# 数学 与
# 挖掘题目性质的突破口

2023 CSP-S 秋令营
feecle6418

洛谷
www.luogu.com.cn

# 数论部分

2023 CSP-S 秋令营
feecle6418

# 数论

若无特殊说明，认为 $p$ 是质数，且所有提到的数均为非负整数。

# 整除

**整除：** 若 $ak = b$，则 $a$ 是 $b$ 的因子，$b$ 是 $a$ 的倍数。

只有两个因子的数叫质数。

质数的判定:

1. 试除法;
2. 费马小定理;
3. 筛法预处理。

# 埃式筛

对于每一个质数，将它的所有非平凡倍数全部标记为合数。

时间复杂度为 $O(n\log\log n)$。

注 1:

$$\sum_{i=1\sim n}\frac{1}{i}=O(\log n)$$

可以据此预处理出 $1\sim n$ 的每个因子。

注 2：该复杂度分析同样适用于区间筛。

# 例题：P7960

输入一个数，输出它的上一个，不含有任何含 7 因数的数。

含 7 是指有一位是 7。

至多 $2 \times 10^5$ 组询问。输入的数不超过 $10^7$。

本题中，可以用类似埃式筛的做法，达到 $O(n \log\log n)$ 的理论复杂度。

# gcd

$\gcd(a,b)$ 表示 $a,b$ 的最大公因子。

$\operatorname{lcm}(a,b)$ 表示 $a,b$ 的最小公倍数。

做法：辗转相除。

gcd/lcm 与 min/max 有类似的性质（幂等律）。

特别地，若 $a$ 不整除 $b$，则 $\gcd(a,b) \le \frac{a}{2}$。

若 $a \ne b$，则 $\gcd(a,b) \le |a-b|$。

# 欧拉函数

欧拉函数的定义是：$\le x$ 且与 $x$ 互质的正整数个数。

# 例题： P2568

给定 $n$，求有几对正整数 $x, y$ 满足 $1 \leq x, y \leq n$，且 $\gcd(x, y)$ 是质数。$n \leq 10^7$。

# 线性筛

如果对于每个数，我们只标记它 1 次，那就可以做到线性了。

我们希望：每个合数 $x$ 都是被其最小质因子 $p$ 和 $q = x/p$ 标记到的。只需保证 $q$ 中不含有比 $p$ 小的因子。

因此，我们将 $1 \sim n$ 分别作为 $q$，每次枚举从小到大枚举质因子 $p$，若 $p|q$，则在这次标记完之后退出。

线性筛可以筛很多其它函数。例如：莫比乌斯函数，欧拉函数，因数个数，最小质因子，最大质因子，质因子次幂最大值，质因子次幂最小值，次幂（特别地，逆元）……

# 例题： CF1548B / CF1458A / CF1834E

给定 $n$ 和一个长度为 $n$ 的数组 $a$，求一个**最长**的区间 $[l, r]$，使得存在 $m \geq 2$ 和 $k$，对于所有 $l \leq i \leq r, a_i \equiv k \pmod{m}$（即**区间内所有数对 $m$ 取模余数相等**），输出最长区间长度（区间长度定义为 $r - l + 1$）。

$$n \leq 2 \times 10^5, a_i \leq 10^9$$

给定长度为 $n$ 的序列 $a$,长度为 $m$ 的序列 $b$.

现在你需要对于所有整数 $j$ 满足 $1 \leq j \leq m$,求出 $\gcd(a_1 + b_j, a_2 + b_j, \ldots, a_n + b_j)$ 的值.

$$n \leq 2 \times 10^5, a_i, b_j \leq 10^{18}$$

给定一个长度为 $n$ 的序列 $a$，一个正整数 $x$ 被称为「可爱的」，当且仅当在序列 $a$ 中，不存在一个子段所有元素的 最小公倍数 等于 $x$。求最小的「可爱数」。

