随机化选讲

isaunoya

Hangzhou Dianzi University

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

目录

随机枢轴
随机指纹
见证抽样
选讲：指纹与抽样
选讲：交互
选讲：顺序与搜索
附录：构造与证明

P1923：求第 k 小的数

P1923 求第 k 小的数：给定 n 个数，求它们排序后下标为 k 的元素。$n \leq 5 \times 10^6$，k 从零开始计数，允许重复值。

P1923：求第 k 小的数

P1923 求第 k 小的数：给定 n 个数，求它们排序后下标为 k 的元素。$n \le 5 \times 10^6$，k 从零开始计数，允许重复值。

在当前区间均匀随机取枢轴 v，原地划分为三段，前两段长度分别为 a, b：

小于 v	等于 v	大于 v
$[0, a)$	$[a, a+b)$	$[a+b, m)$
继续左段	立即返回 v	转到右段，秩减 $a+b$

段内不必排序，目标所在的段已经确定。相等元素被整段排除，大量重复值也不会造成逐个剥离。

P1923：期望线性不等于最坏线性

当前段长为 m。枢轴落在排序后中间一半的概率至少为 $\frac{1}{2}$，此时继续处理的段长至多约 $3\frac{m}{4}$；重复值只会扩大一次排除的相等段。

一次收缩前，每轮命中概率仍至少为 $\frac{1}{2}$。连续失败 j 轮的概率至多 2^{-j}，所以等待轮数的期望至多 $1+\frac{1}{2}+\frac{1}{4}+\ldots=2$。

这一阶段每轮扫描至多 m 个数，期望代价为 $O(m)$。再按收缩阶段累加：

$$n+\left(\frac{3}{4}\right)n+\left(\frac{3}{4}\right)^2 n+\ldots=O(n).$$

答案始终正确，随机性只影响耗时；这种算法称为 Las Vegas 算法。最坏仍可能 $O(n^2)$。循环维护工作区间、原地三路划分，额外空间 $O(1)$。

附录：三路划分不变量

随机算法的两种保证

输入固定以后，程序使用随机数决定执行路径。随着随机选择而变化的量，例如运行步数、抽中的位置或输出答案，就是这里要分析的随机变量。

类型	答案	工作量
Las Vegas	每种随机选择下都正确	通常分析期望耗时
Monte Carlo	允许有界的小概率错误	可以预先限制运行轮数

随机选择枢轴，可能让选择算法多做几次划分；随机指纹相等，则可能让两个不同对象被误判。前者分析**工作量**，后者先定义**失败事件**。

主讲与选讲

主讲部分	核心问题
随机枢轴	随机性影响耗时，如何累计期望工作量？
随机指纹	如何把相等条件压缩，又控制全程误判？
见证抽样	抽中一个见证，如何覆盖并验证最优答案？

主讲需要前缀和、基本概率与常见数据结构；超集计数会在对应例题中给出递推。

选讲题目 按所需数据结构选择。交互题 讨论抽样、还原与询问预算。顺序与搜索 讨论集中界、随机试探与局部修复；构造与证明附录 可按需跳转。

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

矩阵乘法验证

给出 $n \times n$ 矩阵 A, B, C，判断 $AB = C$。

矩阵乘法验证

给出 $n \times n$ 矩阵 A, B, C，判断 $AB = C$。

直接计算整个乘积需要通常的 $O(n^3)$ 时间；如果只要求发现不相等，可以随机投影到一列。

取独立均匀的零一向量 r，利用乘法结合律，计算

$$u = Br, \quad v = Au, \quad w = Cr.$$

只做三次矩阵乘向量，耗时 $O(n^2)$。若 $v \neq w$，一定有 $AB \neq C$；若相等，则还需控制未发现差异的概率。这里不能先计算 AB，否则失去了投影的意义。

Freivalds：单轮误判概率

令 $D = AB - C \neq 0$，取一行中的非零系数 $D_{i,j} = d$。固定 r 的其他坐标后，该行内积只剩独立比特 r_j 未确定：

$r_j = 0$	$r_j = 1$
内积为 z	内积为 $z + d$

因为 $d \neq 0$，两个分支不可能同时为零。每个分支概率为 $\frac{1}{2}$，故单轮漏掉差异的概率至多 $\frac{1}{2}$。

独立重复 t 轮，全部漏检的概率至多 2^{-t}，时间 $O(tn^2)$。这里使用精确运算；若把整数矩阵另取模数，还需防止非零差被模数抹成零。

独立重复与全程误差

若一次矩阵检查允许误差 ε，独立重复

$$t = \left\lceil \log_2 \left(\frac{1}{\varepsilon} \right) \right\rceil$$

轮即可。若全程有 Q 次检查，则取 $t = \left\lceil \log_2 \left(\frac{Q}{\varepsilon} \right) \right\rceil$，总漏检概率至多 $Q2^{-t} \le \varepsilon$。

每轮只要发现不同即可拒绝；全部通过才接受。独立重复得到乘积界，不同查询之间则使用并集界，不要求查询事件独立。

以下输入在随机种子产生前固定；同一张指纹表上的多次查询，共同计入全程误差。一般失败率的轮数公式 与 依赖历史的查询 见附录。

ABC367F：Rearrange Query

ABC367F Rearrange Query：给定长度为 n 的数组 A, B，每次问 $A[l..r]$ 能否重排成 $B[L..R]$。$n, q \leq 2 \times 10^5$，所有值属于 $[1, n]$。

ABC367F：Rearrange Query

ABC367F Rearrange Query：给定长度为 n 的数组 A, B，每次问 $A[l..r]$ 能否重排成 $B[L..R]$。$n, q \le 2 \times 10^5$，所有值属于 $[1, n]$。

逐值比较频次要 $O(n)$，能否把整张频次表压成一个数？

选素数 $p > n$，为每个值 x 独立均匀生成 $h_x \in [0, p)$，计算时对 p 取模。两数组共用同一张权值表，分别计算权值前缀和。

多重集哈希：一个未固定的随机权值

以下在素域 $\mathbb{F}_p$ 中计算。给每个值 x 独立均匀权值 w_x，多重集指纹为 $H = \sum_x c_x w_x$。

两个计数向量不同，其差中必有一个非零系数。选取对应的 w_x，固定其余权值，则

$$\Delta = c w_x + z, \quad c \neq 0.$$

只要 p 大于可能的计数差，c 在模 p 后仍非零。素数模下非零数有乘法逆元，因此 $w_x \mapsto c w_x + z$ 将 p 个余数一一对应地排列；恰好一个取值 $w_x = -\frac{z}{c}$ 令差为零，碰撞概率为 $\frac{1}{p}$。

矩阵验证保留一个未确定比特，这里保留一个未确定域元素；两者都由一个非零系数控制误判。

ABC367F：前缀差比较多重集指纹

两数组使用相同的权值表。先检查区间长度，再比较权值前缀差：

$$H_A(r) - H_A(l-1) \equiv H_B(R) - H_B(L-1) \bmod p.$$

相同多重集一定通过。不同多重集的计数差满足 $|c_x| \le n < p$，因此单次误判概率为 $\frac{1}{p}$。

本题可取 $p = 2^{61} - 1$，使用两套独立权值，要求两份指纹同时相等。固定输入的 q 次查询，全程误差至多 $\frac{q}{p^2}$。

预处理 $O(n)$，每次查询 $O(1)$，空间 $O(n)$。权值和始终按模 p 计算；这是只可能把“不相等”误判成“相等”的单侧错误。

官方题解

ABC367F：异或无法保留成对出现的次数

考虑同样长的两个区间 $[1,1]$ 与 $[2,2]$。它们不能互相重排，但若用异或：

$$h_1 \oplus h_1 = 0 = h_2 \oplus h_2.$$

无论怎样生成随机权值，这两个区间都会相等；多重复几轮也无济于事。

使用域上的加法，比较的是 $2h_1$ 与 $2h_2$。只要 $p > n \geq 2$，系数二不为零，错误才成为概率至多 $\frac{1}{p}$ 的偶然碰撞。

先检查指纹是否保留题目需要的信息，再讨论随机性有多强。

保序哈希：多项式的随机求值

素数 p 的模运算构成有限域 $\mathbb{F}_p$：加、减、乘都对 p 取模，每个非零数有乘法逆元。在域上，非零的 d 次多项式至多有 d 个根。

将一个定长字符串写成多项式，均匀随机取 $r \in \mathbb{F}_p$：

$$H(a_1, ..., a_m) = \sum_{i=1}^{m} a_i r^{m-i}.$$

两个不同字符串相减，得到次数至多 $m-1$ 的非零多项式；只要字符编码在模 p 后仍不同，碰撞就是“随机底数恰好落在某个根上”，概率至多 $\frac{m-1}{p}$。不同长度应另行比较长度。

保序哈希：前缀截取与位置对齐

预处理幂 $R_i = r^i$ 和前缀 $P_0 = 0$、$P_i = P_{i-1} r + a_i$，所有运算在 $\mathbb{F}_p$ 中进行。则

$$H(l, u) = P_u - P_{l-1} R_{u-l+1}.$$

减去左边部分以后，留下的是 $a_l r^{u-l} + \ldots + a_u$，指数从区间长度减一降到零，与原下标无关。因此同长度的两个区间可以直接比较，不必计算逆元。

预处理 $O(n)$，每次截取 $O(1)$。多组独立底数分别维护前缀与幂表；负数取模需规范到同一剩余类，模乘不能先在过窄的整数类型中溢出。

指纹与需要保留的信息

序列

$\sum_i a_i r^{m-i}$

交换位置通常改变指纹

多重集

$\sum_x c_x w_x$

只关心各值出现次数

奇偶集合

$\bigoplus_x w_x$

同一项出现两次抵消

同样是“哈希”，保留的信息并不相同。先把目标条件写成等式，再挑一个能检测非零差异的随机量。

P8819：星战

CSP-S 2022 星战：维护一张有向图，支持关闭或恢复一条边，以及关闭或恢复某个点的全部入边。$n, m, q \leq 5 \times 10^5$。

每次操作后，判断是否同时满足：每个点恰有一条可用出边；从每个点出发，都能沿可用边无限行走。

星战：出度条件与线性指纹

每点恰有一条可用出边时，有限图上从任意点出发都会进入环，能够无限行走，因此只需检查所有出度是否为一。

取素数 $p > m$，给每个顶点独立均匀权值 $w_u \in \mathbb{F}_p$。对两套独立权值分别维护以下量：

$$b_v = \sum_{(u,v)\ \text{激活}} w_u, \quad H = \sum_v b_v, \quad H_0 = \sum_u w_u.$$

于是 $H - H_0 = \sum_u (d_u - 1) w_u$。如果所有出度都为一，等式必然成立；否则计数差向量非零，可以用随机线性指纹判定。

星战：按终点维护入边的权值和

预存每个终点的原始入边权值和 b_v^0。单边激活或失活时，对 b_v, H 加减 w_u；全部失活时从 H 扣掉旧 b_v 并清零；全部激活时令 $H \leftarrow H - b_v + b_v^0$，再恢复 $b_v = b_v^0$。

本题保证单边删除时该边可用、单边修复时该边失效，所以可以直接加减。全部入边统一操作可以重复，必须使用当前 b_v 与原始和 b_v^0 整体替换。初始化 $O(n+m)$，每次操作 $O(1)$。

非零系数 $d_u - 1$ 的绝对值不超过 $m < p$，因此一组误判至多 $\frac{1}{p}$。要求两组同时通过，固定输入下 q 次操作的全程误差至多 $\frac{q}{p^2}$。

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

P3765：总统选举

总统选举 每次询问区间的严格多数；若不存在，就由给定的 s_i 获胜。公布赢家后，将指定的 k_i 个位置全部改投它。$n, m \leq 5 \times 10^5$，$\sum k_i \leq 10^6$。

P3765：总统选举

总统选举 每次询问区间的严格多数；若不存在，就由给定的 s_i 获胜。公布赢家后，将指定的 k_i 个位置全部改投它。$n, m \le 5 \times 10^5$，$\sum k_i \le 10^6$。

若多数候选存在，它占据区间中超过一半的位置。均匀抽一个位置：

多数候选	其余位置
命中概率 $> \frac{1}{2}$	未命中概率 $< \frac{1}{2}$

抽到候选后精确数票；只有确实超过一半才接受。连续 t 次都未抽中多数，概率小于 2^{-t}；未找到严格多数时，使用题目给定候选。

总统选举：抽到候选后仍要精确数票

为每种候选维护支持位置的有序集合。两次秩查询精确数票，一次改票删除旧位置、插入新位置；两者均为 $O(\log n)$。

若共有 q 次预选与 u 次改票，最后还要对全体选民统计一次多数。每次使用新的抽样，给定此前历史，漏掉当前多数的概率仍小于 2^{-t}。

若全程允许误差 ε，取

$$t = \left\lceil \log_2 \left(\frac{q+1}{\varepsilon} \right) \right\rceil,$$

总误差不超过 ε，总工作量为 $O(n \log n + (qt+u) \log n)$。每个抽样候选都需精确验证；最后全体统计未找到多数时输出 -1。下一页给出省去这 t 倍工作的确定性方案。

总统选举：确定性多数候选

线段树的每个节点保存一个候选与余额。合并时，相同候选的余额相加；不同候选相互抵消，只留下余额较大者及两者之差，差为零则为空。

这个过程等价于不断删除两个不同的值。严格多数的数量大于所有其他值之和，因此无论怎样抵消，它都不会被完全删掉。区间合并得到的候选仍需用位置平衡树精确验证，非空摘要不代表一定存在多数。

每次改票同时更新候选线段树与位置树。查询候选和验证均为 $O(\log n)$，计入建树后总时间 $O((n+q+u)\log n)$、空间 $O(n)$，答案完全确定；最后一次全体统计使用同样流程。

见证抽样与精确验证

固定一个最优答案。若有至少 α 比例的位置能够“见证”它，就从这些位置中均匀抽样。抽中的位置不必直接等于答案，只需将答案限制到一个可处理的小候选集。

每轮只做三件事：抽位置、枚举这个位置所允许的候选、精确验证候选。验证保证输出可行，随机性只影响是否漏掉更好的解。

如果抽到任意见证就能枚举出固定的最优答案，独立抽样 t 轮后漏解概率至多 $(1-\alpha)^t$。一次命中已经足够，不需要命中所有最优答案，也不需要不同候选相互独立。

CF1523D：Love-Hate

CF1523D Love-Hate：$n \le 2 \times 10^5$ 行、$m \le 60$ 列的零一矩阵，每行至多 $p \le 15$ 个一。求最大的列集合，使至少 $\lceil \frac{n}{2} \rceil$ 行在这些列上全为一。

CF1523D：Love-Hate

CF1523D Love-Hate：$n \le 2 \times 10^5$ 行、$m \le 60$ 列的零一矩阵，每行至多 $p \le 15$ 个一。求最大的列集合，使至少 $\lceil \frac{n}{2} \rceil$ 行在这些列上全为一。

固定一个最优集合 S^*。至少一半的行包含它，均匀抽一行，命中这样一行的概率至少为 $\frac{1}{2}$。

设抽中行的一的位置集合为 P，大小为 $s \le 15$。若抽到了见证，$S^* \subseteq P$；我们只需在 P 的 2^s 个子集中寻找最优。

CF1523D：超集计数一次验证全部候选

把每一行投影到 P，得到 s 位掩码 M，统计 $f[M]$。初始化 $C = f$，然后依次处理每个比特 b：

$$C[S] \leftarrow C[S] + C[S \cup \{b\}] \quad (b \notin S).$$

每处理一个比特，就允许这一位出现在原行中而不出现在候选中。处理完所有位后，$C[S] = \sum_{M \supseteq S} f[M]$，恰好是包含 S 的行数。

在 $C[S] \geq \lceil \frac{n}{2} \rceil$ 中取置位数最多的掩码，还原到原来 m 列；保留多轮中最大的可行解。

CF1523D：见证命中与复杂度

只要抽中的行包含某个固定最优集合 S^*，本轮枚举就会检查它；所有输出又都经过精确计数，所以不会出现不可行答案。

允许单组漏解概率 ε 时，取 $R = \left\lceil \log_2\left(\frac{1}{\varepsilon}\right) \right\rceil$ 轮。多组有总预算时，对各组的 2^{-R} 求和。

