

青少年软件编程（C 语言）等级考试试卷（九级）

1 图像分割

图像分割实际上是图的连通分量划分问题。将每个像素看成图中的一个顶点，顶点之间的边有权重，是相邻两个像素之间的区分度（非负整数）。图像分割的目标就是将图像构成的图分割成若干个不相交的连通分量，使得同一个分量中的像素都相似，而不同分量中的像素不相似。

连通分量的定义如下：

- 连通分量是一个互相连通的顶点的集合；
- 任何两个连通分量之间没有共同的顶点；
- 任何两个连通分量 C_1 和 C_2 之间的区分度 $D(C_1, C_2)$ 大于这两个连通分量的内聚度 H ；
- 区分度 $D(C_1, C_2)$ 定义为连接 C_1 和 C_2 的所有边权重中的最小值，如果不存在连接边，则这个值定义为无穷大；
- 连通分量 C 的内聚度 $H(C)$ 定义为 C 的最小生成树中的最大边权重，外加一个函数值 $f(C)=c/|C|$ ，其中 c 是一个正整数常量， $|C|$ 是 C 的规模，即顶点个数；
- 如果一个顶点集合还能被划分为两个及以上的连通分量，则这个集合不能算做是一个连通分量。

你的任务就是列出所有连通分量。

时间限制：1000

内存限制：65536

输入

第一行输入三个整数： N_v ($1 < N_v \leq 1000$) 为顶点数量（顶点从 0 到 $N_v - 1$ 编号）； N_e 为边的数量； c 是函数 $f(C)$ 中的那个常数。随后 N_e 行，每行按以下格式给出一条边：
顶点 1 编号 顶点 2 编号 权重 注意：题目保证每个像素点有不超过 8 个相邻的像素点。
常数和所有的权重均不超过 1000。

输出

每行输出一个连通分量。每个连通分量的顶点要在一行中按编号升序输出，数字间以 1 个空格分隔，行首尾不得有多余空格。连通分量按照其首个顶点的升序输出。

样例输入

样例#1：

10 21 100

0 1 10
0 3 60
0 4 90
1 2 90
1 3 50
1 4 200
1 5 86
2 4 95
2 5 5
3 4 95
3 6 15
3 7 101
4 5 500
4 6 100
4 7 101
4 8 101
5 7 300
5 8 50
6 7 90
7 8 84
7 9 34

样例#2：

7 7 100
0 1 10
1 2 61
2 3 50
3 4 200
4 5 82
5 0 200
3 6 90

样例输出

样例#1：

0 1 3 6

2 5 8

4

7 9

样例#2：

0 1

2 3 6

4 5

2 迎风一刀斩

迎着一面矩形的大旗一刀斩下，如果你的刀够快的话，这笔直一刀可以切出两块多边形的残片。反过来说，如果有人拿着两块残片来吹牛，说这是自己迎风一刀斩落的，你能检查一下这是不是真的吗？

注意摆在你面前的两个多边形可不一定是端端正正摆好的，它们可能被平移、被旋转（逆时针 90 度、180 度、或 270 度），或者被（镜像）翻面。

这里假设原始大旗的四边都与坐标轴是平行的。

时间限制：1000

内存限制：65536

输入

输入第一行给出一个正整数 N (≤ 20)，随后给出 N 对多边形。每个多边形按下列格式给出： $k\ x_1\ y_1\ \dots\ x_k\ y_k$ 其中 k ($2 < k \leq 10$) 是多边形顶点个数； (x_i, y_i) ($0 \leq x_i, y_i \leq 10^8$) 是顶点坐标，按照顺时针或逆时针的顺序给出。注意：题目保证没有多余顶点。即每个多边形的顶点都是不重复的，任意 3 个相邻顶点不共线。

输出

对每一对多边形，输出`YES`或者`NO`。

样例输入

8

```
3 0 0 1 0 1 1
3 0 0 1 1 0 1
3 0 0 1 0 1 1
3 0 0 1 1 0 2
4 0 4 1 4 1 0 0 0
4 4 0 4 1 0 1 0 0
3 0 0 1 1 0 1
4 2 3 1 4 1 7 2 7
5 10 10 10 12 12 12 14 11 14 10
3 28 35 29 35 29 37
3 7 9 8 11 8 9
5 87 26 92 26 92 23 90 22 87 22
5 0 0 2 0 1 1 1 2 0 2
4 0 0 1 1 2 1 2 0
4 0 0 0 1 1 1 2 0
4 0 0 0 1 1 1 2 0
```

样例输出

```
YES
NO
YES
YES
YES
YES
NO
YES
```

3 代码查重算法

本题请你实现一个简化版的代码查重算法。算法步骤如下：

第 1 步：单词切分。即将代码中连续无间断的若干字母和数字组成的子串定义为一个“单词”。

第 2 步：保留关键词。这里简化为只考虑给定集合 $S=\{\text{main, return, if, else, for, do, while, int, float, double}\}$ 中的词。将代码中除关键词之外的单词都统一替换为单个字母

`x`。

第 3 步：去掉代码中所有白字符和注释 —— 删除注释比较繁琐，所以本题简化为只删除空格和回车即可。

第 4 步：不妨将前 3 步处理好的两段代码记为 C_1 和 C_2 ，它们对应的字符串长度分别为 L_1 和 L_2 。初始化当前重复代码长度 $L=0$ ，记重复代码长度阈值为 L_T 。重复下述步骤：

