

nth 解题报告

王又嘉

题目描述

Alice 和 Bob 正在与 Eve 玩一个游戏. Eve 有一个非负整数集(可重) S , 且所有数均小于 $M = 2^{21}$, 他将这个集合分为两部分 A 和 B , 并分别告诉了 Alice 和 Bob. 为了信息传输的方便, 他保证了每种数在一个部分内至多出现一次. 他给了 Alice 和 Bob 一个任务: 求出 S 中第 c 小的数. 但这过于简单, 因此 Eve 要求他们只能传输很少 bit 的信息.

数据范围

传输信息不超过 90 bits 可以获得满分.

解题过程

算法一

Alice 向 Bob 传输整个 bitset, 需要 2^{21} bit. Bob 可以得到完整集合计算出答案.

算法二

二分答案, Alice 每次用 21 bit 传输不超过 mid 的个数, Bob 计算出应该进入哪一边后告诉 Alice. 共计需要询问 21 次, 总计 $21 \times (21 + 1)$ bit.

算法二的改进

二分等价于依次确定每一位. Alice 从高到低枚举每一位, 设已经确定了一个前缀为 pre , 则 Alice 将前缀为 $pre + 0$ 的数的个数传输给 Bob, 则第 i 轮 Alice 只需要传输 $22 - i$ bit, 总计 $21 + \dots + 1 + 21$ bit.

算法三

我们尝试将问题简化. 不妨设 $a \leq b$, 两集合分别为 $A_1 \leq A_2 \leq \dots \leq A_a$ 和 $B_1 \leq B_2 \leq \dots \leq B_b$. 对 c 的大小分类讨论.

- 若 $1 \leq c \leq a$, 则原问题等价于在 $A_{1\dots c}, B_{1\dots c}$ 中求出第 c 小的数.
- 若 $a < c \leq b$, 则原问题等价于在 $\{0\} \cup A_{1\dots a}, B_{c-a\dots c}$ 中求出第 $a + 2$ 小的数.
- 若 $b < c \leq a + b$, 则原问题等价于在 $A_{c-b\dots a}, B_{c-a\dots b}$ 中求出第 $a + b - c + 2$ 小的数.

Alice 和 Bob 分别可以根据 `init` 得到的信息进行化简.

三种情况均满足: $|A| = |B|, c = |A| \vee c = |A| + 1$. 即: 两人集合大小相同, 且所求为低中位数或高中位数. 设 $n = |A| = |B|$. 接下来可以采取这样的策略转化为更小的子问题:

若 $n \leq 2$, 则一人向另一人发送所有数即可. 否则,

传输 Alice 的低中位数(第 $m = \lfloor \frac{n+1}{2} \rfloor$ 小)和 Bob 的高中位数(第 $M = \lceil \frac{n+1}{2} \rceil$ 小).

- 若 $A_m < B_M$, 则总的集合中不超过 B_M 的数至少有 A 中 m 个和 B 中 M 个共 $m + M = n + 1$ 个, 不小于 A_m 的数至少有 A 中 $n + 1 - m$ 个和 B 中 $n + 1 - M$ 个共

$n + 1$ 个. 因此高低中位数均落在 $[A_m, B_M]$ 内, 因此删去 $A_1 \dots m-1$ 和 $B_{M+1} \dots n$, 全集的高低中位数均不改变. 此时两人集合大小仍相同. 若原本求的是低(高)中位数, 删除后求的仍是低(高)中位数.

- 若 $A_m = B_M$, 则该数即为答案.
- 若 $A_m > B_M$, 则同理, 可删去 $A_{m+1} \dots n$ 和 $B_1 \dots M-1$.

这样每一轮操作可使 n 的问题转化为不超过 $\lceil \frac{n+1}{2} \rceil$ 的问题, 故至多 21 轮交换可使 $n \leq 2$, 总计 $(21 \times 21 \times 2 + 21 \times 2)$ bit.

算法三改进(正解)

首先容易证明即使集合 A 内有重复元素, 上述算法也是正确的, 只需给相同的数安排顺序.

注意到算法三的简洁性: 在传输中位数外不需要任何操作. 我们运用比较两个数只需要从高位开始比的性质.

我们有上述算法的推论: 若 A_m 和 B_M 均在区间 $[L, R]$ 内, 则答案也在 $[L, R]$ 内.

另有: 若已知答案在区间 $[L, R]$ 内, 则将集合内的数向 $[L, R]$ 取整, 即将 x 映射到
$$\begin{cases} L & x < L \\ x & L \leq x \leq R \\ R & x > R \end{cases}$$

不改变答案.

因此有如下算法:

从高位到低位依次比较两个数,

- 若这一位不同, 则可按上述算法删去一半的数.
- 若这一位相同, 则由推论, 可以将所有数取整到到该位的前缀都完全相同的前缀.

因此, 每比较一位, 要么可以将集合大小减半, 要么可以将值域减半. 因此只需 $2 \log M$ 轮即可简化至平凡情况. 总传输位数为 $4 \log M$.

参考资料

Eyal Kushilevitz and Noam Nisan. Communication Complexity. Cambridge University Press, New York, NY, USA, 1997. ISBN 0-521-56067-5.