输入读取 $O(nm)$；每轮投影 $O(np)$、超集 DP $O(p2^p)$，总时间为

$$O(nm + R(np + p2^p)).$$

空间 $O(n + 2^p)$。原行用 64 位整数保存，投影只检查样本的一的位置；全零样本对应空集。门槛为 $\lceil \frac{n}{2} \rceil$。重试次数与总工作量由同一个 R 控制。

随机化与确定性方案的选择

随机枢轴保持答案正确，分析期望工作量；指纹压缩判等，分析非零差恰好消失的概率；见证抽样覆盖候选，靠精确验证保证输出可行。

选择算法时，同时写出每轮代价、单轮失败事件与重复次数，再比较确定性替代的总工作量。能验证候选可行，不等于已经证明它最优。

至此完成主讲内容。后续按需要选择 数据结构与抽样应用、随机顺序的方法边界 、交互题 或 构造与证明。

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

选讲题目与前置知识

所需基础	对应主题
分治、前缀和	候选区间枚举、带位置关系的哈希
线段树、滑动窗口	动态指纹、模次数的前缀状态
树状数组、动态树	随机比特检验、跨割与虚子树摘要
质因数分解、因子格	最优公因子、最优质因子
图算法、状态 DP	判桥、抽取割边、随机限制访问顺序、染色约束

这些题复用主讲中的失败事件分析，各自补充所需结构。返回主讲安排

CF1175F：The Number of Subpermutations

CF1175F The Number of Subpermutations 统计多少个连续子区间恰好是 $1, 2, \ldots, m$ 的一个排列，其中 m 为区间长度。

CF1175F：The Number of Subpermutations

CF1175F The Number of Subpermutations 统计多少个连续子区间恰好是 $1, 2, \ldots, m$ 的一个排列，其中 m 为区间长度。

给每个值独立随机权值，预处理原数组的哈希前缀和与值域 $[1, m]$ 的哈希前缀和。一个长度为 m 的区间若哈希相同，就以高概率具有目标多重集。

逐一检查全部区间仍需 $O(n^2)$ 次判定，可以用分治减少候选：设当前区间分成 $[L, M]$ 与 $[M+1, R]$，递归统计不跨中点的答案，只需处理跨中点的子排列。

CF1175F：最大值与区间端点

1. 枚举左端点 l，令 $v=\max(a_l,\ldots,a_M)$。若整个区间的最大值来自左半段，长度只能为 v，所以右端点必须是 $r=l+v-1$。检查它是否跨中点、是否在 R 内、右半段最大值是否小于 v，再比对哈希。
2. 对称地枚举右端点 r，令 $v=\max(a_{M+1},\ldots,a_r)$，候选左端点唯一为 $l=r-v+1$；要求左半段最大值小于 v。

合法排列的最大值只出现一次，恰好落入其中一类，因此不会漏计或重复。左右最大值可在线扫描并保存成数组，每层总工作量 $O(n)$，总时间 $O(n\log n)$。长度为一的区间只有元素为 1 时合法。

CF1175F：候选判定与累计误差

为每个值生成两套独立的域权值，取素数 $p > n$。原数组前缀和给出候选区间指纹，值域前缀和给出 $1, \ldots, m$ 的目标指纹；两组都相同才计入答案。

合并时预存左半后缀最大值与右半前缀最大值。候选另一端已经由长度唯一确定，因此跨中点检查、范围检查和两边最大值比较均为 $O(1)$。

设实际检查的候选数为 $M = O(n \log n)$。它们由原数组的最大值确定，不依赖随机权值；不同多重集被误计的全程概率至多 $\frac{M}{p^2}$。计数用 64 位整数，候选越过当前分治边界时直接跳过。

合肥 D：Balanced Array

2023 合肥 D：Balanced Array 的一个前缀 $a_1, \ldots, a_l$ 合法，当且仅当存在 $k \geq 1$、$2k < l$，使所有 $1 \leq i \leq l - 2k$ 满足

$$a_i + a_{i+2k} = 2a_{i+k}.$$

判断数组的每一个前缀是否合法。

合肥 D：Balanced Array

2023 合肥 D：Balanced Array 的一个前缀 $a_1, \ldots, a_l$ 合法，当且仅当存在 $k \geq 1$、$2k < l$，使所有 $1 \leq i \leq l - 2k$ 满足

$$a_i + a_{i+2k} = 2a_{i+k}.$$

判断数组的每一个前缀是否合法。

将三段的多项式哈希对齐到相同长度，这组条件可以判定为

$$H(1, l-2k) + H(2k+1, l) = 2H(k+1, l-k) \bmod p.$$

这里使用保序的多项式哈希。设 $|a_i| \leq A$，差多项式的每个系数绝对值至多 $4A$；选素数 $p > 4A$，才能保证非零整数系数不会被模数抹成零。

平衡数组：固定间隔后，合法前缀单调

还剩下如何枚举 k。固定 k 后，合法性对前缀长度单调：若长度 l 已有一个不满足的等式，任何更长前缀都会包含它。于是先检查最短可能长度 $2k+1$，再二分最大合法长度 L_k。

这个 k 能使所有 $l \in [2k+1, L_k]$ 合法，用差分数组标记这段长度区间。枚举 $k \le \left\lfloor \frac{n-1}{2} \right\rfloor$ 后做前缀和，即可回答所有前缀。时间 $O(n \log n)$，空间 $O(n)$。

每个不合法比较对应次数至多 $n-1$ 的非零多项式，随机底数的误判概率至多 $\frac{n-1}{p}$。在首次误判前，执行路径与精确二分相同，因此只需对这条精确路径上的 $O(n \log n)$ 次比较取并集界，无须假设各次判定独立。

题目 · 原提交

P3792：连续整数重排判定

P3792 由乃与大母神原型和偶像崇拜 支持单点修改，查询区间能否重排成一段连续整数。

P3792：连续整数重排判定

P3792 由乃与大母神原型和偶像崇拜 支持单点修改，查询区间能否重排成一段连续整数。

设区间长为 m，最小、最大值为 x, y。首先必须满足 $y - x + 1 = m$，但这不能排除重复值。还需要比对区间多重集与 $x, x+1, \ldots, y$ 的指纹。

一种无需枚举整个值域的方法，是取随机底数 $r \notin \{0, 1\}$，令值 v 的权值为 r^v。目标区间的指纹由等比数列求和得到：

$$\sum_{v=x}^{y} r^v = r^x \frac{r^m - 1}{r - 1} \bmod p.$$

P3792：次数上界与误差预算

线段树维护区间最值与指纹和。值域上界为 V，单组操作 $O(\log n + \log V)$，其中 $\log V$ 来自模幂；使用 t 组时，操作 $O(t(\log n + \log V))$、空间 $O(tn)$。

取素数 $p > \max(n, 2)$，防止非零计数差被抹零。极差检查通过后，所有值都在 $[x, x + m - 1]$ 内；差多项式除以非零的 r^x，次数至多 $m - 1$。

底数从 $\mathbb{F}_p$ 去掉零、一后的 $p - 2$ 个值中均匀选取，单次误判至多 $\frac{m-1}{p-2}$。取 t 个独立底数，固定输入下 q 次比较的全程误差至多

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

按目标误差选择参数；仅更换固定模数不代表事件独立。

CF1418G：Three Occurrences

CF1418G Three Occurrences 统计每个出现过的值都恰好出现三次的区间。

CF1418G：Three Occurrences

CF1418G Three Occurrences 统计每个出现过的值都恰好出现三次的区间。

先把“恰好三次”拆成两件事：所有次数都是三的倍数，且没有值出现超过三次。前一件事用随机指纹，后一件事用滑动窗口。

令前缀 $[1,i]$ 中值 x 的次数模三为 $c_x(i)$。选素数 $p>2$，分别用两套独立均匀权值维护指纹对 $H_i=(H_i^1,H_i^2)$，其中

$$H_i^j=\sum_x c_x(i)w_x^j \bmod p,\quad j=1,2.$$

CF1418G：滑动窗口排除出现六次的情况

一个值连续出现时，对指纹的增量循环为 $w_x, w_x, -2w_x$。区间 $[l, r]$ 中所有次数为三的倍数，等价于两个前缀的完整余数向量相等，进而可以通过 $H_{l-1} = H_r$ 判定。

扫描右端点 r，若当前值最近四次出现的位置为 $p_1 < p_2 < p_3 < p_4 = r$，则合法左端点必须大于 p_1。将全局左边界维护为这些限制的最大值，在哈希表中只保留允许的前缀下标 $l - 1$。查询与 H_r 相同的前缀数量，再插入当前前缀。

CF1418G：允许的前缀下标

处理右端点 r 时，字典应当只保存下标在 $[L-1, r-1]$ 中的前缀，其中 L 为最小合法左端点。

初始字典放入空前缀 H_0 一次。令 L 为最小合法左端点，及时删除所有下标小于 $L-1$ 的旧前缀；先查询 H_r 的出现次数，再插入 H_r，避免把空区间计入答案。答案用 64 位整数。

字典以指纹对为键；哈希表期望时间 $O(n)$、空间 $O(n)$，平衡树时间为 $O(n \log n)$。

CF1418G：前缀比较的全程误差

每个余数向量坐标属于 $\{0,1,2\}$。两个不同向量的差有一个非零坐标，绝对值至多二，因此在 $p>2$ 的域中仍非零。两套独立权值下，一对不同向量的指纹同时相等的概率为 $\frac{1}{p^2}$。

虽然字典只执行 $O(n)$ 次操作，累计计数隐含比较了许多前缀对。共有 $n+1$ 个前缀，固定输入下的全程碰撞概率至多

$$\frac{\binom{n+1}{2}}{p^2}.$$

窗口限制只会减少需要比较的前缀对；不能仅按字典操作次数把误差预算写成 $\frac{n}{p^2}$。

CF1746F：Kazaee

CF1746F Kazaee 支持单点修改，查询区间中每个值的出现次数是否均为 k 的倍数。

CF1746F：Kazaee

CF1746F Kazaee 支持单点修改，查询区间中每个值的出现次数是否均为 k 的倍数。

每轮给每个值 x 独立分配一个随机比特 $b_x \in \{0,1\}$。维护区间的 $S = \sum_x c_x b_x$，检查它能否被 k 整除。真正满足条件时，每轮都会通过。

CF1746F：独立重复，逐轮重放省空间

若存在 c_x 不能被 k 整除，固定其他比特后，$b_x = 0$ 和 $b_x = 1$ 得到的两个和模 k 不同，最多一个能通过。因此单轮误判概率至多 $\frac{1}{2}$，重复 t 轮后至多 2^{-t}。

每轮用树状数组支持单点修改和区间求和，总时间 $O(t(n+q)\log n)$。

设初始值及全部修改中共有 $D \le n+q$ 种值。并行维护 t 轮时，树状数组占 $O(tn)$，随机比特表占 $O(tD)$，合计 $O(t(n+D))$。逐轮重放则只需一棵树、一轮比特表和操作记录，空间为 $O(n+q)$。

CF1746F：重放顺序与错误预算

先收集初始数组和所有修改中出现的值，离散化后分配随机比特。同一轮内，同一个值始终使用同一个比特；修改位置时，在树状数组中加入新旧比特之差。

每轮从原始数组重新开始，依次重放全部操作，不能把上一轮结束后的数组当作新起点。询问一旦被某轮否定，最终答案就为否；其他询问仍需完成余下轮次。

若真正执行的询问为 Q 次，全程错误概率至多 $Q2^{-t}$，所以取 $t \geq \left\lceil \log_2\left(\frac{Q}{\varepsilon}\right) \right\rceil$ 足够。长度不是 k 的倍数时可以先判否；$k=1$ 时必然成立。

UOJ207：共价大爷游长沙

共价大爷游长沙 维护一棵动态树和一个点对集合。树边可以断开后重新连接；点对可以加入、删除；询问某条树边是否被所有点对的路径经过。

UOJ207：共价大爷游长沙

共价大爷游长沙 维护一棵动态树和一个点对集合。树边可以断开后重新连接；点对可以加入、删除；询问某条树边是否被所有点对的路径经过。

给每个活动点对 $e = (u, v)$ 独立的随机 64 位权值 w_e，在两个端点各异或一次；另维护所有活动点对权值的异或和 H。不同点对记录即使端点相同，也使用不同的记录编号和权值，删除时复用加入时保存的权值。

UOJ207：跨割摘要应等于全部点对摘要

删去询问边后，树分成 A, B。对 A 内所有点的权值异或，两个端点都在 A 的点对出现两次抵消，都在 B 的出现零次，恰好剩下跨割点对的权值异或 X_A。

一条点对路径经过这条边，当且仅当它的端点跨割。因此所有路径都经过的必要条件是 $X_A = H$。如果有不跨割的点对，两者的差就是这些点对随机权值的异或；对固定集合，误判概率为 2^{-64}。空点对集合时条件自然成立。

UOJ207：access 时同时更新虚子树异或

用支持虚子树信息的 Link-Cut Tree。每个点维护自身权值、所有虚儿子贡献的异或，以及辅助树左右儿子和虚贡献合起来的总异或。

access 改变重儿子时，旧右儿子的总异或加入虚贡献，新右儿子的总异或从虚贡献移除；异或的加入与移除是同一个操作。随后正常 pushup。点对增删只需对两个端点执行 access，修改自身权值并向上更新。

UOJ207：暂时断边，查询整棵连通块

查询边 (u,v) 时暂时 `cut(u,v)`，将 u 所在连通块换根并访问 u，其维护的整棵树异或即为 X_A；比较后 `link(u,v)` 恢复。换边操作直接使用同样的断边、连边接口。

每次操作仅执行常数次动态树操作，摊还时间 $O(\log n)$，空间为 $O(n+q)$，其中点对记录占用 $O(q)$。普通只维护路径信息的 LCT 不够，关键是 `access` 时同步维护虚子树贡献。

若最多有 Q 次查询，沿精确执行路径取并集界，全程误判至多 $Q2^{-64}$。需要更低预算时，可用两套独立 64 位权值，要求两份摘要都相等。

返回选讲题目

CF364D：Ghd

CF364D Ghd：在 $n \le 10^6$ 个正整数中，求能整除至少 $\lceil \frac{n}{2} \rceil$ 个数的最大整数 g。数值上界 $A = 10^{12}$，重复值按位置计数。

CF364D：Ghd

CF364D Ghd：在 $n \le 10^6$ 个正整数中，求能整除至少 $\lceil \frac{n}{2} \rceil$ 个数的最大整数 g。数值上界 $A = 10^{12}$，重复值按位置计数。

固定一个最优答案 g^*。均匀抽取一个位置，其值 x 被 g^* 整除的概率至少为 $\frac{1}{2}$。抽中后，g^* 必在 x 的因子中。

因此抽样后只需求“x 的所有因子中，哪些整除了足够多的数组元素”。**不能为每个因子都扫描一遍整个数组。**

CF364D：gcd 频次与因子格累加

对每个数组元素计算 $d_i = \gcd(a_i, x)$，统计每个 x 的因子作为 d_i 出现了多少次。

分解 $x = \prod_{j=1}^{k} p_j^{e_j}$，每个因子对应指数向量 $(b_1, \ldots, b_k)$。将这些频次放入大小 $D = \prod_{j=1}^{k} (e_j + 1)$ 的多维数组。

对每一维，从大指数向小指数累加：固定其他坐标，依次令 $b_j = e_j - 1, \ldots, 0$，执行 $C[b] += C\left[b + \mathrm{unit}_j\right]$。全部维度完成后，$C[b]$ 就是所有倍于该因子的 gcd 数量。

CF364D：精确验证，随机性只影响漏掉最优

因为 g 是 x 的因子，有

$$g \mid a_i \quad \Leftrightarrow \quad g \mid \gcd(a_i, x).$$

因子格累加得到的正是整除计数。取 $C[g] \geq \lceil \frac{n}{2} \rceil$ 的最大因子，并与此前轮次的答案取最大值。答案始终可行，初始为一。

独立抽样 R 轮，未抽中任何最优答案见证的概率至多 2^{-R}。抽中一次就枚举并验证了 g^*，此后不会丢失它。

CF364D：因子计数的完整工作量

试除分解 x 最坏 $O\left(\sqrt{A}\right)$。生成因子并排序，每个因子同时记录其指数数组下标；对 n 个 gcd 用二分找到下标，再累加频次。