$$n \leq 3 \times 10^5, a_i \leq 10^9$$

# 同余与逆元

若 $m \mid a-b$，则 $a \equiv b \pmod{m}$。在模 $m$ 意义下，可以进行加减乘运算。

Bezout 定理：对于 $a, b$，$ax+by=d$ 有整数解当且仅当 $\gcd(a,b) | d$。

等价于 $ax+by=1$ 有整数解当且仅当 $a, b$ 互质。

也就是说，$a$ 在模 $m$ 意义下有逆元当且仅当 $\gcd(a,m)=1$。逆元可以用来进行模意义下除法运算。

# exgcd

如果能找到 $ax+by=d$ 的一组解，那么 $x$ 就是 $a$ 模 $b$ 意义下的逆元了。

如何找到这个不定方程的一组解？

递归地解决这个问题：当 $a=d, b=0$ 时，解为 $x=1, y=0$。

若已求出 $bx'+(a \bmod b)y'=d$ 的一组解，则：

$$\begin{aligned} ax+by &= bx'+(a \bmod b)y' \\ &= bx'+\left(a-b\times\left\lfloor\frac{a}{b}\right\rfloor\right)y' \\ &= ay'+b\left(x'-\left\lfloor\frac{a}{b}\right\rfloor y'\right) \end{aligned}$$

可以看出，$x=y'$，$y=x'-\left\lfloor\frac{a}{b}\right\rfloor y'$ 就是原式的一组解。

由欧几里得算法，最终总会递归到 $a=d, b=0$。

可以证明，最终找到的解的绝对值大小不超过 $2\max(a,b)$。

# exgcd 与 excrt

求出二元一次不定方程的一组特解，随即可以得到其通解的表达式。

使用 exgcd，也可以把两个同余方程 $x \equiv a_i \pmod{m_i}, i = 1,2$ 合并为新的，模 $\mathrm{lcm}(m_1, m_2)$ 的同余方程。

合并两个同余方程组：

$$\begin{cases} x \equiv a_1 \bmod m_1 \\ x \equiv a_2 \bmod m_2 \end{cases}$$

把 x 表示为 $a_1 + pm_1$ 和 $a_2 - qm_2$，那么有：

$$a_1 + pm_1 = a_2 - qm_2$$
$$pm_1 + qm_2 = a_2 - a_1$$

直接 exgcd 求解即可。

# 费马小定理与（拓展）欧拉定理

当 $p$ 是质数时，$a^p \equiv a \pmod p$。

如果 $a \neq 0$，则 $a^{p-2} \equiv a^{-1} \pmod p$。

如果 $\gcd(a,p) = 1$，则 $a^b \equiv a^{b \bmod \varphi(p)} \pmod p$。

如果 $b > \varphi(p)$，则 $a^b \equiv a^{b \bmod \varphi(p) + \varphi(p)} \pmod p$。

# 例题：P4588

维护一个数，初始为 1，支持乘以一个正整数，除以一个正整数（保证除完之后还是整数），输出这个数模 $M$ 的值。

操作数 $10^6$，$M \le 10^9$。

# 例题：P4588

分开维护与 $M$ 互质的部分和 $M$ 的质因子部分。

# 例题：幂塔方程（弱化版）

给定 $n, p$，保证 $p$ 是质数，解方程 $x^x \equiv n \pmod p$ 。

求出任何一个解就可以了。

$0 < n < p \leq 10^{18}$，需要满足 $x \leq 2^{125}$。

# 例题：幂塔方程（弱化版）

指数和底数，分别对应两个模数互质的同余方程。

# 例题：P8338

对于排列 $A$，定义 $v(A)$ 为最小的 $x$ 使得 $A^x = A$。

给定排列 $S$，定义 $f(i,j)$，若存在 $x$ 使得 $S_i^x = j$，则 $f(i,j) = 0$；否则令 $S'$ 表示 $S$ 交换 $i,j$ 两个元素形成的排列，$f(i,j) = v(S)$。

