

# 全国青少年信息学奥林匹克竞赛

## CCF NOI 2022

### 第一试

时间：2022 年 8 月 23 日 08:00 ~ 13:00

题目名称	众数	移除石子	树上邻域数点
题目类型	传统型	传统型	交互型
目录	major	stone	count
可执行文件名	major	stone	count
输入文件名	major.in	stone.in	count.in
输出文件名	major.out	stone.out	count.out
每个测试点时限	1.0 秒	1.0 秒	3.0 秒
内存限制	1024 MiB	1024 MiB	2048 MiB
测试点数目	20	20	20
测试点是否等分	是	是	是

提交源程序文件名

对于 C++ 语言	major.cpp	stone.cpp	count.cpp
-----------	-----------	-----------	-----------

编译选项

对于 C++ 语言	-O2 -std=c++14 -static
-----------	------------------------

**注意事项（请仔细阅读）**

1. 文件名（程序名和输入输出文件名）必须使用英文小写。
2. C++ 中函数 main() 的返回值类型必须是 int，程序正常结束时的返回值必须是 0。
3. 因违反以上两点而出现的错误或问题，申诉时一律不予受理。
4. 若无特殊说明，结果的比较方式为全文比较（过滤行末空格及文末回车）。
5. 选手提交的程序源文件必须不大于 100KB。
6. 程序可使用的栈空间内存限制与题目的内存限制一致。
7. 只提供 Linux 格式附加样例文件。
8. 禁止在源代码中改变编译器参数（如使用 #pragma 命令），禁止使用系统结构相关指令（如内联汇编）和其他可能造成不公平的方法。

## 众数 (major)

### 【题目描述】

对于一个序列，定义其众数为序列中出现次数严格大于一半的数字。注意该定义与一般的定义有出入，在本题中请以题面中给出的定义为准。

一开始给定  $n$  个长度不一的正整数序列，编号为  $1 \sim n$ ，初始序列可以为空。这  $n$  个序列被视为存在，其他编号对应的序列视为不存在。

有  $q$  次操作，操作有以下类型：

- 1  $x\ y$ ：在  $x$  号序列末尾插入数字  $y$ 。保证  $x$  号序列存在，且  $1 \leq x, y \leq n + q$ 。
- 2  $x$ ：删除  $x$  号序列末尾的数字，保证  $x$  号序列存在、非空，且  $1 \leq x \leq n + q$ 。
- 3  $m\ x_1\ x_2 \dots x_m$ ：将  $x_1, x_2, \dots, x_m$  号序列顺次拼接，得到一个新序列，并询问其众数。如果不存在满足上述条件的数，则返回  $-1$ 。数据保证对于任意  $1 \leq i \leq m$ ， $x_i$  是一个仍然存在的序列， $1 \leq x_i \leq n + q$ ，且拼接得到的序列非空。**注意：不保证  $x_1, \dots, x_m$  互不相同，询问中的拼接操作不会对后续操作产生影响。**
- 4  $x_1\ x_2\ x_3$ ：新建一个编号为  $x_3$  的序列，其为  $x_1$  号序列后顺次添加  $x_2$  号序列中数字得到的结果，然后删除  $x_1, x_2$  对应的序列。此时序列  $x_3$  视为存在，而序列  $x_1, x_2$  被视为不存在，在后续操作中也不会被再次使用。保证  $1 \leq x_1, x_2, x_3 \leq n + q$ ， $x_1 \neq x_2$ ，序列  $x_1, x_2$  在操作前存在，且在操作前没有序列使用过编号  $x_3$ 。

### 【输入格式】

从文件 `major.in` 中读入数据。

输入的第一行包含两个正整数  $n$  和  $q$ ，分别表示序列的个数和操作的次数，保证  $n \leq 5 \times 10^5$ ， $q \leq 5 \times 10^5$ 。

接下来  $n$  行，第  $i$  行表示编号为  $i$  的序列。每一行的第一个正整数  $l_i$  表示初始第  $i$  号序列的数字个数，接下来有  $l_i$  个非负整数  $a_{i,j}$  按顺序表示序列中的数字。假定  $C_l = \sum l_i$  代表输入序列长度之和，则保证  $C_l \leq 5 \times 10^5$ ， $a_{i,j} \leq n + q$ 。