一轮时间为

$$O\left(\sqrt{A}+n(\log A+\log D)+D\log D+kD\right),$$

其中 kD 是因子格累加，不能写成 nD。总时间乘以 R；总空间 $O(n+D)$。实现用 64 位整数；相同样本值可跳过，避免重复计算，但概率估计仍按原始独立抽样轮次。

CF1305F：Kuroni and the Punishment

CF1305F Kuroni and the Punishment 允许对正整数做加一或减一，要求最后仍为正数且整体 gcd 大于一，求最少操作次数。

CF1305F：Kuroni and the Punishment

CF1305F Kuroni and the Punishment 允许对正整数做加一或减一，要求最后仍为正数且整体 gcd 大于一，求最少操作次数。

把所有数变成偶数，代价至多为 n，所以最优解的总代价不超过 n。如果超过一半的位置都花费至少两次，就会超过这个上界。因此至少一半的位置在某个最优解中只改了零次或一次。

设最优结果的公共质因子为 p。从这样的一个位置抽到原值 x，则 p 必然整除 $x-1, x, x+1$ 中某个正整数。随机选位置，分解这三个数，将所有质因子加入候选集即可。

CF1305F：枚举抽中位置附近的质因子

对候选质数 p，独立计算每个数到最近正倍数的距离。若 $a_i < p$，只能增加到 p；否则代价是 $\min(a_i \bmod p, p - a_i \bmod p)$。对所有候选取最小值。

每个候选都对应真实可行方案，不会得到比最优解更小的假答案。抽样 t 次漏掉固定最优质因子的概率至多 2^{-t}。将候选质数去重，设有 C 个；试除分解时总时间为 $O\left(t\sqrt{A+1} + nC\right)$，其中 $A = \max_i a_i$。

初始以全部变偶数的代价为当前答案；分解时跳过零、一。评估一个候选时，部分代价超过当前答案即可停止，因为剩下的代价均非负。

CF1986F：Non-academic Problem

CF1986F Non-academic Problem：给定连通无向简单图，删除恰好一条边，使仍然互相可达的无序点对数量最少。

$2 \le n \le 10^5$，$n-1 \le m \le 10^5$；全部测试的点数、边数之和分别不超过 2×10^5。

CF1986F：Non-academic Problem

CF1986F Non-academic Problem：给定连通无向简单图，删除恰好一条边，使仍然互相可达的无序点对数量最少。

$2 \le n \le 10^5$，$n-1 \le m \le 10^5$；全部测试的点数、边数之和分别不超过 2×10^5。

只有删除桥才会改变连通性。若一条桥把图分成大小为 s 与 $n-s$ 的两部分，可达点对减少 $s(n-s)$。于是需要找出桥，并计算对应一侧的大小。

CF1986F：用生成树描述每一个候选割

用并查集选出一棵生成树 T，任取根。非树边本身不可能是桥；对树边 $(\mathrm{parent}(v), v)$，令 S_v 为 v 的子树。

删去这条树边后，生成树分成 S_v 与其补集。原图仍然连通，当且仅当存在一条非树边跨过这个割。

没有跨割非树边

树边是桥

删除后，两侧失去连接

存在跨割非树边

树边在某个环上

非树边提供替代路径

因此，不必恢复每个环，只需检测每个子树割是否被非树边跨过。

CF1986F：端点异或，子树汇总

给每条非树边 $e=(a,b)$ 独立的均匀 64 位权值 w_e，分别异或到端点 a,b。从叶到根汇总，得到子树摘要 H_v。

$$H_v = \bigoplus_{e \text{ 恰有一个端点在 } S_v} w_e.$$

内部边出现两次抵消，外部边出现零次，恰好留下跨割的非树边。如果没有跨割边，$H_v=0$；若存在跨割边，固定其他权值后，剩下一个均匀权值使 $\Pr[H_v=0]=2^{-64}$。

端点异或已经自动完成路径上的覆盖与抵消，不需要查询 LCA。这里检测的是固定输入中预先确定的 $n-1$ 个割，可以直接取并集界。

CF1986F：计算答案与全程误差

同时汇总子树大小 s_v，在所有非根且 $H_v = 0$ 的节点中，取最大的 $s_{v(n-s_v)}$，记为 D；候选为空时取 $D = 0$。

$$\text{答案} = \binom{n}{2} - D.$$

真桥一定被保留；错误只可能来自非桥的摘要恰好为零。一组测试的错误概率至多 $(n-1)2^{-64}$，多组测试按全部候选割的数量累加。

生成树与汇总共用 $O(m\alpha(n)+n)$ 时间、$O(n+m)$ 空间。确定性替代是 DFS 的 lowlink；随机指纹在这里提供的是另一种简洁的图结构判定。

isaunoya 的原讨论 · 原提交

Graph Cuts

Graph Cuts 维护顶点集合 U，要求找到一条两端分别位于 U 内外的现存边，并删除它。

Graph Cuts

Graph Cuts 维护顶点集合 U，要求找到一条两端分别位于 U 内外的现存边，并删除它。

给边编号。对每个点，将关联边的编号异或起来；再对 U 内所有点的摘要做异或。内部边出现两次，完全抵消，留下的恰好是割边编号的异或和。

但多条割边异或后的值不一定是一条边。解决方法是建立多个抽样层，第 j 层以概率 2^{-j} 独立保留各条边，每层都维护同样的顶点摘要与集合摘要。

Graph Cuts：内部边抵消，跨割边留下摘要



$$H(U) = \mathrm{id}(AC) \oplus \mathrm{id}(BD).$$

摘要还不是答案：抽样到单边后，仍要验证编号、存活状态和两端是否跨割。

Graph Cuts：找到恰好留下单条割边的层

设初始边数为 $m \geq 1$，共设 $J = 1 + \lceil \log_2 m \rceil$ 层，第 j 层以概率 2^{-j} 保留各边，$0 \leq j < J$。

固定图与割，设有 $c > 0$ 条跨割边。$c = 1$ 时第零层必然保留唯一边；$c \geq 2$ 时取 $j = \lceil \log_2 c \rceil$，得到 $\frac{1}{2c} \leq p \leq \frac{1}{c}$。这一层恰好保留一条割边的概率满足

$$cp(1-p)^{c-1} \geq \frac{1}{2}\left(1 - \frac{1}{c}\right)^{c-1} \geq \frac{1}{2\mathrm{e}}.$$

对这个固定割，独立重复 t 组后仍未找到割边的概率至多

$$\left(1 - \frac{1}{2\mathrm{e}}\right)^{t}.$$

Graph Cuts：验证候选，并明确自适应边界

候选编号必须有效、对应边尚未删除、两端确实跨割。即使多条边的异或偶然得到另一条合法割边，也可直接返回；验证保证输出合法。

共维护 t 组、每组 J 层。点进出集合时更新集合摘要；删边时更新两端摘要及对应集合摘要。每次操作维护 tJ 份记录，时间 $O(tJ)$，顶点摘要空间 $O(tnJ)$。

漏检界只针对固定图与固定割。原题删除找到的边，后续图依赖此前的随机输出；即使点集操作预先固定，也不能直接沿用固定割的证明。

这里展示抽样摘要如何恢复割边，不将其作为原题的全程概率保证。确定性处理见 根号分治课件中的 Graph Cuts。

P8817：假期计划

CSP-S 2022 假期计划 要从家出发，依次访问四个不同景点后回家，每段路允许经过至多 $k+1$ 条边，最大化四个景点的权值和。

P8817：假期计划

CSP-S 2022 假期计划 要从家出发，依次访问四个不同景点后回家，每段路允许经过至多 $k+1$ 条边，最大化四个景点的权值和。

先从每个点 BFS，得到两点之间能否在限制内到达。随后随机排列除家以外的顶点，要求四个景点在这个排列中的位置严格递增。

设 $d_{t,v}$ 表示访问了 t 个景点、最后在 v 的最大分数。按排列顺序处理顶点，对每个可达且排在前面的 u，转移 $d_{t,v}=\max\left(d_{t,v},d_{t-1,u}+a_v\right)$。初始景点需要从家可达，第四个景点需要能返回家。递增顺序保证四个景点互不相同。

假期计划：一条路线有正反两种可用顺序

固定一条最优路线，其四个景点有 $4!$ 种相对顺序。正向和反向都可接受，所以单轮命中概率至少 $\frac{2}{4} \neq \frac{1}{12}$。重复 t 次、保留最大答案，失败概率至多 $\left(\frac{11}{12}\right)^t$。

预处理 $O(n(n+m))$，每轮 DP 为 $O(n^2)$。例如 100 次仍约有 1.66×10^{-4} 的漏解上界，应根据总测试量与时间预算选择次数，而不能仅凭一次通过判断可靠性。

接续：保留三个端点候选的确定性算法

假期计划：每个中间点保留三个端点

记两点距离不超过 $k+1$ 为可达，家为 s。对每个中间景点 b，收集满足 $a \neq s, b$、s 可达 a、a 可达 b 的端点 a，只保留权值最大的三个，记为 T_b。

枚举不同的中间景点 b, c，要求二者可达。枚举 $a \in T_b, d \in T_c$，检查 a, b, c, d 两两不同，再用四点权值和更新答案。

每对 b, c 至多检查九对端点。距离预处理为 $O(n(n+m))$，构造候选与枚举均为 $O(n^2)$，空间 $O(n^2)$，无需随机重试。

假期计划：三个候选为何足够

固定最优路线 s, a, b, c, d, s。在 b 的端点候选里，除了已经排除的 s, b，只需避开 c, d 两点。

若 a 不在前三名中，前三名至少有一个不等于 c, d，且权值不小于 a，可将它替换为 a'。随后在 c 的前三名中避开 b, a'，同样能把 d 替换为不更差的 d'。

替换后仍是合法路线，并且四个端点全部来自所保留的集合。因此某个最优方案一定会被枚举；候选不足三个时保留全部即可。

CF1310D：Tourist

CF1310D Tourist 寻找从点 1 出发又回到点 1 的恰好 k 步最小代价行走，并要求所用边不形成奇环。边权有方向，不能把两方向权值视为相同。

$n \le 80$，$2 \le k \le 10$，且 k 为偶数。

CF1310D：Tourist

CF1310D Tourist 寻找从点 1 出发又回到点 1 的恰好 k 步最小代价行走，并要求所用边不形成奇环。边权有方向，不能把两方向权值视为相同。

$n \le 80$，$2 \le k \le 10$，且 k 为偶数。

不含奇环的无向底图可以黑白染色。随机给各点二染色，只保留异色之间的有向边，然后做恰好 k 步的最短路 DP：

$$d_{t,v} = \min_{u:\mathrm{color}(u)\neq\mathrm{color}(v)} \left(d_{t-1,u} + w_{u,v}\right).$$

CF1310D：只需染对固定答案涉及的点

初始只有 $d_{0,1}=0$，答案是 $d_{k,1}$。每轮产生的行走都在二分图上，因此一定满足无奇环条件；并不需要额外要求点互不相同。

固定一个最优行走，它用到的不同顶点数 v 不超过 k，所用底图连通且二分。随机染色与它的两种合法染色之一吻合的概率为 $2^{1-v}\geq 2^{1-k}$。重复 t 次后的漏解概率至多 $\left(1-2^{1-k}\right)^t$。每轮 $O\left(k\times n^2\right)$，小 k 才能承受这个指数因子。

返回选讲题目

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

交互题选讲

前置知识	题目与内容
链表	CF843B：抽样定位，再沿链表搜索
二分、最大公因子	CF1114E：恢复等差数列
位运算	CF1364E：寻找零元素，还原排列
条件期望	CF2074E：自适应交互中的随机收缩
图的度数与割	Gym103469E：用随机割检验欧拉回路

先区分随机性影响的是答案还是询问次数，再把概率界与题目的询问上限对应起来。进阶：树根定位与在线考试

返回主讲安排

交互题进阶

前置知识	题目与内容
树上路径、概率尾界	CF1438F：由出现频率寻找树根
差分、条件期望	CF1705F：从总分中恢复每一道答案

前一题的固定预算下允许小概率选错候选；后一题对每次执行都保证正确和询问上界，随机性只改善平均工作量。

基础交互题 • 返回主讲安排

CF843B：Interactive LowerBound

CF843B Interactive LowerBound：有 n 个节点的严格递增单链表，下标次序与链表次序无关。已知头节点下标 s 和目标值 x。

询问一个下标，得到该节点的值与后继下标；尾节点的后继为 -1。求链表中不小于 x 的最小值，不存在则输出 -1。

$n \le 50000$，最多允许 1999 次节点询问。

CF843B：Interactive LowerBound

CF843B Interactive LowerBound：有 n 个节点的严格递增单链表，下标次序与链表次序无关。已知头节点下标 s 和目标值 x。

询问一个下标，得到该节点的值与后继下标；尾节点的后继为 -1。求链表中不小于 x 的最小值，不存在则输出 -1。

$n \le 50000$，最多允许 1999 次节点询问。

下标无序，不能按下标二分；从头遍历又可能耗尽预算。若已找到一个值小于 x、且离答案很近的节点，后续只需沿链表前进。

CF843B：随机样本确定搜索起点

先询问头节点；若它的值已不小于 x，直接结束。否则随机抽取若干不同下标，与头节点一起保留其中**值小于 x 且最大**的节点 u。

链表前段	**靠近答案的连续一段** $\ldots \to u \to \ldots$	首个 $\geq x$

从 u 沿后继走到第一个值不小于 x 的节点，或走到链尾。抽中的节点不必恰好是答案，只需落入答案前的一段。

缓存已问节点；一次询问同时获得值和后继。没有答案时，同样希望抽中链尾附近的节点。

CF843B：命中末段的概率

$n \le 1999$ 时可以问完所有节点。否则分配 $1 + 999 + 999$ 次询问：头节点、其余下标中的 999 个均匀无放回样本、至多 999 次后继访问。

若小于 x 的节点至多 999 个，从头出发已足够；否则，把其中最后 999 个节点记作 W。命中 W 后必能在剩余预算内结束。

$$\Pr[\text{未命中 } W] \le \left(1 - \frac{999}{n-1}\right)^{999} \le \exp\left(-\frac{999^2}{n-1}\right) < 2.2 \times 10^{-9}.$$

这个界来自**好下标的数量**。预算耗尽仍未定位时，算法可能答错，但不能继续询问；此时只能提交已见候选或 -1。

CF1114E：Arithmetic Progression

CF1114E Arithmetic Progression：一个长度为 n 的无序数组，排序后是公差 $d > 0$ 的整数等差数列。求首项与公差。

可以询问某个下标的值；也可以给定 x，询问数组中是否存在严格大于 x 的数。

$2 \le n \le 10^6$，所有值在 $[0, 10^9]$ 内，两类询问合计至多 60 次。

CF1114E：Arithmetic Progression

CF1114E Arithmetic Progression：一个长度为 n 的无序数组，排序后是公差 $d > 0$ 的整数等差数列。求首项与公差。

可以询问某个下标的值；也可以给定 x，询问数组中是否存在严格大于 x 的数。

$2 \le n \le 10^6$，所有值在 $[0, 10^9]$ 内，两类询问合计至多 60 次。

比较询问可以二分最大值。最大值确定后，其余信息都可以转化成“相差多少个公差”。

CF1114E：对抽样差值取最大公因子

用至多 30 次比较询问求出最大值 M。再均匀无放回抽取 $t=\min(n,30)$ 个下标，得到 $y_1,\dots,y_t$。

每个样本满足 $M-y_i=k_i d$，其中 $k_i\in\{0,1,\dots,n-1\}$。因此

$$g=\gcd(M-y_1,\dots,M-y_t)=d\times\gcd(k_1,\dots,k_t).$$

若 $\gcd(k_1,\dots,k_t)=1$，就得到 $d=g$，首项为 $M-(n-1)g$。

需要的是样本对应的步数互质；不必碰巧抽到两个相邻项。$n\le 30$ 时查询全部元素，结果确定正确。

CF1114E：错误对应一个共同质因子

当 $n > 30$ 时，抽到的步数是 $0, \ldots, n-1$ 中均匀选择的 30 个不同整数。

若得到的 g 大于真正公差，必有某个质数 $p \le n-1$ 整除全部步数。固定 p，这样的整数只有 $\left\lceil \frac{n}{p} \right\rceil$ 个，故

$$\Pr[g \neq d] \le \sum_{p<n, p \text{ 为质数}} \left(\frac{\left\lceil \frac{n}{p} \right\rceil}{n} \right)^{30}.$$

