

# 《学姐买瓜》解题报告

## 1. 题目大意

$n$  次操作，每次操作如下：

- $1 \ 1 \ r$  加入区间  $[l, r]$  满足  $1 \leq l \leq r \leq m$ .
- $2 \ 1 \ r$  询问区间  $[l, r]$  满足  $1 \leq l \leq r \leq m$ ，询问最大不相交区间数，且选出区间都位于  $[l, r]$  中。

我们称一个区间  $[l_1, r_1]$  位于  $[l_2, r_2]$  中当且仅当  $l_2 \leq l_1 \leq r_1 \leq r_2$ .

对于 20% 的数据，满足  $n, m \leq 2 \times 10^3$ .

对于 50% 的数据，满足  $n, m \leq 8 \times 10^4$ .

对于 100% 的数据，满足  $n, m \leq 3 \times 10^5$ .

## 2. 参考算法

下面“区间”和“线段”的含义相同，“下标”和“位置”的含义相同。

### 2.1 算法一

设询问区间  $[L, R]$ 。对所有位于区间中的线段进行排序，设这些线段是  $[l_1, r_1], [l_2, r_2], \dots, [l_t, r_t]$ ，满足  $r_i \leq r_{i+1} (1 \leq i < t)$ 。

**贪心**，从左到右扫描。假设当前在位置  $i$ ，之前选的最后一条线段是  $[tl, tr]$ ，如果  $l_i > tr$ ，则答案+1同时  $[tl, tr]$  变为  $[l_i, r_i]$ 。

初始化  $[tl, tr]$  为  $[L-1, L-1]$ 。

**证明：**

假设贪心选出的不是最优解：

设贪心选出来的下标是  $\{a_1, a_2, a_3, \dots, a_p\} (1 \leq a_1 < a_2 < \dots < a_p \leq t)$ ，最优解法中字典序（以排序后的下标作为字典序）最小的为  $\{b_1, b_2, \dots, b_q\} (1 \leq b_1 < b_2 < \dots < b_q \leq t)$ ，有  $p < q$ 。

设  $i$  是最小的位置满足在  $\{a\}$  中出现，在  $\{b\}$  中不出现或在  $\{a\}$  不出现，在  $\{b\}$  中出现。则对于  $j < i$ ， $j$  在  $\{a\}, \{b\}$  同时出现或者同时不出现。

如果存在  $k < i$  满足  $k$  在  $\{a\}, \{b\}$  中出现，且对于  $(k, i)$  的下标都不在  $\{a\}, \{b\}$  中出现，则  $[tl, tr] = [l_k, r_k]$ 。如果不存在  $k$ ，则  $[tl, tr] = [L-1, L-1]$ 。

观察贪心的选法， $i$  在  $\{b\}$  中出现而在  $\{a\}$  中不出现意味着  $l_i > tr$ ，所以一定也会在  $\{a\}$  中出现，矛盾。因此  $i$  一定是在  $\{a\}$  中出现而在  $\{b\}$  中不出现。

定义  $h$  满足  $(i, h)$  中的下标都不在  $\{b\}$  出现，且  $h$  在  $\{b\}$  中出现。因为  $p < q$ ，所以  $h$  一定存在。把  $\{b\}$  中的  $h$  替换成  $i$ ，其它位置不变，则可以形成一个新的选法  $\{b'\}$ 。

由  $i < h$ ，有  $r_i \leq r_h$  且  $l_i > tr$ ，则  $\{b'\}$  合法。于是我们得到了一个字典序更小的选法  $b'$ ，且选出区间个数不变。综上所述，贪心选出的就是最优解。

令  $lmax_i$  表示所有区间中右端点为  $i$  的区间左端点的最大值。如果不存在这样的区间， $lmax_i = 0$ 。询问的时候可以利用  $lmax$  方便的实现贪心。

