IOI2026 中国国家集训队集中培训 第三试 C. 解开尘封的序列

IOI2026 中国国家集训队工作组

2025 年 12 月 4 日

问题（解开尘封的序列（sequence））

- 给定长度为 n 的 d 位 p 进制整数序列 a。
- 给定长度为 n 的序列 w 和长度为 p^d 的序列 z。
- 定义 $\mathrm{popcount}_p, \mathrm{and}_p, \mathrm{or}_p, \mathrm{xor}_p$ 分别为 p 进制下的非零位个数、按位取最小值、按位取最大值、不进位加法。
- 对于所有 $0 \le u < p^d$，定义 u 的生成序列 $F(u)$ 如下：对于 $1 \le i \le n$，令 $b_i = A \cdot \mathrm{popcount}_p(a_i \mathrm{and}_p u) + B \cdot \mathrm{popcount}_p(a_i \mathrm{or}_p u) + C \cdot \mathrm{popcount}_p(a_i \mathrm{xor}_p u)$，然后令 $F(u) = \mathrm{sorted}([b_1, \dots, b_n])$。
- 有 q 次询问，每次询问给定 $A, B, C, l_1, r_1, l_2, r_2$，求

$$\left(\sum_{i=l_1}^{r_1} \sum_{j=l_2}^{r_2} z_i w_j F(i)_j \right) \bmod 2^{32}.$$

数据范围：$1 \le n, q \le 3 \times 10^5$。
子任务 $1 \sim 6$：$p = 2$，$d \le 12$；子任务 $7, 8$：$p = 3$，$d \le 5$。

$n \le 5000$，$q \le 5$

$n \le 5000$，$q \le 5$

- 枚举 i, j，暴力计算 $F(i)_j$。

$n \le 5000$，$q \le 5$

- 枚举 i, j，暴力计算 $F(i)_j$。
- 时间复杂度 $O(qp^d n \log n)$，期望得分：5。

$$p = 2$$

$p=2$

- 注意到

$$\operatorname{popcount}(a_i \operatorname{or} u) = \operatorname{popcount}(a_i \operatorname{and} u) + \operatorname{popcount}(a_i \operatorname{xor} u),$$

  故可以将原 b 序列转化为

$$b_i = (A+B)\cdot \operatorname{popcount}(a_i \operatorname{and} u) + (B+C)\cdot \operatorname{popcount}(a_i \operatorname{xor} u).$$

- 由于 $\operatorname{popcount}(x) \le d$，因此对于每一个 u，仅需考虑 $O(d^2)$ 个状态。预处理个数可以做到 $O(2^d d^3) \sim O(4^d)$ 等复杂度，均可以接受。

特殊性质 B：对于所有询问，均有 $l_1 = r_1$

特殊性质 B：对于所有询问，均有 $l_1 = r_1$

- 每次询问暴力将 $O(d^2)$ 个状态排序。
- 时间复杂度 $O(qd^2 \log d)$，期望得分：8。

特殊性质 A：所有询问的给出的三元组 (A,B,C) 均相同

特殊性质 A：所有询问的给出的三元组 (A,B,C) 均相同

- 每一个 u 对应 $O(d^2)$ 个长条加 z_i，共 $O(2^d d^2)$ 次长条加。
- q 次查询一个矩形内的各位置的权值点乘上 w_j 的和。
- 可以使用各种数据结构做到 $O(2^d d^2 \log n + q \log n)$，期望得分：11。

- 考虑将 $a_i \operatorname{and} u$ 与 $a_i \operatorname{xor} u$ 的取值对看作一个点 (x,y)，共 $O(d^2)$ 个点。
- 令 $A'=A+B$，$B'=B+C$。考虑两个点在排序中顺序的变化会在什么时候发生，即满足 $A'x_0+B'y_0 \le A'x_1+B'y_1$ 的 (A',B') 的限制。
- 整理得到 $A'(x_0-x_1) \le B'(y_1-y_0)$，因此实际上只关心 A'/B' 与 $(y_1-y_0)/(x_0-x_1)$ 这两个广义分数在 Stern-Brocot 序列上位置的先后关系。
- $(y_1-y_0)/(x_0-x_1)$ 将这个序列划分成了 $O(d^2)$ 段，其中每段内这 $O(d^2)$ 个点所对应的原式之间的大小关系是一样，故可以一起进行处理。

特殊性质 C：对于所有询问，均有 $l_1 = 0$ 且 $r_1 = p^d - 1$

特殊性质 C：对于所有询问，均有 $l_1 = 0$ 且 $r_1 = p^d - 1$

- 将 A', B' 两部分的贡献拆开，则每个段对应初始的若干次区间加。
- 询问是静态的区间查询，可以使用前缀和简单维护。
- 时间复杂度 $O(2^d d^4)$，期望得分：17。
- 实际上可以分析得更精细：由于 $A', B' \geq 0$，在 Stern-Brocot 序列中 $0/1$ 之前的分数无需考虑，之后只需要在所有分子分母互质的分数上添加一个划分，总段数为

