i，问数字变为 $2^n - 1$ 的期望时间。

$n \leq 20$

Solution

考虑 min-max 反演，由于原本要求的是 n 位中变成 1 的期望时间的最大值，即有：

$$E_{\max}(U) = \sum_{S \subset U} (-1)^{|S|-1} E_{\min(S)}$$

Solution

考虑 min-max 反演，由于原本要求的是 n 位中变成 1 的期望时间的最大值，即有：

$$E_{\max}(U)=\sum_{S\subset U}(-1)^{|S|-1}E_{\min(S)}$$

稍加利用高维前缀和即可求得后者。

Arbitrary Arrangement

给定整数 k 和模意义下的概率 p。初始有一个空序列，每次往末尾添加一个字符，有 p 的概率添加 a，有 $1-p$ 的概率添加 b。当 ab 作为子序列出现了至少 k 次的时候停止，问此时子序列 ab 出现次数的期望。对 10^9+7 取模。$1 \le k \le 1000$。

Solution

令 $f_{i,j}$ 表示当前有 i 个字符 a，且 ab 已经出现了 j 次时停止时答案的期望，则有：

$$f_{i,j} = pf_{i+1}j + (1-p)f_{i,i+j}$$

Solution

令 $f_{i,j}$ 表示当前有 i 个字符 a，且 ab 已经出现了 j 次时停止时答案的期望，则有：

$$f_{i,j} = pf_{i+1}j + (1-p)f_{i,i+j}$$

当 $j > k$ 时答案即为 j，关键问题在于找到 i 的边界。

Solution

令 $f_{i,j}$ 表示当前有 i 个字符 a，且 ab 已经出现了 j 次时停止时答案的期望，则有：

$$f_{i,j} = pf_{i+1}j + (1-p)f_{i,i+j}$$

当 $j > k$ 时答案即为 j，关键问题在于找到 i 的边界。

注意到当 $i = k$ 时 $f_{i,j} = j + k + \frac{p}{1-p}$，于是可以 $O(k^2)$ 解决该问题。

isn

给定一个长度为 n 的序列 a，当 a 不是单调不降时必须删去一个元素，求有多少种不同的操作方案。

$n \leq 2000$

Solution

不妨考虑最终序列的长度，简单容斥可以得到最终的答案。

Solution

不妨考虑最终序列的长度，简单容斥可以得到最终的答案。

可以利用树状数组求出长度为 k 的不降子序列数目。

Atcoder Regular Contest 062F Painting Graphs with AtCoDeer

给定一张 n 个点 m 条边的无向图，要给所有边染上 1 到 k 之间的颜色。两个方案相同，当且仅当可以通过若干次环的 shift 来得到彼此。求本质不同的方案数目。

$n \leq 50, m, k \leq 100$

Solution

显然答案在每个点双是独立的。同时，若一个点双有多个环，实际上所有的置换都是能够被生成出来的。

Solution

显然答案在每个点双是独立的。同时，若一个点双有多个环，实际上所有的置换都是能够被生成出来的。

因此对于环的情况统计上 Polya 的贡献，对于多个环的情况统计上组合数即可。

喂鸽子

有 n 个数，每次随机选一个数加一，问最晚达到 k 的数达到 k 的时间的期望。

$n \leq 50, k \leq 1000$

Solution

直接 DP，令 $f_{i,j}$ 表示前 i 个有效轮有 j 只鸽子已经饱了 (数字达到 k) 的贡献，可以做到 $O(n^2k)$。

一开始有 N 个未知颜色的球，颜色只有两种。每次操作为先取出一个任意颜色的球，放入两种颜色的球各一个，再取出一个任意颜色的球。问 M 次操作后取出球的序列的方案数。$N, M \leq 3000$ 。

Solution

令 $f_{i,j}$ 表示前 i 次操作后剩余 j 个红球时的答案。这样进行 dp 显然会算重，因此考虑统计初始红球数最少的方案，这意味着这样的方案至少有一次达到没有红球的临界状态，而每一种方案也均只会有一种表示方法达到临界状态。因此这样做的话，时间复杂度就优化成了 $O(NM)$ 。