对所有可能的共同质因子求并集界，可得全题范围内错误概率小于 3.5×10^{-8}。询问数始终不超过 60，随机性只影响参数是否恢复正确。

附录：统一误差界

CF1364E：X-OR

CF1364E X-OR：隐藏一个 $0, \ldots, n-1$ 的排列 p。询问两个**不同**下标 i, j，得到 p_i 与 p_j 的按位或。

还原整个排列。$3 \leq n \leq 2048$，最多允许 4269 次询问。排列在交互开始前固定。

CF1364E：X-OR

CF1364E X-OR：隐藏一个 $0, \ldots, n-1$ 的排列 p。询问两个**不同**下标 i, j，得到 p_i 与 p_j 的按位或。

还原整个排列。$3 \le n \le 2048$，最多允许 4269 次询问。排列在交互开始前固定。

若找到 $p_z = 0$ 的位置，询问 (z, i) 就直接得到 p_i。任务可以拆成：先找零，再还原其余元素。

CF1364E：从二进制一较少的数出发

记 $q(i,j)=p_i \text{ OR } p_j$。独立均匀选择不同下标 a,b，直到返回值至多含六个二进制一；取 $u=a$。于是 p_u 也至多含六个一。

询问 u 与所有其他位置，并对这些返回值取按位与，得到 p_u 的准确值 m。

$$m=\bigwedge_{i\neq u} q(u,i)=p_u.$$

每个返回值都含有 p_u 的全部一位；若 $p_u\neq 0$，与零元素的那次询问恰好返回 p_u。若 $p_u=0$，排列中同时存在 $1,2$，其按位与为零，结论仍成立。

保存抽样成功时的那次回答，这一整行还需询问 $n-2$ 次。

CF1364E：候选恰好是全部真子掩码

保留 $i \neq u$ 且 $q(u,i) = m$ 的位置，记作 S。这个条件等价于 p_i 是 m 的真子掩码。

任一子掩码都不超过 $m < n$，所以一定出现在排列中。候选因此包含**全部**真子掩码，而不只是其中一部分。

若 m 含有 k 个一，则

$$|S| = 2^k - 1 \leq 63.$$

$m \neq 0$ 时，零元素始终在 S 中。后续只需在这个很小且结构完整的集合里继续询问。

CF1364E：在子掩码集合中继续下降

当 $|S| \geq 2$，从 S 中均匀选新枢轴 v，询问它与其他候选的按位或。以旧值 m 为初值，对回答取按位与，得到新值 $m' = p_v$。

若 $p_v \neq 0$，集合中的零元素保证读数准确。若 $p_v = 0$，旧值至少含两个一；对应的两个单比特子掩码都在 S 中，也能保证按位与为零。

随后仅保留 $q(v, i) = m'$ 的其他位置。新集合又恰好是 m' 的全部真子掩码。

m' 是 m 的真子掩码，一的数量严格减少。因此至多进行五轮询问下降，随后直接判定零位置；候选数依次受到 $63, 31, 15, 7, 3, 1$ 的限制。

CF1364E：终止与排列还原

候选为空

$$m = 0$$

当前枢轴就是零元素

候选只有一个

$$S = \{z\}$$

唯一真子掩码只能是零

每次枢轴的准确值都已求出，可以保留。找到零位置 z 后，仅对尚未确定的位置询问 (z, i)，即可补全排列。

随机选哪个候选只影响下降路径；只要询问过程完成，恢复出的值一定正确。初始抽样可能反复失败，必须另行分析总询问预算。

CF1364E：询问总数的分解

记初始抽样次数为 R，下降阶段的询问数为 S。首个枢轴之外，记录了 d 个新枢轴或终止零点，同一位置只计一次。

阶段	询问次数
随机找到合格起点	R
补完首个枢轴的整行	$n-2$
在真子掩码中下降	S
用零元素还原未知位置	$n-1-d$

首行已有一次成功抽样的回答；最后已有 $1+d$ 个值确定，因此分别只需 $n-2$ 次与 $n-1-d$ 次。

$$Q = R + (n-2) + S + (n-1-d) = R + 2n - 3 + (S-d).$$

CF1364E：下降阶段的净开销

当前枢轴有 $k \geq 2$ 个一时，候选有 $2^k - 1$ 个。选定新枢轴后，询问其余候选，需要 $2^k - 2$ 次；同时确定了一个新值，使最后还原少问一次。因此本轮对 $S - d$ 的贡献是

$$\left(2^k - 2\right) - 1 = 2^k - 3.$$

一的数量严格下降，初始又至多为六，所以 $k = 6, 5, 4, 3, 2$ 每种至多出现一次。若终止时从唯一候选直接确定零，还会再省一次。于是

$$S - d \leq \sum_{k=2}^{6} \left(2^k - 3\right) = 61 + 29 + 13 + 5 + 1 = 109.$$

代回总数分解，得到

$$Q \leq R + 2n - 3 + 109 = R + 2n + 106.$$

CF1364E：随机起点与询问预算

前面得到 $Q \le R + 2n + 106$，但初始抽样次数 R 没有确定上界，因此仍需分析总数超过 4269 的概率。结合起点位数与后续下降路径的分布，可对全部 $3 \le n \le 2048$ 算出

$$\Pr[Q > 4269] < 4.8 \times 10^{-7}.$$

分析采用各轮独立均匀抽样。统一累计询问次数，已问满 4269 次且仍需询问时停止；这将超限事件转化为相同概率上界内的失败，而不会发出多余询问。

附录：精确计数与有限范围概率核验

CF2074E：Empty Triangle

CF2074E Empty Triangle：平面上有 n 个坐标未知的固定点，任意三点不共线。询问三个下标，交互器返回三角形内部的一个点；若内部为空则返回零。

交互器可以根据此前询问选择返回哪个内部点，但不能改变点的位置。求一个内部没有其他点的三角形。

$3 \le n \le 1500$，每组至多 75 次询问，一次测试至多 20 组。

CF2074E：Empty Triangle

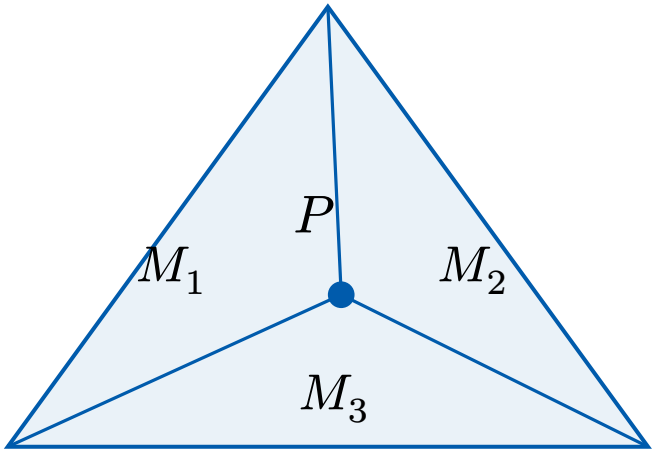
CF2074E Empty Triangle：平面上有 n 个坐标未知的固定点，任意三点不共线。询问三个下标，交互器返回三角形内部的一个点；若内部为空则返回零。

交互器可以根据此前询问选择返回哪个内部点，但不能改变点的位置。求一个内部没有其他点的三角形。

$3 \le n \le 1500$，每组至多 75 次询问，一次测试至多 20 组。

返回一个内部点后，原三角形被分成三个子三角形。三个都可以成为下一步候选，问题在于如何选择。

CF2074E：均匀保留一个子三角形



设原三角形内有 M 个点。交互器返回 P 后，三个子三角形的内部点数满足

$$M_1 + M_2 + M_3 = M - 1.$$

用新的随机数等概率选择其中一个。于是下一步的内部点数期望为 $\frac{M-1}{3}$。

不需要知道坐标或数出各部分的点，也不需要交互器返回一个“居中”的点。

CF2074E：对任意交互历史都成立的收缩

令 M_t 为第 t 次收缩后的内部点数，空三角形视为保持在零。无论交互器如何选择返回点，都有

$$E[M_{t+1} \mid \text{此前历史与返回点}] \le \frac{M_t}{3}.$$

反复取期望，得到 $E[M_t] \le \frac{n-3}{3^t}$。内部非空意味着 $M_t \ge 1$，所以

$$\Pr[M_{75} \ge 1] \le E[M_{75}] \le \frac{1497}{3^{75}} < 2.5 \times 10^{-33}.$$

第 75 次询问若仍返回内部点，再随机保留一个子三角形并直接输出，不发第 76 次询问。至多 20 组时，对各组失败事件再求并集界。

Gym103469E：Eulerian?

Gym103469E Eulerian?：隐藏一张连通无向简单图。询问一个顶点集合，交互器返回该集合内部的边数。

判断图是否存在欧拉回路。交互器不自适应，最多允许 60 次询问。

Gym103469E：Eulerian?

Gym103469E Eulerian?：隐藏一张连通无向简单图。询问一个顶点集合，交互器返回该集合内部的边数。

判断图是否存在欧拉回路。交互器不自适应，最多允许 60 次询问。

连通性已经保证，问题等价于判断是否所有点的度数都为偶数。询问看不到单个顶点的度数，但可以组合成一个割的边数。

Gym103469E：随机割检测奇度点

先问全图边数 m。每轮将各顶点独立等概率分到 A, B，分别询问内部边数 $e(A), e(B)$，得到割边数

$$c(A,B) = m - e(A) - e(B).$$

每条 A 内部的边在度数和中出现两次，跨割边只出现一次，因此

$$c(A,B) \equiv \sum_{v \in A} \deg(v) \bmod 2.$$

所有度数为偶数时，每个割都为偶数。若存在奇度点，固定其他点的分组后，让这个点进出 A 会翻转割的奇偶性；所以一轮恰有 $\frac{1}{2}$ 概率发现奇数割。

Gym103469E：单侧错误与询问预算

发现奇数割就确定不存在欧拉回路；未发现只能给出概率保证。

一次全图询问，加上 29 轮独立随机二分割，共用 $1+2\times 29=59$ 次询问。若存在奇度点，全部漏检的概率为 2^{-29}。

这里的输入固定，才能反复检测同一个度数奇偶向量。每轮重新独立分组，固定其他顶点后的翻转论证也始终成立。

交互时每次输出后刷新缓冲区，收到错误标记立即终止；这些约定也适用于前面的例题。

返回交互题选讲

CF1438F：Olha and Igor

CF1438F Olha and Igor：一棵高度为 h 的完美二叉树有 $n = 2^h - 1$ 个点，$3 \le h \le 18$，标签固定。询问三个不同点 u, v, w，得到以 w 为根时 u, v 的 LCA。

至多询问 $n + 420$ 次，求原树根的标签。

CF1438F：Olha and Igor

CF1438F Olha and Igor：一棵高度为 h 的完美二叉树有 $n = 2^h - 1$ 个点，$3 \le h \le 18$，标签固定。询问三个不同点 u, v, w，得到以 w 为根时 u, v 的 LCA。

至多询问 $n + 420$ 次，求原树根的标签。

返回值是三点的**中位点**：三条两两路径共同经过的唯一点。因此三个参数的顺序不影响回答。

原根的度数为二，其他非叶节点的度数为三。随机询问更常返回的点，并不一定是原根。

CF1438F：中位点的出现频率

删去点 x 后，各连通分支的大小为 s_1, s_2, s_3；若 x 是原根，令 $s_3 = 0$。均匀选三个不同点，使中位点为 x 的无序三元组数为

$$F(x) = s_1 s_2 s_3 + s_1 s_2 + s_2 s_3 + s_3 s_1.$$

不选 x 时，三个点必须来自三个不同分支；选了 x 时，另外两点必须来自两个不同分支。故返回 x 的概率为 $\frac{F(x)}{\binom{n}{3}}$。

原根只有两个分支，必须自身被抽中才可能成为返回值。同一深度的点，其出现概率相同。

CF1438F：根的两个孩子

原根深度为零。深度 d 的点，两棵向下子树大小均为 a_d，向上分支大小为 b_d：

$$a_d = 2^{h-d-1} - 1, \quad b_d = n - 2a_d - 1,$$

$$F_d = a_d^2 b_d + a_d^2 + 2a_d b_d, \quad p_d = \frac{F_d}{\binom{n}{3}}.$$

对 $d \geq 1$，F_d 随深度严格减小。根的孩子与原根的计数差为

$$F_1 - F_0 = \frac{(n+1)^2(n-5)}{32} > 0.$$

所以根的两个孩子是唯一的最大概率点。深度比较的代数证明

CF1438F：固定预算的定位

当 $h \geq 4$，独立均匀询问 420 个三元组；每轮内部的三个点不同，各轮允许重复。统计返回次数，线性扫描选出次数最多的两个点 c_1, c_2。

再对其余每个点 i 询问 (c_1, c_2, i)。**若候选确为原根的两个孩子**，仅当 i 是原根时回答才等于 i；其他点都返回 c_1 或 c_2。

总询问数至多 $420 + (n-2) = n + 418$。候选错误时不保证定位，但仍在扫描后结束，不能追加无上限的随机重试。

$h = 3$ 时只有七个点。枚举全部 $\binom{7}{3} = 35$ 个三元组，两个孩子各出现 13 次、原根出现 9 次；再用至多五次询问定位，总共至多 40 次，确定正确。

CF1438F：出现次数之差

固定一个原根孩子与一个非孩子节点，单轮返回概率分别为 p, q，且 $p > q$。记它们在 T 轮中的出现次数为 X, Y。

一次采样对 $Y - X$ 的增量分别为 $-1, +1, 0$，概率为 $p, q, 1-p-q$。因此对任意 $z > 1$，Markov 不等式给出

$$\Pr[Y \geq X] \leq E\left[z^{Y-X}\right] = \left(1 - p - q + \frac{p}{z} + qz\right)^T.$$

取 $z = \sqrt{\frac{p}{q}}$，得到

$$\Pr[Y \geq X] \leq \left(1 - p - q + 2\sqrt{pq}\right)^T.$$

$q = 0$ 时右侧为 $(1-p)^T$，就是孩子一次也未出现的概率。这里比较同一次采样中的计数差，没有假设两点频数独立。

CF1438F：全程错误概率

只要所有非孩子节点的次数都严格小于两个孩子，选出的候选就一定正确。深度 d 有 2^d 个点；对两孩子与全部竞争点取并集界：

$$\Pr[\text{选错候选}] \le 2 \sum_{0 \le d < h,\, d \ne 1} 2^d \left(1 - p_1 - p_d + 2\sqrt{p_1 p_d}\right)^{420}.$$

将 $4 \le h \le 18$ 的十五种允许高度分别代入，按深度合并，最大上界出现在 $h = 18$，小于

$$4.902 \times 10^{-5}.$$

这个上界同时计入了并列次数可能造成的错误。$h = 3$ 使用上一页的确定性处理。全程询问预算始终成立，概率只影响答案是否正确。

官方题解：随机三元组与候选孩子

CF1705F：Mark and the Online Exam

CF1705F Mark and the Online Exam：有 $n \le 1000$ 道判断题，每次提交一个长为 n 的 T/F 串，得到答对的题数。答案预先固定，要求在至多 675 次查询内提交全对的答案；一旦返回 n，立即结束。

CF1705F：Mark and the Online Exam

CF1705F Mark and the Online Exam：有 $n \le 1000$ 道判断题，每次提交一个长为 n 的 T/F 串，得到答对的题数。答案预先固定，要求在至多 675 次查询内提交全对的答案；一旦返回 n，立即结束。

