

# 《数据与算法》课程实验

## 实验六 哈希表，大量文件查找重复对

### 一、 实验内容

用字符串代替文件，给出大量字符串，要求找出所有两两完全相同的字符串。

### 二、 输入格式

第1行：给出字符串的个数 $N(N < 10^7)$ ；

第2~( $N + 1$ )行：每行一个字符串，每个字符串的长度不超过 $10^6$ 个字符，为了方便起见，字符串中只包含  $a \sim z$  的26个小写字母。每个字符串依次从0到 $N - 1$ 对应一个原始编号，在输入中不会体现，但会在输出中用到。

### 三、 输出格式

若存在两两相同的字符串，依次在每行输出相同字符串所对应的原始编号。具体的：在同一行中，输出的原始编号应当不小于2个，对应的字符串应当相同，并且要按照原始编号从小到大的顺序输出，原始编号之间以空格作为间隔；在不同行之间，输出的编号对应不同的字符串，且每行的第一个原始编号之间也要按照从小到大的顺序输出。

若不存在两两相同的字符串，输出-1。

### 四、 输入输出样例

输入 1:

```
6
string
hash
c
hash
string
hash
```

正确输出 1:

```
0 4
1 3 5
```

错误输出 1.1:

```
1 3 5
0 4
```

错误输出 1.2:

```
4 0
1 3 5
```

输入 2:

```
2
string
hash
```

正确输出 2:

```
-1
```

## 五、 实验要求

- 1) 用C或C++实现，在LambdaOJ上进行提交；
- 2) 共10个测试样例，每个测试样例10分，满分100分；
- 3) 提交次数不限，取有效成绩最高的一次；
- 4) 坚决反对作弊行为，若出现代码雷同，两个人的相应实验都记为0分处理。
- 5) 本次实验截止时间为2015年11月19日23:59，超过截止日期，每迟交一天，有效成绩衰减10%。