接下来  $q$  行，每行若干个正整数，表示一个操作，并按照题面描述中的格式输入。假定  $C_m = \sum m$  代表所有操作 3 需要拼接的序列个数之和，则保证  $C_m \leq 5 \times 10^5$ 。

### 【输出格式】

输出到文件 `major.out` 中。

对于每次询问，一行输出一个整数表示对应的答案，

### 【样例 1 输入】

```
1 2 8
2 3 1 1 2
3 3 3 3 3
4 3 1 1
5 3 1 2
6 4 2 1 3
7 3 1 3
8 2 3
9 3 1 3
10 1 3 1
11 3 1 3
```

**【样例 1 输出】**

```
1 1
2 3
3 -1
4 3
5 -1
```

**【样例 1 解释】**

第一次询问查询序列 1 的众数。由于序列包含两个 1，超过序列长度的一半，因此众数为 1。

第二次询问查询序列 2 的众数。由于序列只包含 3，因此众数为 3。

第三次询问查询序列 3 的众数。此时序列 3 为 (3,3,3,1,1,2)，不存在出现次数大于 3 次的数，因此输出为 -1。

**【样例 2 输入】**

```
1 4 9
2 1 1
3 1 2
4 1 3
5 1 4
6 3 4 1 2 3 4
7 1 1 2
```

```
8 3 2 1 2
9 2 3
10 3 3 1 2 3
11 1 4 4
12 1 4 4
13 1 4 4
14 3 4 1 2 3 4
```

**【样例 2 输出】**

```
1 -1
2 2
3 2
4 4
```

**【样例 2 解释】**

第一次询问查询序列 1, 2, 3, 4 拼接后得到的序列的众数。拼接的结果为 (1, 2, 3, 4), 不存在出现次数大于两次的数, 因此输出为 -1。

第四次询问查询序列 1, 2, 3, 4 拼接后得到的序列的众数。拼接的结果为 (1, 2, 2, 4, 4, 4, 4), 众数为 4。

**【样例 3】**

见选手目录下的 *major/major3.in* 与 *major/major3.ans*。

该样例满足测试点 1 ~ 3 的限制。

**【样例 4】**

见选手目录下的 *major/major4.in* 与 *major/major4.ans*。

该样例满足测试点 11 ~ 12 的限制。

**【数据范围】**

对于所有测试数据, 保证  $1 \leq n, q, C_m, C_l \leq 5 \times 10^5$ 。

$n, q$	$C_m, C_l$	测试点编号	特殊性质 A	特殊性质 B	特殊性质 C
$\leq 300$	$\leq 300$	1,2,3	否	否	是
$\leq 4,000$	$\leq 4,000$	4,5,6,7			
$\leq 10^5$	$\leq 10^5$	8	是	是	
		9		否	
		10	否	是	否
		11,12		否	是
		13			否
$\leq 5 \times 10^5$	$\leq 5 \times 10^5$	14	是	是	是
		15		否	否
		16	否	是	
		17,18		否	是
		19,20			否

特殊性质 A：保证  $n = 1$  且没有操作 4。

特殊性质 B：保证任意时刻任何序列中只有数字 1 和 2。

特殊性质 C：保证没有操作 2。

## 移除石子 (stone)

### 【题目描述】

你正在玩一个名为“移除石子”的小游戏。

有  $n$  堆石子排成一行，第  $i$  堆有  $a_i$  枚，你的任务是通过如下的操作将所有石子移除：

- 操作一：选择一堆石子，将其中的至少 2 枚石子移除；
- 操作二：选择一个连续的编号区间  $[l, r]$  ( $1 \leq l \leq r \leq n$ ) 并满足  $r - l \geq 2$ ，将其中的每一堆石子都恰好移除 1 枚。