Problem

求有多少 n 个点的环套树，满足第 i 个点的度数为给定的 d_i。答案对 10^9+7 取模。

环套树是一个 n 个点、n 条边的简单（无重边、无自环）联通无向图。

$n \leq 5000$

Solution

当问题是树的时候，显然就是一个简单的 prufer 序列的问题。

Solution

当问题是树的时候，显然就是一个简单的 prufer 序列的问题。

当问题变成环套树时，我们可以定义一个新的 prufer 序列，其中最后一个元素是环。

Solution

当问题是树的时候，显然就是一个简单的 prufer 序列的问题。

当问题变成环套树时，我们可以定义一个新的 prufer 序列，其中最后一个元素是环。

唯一的限制是除掉环外的最后一个元素也在环上。

Atcoder Grand Contest 018 E Sightseeing Plan

从三个矩形中选三个点，问从第一个点走到第二个点再走到第三个点的不同的最短路径的条数之和。矩形互不相交，横纵坐标单调增且在 10^6 以内。

Solution

首先考虑如何计算一个点到一个矩形内所有点的方案数之和。通过对行列分别计算可以发现计算的过程就类似于二维差分的过程，这样我们就可以暴力枚举中间点进行计算了。

但仍会超时，于是我们定义一条路径的权值为中间经过的节点总数，将贡献拆成进入和离开的两个点的贡献，就可以得到一个关于边界长度的线性做法了。

Atcoder Grand Contest 019 F Yes or No

有 $n+m$ 个判断题，其中 n 个是对的，m 个是错的。现在将问题随机排列，问最优回答策略期望答对的题数。

$n, m \leq 5 \times 10^5$

Solution

显然我们会回答剩下的多的那个答案。将路径画在平面上，可以发现除了剩余的正确和错误相等的情况之外，我们答对的题目数量一定是 $\max(n, m)$。因此将相等的情况单独统计即可。

取名字太难了

给定一个大小为 N 的正整数集合 $S=\{1,2,\dots,N\}$，定义 $f(n)=\sum_{T\subset S,|T|=n}\prod_{x\in T}x \pmod p$，对于任意一个 $x\in[0,p-1]$，问 $f(n)=x$ 的 $n\in[0,N]$ 的个数，输出对 998244353 取模。

$N\le 10^{18}, p\le 250000$

Solution

即求 $\prod_{i=1}^{N}(x+i)$ 的系数在模 p 下的分布，注意到 $\prod_{i=1}^{p-1}(x+i) \equiv x^{p-1}-1$，问题相当于求 $(x^{p-1}-1)^a$ 和 $\prod_{i=1}^{b}(x+i)$ 在 $\bmod p$ 下的系数分布即可。对于前者我们 Lucas 定理，最后用 fft 合并，总时间复杂度 $O(p\log n)$。

小 Q 的集合

求 n 元集合 S 中选择一个子集 T 的 $|S-T|^k - |T|^k$ 的方差。
$n \le 10^{10^6}, k \le 10^6$，答案对质数 $m \le 10^6$ 取模。

Solution

容易发现期望为 0，那么只要求平方的期望，即
$\sum_{i=0}^{n} \binom{n}{i} (i^k - (n-i)^k)^2$。

Solution

容易发现期望为 0，那么只要求平方的期望，即 $\sum_{i=0}^{n}\binom{n}{i}(i^k-(n-i)^k)^2$。

利用 $lucas$ 定理展开一层，得到 $\sum_{i=0}^{\lfloor\frac{n}{m}\rfloor}\binom{\lfloor\frac{n}{m}\rfloor}{i}\sum_{j=0}^{n \bmod m}\binom{n \bmod m}{j}(j^k-(n-j)^k)^2$。

线性筛或直接快速幂解决即可。

关于组合数取模的一些方法

Compressed Spanning Subtrees

有一棵 $n \le 100$ 的树 T 满足任意点点度不为 2，你每次可以询问一个点集的虚树的大小，在 2550 次内询问出整棵树。

Compressed Spanning Subtrees

有一棵 $n \le 100$ 的树 T 满足任意点点度不为 2，你每次可以询问一个点集的虚树的大小，在 2550 次内询问出整棵树。

先询问出叶子，以任意叶子为根后其他的暴力询问即可。

几排数

有一个 $n \times n$ 的 01 矩阵，每次可以询问一个矩阵是不是这个 01 矩阵的子矩阵，$5n^2$ 次内询问出原矩阵。

几排数

考虑一个朴素的暴力：我们只询问 $1 \times k$ 的矩阵，并在最后将所有行拼起来。我们可以对所有 $1 \times k$ 的矩阵暴力检查能不能拓展，如果能拓展就继续询问，这一部分显然只需要 n^3 次询问。

几排数

考虑一个朴素的暴力：我们只询问 $1 \times k$ 的矩阵，并在最后将所有行拼起来。我们可以对所有 $1 \times k$ 的矩阵暴力检查能不能拓展，如果能拓展就继续询问，这一部分显然只需要 n^3 次询问。

我们可以对这个暴力做一些优化，我们可以对一个串检查所有比它更长的串是否有可行的，或是有比它更短的串不可行的，这样可以拿到更高的分数。

几排数

考虑一个新做法，不妨假设现在有了一些 $1 \times n$ 的串，我们将其所有子串建立一个 AC 自动机，并尝试在每个子串后面增加字符，如果有一个能够增加字符的就暴力拓展一个新的 $1 \times n$ 的串出来，而如果其 fail 指向的点已经不能拓展了我们也不拓展。考虑最终 n 个串构成的广义后缀自动机，容易发现失败的询问次数不会超过 $2n^2$，拓展出所有行需要 $2n^2$，最后的一部分不会超过 n^2，因此总的询问次数也不会超过 $5n^2$。

Thanks!