

注意事项

1. 普通PC（包括我们的OJ服务器）每秒的基本运算次数 $X * 10^8$ ，在绝大多数情况下，估算运行时间只需要考虑数量级，不考虑前面的常数。这一点常用于估算程序运行时间和估算算法所需的复杂度
2. 语言是整数是32位的，其取值范围为 $[-2147483648, 2147483647]$
3. 比较大的数组（如超 10^5 规模）要开在全局变量空间中
4. 碰到**WA (Wrong Answer)**可以自己尝试动手生成一些简单到复杂的样例，测试自己的程序。具体可以通过对stdin和stdout进行重定向实现，如 `freopen("test.in", "r", stdin);` 和 `freopen("test.out", "w", stdout);`

H2-1 抽奖

某次晚会，主办方给观众准备了一个抽奖活动，抽奖规则如下：

主持人会在大屏幕上随机产生一个数字 k ，然后观众通过手机在抽奖系统中输入一个0到10000之间的数字（包括0和10000），此时中奖观众将会是所有观众中输入的第 k 小的人，请输出获奖观众选择的数字。

$$0 < k \leq n \leq 10^7$$

$$0 \leq a_i \leq 10,000$$

10^7 的数据量意味着必须使用线性时间的算法

解决方法：利用快速排序中Partition函数，如果随机到的数排序后刚好在第K个位置上，则直接输出；否则，由于比中心值小的都在左边，比中心值大的都在右边，根据K的位置进一步缩小需执行Partition的空间。

最终复杂度 $n + n/2 + n/4 + \dots + 1 = O(n) < O(n \log n)$

H2-2 数组排序

给定一个长为 n 的int类型数组 A ，将该数组排序。

$$1 \leq n \leq 10,000,000$$

$$0 \leq a_i \leq 10^9$$

使用 $O(n \log n)$ 的排序算法即可，略

E1-1 超过一半的数

首先输入一个正整数 T ，代表存在 T 组样例。

接下来共 T 组样例，每组样例中首先输入 N ，下一行 N 个正整数。

在每组样例中你需要找到是否存在一个“超过一半的数” x （即存在 x 在这 N 个数中出现次数 $\geq (N + 1)/2$ ），若存在这样的 x 请输出 x ，否则输出 -1 。

$$1 \leq N \leq 10^6$$

最简单的想法：

先排序再统计每个数字出现的次数

其他想法：

把数字当key进行哈希，统计各个数字出现的次数

E1-2 排序

首先输入两个正整数 N 和 T

接下来输入一串只包含数字且长度为 N 的字符串，你需要将字符串按 T 分隔开并删除 T ，将剩余字符串认为数字，并将数字进行从小到大排序后输出。

在字符串转换为数字的过程中，需要去除前导0。若字符串全为0则视作0。

如输入"

8 1

01210213

"

首先从字符串"01210213"中剔除"1"，变为"0 2 02 3"，按从小到大排序的结果为"0 2 2 3"

被 T 分割形成的数字大小不超过 10^8

解体思路：

字符串处理，耐心调整自己的算法，使算法可以处理好好每一种情况即可

包括但不限于如下情况

```
7 0
0102030
```

```
7 1
1213141
```

```
7 1
1023121
```

```
7 1
2312410
```

E1-3 逆序对

输入一个长度为 N 的数组，你要求出数组中逆序对的数量（答案对10,000,007取模）

逆序对即对于数组 A ，满足 $i < j$ ， $A_i > A_j$ 的 (i, j) 对

因为输出值可能非常大，请将输出的答案模10,000,007

$$0 < N \leq 1,000,000$$

$$0 \leq A_i \leq 10^8$$

解题思路

1. 逆序对的数目，即对这个数组进行**稳定排序**时进行交换的次数，可以在归并排序的过程中加入交换次数的统计即为答案
2. 并不是在最后对答案进行取模，而是在过程中对答案的每一步加、乘操作进行取模，否则答案可能超出int范围

E1-4 军训排队

现有 n 个学生排成一个固定队伍进行军训，教官小明有一份所有 n 个人的名单（不同的人可能重名）。小明想要在整个队伍中找到一个连续的子队伍，并且满足该子队的所有人恰好有 k 个不重复的名字。请帮小明计算一下一共有多少种可能的子队伍。

$$0 < k < n \leq 10,000,000$$

【输入】

```
5 2
yyyyy iiiii yyyyy iiiii sssss
```

【输出】

```
7
/*
一共7种可能的子队伍为如下：
yyyyy iiiii
iiiiii yyyyy
yyyyy iiiii
iiiiii sssss
yyyyy iiiii yyyyy
iiiiii yyyyy iiiii
yyyyy iiiii yyyyy iiiii
*/
```

解题思路

1. 滑动窗口
2. 问题在于如何高效统计窗口内不重复字符串的个数，且在窗口移动过程中也能高效的增删？**哈希**、**STL中的Map**

E1-EX index K-1

在 T 组数组中，对于每个数组，仅求对该数组从小到大排序后index为 $K - 1$ 的数（index从0开始）

如 $K = 3$ 对数组 3 1 5 9，即求1 3 5 9中index为2的数，为5

$$5 \leq T \leq 100$$

$$1 \leq K \leq N \leq 10^5$$

考虑时间复杂度，是否可以在 $O(n)$ 的时间内求得第K大的数？（提示：快速排序）

解题思路：

与H2-1类似，略