你可以采用任意顺序执行任意多次上述两种操作，直到无法再执行为止。若最后你能将所有石子全部移除则胜利。

你或许已经开始计算起了诸如“有多少种本质不同的操作方式”的问题，但实际玩起来你却发现自己总是在输。因此，你打算玩个小花招：在游戏开始时，你在手里偷偷藏有  $k$  枚石子，在执行所有操作之前你可以且必须将这些石子放入某一堆或某几堆石子中。你期望这会提高自己的胜率，但也清楚这可能会使自己输掉原本可能胜利的游戏。

现在，你可以自由选择一个初始局面进行游戏，具体而言，每个  $a_i$  可以选择  $[l_i, r_i]$  范围内的任意整数。你希望计算出，在多少种初始局面下，自己存在至少一种获胜的方案。由于答案很大，你只需要输出其对  $(10^9 + 7)$  取模的结果。两个初始局面不同，当且仅当存在至少一个  $1 \leq i \leq n$  使得两者的  $a_i$  不相等，注意这里的“初始局面”指的是你放入  $k$  枚石子之前的局面。

### 【输入格式】

从文件 `stone.in` 中读入数据。

本题有多组测试数据。第一行一个正整数  $T$  表示测试数据组数，接下来依次给出每组测试数据。

对于每组测试数据，第一行两个整数  $n, k$ ，分别表示石子堆数和加入的石子个数，接下来  $n$  行，每行两个非负整数  $l_i, r_i$  表示每堆石子初始石子数的范围。

### 【输出格式】

输出到文件 `stone.out` 中。

对于每组数据输出一行一个整数，表示可能获胜的局面数对  $(10^9 + 7)$  取模的结果。

### 【样例 1 输入】

```
1 1
2 4 1
```

```
3 0 1
4 0 1
5 0 1
6 0 1
```

### 【样例 1 输出】

```
1 14
```

### 【样例 1 解释】

共有  $2^4 = 16$  种可能的初始局面，可以证明除了  $(0, 0, 0, 0)$  和  $(1, 0, 0, 1)$  这两种初始局面无法获胜以外，其余初始局面均存在获胜方案。例如，初始局面为  $(1, 0, 1, 0)$  时，你可以将手中的 1 枚石子放入第 2 堆石子，使局面变为  $(1, 1, 1, 0)$ ，再对区间  $[1, 3]$  使用一次操作二即可。

### 【样例 2】

见选手目录下的 *stone/stone2.in* 与 *stone/stone2.ans*。

### 【样例 3】

见选手目录下的 *stone/stone3.in* 与 *stone/stone3.ans*。

### 【样例 4】

见选手目录下的 *stone/stone4.in* 与 *stone/stone4.ans*。

### 【数据范围】

对于 100% 的数据，保证  $T \leq 10$ ， $3 \leq n \leq 1000$ ， $0 \leq l_i \leq r_i \leq 10^9$ ， $0 \leq k \leq 100$ 。

测试点编号	$n \leq$	$k \leq$	特殊条件
1 ~ 3	5	2	$r_i \leq 5$
4 ~ 5	1000	0	$l_i = r_i$
6 ~ 8		100	
9 ~ 11		0	无
12 ~ 13		2	
14 ~ 15		100	$r_i \leq 10$
16 ~ 20			无

## 树上邻域数点 (count)

这是一道交互题。

### 【题目描述】

给出五元组  $(T, I, S_V, S_E, \iota)$ , 其中:

- $T$  是一棵  $n$  个点的有根树  $T = (V, E)$ , 其中  $V$  为  $T$  的点集,  $E$  为  $T$  的边集。树的节点被编号为  $1, 2, \dots, n$ , 其中根节点编号为 1。
- $I$  是一个集合, 集合中的元素称作信息。其中有两个特殊元素: 单位元  $\epsilon$  和非法信息  $\perp$ 。

