

# 丢失的题面 试题讨论

大连市第二十四中学 于纪平

## 摘 要

2015年集训队互测第五场中，一道非传统题的命题思路和题目分析。

## 1 题目大意

传统题，不给出题面，只给出10组输入输出数据。

时间限制1秒，内存限制128M，没有Special Judge。

## 2 总体分析

### 2.1 命题契机

近年来，信息学竞赛的题目类型变得多种多样，被选手们称为“传统题十合一”的逐测试点分析的非传统题也连续在两年的冬令营中出现（WC2014非确定机，和WC2015未来程序），引发了大家很大的关注。

分析这两道题目的本质，“非确定机”是给出程序和输出，求任意一个可能的输入，“未来程序”是给出程序和输入，求它的输出。两道题也有各自的难点，例如“非确定机”的程序是通过黑箱给出的，需要选手手动分析；“未来程序”的输入会使程序运行的非常慢，需要选手做很大的分析优化。

注意我们的三个要素：输入、程序和输出，给出任意两个都可以让选手求出剩下的一个。“非确定机”求的是输入，“未来程序”求的是输出，这道“丢失的题面”就是给出输入和输出求程序。

## 2.2 得分估计

这道题目虽然以传统题的形式评测，其本质却是需要对每组输入输出数据分别分析，应该称之为非传统题。

在这道题上花费的时间越多，得到的分数就会越多，平均每10分钟应该就能得到10分。

## 2.3 集训队互测得分情况

1人90分，3人70分，2人60分，2人50分，1人20分，2人10分，略低于预期情况。

## 2.4 UOJ镜像得分情况

互测结束以后，在Universal Online Judge上进行了镜像比赛，向所有人开放，得分情况如下：2人90分，4人80分，3人70分，2人60分，1人50分，1人40分，4人30分，1人20分，5人10分。

## 2.5 得分总结

对于实际的情况，平均每15~30分钟应该就能得到10分。这也能看出出题人对于分数的估计很可能高于实际情况。

# 3 题目解法及命题思路

## 3.1 测试点1,2,3

这三个测试点的共同点是，输入文件只有一个整数，输出文件是一个二进制或三进制串。

这些测试点的主要命题思路是考察选手的找规律能力和找到正确的分析方法（例如，观察输出串的长度和含有各数字的比例等）。

### 3.1.1 测试点1

输入是22，输出是长度为 $2^{22}$ 的二进制串，其中0和1各占一半。

从头开始读容易发现明显的规律，从而得出输出串的构造方式：初始串是”0”，每次把串取反接在尾部。

### 3.1.2 测试点2

输入是33，输出是长度为3524578的二进制串，其中0和1的比例大约是黄金分割比。

同时注意到3524578是第33个斐波那契数，可知这是斐波那契字符串：初始串是”0”，每次扫一遍串，把”0”变成”1”，把”1”变成”01”。

### 3.1.3 测试点3

输入是12，输出是长度为531452的三进制串，三种数出现的次数大致相同，但是0稍微多一点。

注意到 $531452 = 3^{12} + 11$ ，经观察发现输出串包含了所有的长度为12的三进制串作为子串且是字典序最小的。

按照这个规则暴搜+剪枝就可以通过了。

## 3.2 测试点4,5,6

这三个测试点的共同点是，输入文件和输出文件的格式相同，都是第一行一个数 $n$ ，接下来 $n + 1$ 行每行一个数。其中测试点4和5的输入输出文件还具有对称性。

这些测试点的主要命题思路是考察选手分析多项式的能力和对多项式的计算。

大部分的数看起来很随机而且分布在某个数以内，很可能是模某个数的运算。容易得出模数是104857601。

而 $104857601 = 25 \times 2^{22} + 1$ ，这几个测试点很可能与FFT有关。

### 3.2.1 测试点4

给出一个多项式，要给它平方。

我们可以套用经典的多项式乘法计算。

### 3.2.2 测试点5

给出一个多项式，要给它开方。

我们可以套用经典的多项式开方计算。

### 3.2.3 测试点6

给出一个多项式，要给它开立方。

由于模数104857601的特殊性质，这个是可以直接开三次方的，但是如果不会FFT怎么办呢？

.....

注意到测试点4的输出实际上是 $(x+1)^{262144}$ ，直接算组合数就可以了。

测试点5的输出实际上是 $(x-1)^{131072}$ ，也是算组合数。

测试点6的输入和输出看起来非常随机，但是我们可以猜测输出文件是 $(ax+b)^{177147}$ 。

根据第一项和最后一项，可列方程

$$a^{177147} \equiv 22131490, b^{177147} \equiv 43392819 \pmod{104857601}$$

暴力解得

$$a = 233333333, b = 333333333$$

这个答案一看就很靠谱，代入验证成立。

这样就可以通过测试点4,5,6了。

## 3.3 测试点7,8,9

这三个测试点的共同点是，输入输出格式非常像图上的数据结构题，其中测试点7没有边权，测试点8是一棵树。

这些测试点的主要命题思路是考察选手通过数据猜测题面的能力以及简单数据结构的应用。

### 3.3.1 测试点7

点数和边数都是 $10^5$ ，没有边权，看起来非常随机。每个询问的答案只可能是0或 $+\infty$ 。

观察发现，如果两个点连通，答案就是0，否则答案是 $+\infty$ 。用并查集或Floodfill就可以做了。

### 3.3.2 测试点8

是 $10^5$ 个点的随机生成的树，边上带权，权也是随机生成的。每个询问的答案看起来都比较大。

观察发现，答案就是两点间树上路径边权最大值。由于数据随机所以可以不用高级数据结构维护。

### 3.3.3 测试点9

格式与测试点8相同，但不是树而是一般的图。答案看起来也没有那么大，而且出现了不连通的情况。

询问的依然是路径问题，但是不连通时答案是 $+\infty$ ，可知求的是某个量的最小值。结合测试点8分析可知求的是两点间路径边权最大值的最小值。

将所有边从小到大排序后依次加入按秩合并不路径压缩的并查集，对于每次询问二分答案在并查集中查询即可。

（当然，也可以用一些小的技巧去掉二分，不过这道题不卡时间，所以怎么做都可以）

## 3.4 测试点10

测试点10要求程序能输出自己。这个任务在完成前9个测试点的基础上是很难的。

但是由于题目的评测方式是传统评测方式，所以直接无视输入输出文件里面奇怪的英文，把输出文件原样打出来就可以了。

## 4 总结

在互测当中出了这样一道题，是想引发大家对于非传统题命题方向的思考，希望以后会有更多更新颖的非传统题涌现出来。

## 5 致谢

感谢中国计算机学会提供的交流平台。

感谢国家集训队教练余林韵和陈许旻的热情帮助。

感谢我的教练张新钢老师多年来的培养。

感谢参加互测的集训队员和参加镜像的全国各地的OIer。