$$\sum_{i=0}^{d}\sum_{j=0}^{d}[\gcd(i,j)=1]-1=2\sum_{i=1}^{d}\varphi(i),$$

当 $d = 12$ 时共 92 个段。

$n, q \le 3 \times 10^5$

$n, q \le 3 \times 10^5$

- 继续考虑特殊性质 C 的做法，将 A', B' 两部分的贡献拆开。
- 考虑按照求和的第一维，即 $0 \sim 2^d - 1$ 进行扫描线。
- 所有贡献均为区间加的形式，而所有查询均为区间带权求和，其中权值序列固定为 w。
- 由于修改次数为 $O(2^d d^4)$，查询次数为 $O(q)$，因此可以用 $O(1) - O(\sqrt{n})$ 的分块进行维护。
- 时间复杂度 $O(2^d d^4 + q\sqrt{n})$，期望得分：73。
- 若复杂度略高或常数较大可能只能获得 56 分。

$p = 3$

$p = 3$

- $p = 3$ 时，
  $$\operatorname{popcount}(a_i \operatorname{or} u) = \operatorname{popcount}(a_i \operatorname{and} u) + \operatorname{popcount}(a_i \operatorname{xor} u)$$
  不再恒成立。
- 此时对于每一个 u，有 $O(d^3)$ 个状态需要考虑，可以在 $O(3^d d^4) \sim O(9^d)$ 的时间复杂度内预处理。
- 类似 $p = 2$ 的性质，猜测本质不同的排序仍然很少。
- 考虑取出 $d = 5$ 时所有 popcount 可以取到的三元组，共 56 个。
- 抽样很多 (A, B, C) 统计排序，对排序结果哈希统计本质不同的个数。
- 可以发现，本质不同的排序只有 1242 个。

- 按照 $p=2$ 的方法，对 $0 \sim 3^d-1$ 一维扫描线，另一维数据结构维护即可。
- 由于修改操作的次数较少，将分块换为树状数组可能速度更快。两种算法均可通过。期望得分：27。
- 如果取出全部 216 个三元组进行排序统计，将得到 21678 个本质不同的序列；如果取出 $x \le y$ 的三元组进行排序统计，将得到 8044 个本质不同的序列。由于状态数过多，可能无法通过，但可能可以通过特殊性质 C。

- 以下为本质不同的排序数量的一个不是很紧的 bound。

定义

对于一个有序有限点集 $S=\{s_1,\ldots,s_n\}\subseteq\mathbb{R}^d$，令集合

$$x(S)=\{s_i-s_j|1\le i<j\le n\}.$$

对于向量 $a\in\mathbb{R}^d$，定义长度为 n 的排列 $p(S,a)$ 如下：

$$p(S,a)_i=\sum_{j=1}^{i}[d_j\le d_i]+\sum_{j=i+1}^{n}[d_j<d_i],$$

其中 $d_i=s_i\cdot a$。
定义 $\mathrm{per}(S)$ 为以下集合：

$$\mathrm{per}(S)=\left\{p(S,a)|a\in\mathbb{R}^d\right\}.$$

引理

$$|\mathrm{per}(S)| \le 2\sum_{i=0}^{d-1}\binom{|x(S)|-1}{i}.$$

证明.

考虑所有本质不同的 $p(S,a)$，只由所有不同的 $s_i \cdot a$ 间的大小关系决定，即由 $x(S)$ 中的元素与 a 的点积的正负性决定。这对应着过原点的 $d-1$ 维超平面 $s \cdot a = 0$ $(s \in x(S))$ 对 S 的一个划分，而 $S \subseteq \mathbb{R}^d$，因此 $|\mathrm{per}(S)|$ 不超过 $|x(S)|$ 个过原点的 $d-1$ 维超平面将 d 维空间划分出的区域数的最大值。

令 $D(d,n)$ 表示 n 个过原点的 $d-1$ 维的超平面将 d 维空间划分出的区域数的最大值。对于第 n 个 $d-1$ 维超平面，若不考虑第 n 个超平面，则有 $D(d,n-1)$ 个区域；若考虑对第 n 个超平面的一侧取交，则构成了一个 $d-1$ 维空间，仅有该空间的每个区域可在原来 $D(d,n-1)$ 个区域的空间中划分出一个新的区域，故 $D(d,n) \le D(d,n-1) + D(d-1,n-1)$。由边界条件 $D(1,n)=2$ 与 $D(d,1)=2$ 可得，

$$D(d,n) \le 2\sum_{i=0}^{d-1}\binom{n-1}{i},$$

则该引理得证。 $\square$

- 回到原问题，有 $|x(S)| = O(d^3)$，因此

$$|\mathrm{per}(S)| \le 2\sum_{i=0}^{2}\binom{|x(S)|-1}{i} = O(d^6).$$

- 故总时间复杂度为 $O(n+3^d d^9 + q\sqrt{n})$ 或 $O(n+3^d d^9 \log n + q\log n)$。

- 注意到 3^d 不大，因此每次询问时也可以直接枚举 i 分别计算贡献。
- 对于所有 1242 个排序预处理前缀和，每次询问时只需二分即可。
- 时间复杂度为 $O(n+3^d d^9+q3^d \log d)$。