取独立均匀随机串 r。将位置分成大小相差至多一的两组 U, V，再令 b 只翻转 r 在 V 中的位置。

查询这两份固定基准，记分数为 S_0, S_1。以后只翻转其中少量位置，再与对应基准分数作差；其他位置的贡献全部抵消。

翻转一位，分数变化只能为 $+1$ 或 -1，分别说明这一位原来答错或答对。

CF1705F：用两种基准区分三个答案

从 U, V 的未解位置中各取一位，先在 r 中翻转它们，得到分数差 δ_0。

若 $\delta_0 = 2$，两位相对 r 都错；若 $\delta_0 = -2$，两位相对 r 都对。一次询问已经确定两个答案。

若 $\delta_0 = 0$，它们相对 r 一对一错，相对 b 则同对或同错。再取第三个未解位置，在 b 中同时翻转三位，得到

$$\delta_1 = 2\alpha + \beta \in \{-3, -1, 1, 3\}, \quad \alpha, \beta \in \{-1, 1\}.$$

α 是配对中每位的翻转贡献，β 是第三位的贡献。四种结果各不相同，两次询问便能确定三个答案。

CF1705F：均衡分组与硬询问上界

每轮先从两组各取一位；需要第三位时，从剩余位置较多的一组取。两组未解数之差始终不超过一。

剩下至多两位时，每组至多一位。用两次基准总分扣除已解位置的贡献，就能直接恢复余位。恢复公式

两类步骤的查询数分别为确定位置数的 $\frac{1}{2}$ 与 $\frac{2}{3}$，局部步骤合计至多 $\left\lfloor \frac{2n}{3} \right\rfloor$ 次。再计入两次基准和最后一次完整答案提交：

$$Q \le 2 + \left\lfloor \frac{2n}{3} \right\rfloor + 1 \le 669 < 675.$$

该界对每个随机串都成立，超预算概率为零。任何查询返回 n，立即停止。

CF1705F：随机化减少平均询问数

固定答案以后，随机基准 r 各位置的匹配状态是独立公平比特。位置的选择只依赖此前的局部差分与两组大小，不利用基准总分选择未解位置。

因此下一对同对或同错的概率为 $\frac{1}{2}$，用一问解两位；一对一错的概率也为 $\frac{1}{2}$，用两问解三位。

忽略提前终止，尚有 m 位时的局部查询期望 G_m 满足（$m \geq 3$）：

$$G_m = \frac{1 + G_{m-2}}{2} + \frac{2 + G_{m-3}}{2}, \quad G_0 = G_1 = G_2 = 0.$$

由此得到 $G_m = \frac{3m}{5} + O(1)$，完整流程满足 $E[Q] \leq \frac{3n}{5} + 3$。随机化改善平均工作量，正确性和 669 次硬界均不依赖运气。

期望递推的独立性与常数项 · 官方题解

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

顺序与搜索选讲

方法	阅读入口
随机顺序	洗牌后的轨迹集中与带宽 DP
随机预处理	CF1559D2：随机合并与确定性收尾
随机局部修复	2-SAT：用势函数分析到达解的时间

分别讨论随机性改变状态分布、压缩剩余规模和选择搜索方向的情形。每种方法都要把成功条件与总工作量写出来。

返回主讲安排

随机顺序的适用范围

本节需要熟悉按物品推进的 DP，以及无放回抽样的集中界。先固定一个最优方案，再分析洗牌后它的前缀轨迹。

必须分别检查三件事：重排是否改变问题、所选带宽是否满足误差目标、保留状态的总代价是否可接受。

选取物品模型 证明截断原理；后面的原题用于比较不同资源维度。现有保守界尚不能同时保证这些原题的时间限制，属于方法与适用边界讨论。

需要完整确定性解法时，可直接阅读 Cardboard Box 的分类算法 与 Brukhovich and Exams 的贪心。返回主讲安排

选取物品与 DP 状态数

选取物品模型。 一个按物品逐个决策的 DP，用 j 表示已选数量，最后要求恰好选 K 个。共有 N 个物品，顺序不影响可行性和目标值。希望减少每层保留的 j 状态数。

选取物品与 DP 状态数

选取物品模型。 一个按物品逐个决策的 DP，用 j 表示已选数量，最后要求恰好选 K 个。共有 N 个物品，顺序不影响可行性和目标值。希望减少每层保留的 j 状态数。

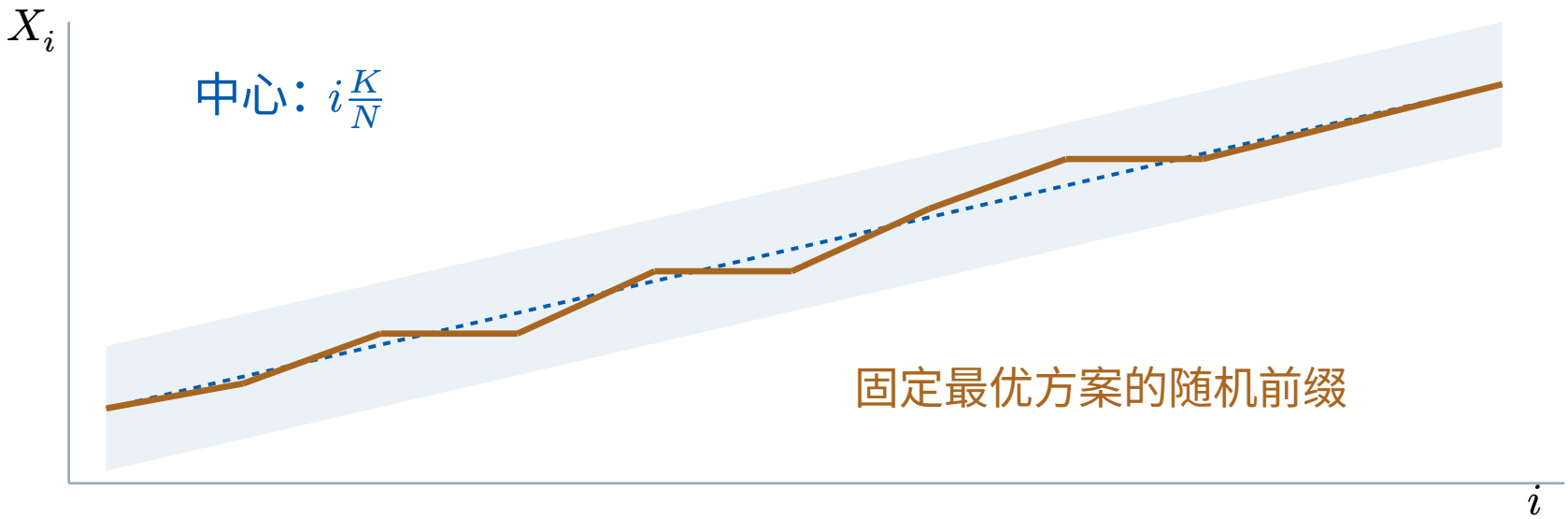
固定任意一个最优解，再均匀打乱全部物品。

这个最优解的 K 个物品仍是同一批物品，只是位置变成均匀随机的 K 元子集。前 i 个位置中它选了 X_i 个，每个物品落入前缀的概率为 $\frac{i}{N}$，故

$$E[X_i] = i\frac{K}{N}.$$

只保留 $\left|j - i\frac{K}{N}\right| \le W$ 的状态。若固定最优解的全部前缀都在带内，受限 DP 就仍含有一个最优解；其余最优解是否被删掉并不影响结论。

洗牌后，最优方案的轨迹集中在均值附近



带宽两侧各留 W 个状态。只有这条轨迹越出阴影，才可能丢掉所固定的最优方案；需要控制的是所有前缀的联合事件。

带宽 DP：对所有前缀一起控制漏解概率

对固定的前缀长度 $i \ge 1$，X_i 是无放回抽样计数。Hoeffding 界给出

$$\Pr\left[\left|X_i - i\frac{K}{N}\right| > W\right] \le 2\exp\left(-2\frac{W^2}{i}\right) \le 2\exp\left(-2\frac{W^2}{N}\right).$$

对至多 N 个前缀取并集界，不要求它们彼此独立，得到

$$\Pr\left[\exists i : \left|X_i - i \times \frac{K}{N}\right| > W\right] \le 2N\exp\left(-2\frac{W^2}{N}\right).$$

若要求右侧不超过 ε，取 $W \ge \sqrt{\left(\frac{N}{2}\right)\ln\left(2\frac{N}{\varepsilon}\right)}$ 即足够。实际状态还要与 $0 \le j \le \min(i, K)$ 相交；更窄带宽的误差需要另行分析。

CF436E：Cardboard Box

CF436E Cardboard Box 有 n 个关卡，不玩、花 a_i 得一星、花 b_i 得两星。求获得至少 w 星的最小代价，并输出每关的星数方案。

$n \le 3 \times 10^5$，$1 \le w \le 2n$，$1 \le a_i < b_i \le 10^9$。

CF436E：Cardboard Box

CF436E Cardboard Box 有 n 个关卡，不玩、花 a_i 得一星、花 b_i 得两星。求获得至少 w 星的最小代价，并输出每关的星数方案。

$n \le 3 \times 10^5$，$1 \le w \le 2n$，$1 \le a_i < b_i \le 10^9$。

若超过 w 星，可将某关两星降为一星、或一星降为零星。由 $a_i < b_i$，每次降星都不增加代价，因此只需求恰好 w 星。

打乱关卡后，用 $d_{i,j}$ 表示前 i 关获得 j 星的最小代价：

$$d_{i,j} = \min\left(d_{i-1,j}, d_{i-1,j-1} + a_i, d_{i-1,j-2} + b_i\right).$$

CF436E：单项范围与尾界常数

固定最优方案，每关的星数是 $0,1,2$ 中的某个固定值。随机排列后，前 i 关的星数和 X_i 满足 $E[X_i]=i\frac{w}{n}$。

将星数除以二，单项就落入 $[0,1]$，原来的偏差 W 对应偏差 $\frac{W}{2}$。代入前面的无放回抽样界：

$$\Pr\left[\left|X_i-i\frac{w}{n}\right|>W\right]\le 2\exp\left(-2\frac{\left(\frac{W}{2}\right)^2}{i}\right)=2\exp\left(-\frac{W^2}{2i}\right).$$

再对至多 n 个前缀取并集界，并用 $i\le n$，得到总漏解概率至多

$$2n\exp\left(-\frac{W^2}{2n}\right).$$

CF436E：误差预算与状态规模

每层保留带宽 W，时间 $O(nW)$，滚动数组空间 $O(W)$。直接保存全部状态来源以输出方案，则需 $O(nW)$ 空间，也可用分治重算恢复。

这里单项范围为 $[0,2]$，所以误差界中的指数常数与零一计数不同，不能直接照抄原来的带宽。

在 $n \le 3 \times 10^5$ 的范围内，还需把满足误差目标的 W 代回 nW，并核算方案恢复成本。这个分析本身没有证明随机截断方案能够通过原题。

接续：按边际代价分类的确定性算法

CF436E：凸关卡与凹关卡

每关从零星到一星、再到两星的边际代价分别为 $a_i, b_i - a_i$。若 $b_i \geq 2a_i$，后一次不便宜，称为凸关卡；否则称为凹关卡。

对凸关卡，把全部边际代价排序，同价时先取第一星。因为第一星不贵于第二星，取排序后的任意前缀都满足先后约束。记前 t 项之和为 C_t，就是凸关卡取得 t 星的最小代价。

凹关卡在最优方案中至多有一个只取一星。若两个凹关卡都只取一星，将它们改为一个两星、另一个零星，代价至多 $\min\left(b_i, b_j\right) < a_i + a_j$。因此只需考虑完整的两星关卡和至多一个单星关卡。

CF436E：凹关卡的奇偶星数

将 h 个凹关卡按 b_i 递增排序，记 $B_k = \sum_{i=1}^{k} b_i$。取得 $2k$ 星时，直接选最便宜的 k 个完整关卡，代价 $J_{2k} = B_k$。

取得 $2k+1$ 星时，枚举单星关卡的位置。有两种情况：它在前 k 个完整关卡之外；或它原本在其中，需要补入第 $k+1$ 个完整关卡。因此

$$J_{2k+1} = \min\left(B_k + \min_{i>k} a_i, B_{k+1} + \min_{i\le k}(a_i - b_i) \right).$$

空集合的最小值记为正无穷。预处理 a_i 的后缀最小值与 $a_i - b_i$ 的前缀最小值，并记录取得最小值的原编号，便可 $O(1)$ 求每个 J_r。

CF436E：合并代价并恢复方案

枚举凸关卡取得的星数 t，只保留凹关卡能承担的 $0 \le w - t \le 2h$，答案为

$$\min_{t}(C_t + J_{w-t}).$$

恢复时，凸关卡取边际代价排序中的前 t 项，每出现一次原编号就为该关增加一星。凹关卡取前 k 个完整关卡；奇数星数则按达到最小值的分支选择单星关卡，必要时以第 $k+1$ 个完整关卡替换它。

两次排序之外均为线性扫描，总时间 $O(n \log n)$、空间 $O(n)$。代价前缀和与无穷哨兵均使用 64 位整数。最终按原编号输出每关的零、一、二星方案。

返回随机顺序

CF739E：Gosha is hunting

Gosha is hunting 有 a 个普通球和 b 个高级球。对第 i 个目标，两种球的独立捕获概率分别为 p_i, u_i；每种球对同一目标至多用一次，最大化捕获总数的期望。

正式范围 $n \le 2000$、$0 \le a, b \le n$、$p_i, u_i \in [0, 1]$。

CF739E：Gosha is hunting

Gosha is hunting 有 a 个普通球和 b 个高级球。对第 i 个目标，两种球的独立捕获概率分别为 p_i, u_i；每种球对同一目标至多用一次，最大化捕获总数的期望。

正式范围 $n \le 2000$、$0 \le a, b \le n$、$p_i, u_i \in [0, 1]$。

由期望的线性性，每个目标有四种选择：不用球贡献零，只用第一种贡献 p_i，只用第二种贡献 u_i，两种都用贡献 $p_i + u_i - p_i \times u_i$。最后一项来自“至少有一次成功”的概率，不能直接把两次成功相加。

CF739E：两种资源围绕各自均值截断

令 $d_{i,j,k}$ 表示前 i 个目标使用两种球各 j,k 个时的最大期望。每层从 (j,k) 向 (j,k)、$(j+1,k)$、$(j,k+1)$、$(j+1,k+1)$ 转移，加上对应贡献。球用得更多不会损害答案，故可以补足到恰好 a,b 个，最终取 $d_{n,a,b}$。

先将目标整体随机排列，再只保留

$$\left|j-i\times\frac{a}{n}\right|\le W_1,\quad \left|k-i\times\frac{b}{n}\right|\le W_2.$$

CF739E：分别估计概率，再取并集界

固定一个用满资源的最优方案，两种球各自的使用位置都是固定子集，洗牌后分别满足前缀集中界。两者不需要相互独立：分别估计偏离概率，再取并集界即可。总漏解概率至多

$$2n\exp\left(-2\frac{W_1^2}{n}\right)+2n\exp\left(-2\frac{W_2^2}{n}\right).$$

每层至多 $O(W_1\times W_2)$ 个状态，时间 $O(nW_1\times W_2)$，滚动数组空间 $O(W_1\times W_2)$。每层清空新数组，不可达状态置负无穷，两个坐标都要与实际可用数量相交。$a=0$ 或 $b=0$ 时退化为一维。输出允许 10^{-4} 的绝对或相对误差，这是浮点精度要求，与随机截断漏解概率不同。

CF739E：误差目标与状态规模

把两项失败概率各控制在 $\frac{\varepsilon}{2}$ 内，可取

$$W_1 = W_2 = \left\lceil \sqrt{\left(\frac{n}{2}\right)\ln\left(4\frac{n}{\varepsilon}\right)} \right\rceil .$$

代回时间界 $O(nW_1W_2)$，得到 $O\left(n^2 \ln\left(\frac{n}{\varepsilon}\right)\right)$ 级别的状态处理量，实际还需与资源上限裁剪。它给出了严格误差保证，但没有证明满足原题时限。

