

# 消失的源代码 解题报告

东营市胜利第一中学 姜志豪

## 1 试题来源

2016年集训队互测

## 2 试题大意

给出只能对小数据进行计算的可执行程序，给出输入文件，求输出文件。  
共10种不同的数据。

## 3 算法介绍

我们并不知道每一个测试点的输入和输出的关系，所以需要先确定它们之间的关系。

通过输入文件，可以得出输入格式。构造一些特殊的小数据，就能够发现输入与输出的关系。然后写出相应的程序，求出输出文件。

### 3.1 测试点1,2,3

这三个测试点的共同点是，代码实现非常简单，关键在于发现输入与输出之间的关系。

#### 3.1.1 测试点1

输入和输出是等长的小写字母串。当我们修改输入字符串的某个位置时，输出字符串也只是同样的位置发生了改变。并且，输入串中两个位置的字符相同时，输出串中对应位置的字符也相同。

我们猜测，存在一种小写字母到小写字母的映射 $f(x)$ ，满足输入串中的字符 $x$ 变成 $f(x)$ 后就是输出串。

对于每一个小写字母 $x$ ，都求出 $f(x)$ ，就可以由输入串求出输出串了。

### 3.1.2 测试点2

输入和输出都是一个整数，分别用 $x$ 、 $y$ 表示。第一组数据 $x = 20$ ，可以用lost求出 $x = 1 \sim 20$ 时的 $y$ 。

可以发现，多数情况下， $y$ 随着 $x$ 的增加而增加，但有时 $y$ 会从一个很大的数变成一个很小的数。我们猜测是进行了取模操作。 $x = 1 \sim 10$ 时，没有进行过取模操作，我们先只对这一段进行分析。

$x$ 越大， $x$ 增加1后 $y$ 的增加量越大，我们对 $y$ 进行差分，寻找 $y$ 的增加量有什么规律。我们发现 $y$ 的增加量是一个等差数列，所以 $y$ 是一个关于 $x$ 的二次函数值，可以求出它们之间的关系是 $y = 2016x^2 + 4x + 10$ 。

如果一开始就猜测到 $y$ 是关于 $x$ 的多项式函数值，就不必进行以上分析，可以直接解出它们之间的关系了。

现在，我们也能求出 $x = 11 \sim 20$ 时 $y$ 的值了，可以求出模数是233333。

### 3.1.3 测试点3

输入和输出都是一个整数，分别用 $x$ 、 $y$ 表示。第一组数据 $x = 50000$ ，此时 $y$ 仅为126。并且，随着 $x$ 的增加， $y$ 单调不下降。

我们可以求出 $y$ 在什么时候发生变化。或者更方便一些，只求 $y$ 什么时候从9变成10，从99变成100。因为10和100比较特殊，更容易发现规律。

$x$ 从314变成315时， $y$ 从9变成10。 $x$ 从31415变成31416时， $y$ 从99变成100。我们联想到 $\pi$ 。

我们猜测 $x$ 是圆的面积， $y$ 是圆的半径下取整。通过尝试其他输入，可以验证我们的猜测是正确的。

## 3.2 测试点4,5,6

这三个测试点的共同点是，发现需要输出什么之后，需要用数据结构维护信息来求解。不过即使是暴力程序，也能够在此时间内运行结束。

### 3.2.1 测试点4

输入是一个图，输出是一个整数。

输入是完全图、树等连通图时，输出就是点数的平方。

input中的第一组输入数据不是连通图，它的输出是每个连通块点数的平方和。

再尝试其他输入，可以确定：输出是每个连通块点数的平方和。

可以通过并查集或对图进行dfs求解。

### 3.2.2 测试点5

输入是一棵有边权的树，多次询问，每次询问是一条树上的路径，然而输出只有一个数。

我们把询问改为一次，发现询问的是路径长度。

再把询问改为两次，发现输出的是两次询问答案的异或值。询问改为多次，输出仍然是异或值。

所以，询问是路径长度，输出是每次询问结果的异或和。

可以写树上倍增或树链剖分线段树。不过既然是提交答案题，当然也可以直接写暴力。

### 3.2.3 测试点6

和测试点5类似，输入是一棵有边权的树，多次询问，每次询问一条路径，输出一个数。

通过尝试，可以发现：询问是路径上最小的边权，输出是每次询问结果的异或和。

把测试点5的程序稍微改一改就可以了。

## 3.3 测试点7,8,9

这三个测试点的共同点是，发现需要输出什么之后，需要使用合适的算法求解，暴力程序难以在较短时间内运行结束。

### 3.3.1 测试点7

输入只有两个数 $n$ 、 $m$ ，输出一个数。

可以想到，输出是某种二维前缀和。也就是输出为

$$\sum_{i=1}^n \sum_{j=1}^m f(i, j)$$

对输出进行差分，发现 $f(i, j) = \gcd(i, j)$ 。

也就是说，求的是

$$\sum_{i=1}^n \sum_{j=1}^m \gcd(i, j)$$

可以通过容斥求解。

### 3.3.2 测试点8

输入一个长度为 $n$ 的序列，输出一个整数。

通过在序列结尾增删数字、构造特殊序列等方法，可以发现，输出的是本质不同的非空连续子序列数。

可以用后缀数组等字符串算法求解。

### 3.3.3 测试点9

输入 $n$ 个数对，输出一个保留三位的小数。

$n = 1$ 时，输入无效。 $n = 2$ 时，输出是输入的两个点的距离。

尝试其他数据，观察发现，输出是平面上最近点对的距离。

可以通过分治求解。

## 3.4 测试点10

输入是一个字符串。无论输入是什么，输出都是“invalid input!”。题目保证，input中的第一组数据是可以直接用lost程序求出的。所以，我们大胆猜测，这个测试点的输出就是“invalid input!”。提交试一试，是能够通过这个测试点的所有数据的。

其实可以发现，其他测试点的输入规模过大时，输出是“invalid input.”。而这个测试点的输出是“invalid input!”。这就足以证明“invalid input!”是正确输出而不是无效输入时的输出。

#### 4 预期得分

在这道题上用的时间越多，得到的分数就会越多。

预计平均分50 ~ 60分，最高分80 ~ 100分。