除以上特殊元素外, 一般信息具有点集合和边集合两个属性。特别的, 单位元只有边集合的属性, 非法信息  $\perp$  没有以上两种属性。

- 对于信息  $o \in I \setminus \{\epsilon, \perp\}$ ,  $o$  的点集合是  $V$  的一个二元子集, 记作  $S_V(o)$ , 满足  $S_V(o) \subseteq V$  且  $|S_V(o)| = 2$ 。其中, 两个集合  $A, B$  的差  $A \setminus B$  被定义为  $A \setminus B = \{x \in A \mid x \notin B\}$ 。
- 对于信息  $o \in I \setminus \{\perp\}$ ,  $o$  的边集合是  $E$  的一个子集, 记作  $S_E(o)$ , 满足  $S_E(o) \subseteq E$ 。规定单位元的边集合为空, 也即  $S_E(\epsilon) = \emptyset$ 。
- 对于树上的任何一条边  $e \in E$ , 记  $e = (u, v)$ , 存在一个关于  $e$  的信息  $\iota(e) \in I$ , 它以其端点为点集合、自身为边集合, 即  $S_V(\iota(e)) = \{u, v\}$ ,  $S_E(\iota(e)) = \{e\}$ 。

信息有两种合并的方式, 分别记作  $R$  和  $C$ 。对于  $\forall a, b \in I$ , 记  $r = R(a, b), c = C(a, b)$ , 满足  $r, c \in I$ , 则:

- 单位元和任何信息合并都得到对方。也即, 如果  $a = \epsilon$ , 那么  $r = c = b$ ; 如果  $b = \epsilon$ , 那么  $r = c = a$ 。
- 非法信息和任何信息合并都得到非法信息。也即, 如果  $a = \perp$  或者  $b = \perp$ , 那么  $r = c = \perp$ 。
- 对于剩下的情况, 如果两个信息的边集合的交集非空, 或者点集合的交集的大小不为 1, 则合并得到非法信息。也即, 如果  $S_E(a) \cap S_E(b) \neq \emptyset$  或  $|S_V(a) \cap S_V(b)| \neq 1$ , 则  $r = c = \perp$ 。
- 否则, 有

$$S_E(r) = S_E(c) = S_E(a) \cup S_E(b),$$

$$S_V(r) = S_V(a),$$

$$S_V(c) = S_V(a) \oplus S_V(b).$$

其中  $\oplus$  表示集合的对称差运算, 也即  $A \oplus B = (A \cup B) \setminus (A \cap B)$ 。



定义  $T$  中两个点的树上距离为树上以两个点为端点的唯一简单路径经过的边数。

给出评分参数  $M$  和  $q$  次询问，每次询问给出树上的一个点  $u$  和一个非负整数  $d$ 。记点集  $X$  为  $T$  中所有与  $u$  的树上距离不超过  $d$  的点构成的集合，又记边集  $Y = \{(a, b) \in E \mid a, b \in X\}$  为  $X$  内部的边集。可以证明，从  $\epsilon$  和所有  $\iota(e)$  ( $e \in E$ ) 出发，总是能通过有限次  $R, C$  的调用得到信息  $o \neq \perp$  满足  $S_E(o) = Y$ 。

每组询问中，你需要在  $R$  和  $C$  的调用次数总和不超过  $M$  的限制下构造出一个满足这样的要求的信息  $o$ 。特别地，如果  $d = 0$ ，则直接返回单位元  $\epsilon$  即可。

### 【实现细节】

请确保你的程序开头有 `#include "count.h"`。

头文件 `count.h` 中实现了如下内容：

1. 定义了信息对应的数据类型 `info`；
2. 定义了  $\epsilon$  所对应的 `info` 类型常量 `emptyinfo`，你可以在程序中直接使用。
3. 定义并实现了以下两个信息合并函数，你可以在程序中直接调用：

```
1 info MR(info a, info b);
2 info MC(info a, info b);
```

- 两个函数分别返回  $R(a, b)$  与  $C(a, b)$  对应的信息。

你需要保证调用  $R(a, b)$  或  $C(a, b)$  时结果不为  $\perp$ ，否则程序可能会出现异常行为。

4. 定义并实现了判定一个信息是否为单位元的函数，你可以在程序中直接调用：

```
1 bool isempty(info a);
```

- 这个函数返回真当且仅当  $a$  为单位元。

可以查看参考交互库了解更多实现细节。

你不需要，也不应该实现主函数。你需要实现如下几个函数：

```
1 void init(int T, int n, int q, vector<int> fa, vector<info> e, int M);
```