因此本节将它用于说明二维截断怎样分配错误预算。采用更窄的经验带宽时，不能继续沿用此处的误差结论。

返回随机顺序的适用范围

CF1799F：Halve or Subtract

Halve or Subtract 对每个数至多执行一次向上取整减半、一次减去 b 并截到零；两种操作的总次数分别不超过 k_1, k_2，最小化最后的元素和。

$n \le 5000$，所有测试的 $\sum n \le 5000$。

CF1799F：Halve or Subtract

Halve or Subtract 对每个数至多执行一次向上取整减半、一次减去 b 并截到零；两种操作的总次数分别不超过 k_1, k_2，最小化最后的元素和。

$n \le 5000$，所有测试的 $\sum n \le 5000$。

同一个元素有四种资源消耗，留下的最小值分别为

$$c_{0,0}(x) = x, \quad c_{1,0}(x) = \left\lceil \frac{x}{2} \right\rceil,$$

$$c_{0,1}(x) = \max(x - b, 0),$$

$$c_{1,1}(x) = \max\left(\left\lceil \frac{x}{2} \right\rceil - b, 0 \right).$$

CF1799F：减半先于减法

两种操作都使用时，先减半不会更差。因为 b 为非负整数，

$$\max\left(\left\lceil\frac{x}{2}\right\rceil-b,0\right)\le\left\lceil\frac{\max(x-b,0)}{2}\right\rceil.$$

左边等于 $\max\left(\left\lceil\frac{x-2b}{2}\right\rceil,0\right)$，右边等于 $\max\left(\left\lceil\frac{x-b}{2}\right\rceil,0\right)$，由单调性即可比较。因此无需保留两种操作次序。

CF1799F：四种局部选择的转移

令 $d_{i,j,k}$ 为处理前 i 个数、使用两种操作各 j,k 次的最小剩余和。四种转移统一写成

$$d_{i,j,k}=\min_{u,v\in\{0,1\}}\left(d_{i-1,j-u,k-v}+c_{u,v}(a_i)\right).$$

初始只有 $d_{0,0,0}=0$，其余状态为正无穷。每次使用旧层向清空的新层转移，避免同一个元素被重复使用；负下标、超过操作总次数的状态不参与转移。还可删去即使用满剩余元素也不能补足操作次数的状态。完整 DP 的时间为 $O(nk_1k_2)$、空间 $O(k_1k_2)$；限制带宽时，只改变每层遍历的合法范围，转移与初始化不变。

CF1799F：四种选择进入同一个二维 DP

所有操作都不会增大数值，可以补足使用次数到 k_1, k_2。打乱整个元素序列后，围绕 $\left(i \times \frac{k_1}{n}, i \times \frac{k_2}{n}\right)$ 保留二维带，概率与复杂度分析同上一题。终态为 (k_1, k_2)，无穷哨兵与代价和使用 64 位整数。

若两维都取带宽 W，时间为 $O(nW^2)$。按保守并集界取

$$W = \Theta\left(\sqrt{n \log\left(\frac{n}{\varepsilon}\right)}\right)$$

时，上界达到 $O\left(n^2 \log\left(\frac{n}{\varepsilon}\right)\right)$；这并未给出原题 1 秒限制下的可用预算。缩窄带宽会改变漏解概率，必须重新证明或明确作为经验参数。

CF1891E：Brukhovich and Exams

Brukhovich and Exams：将至多 k 个位置改为零，最小化相邻 gcd 等于一的边数。

CF1891E：Brukhovich and Exams

Brukhovich and Exams：将至多 k 个位置改为零，最小化相邻 gcd 等于一的边数。

记 $v_i(0) = a_i, v_i(1) = 0$，状态 $d_{i,j,t}$ 表示前 i 项选了 j 个清零位置、末项状态为 t 时的最小代价。

$$d_{i,j,t} = \min_{s\in\{0,1\}} \left(d_{i-1,j-t,s} + [\gcd(v_{i-1}(s), v_i(t)) = 1]\right).$$

首项两种状态代价均为零，分别消耗零次、一次操作；最终对 $0 \le j \le k$ 和末项状态取最小。完整 DP 为 $O(nk)$ 时间、$O(k)$ 空间。

这里相邻关系属于题目条件，打乱位置会改变问题。围绕 $i\frac{k}{n}$ 截断状态，需要另一份依据。

CF1891E：最优轨迹可以偏离均值线

令 $n = 8k$，前 $2k+1$ 项为 $2, 3, 2, 3, \ldots, 2$，其余全为二。前段有 $2k$ 条坏边，后段没有。

清零前段的全部 k 个三，坏边全部消失。要用 k 次消掉 $2k$ 条边，每次必须消掉两条且不重复，因此这也是唯一的零代价选择。

在前缀 $i = 2k+1$ 处，它已用完 k 次操作，而平均线只有 $\frac{2k+1}{8}$。两者偏差为

$$k - \frac{2k+1}{8} = \frac{6k-1}{8}.$$

小于这个偏差的带宽会删掉所有最优解。随机顺序的集中界不能用于固定顺序；本题可改用 后续的确定性贪心。

CF1891E：非一值坏边链的收益

利用相邻边的结构，可以得到确定性贪心。把 gcd 为一的相邻边称为坏边。

对于值不为一的位置，清零会删掉与其他非一值之间相邻的坏边；但它不能删掉通向值一位置的坏边。对于连续的一段一，清零内部相邻的两个一才会消掉它们之间的边，边界清零一个一就能消掉与外部非一值之间的坏边。因此两部分的收益可以分别计算。

在非一值组成的区域，只看连续的坏边。一个有 e 条边的路径，选择 t 个顶点清零，最多消掉 $\min(2t, e)$ 条边；交替选择内部顶点可以达到这个上界。所以它提供 $\lfloor \frac{e}{2} \rfloor$ 次收益为二的操作，若 e 为奇数，再提供一次收益为一的操作。

CF1891E：填完内部一段一，会多消掉一条边

长度为 l 的一段一，若两边都有非一值：清零 $t < l$ 个时，贴着一侧边界清零可以消掉 t 条边；全部清零时可以消掉 $l + 1$ 条边。相当于每次先得到一分，再在花满 l 次时额外得到一分。若一段一贴着整个数组的边界，则每次只有一分，没有最后的奖励。

因此先取尽能承受的“花一次得两分”的操作，再用剩余次数，从短到长填满两侧有边界的一段一，尽可能多获得完成奖励。其他操作各有一分，最终收益不能超过初始坏边总数。

CF1891E：两分优先，剩余预算填最短段

设非一值坏边链给出的两分操作总数为 D，初始坏边数为 E。取 $t = \min(k, D)$，剩余次数 $r = k - t$；设用这 r 次最多能填满 c 个内部一段，则答案为

$$\max(0, E - 2t - r - c).$$

按段长排序，用前缀和求出 c。交换论证很直接：两分操作每花一次就有额外一分，优于任何长度至少二的完成奖励；在剩下的完成奖励中，短段用更少次数取得同一额外收益。

若整个数组都是一，则没有外部边界，单独返回 $n - 1 - \min(n - 1, \max(k - 1, 0))$。原数组中的零按“非一值”处理，使用 $\gcd(0, 0) = 0$。总时间 $O(n \log n)$，空间 $O(n)$。

返回 CF1891E 的带宽边界

CF1559D2：Mocha and Diana

CF1559D2 Mocha and Diana：两张森林使用相同的 n 个顶点。每次向两张森林同时加入同一条边，要求两者始终无环。求最多能加入多少条边，并输出方案。

$n \le 10^5$，两张森林的初始边数分别为 $m_0, m_1 < n$。

CF1559D2：Mocha and Diana

CF1559D2 Mocha and Diana：两张森林使用相同的 n 个顶点。每次向两张森林同时加入同一条边，要求两者始终无环。求最多能加入多少条边，并输出方案。

$n \le 10^5$，两张森林的初始边数分别为 $m_0, m_1 < n$。

原做法先独立抽取两个顶点，能在两图同时加入就合并，固定尝试 $T = 2 \times 10^6$ 次。随后，当两图的分量数都不超过 10^4 时，用位集完成剩余搜索。

随机阶段每条边都合法。它承担的是压缩剩余规模；是否答对还取决于后续收尾能否执行。

CF1559D2：位集收尾为何能补全

对每个当前分量维护“分量外顶点”的位集。对于顶点 i，两张图对应位集的交集，正好给出能与 i 合法连边的顶点。

合并两个分量时，新分量外集合是两个旧集合的交集。维护当前根的位集，就能始终准确表示连通关系。

依次处理 i，枚举交集中 $x > i$ 的位置；真正加入前再次检查两个并查集。即使枚举的是本轮开始时的快照，也只会多检查已经失效的候选：合并只会使合法点对减少。

因此，只要进入位集阶段，就能消除全部合法点对并达到最优。为何极大方案也是最大方案

CF1559D2：收尾条件可能没有满足

下面构造针对“剩余分量数都不超过 10^4”这个入口条件。取 $n = 100000$，指定顶点 x：

- 第一张森林：一棵包含 x 的 90000 点树，加上 10000 个孤点。
- 第二张森林：孤点 x，加上其余 99999 个点组成的一棵树。

初始分量数为 10001 与 2。合法边恰好是 x 到第一张图的孤点，共 10000 条；加入任意一条就使第二张图连通，达到最优值一。

若随机阶段一直没有加入边，分量数保持不变，位集阶段被跳过，程序输出零条边。

CF1559D2：这个构造的失败概率

原程序每次独立选择 u, v，允许二者相同。按均匀顶点模型，命中任一合法有向点对的概率为

$$p = \frac{2 \times 10000}{100000^2} = 2 \times 10^{-6}.$$

连续失败不会改变森林。因此，两百万次都没有成功的概率恰为

$$(1-p)^{2\times 10^6} \approx 1.83\%.$$

这个概率对应上一页的具体输入，不是对所有输入的平均失败率。实际取模抽样有微小偏差，也不能把一次通过测试视为概率证明。

多个合并的等待还会继续累积，不能把这个单步例子当作整体失败率上界。

附录：更强构造与尾界 • 原讨论

CF1559D2：位运算之外仍有候选枚举

设进入收尾时分量数为 K_0, K_1，$k = \min(K_0 - 1, K_1 - 1)$。固定宽度 $M = 100000$，每个位集占 $W = \lceil \frac{M}{64} \rceil$ 个机器字。

设位集快照中实际枚举了 C 个点对。每轮若实际枚举到候选，首个候选必能加入，因此 $C \le nk$。位集交集虽能按字并行，随后对每个候选的并查集检查仍要单独计入。

以 $D(n)$ 表示一次并查集操作的摊还代价，总时间上界为

$$O((m_0 + m_1 + T + n + C)D(n) + (n + K_0 + K_1 + k)W).$$

原代码只有路径压缩，可保守取 $D(n) = O(\log n)$；加入按大小合并后，$D(n) = O(\alpha(n))$。

CF1559D2：固定宽度与空间开销

两组各 10000 行的位集，即使实际 n 较小，每行仍然按 100000 位计算。按常见 64 位字布局，它们共预留

$$2 \times 10000 \times \left\lceil \frac{100000}{64} \right\rceil \times 8 \quad \text{字节},$$

约 238.5 MiB，另有并查集、答案和临时空间；实际驻留内存取决于访问了哪些行。

若保留位集方案，可以按实际 n, K_0, K_1 分配宽度和行数，并使用按大小合并。但这些修改不解决收尾入口被跳过的问题，也不能抹去候选枚举的最坏开销。

更直接的改进是让确定性收尾适用于任意剩余分量数。

CF1559D2：公共顶点与分量代表

固定顶点 r。依次检查其他顶点 v：若 r, v 在两张森林中都不连通，就加入 (r, v)。

处理完后，每个顶点至少在一张森林中与 r 连通；后续合并不会破坏这个性质。

从第一张森林中每个不含 r 的分量各取一个代表，组成 A；第二张森林同样得到 B。

- 每个 $a \in A$ 必在第二张森林中与 r 连通。
- 每个 $b \in B$ 必在第一张森林中与 r 连通。

所以任取 $a \in A, b \in B$，边 (a, b) 在两图中都合法。加入后，它分别将对应的两个分量并入各自含 r 的分量。

CF1559D2：确定性收尾与最优值

不断配对 A, B 的代表，直到其中一个集合为空。此时至少一张森林已连通，加入边数恰好达到上界

$$\min(c_0 - 1, c_1 - 1), \quad c_0 = n - m_0, \quad c_1 = n - m_1.$$

上述结构还说明：若两张森林都未连通，就一定存在合法点对。因此位集阶段找到的极大方案同样是最大方案。

保留前面的 T 次随机尝试，再用这个构造收尾，总时间为 $O((T + n + m_0 + m_1)\alpha(n))$、空间 $O(n)$，且始终正确。随机阶段已加入的每条边都使两个分量数各减一，不会损失最优值。

也可以直接省去随机预处理，得到 $O((n + m_0 + m_1)\alpha(n))$ 的确定性做法。

随机局部修复：2-SAT

有 n 个布尔变量与 m 个子句，每个子句含两个文字，例如 $(x_i \vee \neg x_j)$。寻找使全部子句成立的赋值。

这里讨论 随机游走算法，不预设公式一定有解。

随机局部修复：2-SAT

有 n 个布尔变量与 m 个子句，每个子句含两个文字，例如 $(x_i \vee \neg x_j)$。寻找使全部子句成立的赋值。

这里讨论 随机游走算法，不预设公式一定有解。

从任意赋值出发。如果还有不满足的子句，任取一个，并等概率翻转其中一个文字对应的变量；随后继续检查。

一次修复可能破坏其他子句，满足的子句数量未必单调增加。分析时需要换一个衡量进展的量。

2-SAT：朝某个固定解的随机游走

假设存在一个满足赋值 $a^\star$，令 D_t 为当前赋值与它不同的变量数量。这个量用于证明，算法不需要知道 $a^\star$。

当前不满足的子句，在 $a^\star$ 中至少有一个文字为真。因此，两个候选变量中至少有一个与 $a^\star$ 不同；随机翻转后，距离减一的概率至少为 $\frac{1}{2}$。

$D_t - 1$	D_t	$D_t + 1$
概率至少 $\frac{1}{2}$	当前距离	概率至多 $\frac{1}{2}$

$D_t = n$ 时下一步一定减一；$D_t = 0$ 时已经找到解。允许暂时走远，但总体不会比这条反射随机游走更慢。

2-SAT：期望步数与有限预算

令势函数 $\Phi(d) = d(2n - d)$。对于尚未找到解的状态，无论选择哪个不满足的子句，都有

$$E\big[\Phi(D_{t+1}) \mid \text{已有历史}\big] \le \Phi(D_t) - 1.$$

势函数非负且至多 n^2，于是找到解所需的期望翻转次数至多 n^2。由 Markov 不等式，执行 $2n^2$ 次翻转仍未找到解的概率至多 $\frac{1}{2}$。

使用新的随机选择独立运行 r 轮，全部失败的概率至多 2^{-r}。找到的赋值可以精确验证；未找到时，若输出无解，就必须保留这一单侧误差说明。

附录：势函数的单步变化

2-SAT：翻转次数不等于实现时间

若每次扫描全部 m 个子句来寻找冲突，一轮 $2n^2$ 次翻转的时间是 $O(mn^2)$，r 轮共 $O(rmn^2)$。维护变量出现位置与未满足子句集合可以减少实际更新量，但不能直接把完整运行时间写成 $O(n^2)$。

2-SAT 本身有 $O(n+m)$ 的强连通分量算法。这里的价值在于：局部修复可以借助“到某个固定解的距离”建立成功率证明，而不是仅看当前得分是否上升。

对含三个文字的子句，同一论证只保证 $\frac{1}{3}$ 概率靠近固定解；二次步数界不能直接照搬。模拟退火或一般随机爬山也需要各自的分析，不能由这份证明获得保证。

返回顺序与搜索选讲

目录

随机枢轴

随机指纹

见证抽样

选讲：指纹与抽样

选讲：交互

选讲：顺序与搜索

附录：构造与证明

附录索引

内容	阅读要求与入口
基础补充	三路划分、一般误差预算、随机源、条件概率
交互题概率界	等差数列、子掩码下降、树上频率
几何与数论构造	线性代数与归一化、完全乘性函数、有限枚举
代数摘要	括号消去、矩阵逆与动态可达性
搜索与期望	分组完备性、随机游走、预处理等待

