

P2058 海港

NOIP 普及组 2016

题目描述:

lester是一个海港的海关工作人员，每天都有许多船只到达海港，船上通常有很多来自不同国家的乘客。lester对这些到达海港的船只非常感兴趣，他按照时间记录下了到达海港的每一艘船只情况；对于第 i 艘到达的船，他记录了这艘船到达的时间 t_i (单位：秒)，船上的乘客数 k_i ，以及每名乘客的国籍 $x(i,1), x(i,2), \dots, x(i,k)$;

lester统计了 n 艘船的信息，希望你帮忙计算出以每一艘船到达时间为止的24小时(24小时=86400秒)内所有乘船到达的乘客来自多少个不同的国家。

形式化地讲，你需要计算 n 条信息。对于输出的第 i 条信息，你需要统计满足 $t_i - 86400 < t_p \leq t_i$ 的船只 p ，在所有的 $x(p,j)$ 中，总共有多少个不同的数

输入输出格式:

输入：第一行输入一个正整数 n ，表示lester统计了 n 艘船的信息。

接下来 n 行，每行描述一艘船的信息：前两个整数 t_i 和 k_i 分别表示这艘船到达海港的时间和船上的乘客数量，接下来 k_i 个整数 $x(i,j)$ 表示船上乘客的国籍

保证输入的 t_i 是递增的，单位是秒；表示从lester第一次上班开始计时，该船在第 t_i 秒到达
 $1 \leq n \leq 10000$, $1 \leq t_i \leq 10^9$, k_i 的总和不超过300000, $1 \leq x(i,j) \leq 100000$

输出：输出 n 个整数表示第 i 艘船到达后的统计信息

样例输入:

```
3
1 4 4 1 2 2
2 2 2 3
10 1 3
```

样例输出:

```
3 4 4
```

破题

→ 从时间轴上来看，本质上是一个区间统计问题

统计各长度为86400的区间上的信息（各国籍的人数）

→ 区间问题

随机统计：线段树（CS103）

有序统计：扫描线思路（CS200）

稀疏情况：离散化（CS102）



建模

→ 哪些是关键时间点？

t_i (到达)

$t_i + 86400$ (离开)

虽然题目没有说离开，但是你可以理解为24小时候这船人就离开了

→ 对于这样的两类事件的模拟，普遍的方法是用优先队列

(还记得银行排队问题么)

→ 但是此题比较特殊，可以更简化一些

因为逗留（办理）的时间是固定的，也就是离开与到达的顺序一致
这样子就不需要维护优先队列了，因为随时知道下一个离开的是谁

解答

```
#define DAY 86400
#define MAX 100000
```

```
using namespace std;
int n, exist[MAX], t[MAX];
vector<int> x[MAX];
```

→ 复杂度: $O(n+k)$

```
int main(){
    cin>>n;
    memset(exist,0,sizeof(int));
    int tot=0,past=0,a,k;
    t[0]=0;
    for(int i=1;i<=n;i++){
        cin>>t[i]>>k;
        for(int j=0;j<k;j++){ // 到达
            cin>>a; // 依次输入国籍
            x[i].push_back(a);
            exist[a]++;
            tot+=exist[a]==1; // 多一个新国家
        }
        for(;t[i]-t[past]>=DAY;past++){ // 离开
            for(int j=0;j<x[past].size();j++){
                exist[x[past][j]]--;
                tot-=exist[x[past][j]]==0;
            }
        }
        cout<<tot<<endl;
    }
    return 0;
}
```

扫描线

- 计算过程可以看做一个延时间轴移动的辅助线
- 依次处理遇到的事件
- 这种思路称为扫描线

一维情况下，扫描线实现都是比较简单的（排序，递推，优先队列）

二维情况下，还有更复杂的扫描线算法（CS103+）