- $T$  表示测试点编号， $n$  表示树的点数， $q$  表示询问数， $M$  表示该测试点的评分参数。
- $fa$  和  $e$  的长度均为  $n - 1$ 。对于  $0 \leq i < n - 1$ ， $fa[i]$  和  $i + 2$  为第  $i$  条边  $e_i$  的两个端点， $e[i]$  为题目描述中提到的  $\iota(e_i)$  所对应的 `info` 类型元素。数据保证  $fa[i]$  小于  $i + 2$ 。

```
1 info ask(int u, int d);
```

- 给出一个询问，参数的意义见题目描述。你需要在函数结束时返回一个满足题设条件的信息。

最终测试时, 在每个测试点, 交互库会恰好调用一次 `init` 函数, 随后调用  $q$  次 `ask` 函数。交互库会使用特殊的实现方式, 单个 `info` 类型的变量会恒定消耗 12 字节内存, 这与下发的参考交互库不同。为保证程序运行时内存使用在题目限制内, 你需要保证运行过程中没有过多的 `info` 类型变量同时存在。

保证在满足调用次数限制且不进行 `isempty` 函数调用的情况下, 最终测试的交互库运行所需的时间不超过 0.6 秒, 交互库本身所消耗的内存不超过 16 MiB。保证在只执行  $10^8$  次 `isempty` 函数调用的情况下, 最终测试的交互库运行的时间不超过 0.25 秒。

在下发文件中包含一个名为 `count.cpp` 的文件, 作为示例程序, 选手可以在此基础上继续实现本题。在下发文件中还额外包含一个名为 `count_backup.h` 的备份文件, 我们保证其与 `count.h` 文件完全相同。

### 【测试程序方式】

本题目录下提供了两个交互库的参考实现 `grader.o`, `checker.o`, 其为两个不同的交互库编译产生的可链接文件。最终测试时所用的交互库实现与该实现有不同, 因此选手的解法不应依赖交互库的具体实现, 同时也不应该依赖 `count.h` 中 `info` 类型的具体实现。

你需要修改下发的 `count.h` 来帮助进行链接。具体的, 在将源代码 `count.cpp` 和程序 `grader.o` 进行链接的时候, 你需要注释掉 `count.h` 代码的第 5 行, 并保留第 4 行的代码。链接 `checker.o` 方法类似, 需要注释掉 `count.h` 代码的第 4 行, 并保留第 5 行的代码。选手可以对 `count.h` 的实现自行修改来实现不同程序的编译。

修改后, 选手可以在本题目录下使用如下命令编译得到可执行程序:

```
1 g++ count.cpp -c -O2 -std=c++14 -lm && g++ count.o grader.o -o  
   count  
2 g++ count.cpp -c -O2 -std=c++14 -lm && g++ count.o checker.o -o  
   count
```

其中第一行命令会编译当前 `count.cpp` 后与 `grader.o` 链接起来, 生成可执行文件 `count`, 第二行命令则会编译当前 `count.cpp` 后与 `checker.o` 链接起来, 生成可执行文件 `count`。

按上述方法编译得到的可执行文件 `count`, 其运行方式如下:

- 可执行文件将从标准输入读入以下格式的数据:
  - 第一行四个整数  $id, n, q, M$ , 分别表示测试点编号、树的点数、询问数和评分参数;
  - 第二行  $n-1$  个整数  $p_2, p_3, \dots, p_n$ , 分别表示 2 至  $n$  的父亲节点编号, 在本地调试时你需要保证  $\forall i \in [2, n], p_i < i$ ;
  - 接下来  $q$  行每行两个整数  $u, d$ , 描述一次询问。