对于所有 $(i,j)$，求 $f(i,j)$ 之和（模大质数）。

先对题目进行转化。

1. $v(A)$ 是什么？
2. 交换两个元素后，$v(A)$ 怎么变？

# 例题：P8338

前置知识：$m$ 个正整数和为 $n$，不同的正整数个数只有 $O(\sqrt{n})$。

我们首先有一个 $O(n\log n)$（实际上跑不满，能过）的解法。

在 $v(A)$ 改变时，暴力维护其每个质因子幂次的变化，并重新统计最大值。

# 例题：P8338

考虑现在合并的是 $a,b$ 这两个环，设此时的 lcm 为 $l$，先前的 lcm 为 $g$，我们来分析是否有好方法可以通过 $g$ 算出 $l$。

考虑 $v_p(l) = \max(v_p(a+b), v_p(others))$，进行分类讨论：

1. 若 $v_p(a) \neq v_p(b)$，则 $v_p(a+b) = \min(v_p(a), v_p(b))$。
2. 若 $v_p(a) = v_p(b)$，则 $v_p(a+b) \geq \max(v_p(a), v_p(b))$。

如果我们设 $f(a) = \max(v_p(a) - v_p(sc), 0)$（$sc$ 是次幂次大值），那么有：

$$v_p(l) = \max\left(v_p(g), v_p(a+b)\right) - f(a) - f(b)$$

回到 lcm 的形式，我们可以：

$$l = \frac{\mathrm{lcm}(g, a+b)}{f^*(a)f^*(b)}$$

其中 $f^*(a)$ 是对于每个质数，$p^{-f(a)}$ 的积。

预处理 $f^*$，并线性筛 lcm，就得到了严格线性的做法。（Bonus：P9135）

# 组合数学部分

2023 CSP-S 秋令营
feecle6418

www.luogu.com.cn

# 组合数

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

$$(x+y)^n = \sum \binom{n}{i} x^i y^{n-i}$$

组合数有递推公式。

$$\binom{n}{m} \equiv \binom{n \bmod p}{m \bmod p} \times \binom{n/p}{m/p} \bmod p$$

考虑方程 $x_1 + x_2 + \cdots + x_n = k$，其正整数解的个数为 $\binom{k-1}{n-1}$。

非负整数解的个数？$x_i \geq a_i$ 解的个数？

# 其它经典问题

错排问题：$D_n = (n-1)(D_{n-1} + D_{n-2})$

第一类斯特林数：$1 \sim n$ 的排列有 $m$ 个环。

第二类斯特林数：$n$ 个不同球放进 $m$ 个相同盒子。

树上拓扑序计数：父亲的数小于儿子，求方案数。

配对：$2n$ 个人配 $n$ 对，有几种方法。

# 例题： P2480

给定 $n, g$， 求

$$g^{\sum_{d|n}\binom{n}{d}} \bmod 999911659$$

其中 $n, g \leq 10^9$。

# 卡特兰数

长为 $2n$ 的合法括号串数量被称为卡特兰数。

$$Ca(n)=\sum_{i=1}^{n}Ca(i-1)Ca(n-i)$$

从 $(0,0)$ 出发，不能跨过 $y=x$，走到 $(n,n)$。有几种走法？

反射容斥：将所有 $(0,0)$ 到 $(n,n)$ 的非法路径与 $(0,0)$ 到 $(n-1,n+1)$ 的路径双射起来。

据此可以得到：$Ca(n)=\binom{2n}{n}-\binom{2n}{n-1}$

# 容斥原理和二项式反演

所有条件都满足的 =钦定满足一个条件的 –钦定满足两个条件的 +钦定满足某三个条件的 - ⋯

需要注意“钦定”的含义。

若

$$f(i)=\sum_{j=i}^{n}\binom{j}{i}g(j)$$

则

$$g(i)=\sum_{j=i}^{n}(-1)^{j-i}\binom{j}{i}f(j)$$

# 容斥原理的应用（一）

方程 $x_1+x_2+\cdots+x_n=k$，$1\le x_i\le t$ 解的个数?