各部分独立使用；矩阵摘要需要有限域与线性代数，构造题需要对应的数论或几何基础。

返回主讲安排 · 返回选讲题目 · 返回随机顺序

P1923：原地三路划分

当前工作区间采用半开区间 $[L,R)$。始终把它分成四段：

$$[L,s)<v,\quad [s,i)=v,\quad [i,t)\ \text{未处理},\quad [t,R)>v.$$

初始 $s=i=L,t=R$。若 $a_i<v$，与 a_s 交换并同时增加 s,i；若相等，只增加 i；若大于，先减少 t 再交换 a_i,a_t，此时新换来的数仍要检查，不能增加 i。

当 $i=t$ 时划分完成。将 k 保持为原数组中的目标下标：若 $k<s$，令 $R=s$；若 $k\geq t$，令 $L=t$；否则返回 v。每次只保留目标所在的一段。

返回主讲：随机枢轴

从错误预算确定重复次数

设一次判定的失败概率至多为 $0 < \rho < 1$，全程最多作 Q 次判定，允许总失败概率 ε。由 $Q\rho^t \le \varepsilon$ 得

$$t \ge \left\lceil \frac{\ln\left(\frac{Q}{\varepsilon}\right)}{\ln\left(\frac{1}{\rho}\right)} \right\rceil .$$

这里的 Q 是可能影响答案的比较数：逐项查询、二分比较、前缀对计数可能对应不同的数量。多组测试使用相同预算时，也要把它们计入全程。

若单轮界本身不小于一，增加轮数没有意义，应先改进单轮算法或参数。$\rho = 0$ 时无需重复；有限轮 Monte Carlo 不能由这条公式得到零错误保证。

期望耗时与运行预算

若一次独立尝试以概率至少 α 找到可验证的答案，等待成功的期望尝试次数至多 $\frac{1}{\alpha}$；每轮耗时至多 C 时，期望总耗时至多 $\frac{C}{\alpha}$。

对于总耗时为 T 的 Las Vegas 算法，$E[T] \le C$ 并不表示每次都在 C 内结束。Markov 不等式给出

$$\Pr[T > \lambda C] \le \frac{1}{\lambda}.$$

若中断后可以从头独立重试，每轮给 $2C$ 的预算，连续 t 轮均超时的概率至多 2^{-t}。耗尽预算只能报告未完成；不能把超时或没找到候选当作无解。

随机种子与均匀抽样

理论中的每轮试验使用新的独立随机量。实现通常初始化一个伪随机数引擎，随后持续取样；不要在循环内按时间反复设种子，短时间内可能重复同一段随机序列。

均匀抽下标使用均匀整数分布；直接把随机整数对区间长取模，除非整除生成范围，否则存在偏差。均匀洗牌可用 Fisher–Yates：依次将末尾位置与未确定部分中的均匀随机位置交换。

调试时固定并记录种子，使错误可以复现；评估时更换种子。伪随机引擎是实现模型，不会自动提供数学上的完全独立性；公开固定种子也不能抵御按该种子构造的输入。

输入依赖与条件概率

若输入预先固定，随机性只在算法内部，通常可以先固定一个答案或一条精确执行路径再分析。若询问会根据之前的正确答案生成，首次误判前的路径仍与精确算法相同，可对这条路径上的失败事件取并集界。

另一种情形是算法返回一个随机选择的合法对象，并据此改变下一步输入。即使此前从未答错，未来对象也可能依赖旧随机量；固定对象的概率界不能直接复用。

若每轮使用新随机量，且对任意已有历史都有条件失败概率至多 ρ，则连续失败概率仍至多 ρ^t。关键是这个条件界，而非仅仅更换种子或隐藏种子。

返回主讲安排

CF1114E：公差误差界

对 $n \geq 31$，质数 2 的倍数比例至多 $\frac{16}{31}$；任何质数 $p \geq 3$ 的倍数比例至多 $\frac{11}{31}$。两者都可以由 $\left\lceil \frac{n}{p} \right\rceil$ 的分段单调性验证。

无放回抽取 30 个位置，全部落入某个大小为 m 的集合的概率为

$$\frac{\binom{m}{30}}{\binom{n}{30}} \leq \left(\frac{m}{n}\right)^{30}.$$

可能的质数不到 10^6 个，因此一个保守的统一界为

$$\Pr[g \neq d] \leq \left(\frac{16}{31}\right)^{30} + 10^6 \left(\frac{11}{31}\right)^{30} < 3.5 \times 10^{-8}.$$

返回 CF1114E

CF1364E：下降路径的询问数

首个枢轴有 k 个一时，令 $A_k = S - d$。$k = 0$ 时零已经找到；$k = 1$ 时唯一候选直接判零，无需询问却多确定一个值。因此

$$A_0 = 0, \quad A_1 = -1.$$

$k \geq 2$ 时，从 $2^k - 1$ 个真子掩码中选新枢轴，它含 $j < k$ 个一的概率为 $\frac{\binom{k}{j}}{2^k - 1}$。本轮问 $2^k - 2$ 次、确定一个新值，故

$$A_k = 2^k - 3 + A_j \quad \text{以概率} \quad \frac{\binom{k}{j}}{2^k - 1}.$$

按原代码只标记枢轴与终止零点，总数恰为 $Q = R + 2n - 3 + A_k$。正文的逐项计数 给出上界；这里保留完整分布，用于计算超限概率。

CF1364E：起点位数与几何等待

令 $c_{n,k}$ 为满足 $x \neq y$、$0 \le x, y < n$、$\mathrm{popcount}(x \ \mathrm{OR}\ y) \le 6$ 且 $\mathrm{popcount}(x) = k$ 的有序数对数量。

一次抽样成功的概率与成功后首端位数分布分别为

$$s_n = \frac{\sum_{k=0}^{6} c_{n,k}}{n(n-1)}, \quad w_{n,k} = \frac{c_{n,k}}{\sum_{j=0}^{6} c_{n,j}}.$$

R 服从参数为 s_n 的几何分布，并与成功数对的条件分布独立。

CF1364E：按起点与下降路径计算尾概率

已知 $Q = R + 2n - 3 + A_k$。给定 $A_k = a$，超限等价于

$$R > 4272 - 2n - a, \quad 4272 = 4269 + 3.$$

几何分布满足 $\Pr[R > h] = (1 - s_n)^h$。记 $f_{k(a)} = \Pr[A_k = a]$，按起点位数 k 和下降净开销 a 加权，得到

$$\Pr[Q > 4269] = \sum_{k=0}^{6} w_{n,k} \sum_{a} f_{k(a)} (1 - s_n)^{4272-2n-a}.$$

CF1364E：核验全部允许规模

$c_{n,k}$ 可按 n 递增枚举数对，f_k 由前面的有限递推用有理数准确计算。对最终幂与加法使用向外舍入，就能得到概率的上下界。

全部 2046 个允许规模的最大值出现在 $n = 2048$，严格小于 4.70258×10^{-7}。正文使用稍宽的 4.8×10^{-7}。

这是对有限分布的计算证明；随机模拟用于检查恢复过程，不能替代尾概率上界。合并重复询问的缓存只会降低询问次数。

实现时采用均匀整数抽样，并为所有询问共用计数器。概率结论基于独立均匀随机模型，不把取模抽样或固定伪随机序列直接视为这个模型的证明。

返回 CF1364E

CF1559D2：补完条件之前的随机匹配

采用独立均匀抽取两个顶点的模型。将 $n = 10^5$ 个点分为 r, A, B，其中 $|A| = 54999$、$|B| = 45000$。

两图分别以 r 连接全部 B 或全部 A，其余点孤立。合法边恰在两侧孤点之间；每次成功各消去一个。

当 B 剩下 j 个孤点时，A 还剩 $j + 9999$ 个。两图分量数为 $(10000 + j, j + 1)$，所以 $j > 0$ 时始终跳过补完。

$$p_j = \frac{2j(j + 9999)}{n^2}, \quad W = \sum_{j=1}^{45000} G_j.$$

G_j 是从 j 降到 $j - 1$ 的独立几何等待，包含成功次，参数为 p_j。达到最优的 45000 条边，需要全部等待完成。

CF1559D2：固定预算下的失败下界

独立几何等待的期望与方差为

$$\mu=\sum_{j=1}^{45000}\frac{1}{p_j},\quad \sigma^2=\sum_{j=1}^{45000}\frac{1-p_j}{p_j^2}.$$

用有理数求和并向外舍入，得到 $\mu\ge 4,793,884.771$、$\sigma^2\le 410,846,421,207$。

取 $T=2\times 10^6$。Cantelli 不等式给出

$$\Pr[W\le T]\le\frac{\sigma^2}{\sigma^2+(\mu-T)^2}<0.050002.$$

故这个构造在固定预算下失败的概率至少为 94.99%；没有完成时又会跳过补完。返回 CF1559D2

CF1438F：最大概率点的比较

令 $t_d = 2^{h-d-1} = a_d + 1$，则

$$F_d = t_d^2(n + 2 - 2t_d) - n.$$

对相邻非根层 $1 \le d < h - 1$，有 $t_d \le \frac{n+1}{4}$，且 $t_{d+1} = \frac{t_d}{2}$，因此

$$F_d - F_{d+1} = \left(\frac{t_d^2}{4}\right)(3(n + 2) - 7t_d) > 0.$$

再代入原根与孩子的分支大小，得到 $F_1 - F_0 = \frac{(n+1)^2 \times (n-5)}{32} > 0$。故两个孩子严格优于根及所有更深节点。

返回候选孩子

CF1705F：最后两位由基准总分恢复

记 r 在两组中原本答对的数量为 c_U, c_V。两次基准给出 $S_0 = c_U + c_V$ 与 $S_1 = c_U + |V| - c_V$，解得

$$c_U = \frac{S_0 + S_1 - |V|}{2}, \quad c_V = \frac{S_0 - S_1 + |V|}{2}.$$

每确定一个答案，就知道它在 r 中是否答对。将这些已知贡献分别从 c_U, c_V 中扣去；最后每组至多一位，余下正确数只能为零或一，剩余答案唯一确定。

恢复只使用已经得到的分数。完整答案仍须作为一次查询提交，并读到 n 后结束。

返回询问上界

CF1705F：局部决策下的期望递推

完整历史包含 S_0, S_1，不能直接声称余位独立；选位规则却不读取这两项全局信息，只依赖此前局部结果与组大小。

按选位规则使用的信息逐步暴露随机比特，未访问位置仍独立公平，下一对的两种分支各占 $\frac{1}{2}$。

令 $H_m = G_m - \frac{3m}{5}$。递推化为

$$H_m = \frac{H_{m-2} + H_{m-3}}{2}.$$

初值为 $0, -\frac{3}{5}, -\frac{6}{5}$，归纳可得 $-\frac{6}{5} \le H_m \le 0$，故 $E[Q] \le \frac{3n}{5} + 3$。提前答对全部题目只会减少查询。

返回期望分析

2-SAT：势函数的单步变化

令 $\Phi(d) = d(2n - d)$，对于 $1 \le d < n$，距离分别减一、增一时，势函数的变化为

$$\Phi(d-1) - \Phi(d) = -2(n-d) - 1,$$

$$\Phi(d+1) - \Phi(d) = 2(n-d) - 1.$$

前者不大于后者，且二者平均为 -1。距离减一的条件概率至少为 $\frac{1}{2}$，所以条件期望变化至多为 -1；$d = n$ 时必然减一，变化也为 -1。

令 T 为首次找到任意满足赋值的时刻。对停止后的过程求和，得到 $E[\min(T, t)] \le \Phi(D_0) \le n^2$；令 t 趋于无穷，就有 $E[T] \le n^2$。

返回随机局部修复

Gym104207H：Equidistance

Equidistance 原题 PDF，第 11 页 给出 N 维空间中 M 个两两距离为一的点，要求再加入尽可能多的点，仍保持所有点对距离为一。

Gym104207H：Equidistance

Equidistance 原题 PDF，第 11 页 给出 N 维空间中 M 个两两距离为一的点，要求再加入尽可能多的点，仍保持所有点对距离为一。

C 个两两等距点的仿射维数为 $C-1$：以其中一个点为原点，其余差向量的 Gram 矩阵对角线为一、非对角线为 $\frac{1}{2}$，是正定矩阵。因此至多有 $N+1$ 个点，还能加入的数量为 $N+1-M$。

Gym104207H：重心半径与新点高度

当前有 C 个两两距离为一的点。记重心为 g，令 $q_i = p_i - g$，则 $\sum_i q_i = 0$。对任意 i，展开距离平方并求和：

$$C - 1 = \sum_j \|q_i - q_j\|^2 = C\|q_i\|^2 + \sum_j \|q_j\|^2.$$

最后一项与 i 无关，所以所有 $\|q_i\|^2$ 相同。记它为 ρ^2，就有 $C - 1 = 2C\rho^2$，即 $\rho^2 = \frac{C-1}{2C}$。

取与当前仿射空间正交的单位向量 u，设新点为 $g + hu$。它到旧点的距离平方为 $\rho^2 + h^2$，令其为一，得到

$$h^2 = \frac{C+1}{2C}, \quad p_{\text{new}} = g + \sqrt{\frac{C+1}{2C}}u.$$

Gym104207H：归一化与残差检查不能省

因为两项正交，它到任意旧点的距离平方为 $\frac{C-1}{2C}+\frac{C+1}{2C}=1$。

先对 p_i-p_1 建立正交基。随机取一个向量，减去它在所有已有基向量上的投影；若剩余范数过小，就重新取样，否则单位化作为 u。加入新点后把 u 加入基，再重复，直到共有 $N+1$ 个点。

只要当前基未占满空间，连续随机分布下得到零余向量的概率为零；浮点计算要重复正交化并检查最终距离，而不是靠“很可能线性无关”跳过数值处理。增量维护正交基与重心，总时间 $O(N^3)$、空间 $O(N^2)$。

CCPC Final L：完全乘性函数

CCPC Final 2022 L：构造 $f:\mathbb{N}\to\{-1,1\}$，满足 $f(1)=1$、$f(ab)=f(a)f(b)$ 以及 $\sum_{i=1}^{n} f(i)=k$。$0\le k\le n\le 10^6$，所有测试的 n 之和不超过 2×10^6。

CCPC Final L：完全乘性函数

CCPC Final 2022 L：构造 $f:\mathbb{N}\to\{-1,1\}$，满足 $f(1)=1$、$f(ab)=f(a)f(b)$ 以及 $\sum_{i=1}^{n} f(i)=k$。$0\le k\le n\le 10^6$，所有测试的 n 之和不超过 2×10^6。

所有值由质数的符号决定：筛出最小质因子 p 后，递推 $f(i)=f(p)f\left(\frac{i}{p}\right)$。

若 n,k 奇偶不同则无解；否则只需让恰好 $r=\frac{n-k}{2}$ 项取负一。

大于 $\frac{n}{2}$ 的质数只影响自身，可以独立调整一项。但这样的自由度有限；把分界移到 $\sqrt{n}$，仍能保持独立性。

完全乘性函数：大质数对应不相交的组

一个不超过 n 的整数至多含一个大于 $\sqrt{n}$ 的质因子，且这个质因子的指数只能为一。因此不同大质数的倍数集合互不重叠。

当 $n \geq 200$ 时，先把所有小质数赋为一。将大质数 p 设为负一，就恰好使其 $w_p = \left\lfloor \frac{n}{p} \right\rfloor$ 个倍数取负一。问题变成选一些权重，使总和为 r。

在本题范围内，这些权重能连续表示 $\left[0, \left\lfloor \frac{n}{2} \right\rfloor\right]$：单位权重足够多，后续权重不会制造空隙。下面的覆盖条件说明为什么可以确定性构造，而无需随机猜质数符号。

完全乘性函数：连续覆盖与倒序贪心

若已有权重能表示 $[0, R]$，新权重满足 $w \le R + 1$，加入后就能表示 $[0, R + w]$。从单位权重开始递增加入，始终保持这一性质。

构造时反过来按权重递减处理：若 w 不超过剩余目标，就取它并扣去 w；否则跳过。剩下的小权重能够覆盖剩余目标，因此不会走入死路。

$n < 200$ 时，枚举至多六个小质数的符号，再对独立大质数做有符号背包，能完整判断是否存在方案。

附录：覆盖条件的验证与小规模构造

完全乘性函数：权重覆盖的两个条件

在有限范围 $200 \le n \le 10^6$ 内，筛法可逐项验证：

$$\sum_{p>\sqrt{n},p \text{ 为质数}} \left\lfloor \frac{n}{p} \right\rfloor \ge \left\lfloor \frac{n}{2} \right\rfloor,$$

$$\pi(n) - \pi\left(\left\lfloor \frac{n}{2} \right\rfloor\right) \ge \lfloor \sqrt{n} \rfloor - 1.$$