- 读入之后，交互库会进行测试。如果你的程序不满足交互库限制，其会在输出中返回对应的错误信息。否则，对于链接的可执行文件，其输出如下：
  - 总共一行三个整数  $C_1, C_2, C_3$ ，其中：
    - \*  $C_1$  表示程序在 `init` 函数中调用交互库函数的总次数；
    - \*  $C_2$  表示程序在运行过程中调用交互库函数的总次数；
    - \*  $C_3$  表示程序在  $q$  次 `ask` 函数中调用交互库函数的次数的最大值。
    - \* 对于上述三个统计量，我们只会计入 `MR`、`MC` 函数的调用次数，而不会计入 `isempty` 函数的调用次数。
- 在链接不同文件的时候，其能够进行的检查也不同，具体的：
  - `grader.o`：其在运行时不会检查 `ask` 函数返回的信息是否正确，但可以帮助选手判断交互操作是否符合要求。这份程序运行时间最接近评测时的交互库，因此选手可以利用该程序测试运行速度，但不保证程序正确性。
  - `checker.o`：其在运行时检查 `ask` 函数返回的信息是否正确，也可以帮助选手判断交互操作是否符合要求。同时其会检查 `ask` 函数返回的信息是否正确。这份程序可以进行答案正确性的检查。

选手在调试时需要保证输入可执行文件 `count` 的数据满足上述输入格式，否则不保证输出结果正确。

### 【样例 1】

见选手目录下的 `count/count1.in` 与 `count/count1.ans`。

### 【样例 2】

见选手目录下的 `count/count2.in` 与 `count/count2.ans`。  
该组样例满足数据范围中的特殊性质 A。

### 【样例 3】

见选手目录下的 `count/count3.in` 与 `count/count3.ans`。  
该组样例满足数据范围中的特殊性质 B。

### 【样例 4】

见选手目录下的 `count/count4.in` 与 `count/count4.ans`。

**【评分方式】**

最终评测只会收取 `count.cpp`，修改选手目录下其他文件不会对评测结果产生影响。注意：

- 未初始化的 `info` 类型的变量不保证是 `emptyinfo`。
- 请不要尝试访问或修改 `info` 类型的成员变量，否则将被视为攻击交互库。
- 请不要在 `init` 函数调用之前调用 `MR` 和 `MC` 函数，否则可能发生未定义行为。
- 你只能访问自己定义的变量和交互库返回的 `info` 类型变量，尝试访问其他空间将可能导致编译错误或运行时错误。

本题首先会受到和传统题相同的限制，例如编译错误会导致整道题目得 0 分，运行时错误、超过时间限制、超过空间限制等会导致相应测试点得 0 分等。

在上述条件以外，在一个测试点中，若程序执行了非法的函数调用或询问操作中给出了错误回答，该测试点将会获得 0 分。否则，记  $C_1, C_3$  分别表示你的程序在 `init` 函数中调用交互库函数的次数，和你的程序在所有  $q$  次 `ask` 函数中调用交互库函数的次数的最大值。如果  $C_1 \leq 3 \times 10^7$  且  $C_3$  不超过该测试点的评分参数  $M$ ，你将获得该测试点的分数，否则你无法获得该测试点的分数。注意：计算  $C_1, C_3$  时只会计入 `MR`、`MC` 函数的调用次数，而不会计入 `isempty` 函数的调用次数。

**【数据范围】**

对于所有测试点， $1 \leq n \leq 2 \times 10^5$ ， $1 \leq q \leq 10^6$ ；每组询问中，有  $1 \leq u \leq n$ ， $1 \leq d \leq n - 1$ 。

测试点	$n =$	$q =$	特殊性质	$M =$
1	$10^3$	$10^4$		500
2	2,000			
3,4	$10^5$	$10^6$	A	5
5,6	$6 \times 10^4$	$6 \times 10^4$	B	50
7				5
8	$10^5$	$10^5$		
9	7,500	$5 \times 10^4$	C	500
10	$10^4$			
11	15,000			
12	$2 \times 10^4$			50
13	25,000			5
14	$3 \times 10^4$	$10^5$		
15	$6 \times 10^4$	D		
16				
17	$8 \times 10^4$	$10^6$		
18	$10^5$			
19	$1.5 \times 10^5$			
20	$2 \times 10^5$		1	

特殊性质 A：保证  $\forall i \in [1, n-1]$ ，编号为  $i+1$  的点的父节点为  $i$ 。

特殊性质 B：保证所有询问均满足  $u = 1$ 。

特殊性质 C：保证所有询问均满足  $d \leq 100$ 。

特殊性质 D：保证所有询问均满足  $d \geq 1000$ 。