**P5505**：有 $n$ 种球，第 $i$ 种有 $a_i$ 个，同种球完全相同。要把这些球分给 $m$ 个同学，每个同学至少得到一个球，求方案数。

$n,m,a_i\le 1000$，答案对大质数取模。

**ABC235G：**有红绿蓝三种球，分别 $R,G,B$ 个。

你需要将他们中的一些放进 $n$ 个盒子，要求:

1. 每个盒子都有球，
2. 每个盒子一个颜色的球只能放一个。

$n,R,G,B\le 5000000$，答案对大质数取模。

# 求和号处理技巧

$$\sum_i \sum_j f(i,j) = \sum_j \sum_i f(i,j)$$

$$\sum_i \sum_j f(i)g(j) = \sum_i f(i) \times \sum_j g(j)$$

$$\sum_i (f(i) + g(i)) = \sum_i f(i) + \sum_i g(i)$$

最终目的是，分离变量，化整为零。

$$\binom{n}{m} = \binom{n-1}{m-1} + \binom{n-1}{m}$$

$$\binom{n}{m} = \frac{n}{m}\binom{n-1}{m-1}$$

$$\binom{n}{m} = \binom{n}{n-m}$$

$$\sum_i \binom{n}{i}\binom{m}{k-i} = \binom{n+m}{k}$$

$$\sum_i \binom{i}{n}\binom{k-i}{m} = \binom{k+1}{n+m+1}$$

# 一些其它技巧

**算贡献**：算每个元素对答案的贡献。

**先列式再优化**：列出式子，总比对着题目瞎想好。

**合理拆分求和号**：如，给定一个序列，多次询问，每次给定一个区间 $[l,r]$，询问其 所有子区间的和 的和。

# 容斥原理的应用（二）

容斥系数有时可以放进 dp 过程。

**ARC101C**：给出一棵大小为偶数 $n$ 的树，将树上的点两两配对，问有几种配对方法满足每条边都至少被一对配对的点跨过。

$n \leq 5000$，答案对大质数取模。

**P4099**：给定一棵树，边有方向，求拓扑序数量。

$n \leq 5000$，答案对大质数取模。

# 例题： P7961

给定 $n, m, k$ 和 $v_0 \sim v_m$。

有一个整数序列 $a_1, \dots, a_n$。定义其权值为 $\prod v_{a_i}$。

如果 $popcount(\sum 2^{a_i}) \le k$，就说 $\{a\}$ 合法。

求所有合法 $a$ 的权值和，对大质数取模。

$$n \le 100, m, k \le 30$$

# 例题：CF1542E2

设 $p,q$ 都是长度为 $n$ 的排列，且 $p$ 字典序小于 $q$，逆序对数大于 $q$。

计算有几对 $(p,q)$ 满足条件。

$$n \le 500$$

部分分：$n \le 50$

# 例题：CF1542E2（一）

本题解法很多。这里只说一种。

首先，如果 $p, q$ 在第一位相同，这一位无论是什么都无关紧要，可以递归到子问题。

如果 $p, q$ 在第一位就不同，我们可以枚举 $p_1 = x, q_1 = y\ (x < y)$。

此时，$p_1$ 和后面的数构成了 $x - 1$ 个逆序对，$q_1$ 构成了 $y - 1$ 个。

而 $p$ 的总逆序对数就是 $(x - 1)$ 加上后 $n - 1$ 个元素内部的逆序对数 $A$，$q$ 同理，设 $q$ 后 $n - 1$ 个元素内部的逆序对数为 $B$。

这可以写成一个关于 $A, B$ 的不等式。

预处理出 $f(n, m)$ 表示长为 $n$ 排列，$m$ 个逆序对的方案数，用这个式子就可以算答案了。

# 例题：CF1542E2（一）

首先，把枚举 $x,y$ 可以换成枚举 $x-y$。

对 $A,B$ 的不等式可以换成只枚举 $A,B$ 中的一个，然后另一个前缀和。

最后，可以把 $x-y$ 处的贡献拆开。

这样最终能做到 $O(n^3)$，可以通过。

# 例题：CF1542E2（二）

也可以不直接优化前面的做法，而是直接把后半部分换成一个 dp。

设 $dp(i,j)$ 表示有多少对长度为 $i$ 的排列，逆序对数之差为 $j$。

这样就很容易做到总共 $O(n^3)$，比之前的推导简单得多。

启示：合理选择 dp 和硬算式子。

# 例题：ARC102C

有一个可重集，大小为 $n$，每个元素都是 $[1,k]$ 的正整数。对于每个 $2 \le s \le 2k$，求出有多少种这样的可重集，使得不存在两个不同的数加起来是 $s$。对大质数取模。

