
字节跳动 2018 校招后端方向（第二批）

1. [编程题]用户喜好

时间限制：C/C++ 3 秒，其他语言 6 秒

空间限制：C/C++ 256M，其他语言 512M

为了不断优化推荐效果，今日头条每天要存储和处理海量数据。假设有这样一种场景：我们对用户按照它们的注册时间先后来标号，对于一类文章，每个用户都有不同的喜好值，我们会想知道某一段时间内注册的用户（标号相连的一批用户）中，有多少用户对这类文章喜好值为 k 。因为一些特殊的原因，不会出现一个查询的用户区间完全覆盖另一个查询的用户区间（不存在 $L1 \leq L2 \leq R2 \leq R1$ ）。

输入描述：

输入：第 1 行为 n 代表用户的个数 第 2 行为 n 个整数，第 i 个代表用户标号为 i 的用户对某类文章的喜好度 第 3 行为一个正整数 q 代表查询的组数 第 4 行到第 $(3+q)$ 行，每行包含 3 个整数 l, r, k 代表一组查询，即标号为 $l \leq i \leq r$ 的用户中对这类文章喜好值为 k 的用户的个数。数据范围 $n \leq 300000, q \leq 300000$ k 是整型

输出描述：

输出：一共 q 行，每行一个整数代表喜好值为 k 的用户的个数

输入例子 1:

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

输出例子 1:

```
1
0
2
```

例子说明 1:

样例解释:

有 5 个用户，喜好值分别为 1、2、3、3、5，

第一组询问对于标号 $[1, 2]$ 的用户喜好值为 1 的用户的个数是 1

第二组询问对于标号 $[2, 4]$ 的用户喜好值为 5 的用户的个数是 0

第三组询问对于标号[3,5]的用户喜好值为 3 的用户的个数是 2

2. [编程题]手串

时间限制：C/C++ 1 秒，其他语言 2 秒

空间限制：C/C++ 64M，其他语言 128M

作为一个手串艺人，有金主向你订购了一条包含 n 个杂色串珠的手串——每个串珠要么无色，要么涂了若干种颜色。为了使手串的色彩看起来不那么单调，金主要求，手串上的任意一种颜色（不包含无色），在任意连续的 m 个串珠里至多出现一次（注意这里手串是一个环形）。手串上的颜色一共有 c 种。现在按顺时针序告诉你 n 个串珠的手串上，每个串珠用所包含的颜色分别有哪些。请你判断该手串上有多少种颜色不符合要求。即询问有多少种颜色在任意连续 m 个串珠中出现了至少两次。

输入描述：

第一行输入 n, m, c 三个数，用空格隔开。 $(1 \leq n \leq 10000, 1 \leq m \leq 1000, 1 \leq c \leq 50)$ 接下来 n 行每行的第一个数 $num_i (0 \leq num_i \leq c)$ 表示第 i 颗珠子有多少种颜色。接下来依次读入 num_i 个数字，每个数字 x 表示第 i 颗柱子上包含第 x 种颜色 $(1 \leq x \leq c)$

输出描述：

一个非负整数，表示该手链上有多少种颜色不符需求。

输入例子 1:

```
5 2 3
3 1 2 3
0
2 2 3
1 2
1 3
```

输出例子 1:

```
2
```

例子说明 1:

第一种颜色出现在第 1 颗串珠，与规则无冲突。

第二种颜色分别出现在第 1, 3, 4 颗串珠，第 3 颗与第 4 颗串珠相邻，所以不合要求。

第三种颜色分别出现在第 1, 3, 5 颗串珠，第 5 颗串珠的下一个是第 1 颗，所以不合要求。

总计有 2 种颜色的分布是有问题的。

这里第 2 颗串珠是透明的。

3. [编程题]字母交换

时间限制：C/C++ 1 秒，其他语言 2 秒
空间限制：C/C++ 32M，其他语言 64M

【编码题】字符串 S 由小写字母构成，长度为 n 。定义一种操作，每次都可以挑选字符串中任意的两个相邻字母进行交换。询问在至多交换 m 次之后，字符串中最多有多少个连续的位置上的字母相同？

输入描述：

第一行为一个字符串 S 与一个非负整数 m 。($1 \leq |S| \leq 1000, 1 \leq m \leq 1000000$)

输出描述：

一个非负整数，表示操作之后，连续最长的相同字母数量。

输入例子 1:

abcbaa 2

输出例子 1:

2

例子说明 1:

使 2 个字母 a 连续出现，至少需要 3 次操作。即把第 1 个位置上的 a 移动到第 4 个位置。所以在至多操作 2 次的情况下，最多只能使 2 个 b 或 2 个 a 连续出现。

4. [问答题]

题目描述

以下函数使用二分查找搜索一个增序的数组，当有多个元素值与目标元素相等时，返回最后一个元素的下标，目标元素不存在时返回-1。请指出程序代码中错误或不符最佳实践的地方（问题不止一处，请尽量找出所有你认为有问题的地方）

```
int BinarySearchMax(const std::vector<int>& data, int target)
```

```
{
```

```
    int left = 0;
```

```
    int right = data.size();
```

```
    while (left < right) {
```

```
int mid = (left + right) / 2;

if (data[mid] <= target)

    left = mid + 1;

else

    right = mid - 1;

}

if (data[right] == target)

    return right;

return -1;

}
```

5. [问答题]

题目描述

【设计题】今日头条会根据用户的浏览行为、内容偏好等信息，为每个用户抽象出一个标签化的用户画像，用于内容推荐。用户画像的存储、高并发访问，是推荐系统的重要环节之一。现在请你给出一个用户画像存储、访问方案，设计的时候请考虑一下几个方面：

用户画像如何存储

如何保证在线高并发、低延迟地访问

机器宕机、负载均衡问题

如果用户增长很快，在你的方案下，该如何做扩容