前者保证总权重足够；后者保证单位权重足够多。所有单项权重至多为 $\lfloor \sqrt{n} \rfloor$，因此从单位权重开始，可表示区间的右端点始终至少为下一权重减一。

由连续覆盖条件，所有 $r \in [0, \lfloor \frac{n}{2} \rfloor]$ 都能表示。这里使用的是题目给定有限范围内的筛法验证结论，不将它当成对任意 n 都成立的数论定理。

完全乘性函数：枚举全部小质数的符号

当 $n < 200$，不超过 $\sqrt{n}$ 的质数最多六个，可以枚举它们的全部符号，共不超过 64 种。对一个固定选择，先计算只含小质因子的整数的总和 S_0。

对每个大质数 p，它的全部倍数形成独立的一组，贡献为

$$f(p) \times C_p, \quad C_p = \sum_{j=1}^{\left\lfloor \frac{n}{p} \right\rfloor} f(j).$$

完全乘性函数：对独立大质数做有符号背包

这里 $j < \sqrt{n}$，其值已由小质数确定。用一个一维背包记录所有 $\sum_p f(p)C_p$ 的可达值，正负选择各做一次转移，并保存来源。范围只有 $[-n, n]$，找到 $k - S_0$ 后回溯各大质数符号；枚举全部小质数赋值仍无解，才能输出无解。

小规模枚举小质数符号，并用背包判断剩余目标是否可达；大规模按大质数权重贪心构造。小规模时间为 $O\big(2^6 n \times \pi(n)\big)$；大规模筛后构造与生成函数值为 $O(n)$，预处理筛法为 $O(V \log\log V)$。

官方题解：小质数赋值与独立大质数的构造

返回完全乘性函数的构造

返回附录索引

CF1909E：Multiple Lamps

Multiple Lamps 中按下开关 i 会翻转所有编号为 i 倍数的灯，限制形如“按 x 就必须按 y”。至少按一次，要求亮灯数不超过 $\left\lfloor \frac{n}{5} \right\rfloor$。

CF1909E：Multiple Lamps

Multiple Lamps 中按下开关 i 会翻转所有编号为 i 倍数的灯，限制形如“按 x 就必须按 y”。至少按一次，要求亮灯数不超过 $\left\lfloor \frac{n}{5} \right\rfloor$。

把全部开关按下，自动满足所有蕴含限制。灯 j 被翻转的次数等于它的约数个数；约数成对出现，仅完全平方数有奇数个约数。因此亮灯恰好有 $\lfloor \sqrt{n} \rfloor$ 个。

$n \geq 25$ 时由 $\sqrt{n} \leq \frac{n}{5}$ 可知满足要求；$20 \leq n \leq 24$ 时两边取整后也都为四。因此 $n \geq 20$ 直接输出全部开关。

CF1909E：剩下至多 19 个开关，完整枚举

只剩 $n \le 19$。预处理每个开关的灯泡翻转位掩码，枚举非空开关子集 S。对每条限制检查 $x \in S \Rightarrow y \in S$；通过后异或所有选中开关的翻转掩码，统计置位数。找到满足上界的集合即输出；全部枚举失败则确实无解。

用去掉最低位的递推预处理每个子集的灯状态，每个子集只需一次异或。将每个开关要求同时选择的对象压成位掩码，也能按相同方式累计整个子集的要求；检查要求是否包含于 S。小规模部分时间 $O(2^n + n^2 + m)$、空间 $O(2^n + n)$。这是确定性做法，也提醒我们：随机化课里遇到的题，不一定最后需要随机数。

返回附录索引

反向括号：区间消去

反向括号模型。 每种符号有一个逆符号，同类正反符号相邻时可以消去，两种方向都允许。询问一个区间能否完全消去。

反向括号：区间消去

反向括号模型。 每种符号有一个逆符号，同类正反符号相邻时可以消去，两种方向都允许。询问一个区间能否完全消去。

用栈从左到右扫描：若新符号与栈顶互逆，则弹栈；否则压栈。最终得到不可约串。消去顺序不影响结果：不重叠的两次消去可交换，重叠只可能形如 $xx^{-1}x$，两种选择都留下 x。

给符号赋标量及其逆元并不保留这个结构：$aba^{-1}b^{-1}$ 的标量乘积总为一，但没有相邻的一对可以消去。顺序信息不能被交换律抹掉。

反向括号：静态区间的精确前缀状态

令 S_i 为前缀 $[1,i]$ 的不可约串。区间 $[l,r]$ 能完全消去，当且仅当 $S_{l-1}=S_r$：在两边左接 S_{l-1} 的逆串并约简，右边剩下的正是这一段的不可约串。

把栈做成持久化节点，每个节点记录“父节点、栈顶符号”。压栈时，对这对信息查字典并复用已有节点；弹栈时回到父节点。相同不可约串因此总有相同节点编号。

一次扫描至多创建 n 个节点。用平衡树字典预处理 $O(n\log n)$、空间 $O(n)$，每个静态询问只比较两个编号，确定性 $O(1)$。普通持久化栈若不合并相同节点，则不能直接比较编号。

反向括号：动态矩阵摘要的适用边界

有单点修改时，后面的前缀状态可能全部改变。另一种思路是给每种符号分配有限域上的随机可逆矩阵 M_x，逆符号用 M_x^{-1}；线段树按原顺序维护矩阵积。

合法消去保持乘积不变。若区间乘积不是 I，可以确定不能消空；乘积为 I 则可能碰撞。固定矩阵阶数时，修改和查询均为 $O(\log n)$。

矩阵的各元素不是独立指纹。这个动态摘要的具体误差界，取决于有限群、允许的字长和采样分布；在证明单组界之前，不能把四个矩阵元素当成四轮独立检验。静态题可以直接采用上一页的精确算法。

返回附录索引

动态可达性查询

动态传递闭包模型。 维护一张 n 个顶点的有向图，支持插入、删除有向边，以及查询从 u 能否到达 v。

动态可达性查询

动态传递闭包模型。 维护一张 n 个顶点的有向图，支持插入、删除有向边，以及查询从 u 能否到达 v。

给每条有向边独立变量，组成邻接矩阵 A。形式幂级数中

$$(I - A)^{-1} = I + A + A^2 + \cdots.$$

在第 (u, v) 项中，每条行走贡献对应边权的乘积。实际算法不求无穷和，而是给边权随机赋值，再计算有限矩阵的逆。

令 $M = I - A$。$M^{-1}_{u,v}$ 是某个余子式与 $\det(M)$ 之比。若 u 不可达 v，分子恒为零；若可达，仅保留一条简单路径、将其他边变量置零时，分子为该路径边权之积，因此不是零多项式。对 $u = v$ 按可达处理。

传递闭包：多变量求值的概率界

非零多项式 P 的总次数为 d，各变量独立均匀取自 $\mathbb{F}_p$。下面说明 $\Pr[P = 0] \le \frac{d}{p}$ 如何由一元根数界得到。

按变量数归纳。将 P 按最后一个变量展开，最高次数为 t，对应的非零系数多项式为 Q；它的总次数至多 $d - t$。

先确定其余变量。归纳知 $Q = 0$ 的概率至多 $\frac{d-t}{p}$；若 $Q \neq 0$，留下一个 t 次一元多项式，最后一个独立随机变量使它为零的概率至多 $\frac{t}{p}$。

两种失败原因合计给出

$$\Pr[P = 0] \le \frac{d - t}{p} + \frac{t}{p} = \frac{d}{p}.$$

传递闭包：分子与分母的错误事件

将边权独立均匀取自 $\mathbb{F}_p$。$\det(I-A)$ 总次数至多 n，常数项为一；余子式总次数至多 $n-1$，因此

$$\Pr[\det(M)=0]\le\frac{n}{p},\quad \Pr[\text{可达点对的分子为零}]\le\frac{n-1}{p}.$$

对预先固定的操作序列，若涉及 s 个图状态和 q 次查询，同一组随机赋值下，遇到奇异矩阵或误判可达性的概率合计至多 $\frac{sn+q(n-1)}{p}$，不要求这些事件独立。

传递闭包：逆矩阵的秩一更新

设 $R = M^{-1}$。一次修改一条边、一行或一列，都可以写成 $M' = M + uv^T$。当分母非零时，Sherman–Morrison 公式给出

$$R' = R - (Ru)\frac{v^T R}{1 + v^T R u}.$$

矩阵乘向量与外积各需 $O(n^2)$，所以初始化 $O(n^3)$，更新 $O(n^2)$，点对查询 $O(1)$。

分母为零意味着新矩阵不可逆，需要重新赋权并以 $O(n^3)$ 重建。以下给出限制重试次数后的全程误差界。

传递闭包：单边更新的实现

将边 (a,b) 的随机权从 w 改为 w'，记 $\Delta = w' - w$。由于 $M = I - A$，有 $M' = M - \Delta e_a e_b^T$，从而

$$R'_{i,j} = R_{i,j} + \frac{\Delta R_{i,a} R_{b,j}}{1 - \Delta R_{b,a}}.$$

先保存旧矩阵的第 a 列、第 b 行和分母，再更新所有位置；不能边写矩阵边读取已经改变的行列。插边令旧权为零，删边令新权为零。模除法由非零分母的逆元实现。

初始化用带主元交换的模高斯消元求逆。若无非零主元或更新分母为零，先重新为现存边独立赋权，再重新求逆；没有得到可逆矩阵以前，不使用它回答询问。

传递闭包：重建次数与独立赋权

固定完整的操作序列，令 S 为初始状态及每次更新后的状态时刻数，q 为查询总数。每个时刻若需要重建，至多尝试 r 次；全部失败则停止并报告未完成。

预先为每个时刻、每次尝试各分配一张独立随机表，共 Sr 张。每张表为所有可能边独立赋权；插边读取当前表中的权值，删边置零，再次插入同一边时复用该权值。

进入某个时刻时，该时刻的新表仍独立于已有历史。当前图固定，每次求逆失败概率至多 $\frac{n}{p}$，所以全程重建耗尽的概率至多

$$S\left(\frac{n}{p}\right)^{r}.$$

传递闭包：重建后的全程误差

不能直接把“已通过求逆检查的随机表”当作原来的均匀样本。对预分配的每张原始表与每次固定的可达查询分别考虑，分子为零的概率至多 $\frac{n-1}{p}$。

实际使用的表虽然依赖历史，误判事件仍包含在这 Srq 个事件的并集中。因此全程误判或未完成的概率至多

$$S\left(\frac{n}{p}\right)^{r}+\frac{Srq(n-1)}{p}.$$

按允许误差选择 p, r。保守总时间为 $O(Sn^2+Srn^3+q)$ 次域运算；随机表可用到时才生成，空间仍为 $O(n^2)$。该界适用于固定操作序列，不涵盖根据随机输出任意调整输入的交互过程。

返回附录索引

CF1198F：GCD Groups 2

GCD Groups 2：将 $2 \le n \le 10^5$ 个正整数分为两个非空组，两组 gcd 均为一；$a_i \le 10^9$。

CF1198F：GCD Groups 2

GCD Groups 2：将 $2 \le n \le 10^5$ 个正整数分为两个非空组，两组 gcd 均为一；$a_i \le 10^9$。

维护两组当前 gcd g_1, g_2，空组为零。扫描 x：若 $\gcd(g_1, x) = g_1$，将它交给第二组；否则若 $\gcd(g_2, x) = g_2$，交给第一组；两边都能改善时，用独立随机比特选择。

结束后精确验证两组 gcd；失败则重新试验，也可以重新打乱顺序。每轮 $O(n \log V)$。全数组 gcd 不为一时确定无解，但它为一并不充分，例如 $6, 10, 15$ 无法分成所需的两组。

CF1198F：可行分支与概率下界

假设当前分配仍能补成合法解。若 x 对第一组的 gcd 没有影响，把它改交第二组不会破坏第一组，并可能帮助第二组，所以强制选择仍保留一种完成方案。随机分支至少有一边可行，保留可行性的概率至少为 $\frac{1}{2}$。

每个随机分支要么建立一个空组，要么使某个正 gcd 至少减半，因此分支数至多 $2+2\lfloor \log_2 V \rfloor \le 60$。这给出单轮成功率至少 2^{-60} 的保守下界。

这个界远不足以保证少量重试可靠。“只有几十次关键选择”本身不能推出高成功率。验证通过时答案确定正确；有限次失败输出 `NO`，仍依赖对具体搜索更强的概率分析。

下面给出可以确定判无解的算法。

CF1198F：缩小第二组种子的候选

设 k 是任意输入值的不同质因子数上界。十个不同质因子的最小乘积是前十个质数之积，等于 $6469693230 > 10^9$，故本题 $k \le 9$。

若全数组 gcd 不为一，立即判无解；否则固定 a_1 在第一组。

对每个质数 $p \mid a_1$，选一个不被 p 整除的位置，连同位置一组成集合 S。因为每个质因子都被消去，$\gcd(S) = 1$，且 $|S| \le k + 1$。

若补集 gcd 也为一，已经得到答案。否则任何合法划分都必须把 S 中至少一个非一号位置放进第二组：若整个 S 都在第一组，第二组包含于补集，其 gcd 不可能为一。因此只需枚举至多 k 个第二组种子。

CF1198F：固定种子后的因子掩码

固定两组种子 a_1, a_j，分别分解出质因子集合 P, Q。第一组的 gcd 最终为一，当且仅当 P 中每个质因子都被某个加入的数消去；第二组同理。

删除两个种子位置，将剩余元素按原顺序记为 $b_1, \ldots, b_m$。状态 (S, T) 记录两组尚未消去的质因子。记 $f[S, T]$ 为达到此状态所选位置中的最大下标的最小值，初态为 (P, Q)、值零。

这里只选择使 gcd 改善的见证；跳过的元素先不分组。两组都达到 gcd 一后，将剩余元素任意补入都不会破坏答案。因此较早达到相同掩码的状态总能支配较晚状态。

CF1198F：跳到下一个有用的位置

对 $P \cup Q$ 中每个质数 p，预处理 $\mathrm{next}_p(i)$：下标严格大于 i、且元素不被 p 整除的第一个位置，不存在则为无穷。

在状态 (S, T)，令 $i = f[S, T]$。对每个 $p \in S$，取 $h = \mathrm{next}_p(i)$；若存在，把该数交给第一组，并一次性删掉 S 中所有不整除 b_h 的质因子；对 $p \in T$ 对称转移。

每次至少删掉一个因子，按剩余因子总数递减处理状态即可。每次松弛保存前驱状态、所选下标与组号；到达空掩码 $(\varnothing, \varnothing)$ 后沿前驱恢复见证分组。

CF1198F：最早位置转移的完备性

考虑某状态的一种可行扩展，取扩展中第一个能消去未消去因子的元素 x。若它加入第一组并消去质因子 p，则 $\mathrm{next}_p(i)$ 对应的位置 y 不晚于 x。

若 $y = x$，转移直接保留这一步。若 $y < x$，更早的 y 原本没有在另一组中带来必要的 gcd 改善，否则 x 就不是第一个有用元素。可以先将 y 加入第一组，再继续使用原扩展中的 x 和后续元素。

因此只选每个因子的最早可用位置，不会删掉全部可行扩展。若所有种子候选的掩码 DP 都无法到达空状态，才可以确定输出无解。

CF1198F：复杂度与方案输出

一个种子对至多有 $2^{|P|+|Q|} \le 4^k$ 个状态，每状态枚举至多 $2k$ 个因子。后缀跳转表与元素掩码预处理 $O(nk)$，单个种子对的 DP 为 $O\left(k4^k\right)$。

至多枚举 k 个第二组种子，总时间为

$$O\left(n \log V + nk^2 + k^2 4^k + (k+1)\sqrt{V}\right).$$

最后一项来自试除分解种子。各轮复用存储，空间为 $O\left(n(k+1) + 4^k\right)$。恢复两组见证后，将未分配的元素交给任意一组，并再次精确计算两个 gcd；两个固定种子保证两组非空。

确定性做法来源：小种子集合与最早位置 DP

返回附录索引