$$n, k \le 2000$$

提示 1：基本思想：先列式，后计算。列式时，要学会分类、分步。

提示 2：善用递推。

# 例题：P6144

数轴上有 $n$ 条线段。定义一个 线段的子集 的价值 为 其形成的连通块个数的 $K$ 次方。

求所有子集价值之和。

$n \le 10^5, k \le 10$。

部分分：$k = 1$

# 展望

近年 CSP-S/NOIP 中的计数方向的题目，很多时候不仅是列式计算，还要对计数 dp 方法有较为清晰的认知，并且也同时考察了其他结构（例如 NOIP2021 的数位用到了数位 dp 的技巧，NOIP2022 的建造军营考察了缩点）。

# 由题目性质推导出突破口

2023 CSP-S 秋令营
feecle6418

洛谷
www.luogu.com.cn

# 前言

在前面的部分，我们已经讲了很多“一步步优化”而得到最终做法的题目。

事实上，不仅数学题可以一步步优化，其它算法的题目也可以凭借一步一步的组合观察，积少成多，最终推出整道题的解法。

而部分分，有时是很大的提示。题目中过于特殊的限制，有时也是提示。

（不过笔者翻阅了最近几年 NOIP 的部分分，提示作用都不能算很大）

# CF1584F：Strange LCS

有 $n$ 个字符串，每个字符串仅包含小写字母或大写字母。每个字符至多在每个字符串里出现两次。

求它们的 LCS。

$$n \le 10$$

# CF1584F：Strange LCS

普通 dp（记录每个字符串分别匹配到了哪里）算 LCS，很多状态其实是无用的。

如果只记录有用（能对答案造成新贡献）的状态，状态数量立刻缩减下去。

# P8339：[AHOI2022] 钥匙

有 $n$ 座城市，编号为 $1, 2, \ldots, n$。这些城市由 $n-1$ 条无向道路相连，每条无向道路连接两座城市，保证任意两个城市连通。即这 $n$ 座城市构成一棵树。

每座城市都有一件宝物。宝物分为两种：钥匙和宝箱。在一座城市里，要么有一把钥匙，要么有一个宝箱。钥匙和宝箱有不同的颜色，颜色为 $i$ 的钥匙只能打开颜色为 $i$ 的宝箱，打开宝箱后可以获得一枚金币，同时这把钥匙会损坏。

**由于某种特殊的原因，同一种的钥匙最多只有 5 把（同一种颜色的宝箱数量不限）。**

现在小 R 规划了 $m$ 次旅行，第 $i$ 次旅行的起点为 $s_i$，终点为 $e_i$。小 R 从 $s_i$ 沿最短路径走到 $e_i$。当他走到一座有钥匙的城市时，他可以将钥匙放入背包。当他走到一座有宝箱的城市时，如果他有相应颜色的钥匙，那么他就会打开这个宝箱并获得一个金币；如果他没有相应颜色的钥匙，那么他什么都不做（宝箱不能带走）。问每次旅行能获得多少枚金币。

**注意：旅行相互独立，即一次旅行完之后所有的钥匙和宝箱都会恢复到初始状态。**

部分分：每种颜色恰有一个钥匙和一个宝箱。

# P8339：[AHOI2022] 钥匙

如果钥匙宝箱都只有一个，可以算每一对“钥匙-宝箱”对询问的贡献。

拓展到多个的情况，仍然可以进行类似的讨论：钥匙 – 宝箱匹配的过程，可以看成括号匹配。

建虚树，从每个钥匙出发 dfs 一遍。

# P8496：[NOI2022] 众数

有 $n$ 个序列，初始为空，支持：push_back，pop_back，按顺序合并两序列，询问某些序列的并的绝对众数。

部分分：$n=1$；序列中只有 1,2；没有 pop_back。

# P8496：[NOI2022] 众数

部分分告诉了你每个组成部分需要用什么数据结构。

# CF1806F2： GCD Master

给定 $n$ 个正整数 $a_i$，分为 $K$ 组，使得每组 gcd 之和最大。$K$ 是给定的。

$$n \le 10^6, a_i \le 10^{18}$$

部分分：

1. $a_i \le 10^6$ 且互不相同
2. $a_i \le 10^6$

# CF1806F2： GCD Master

本题是循序渐进思考的典范。

*1.* $a_i \le 10^6$ 且互不相同

注意到若 $a, b$ 不同，则 $\gcd(a,b) \le \max(a,b)/2$。所以两个元素个数 $\ge 2$ 的集合，总不如合并成一个 $=1$ 的和一个 $\ge 2$ 的。