时间复杂度  $O(nm)$ ，期望得分20。

## 2.2算法二

注意到如果一个线段  $[l_1, r_1]$  包含  $[l_2, r_2]$  且  $l_1 \neq l_2$  或  $r_1 \neq r_2$ ，那么删去  $[l_1, r_1]$  不会影响最优解。只要选了  $[l_1, r_1]$  就能将其替换为  $[l_2, r_2]$ 。

考虑分块。设块的大小为  $B$ ，对于块  $[p, p + B)$ ，我们保留  $[l_{max_i}, i]$  ( $p \leq i < p + B$ ) 当且仅当  $l_{max_i} > 0$  且  $\forall p \leq j < i, l_{max_j} < l_{max_i}$ ，则删去的区间都不会影响最优解。

设块内保留的区间是  $[l_1, r_1], [l_2, r_2], \dots, [l_t, r_t]$  ( $l_1 < l_2 < \dots < l_t, r_1 < r_2 < \dots < r_t$ )。

定义下标  $i$  的后继为最小的  $j$ ，满足  $l_j > r_i$ ，若不存在  $j$  则认为  $i$  没有后继。（这里就是贪心的算法）

对于每个下标  $i$ ，记录它可以走多少步（不停走到后继直到没有后继）以及它走到没有后继的位置（设为  $C_i$ ）。

每次加入区间进行  $O(B)$  的重构。查询时对散块暴力，从左到右对每个整块进行一次二分，找到最小的  $i$  满足  $l_i > tr$ （没有就跳过该整块），答案加上  $i$  的步数并令  $[tl, tr] = [l_{C_i}, r_{C_i}]$ 。

复杂度  $O(nB + n \times n/B \times \log B) = O(n\sqrt{n \log n})$ ，常数较小，期望得分50。

## 2.3算法三

从全局来维护包含关系： $[l_{max_i}, i]$  会保留当且仅当  $l_{max_i} > 0$  且  $\forall j < i, l_{max_j} < l_{max_i}$ 。用set维护。

设当前保留的区间是  $[l_1, r_1], [l_2, r_2], \dots, [l_t, r_t]$ 。

考虑建出一棵内向树。对于每个  $i$ ，如果  $\exists j, l_j = i$  连  $i \rightarrow r_j + 1$ ，边权为1；否则连  $i \rightarrow i + 1$ ，边权为0。

每次询问  $[L, R]$ ，令一变量  $pos$  初始位于  $L$ 。如果  $pos$  的父亲不超过  $R + 1$ ，那么将  $pos$  变为他的父亲  $fa$ ，同时答案加上  $pos \rightarrow fa$  的边权，重复此操作。

其实上面的过程就是算法一的贪心过程。

用LCT维护好加边、删边、查询路径和即可。复杂度  $O(n \log n)$ ，期望得分100。

## 3.命题总结

此题考察了贪心、分块、LCT等内容，正解的建树思路比较巧妙，且正解的代码长度约为2.2K，属于小型代码。考虑到考试的大多是集训队选手，此题属于简单题。

作者构思了这道题后，在与马耀华学长、谭靖宇同学的交流中会了LCT做法。同时，在与做题无数的claris交流后，发现IOI 2011的Elephants类似于此题  $r = l + k$  ( $k$  固定) 的版本，且该题正解的核心和本题一样，因此本题是改编题。

## 4.参考资料

IOI2011 elephants题面：<https://ioinformatics.org/files/ioi2011problem5.pdf>

IOI2011 elephants提交地址：<https://darkbzoi.tk/problem/2902>

IOI2011 范浩强参赛总结（含elephants简要题解）：[http://blog.sina.com.cn/s/blog\\_4ee63ce90102drjn.html](http://blog.sina.com.cn/s/blog_4ee63ce90102drjn.html)

感谢马耀华学长，谭靖宇同学，claris同学为本题提供的帮助。

感谢中国计算机协会、松鼠提供的评测平台。