(4.1) 求两段代码的**最长公共子序列**，记其长度为 len 。若 $len > L_T$ ，则执行 (4.2)；否则执行 (4.3)。

(4.2) 将最长公共子序列从 C_1 和 C_2 中删除。将当前重复代码长度 L 增加 len 。将 C_1 和 C_2 更新为剩余的字符串。回去执行 (4.1)。

(4.3) 计算相似度。相似度定义为 $\max(L/L_1, L/L_2)$ 。

注 1：给定序列 $X=\{x_1, x_2, \dots, x_n\}$ ，其**子序列**是从该序列中删去若干元素后得到的序列。例如给定 $X=\{H, E, L, L, O\}$ ，则 $\{E, L, O\}$ 、 $\{H, O\}$ 、 $\{L\}$ 等均为其子序列。

注 2：当最长子序列匹配不唯一时，首先优先选择 C_1 最左边的字符匹配；当 C_1 的同一个字符可以匹配多个 C_2 中字符时，同样优先选择 C_2 最左边的字符匹配。这样保证匹配结果是唯一的。

时间限制：1000

内存限制：65536

输入

输入首先在第一行给出重复代码长度阈值 L_T ，为不超过 10 的非负整数。随后分别给出两段代码，格式为：- 首先在一行中给出代码的文件名，为长度不超过 20 的、仅由英文字母和数字组成的字符串。- 随后若干行字符为代码内容，最后以 `\$` 和一个回车结尾。题目保证 `\$` 不是代码内容的一部分（或可以理解为，只要读到 `\$` 和一个回车，其后的内容都不属于这段代码）。题目保证代码内容在经过预处理后不为空，且原始代码不超过 10^3 个字符。

输出

顺次输出每一轮查到的重复代码片段，每轮次占一行。在最后一行中顺序输出：第 1 个文件名、第 2 个文件名、两段代码的相似度（输出小数点后 2 位）。数据间以 1 个空格分隔。

样例输入

```
5
file1
#include<stdio.h>
int main()
{ int tmp=100;
  printf("hello world");
  while(tmp) tmp--;
```

```

    return 0;
}
$
file2
#include<stdio.h>
void main()
{
    int i=10;
    while(i) i--;
    cout<<"hello world"<<endl;
}
$

```

样例输出

```

#x<x.x>main(){intx=x;while(x)x--;x;}
file1 file2 0.78

```

提示

样例解释：第一个文件的文件名为 file1。其代码经过前 3 步处理后得到的 C₁ 为：

```
#x<x.x>intmain(){intx=x;x("xx");while(x)x--;returnx;}
```

第二个文件的文件名为 file2。其代码经过前 3 步处理后得到的 C₂ 为：

```
#xxmain(){intx=x;while(x)x--;x<<"xx"<<x;}
```

file1 预处理后长度为 53，file2 预处理后长度为 46。第一轮查到的长度为 36 的重复代码片段（即两段代码的最长公共子序列）为：

```
#x<x.x>main(){intx=x;while(x)x--;x;}
```

删除重复片段后 C₁ 和 C₂ 更新为：

```
intx("xx");return
```

```
x<<"xx"<<x
```

根据上述更新，第二轮查到的重复代码片段为：

```
x"xx"
```

但是因为这个片段长度为 5，没有超过给定的阈值 5，所以不算。最后的重复代码长度为 36。所以相似度为 $\max(36/53, 36/46)=0.78$ 。

4 网络攻击

n 台计算机用 m 条光纤两两直连成一个网络。每条光纤都是双向连通的，任意两台计算机之间的数据传输可以通过直连光纤或者通过其它计算机中转完成。

现有一群黑客企图攻击这个网络，目标是断开网络中两台主机的连接，使得两机无法互相传输数据。黑客计算出破坏每一条光纤需要花的钱，当然，他们希望这个花销越小越好。也就是说，他们需要的方案是使得破坏所有必要光纤的总花销取到最小值。你的任务是判断一下，达到黑客攻击目标的方法是不是唯一的。

时间限制：1000

内存限制：65536

输入

输入包含多组测试。每组测试的第一行给出 4 个正整数： n ($2 \leq n \leq 800$) 为网络中计算机的数量， m ($\leq 10^4$) 为连接用的光纤条数， A 和 B ($1 \leq A, B \leq n$) 为黑客要断开的两台主机的编号。如果某组测试的第一行给出 4 个 0，则表示输入结束，这组数据不要处理。接下来 m 行给出光纤的连接信息，即光纤两端连接的计算机的编号、以及破坏这条光纤的花销。题目保证所有花销为不超过 10^5 的非负整数，任意两台计算机之间最多有一条光纤，没有一条光纤的两端都连在同一台机器上，并且初始状态下 A 和 B 之间是可以传输数据的。

输出

对每一组测试，如果达到黑客攻击目标的方法是唯一的，就在一行中输出 `UNIQUE`（表示“唯一”），否则输出 `AMBIGUOUS`（表示“有歧义”）。

样例输入

```
4 4 1 2
1 2 1
2 4 2
1 3 2
3 4 1
4 4 1 2
1 2 1
2 4 1
1 3 2
3 4 1
0 0 0 0
```

样例输出

```
UNIQUE
AMBIGUOUS
```