枚举 gcd，选择一些数的贡献为 $\sum a_i - \gcd$，所以选择最小的几个。

# CF1806F2： GCD Master

2. $a_i \le 10^6$。

注意到 $K$ 是给定的。

只保留互不相同的数，可以求出 $K = 1 \sim n$ 的所有答案。

只看对于相同数的合并，肯定优先合并小的。

把互不相同的数的答案和合并相同数的答案再合并一次。

# CF1806F2：GCD Master

3\. $a_i \le 10^{18}$。

注意到若 $a, b$ 不同，则 $\gcd(a,b) \le |a-b|$。这指引我们，gcd 的变化，基本上一定不如选来合并的值的变化。

如果选来合并的最大值为 $M$，当前 gcd 为 $g$，还未选来合并的最小值为 $m$，则把 $M$ 换为 $m$ 的代价不超过 $g-(M-m)$。若 $M-m>g$，则替换总是优的。

特别地，如果 $m<M'$（$M'$ 为选来合并的次大值），$M-m>g$ 一定成立。

# CF1806F2： GCD Master

3. $a_i \le 10^{18}$。

这说明选来合并的数除了最大值，一定是一个前缀。

最大值选哪个呢？只和 $\gcd(a_i, g) - a_i$ 有关。$g$ 只变化 $O(\log V)$ 次，并且 gcd 的复杂度还能均摊分析，总复杂度为 $O(n \log V)$。

# CF1806F2：GCD Master 总结

本题的关键是对 gcd 的性质的理解，以及不少“直观感受”。

大胆猜想、小心验证。

# P7116：[NOIP2020] 微信步数

小 C 喜欢跑步，并且非常喜欢在微信步数排行榜上刷榜，为此他制定了一个刷微信步数的计划。

他来到了一处空旷的场地，处于该场地中的人可以用 $k$ 维整数坐标 $(a_1, a_2, \ldots, a_k)$ 来表示其位置。场地有大小限制，第 $i$ 维的大小为 $w_i$，因此处于场地中的人其坐标应满足 $1 \le a_i \le w_i$（$1 \le i \le k$）。

小 C 打算在接下来的 $P = w_1 \times w_2 \times \cdots \times w_k$ 天中，每天从场地中一个新的位置出发，开始他的刷步数计划（换句话说，他将会从场地中每个位置都出发一次进行计划）。

他的计划非常简单，每天按照事先规定好的路线行进，每天的路线由 $n$ 步移动构成，每一步可以用 $c_i$ 与 $d_i$ 表示：若他当前位于 $(a_1, a_2, \ldots, a_{c_i}, \ldots, a_k)$，则这一步他将会走到 $(a_1, a_2, \ldots, a_{c_i} + d_i, \ldots, a_k)$，其中 $1 \le c_i \le k$，$d_i \in \{-1, 1\}$。小 C 将会不断重复这个路线，直到他走出了场地的范围才结束一天的计划。（即走完第 $n$ 步后，若小 C 还在场内，他将回到第 $1$ 步从头再走一遍）。

小 C 对自己的速度非常有自信，所以他并不在意具体耗费的时间，他只想知道 $P$ 天之后，他一共刷出了多少步微信步数。请你帮他算一算。

部分分：$k=1$；$k=2$；$w_i$ 很小。

# P7116：[NOIP2020] 微信步数

算贡献：算每个步数，有多少个起点能走到。总是一个高位长方体！

注意到长方体的边长变化类似等差数列。据此可以将式子拆开。

# P5024：[NOIP2018] 保卫王国

给定一棵树，每次强制两个点选 / 不选，询问最小权点覆盖。

部分分：树高很低，保证询问的其中一个点固定，保证询问的两个点相邻，保证是一条链。

# Thanks!

洛谷

www.luogu.com.cn