

# 2013 高教社杯全国大学生数学建模竞赛题目

(请先阅读“全国大学生数学建模竞赛论文格式规范”)

## B 题 碎纸片的拼接复原

破碎文件的拼接在司法物证复原、历史文献修复以及军事情报获取等领域都有着重要的应用。传统上,拼接复原工作需由人工完成,准确率较高,但效率很低。特别是当碎片数量巨大,人工拼接很难在短时间内完成任务。随着计算机技术的发展,人们试图开发碎纸片的自动拼接技术,以提高拼接复原效率。请讨论以下问题:

1. 对于给定的来自同一页印刷文字文件的碎纸机破碎纸片(仅纵切),建立碎纸片拼接复原模型和算法,并针对附件 1、附件 2 给出的中、英文各一页文件的碎片数据进行拼接复原。如果复原过程需要人工干预,请写出干预方式及干预的时间节点。复原结果以图片形式及表格形式表达(见【结果表达格式说明】)。

2. 对于碎纸机既纵切又横切的情形,请设计碎纸片拼接复原模型和算法,并针对附件 3、附件 4 给出的中、英文各一页文件的碎片数据进行拼接复原。如果复原过程需要人工干预,请写出干预方式及干预的时间节点。复原结果表达要求同上。

3. 上述所给碎片数据均为单面打印文件,从现实情形出发,还可能有双面打印文件的碎纸片拼接复原问题需要解决。附件 5 给出的是一页英文印刷文字双面打印文件的碎片数据。请尝试设计相应的碎纸片拼接复原模型与算法,并就附件 5 的碎片数据给出拼接复原结果,结果表达要求同上。

### 【数据文件说明】

- (1) 每一附件为同一页纸的碎片数据。
- (2) 附件 1、附件 2 为纵切碎片数据,每页纸被切为 19 条碎片。
- (3) 附件 3、附件 4 为纵横切碎片数据,每页纸被切为  $11 \times 19$  个碎片。
- (4) 附件 5 为纵横切碎片数据,每页纸被切为  $11 \times 19$  个碎片,每个碎片有正反两面。该附件中每一碎片对应两个文件,共有  $2 \times 11 \times 19$  个文件,例如,第一个碎片的两面分别对应文件 000a、000b。

### 【结果表达格式说明】

复原图片放入附录中,表格表达格式如下:

- (1) 附件 1、附件 2 的结果:将碎片序号按复原后顺序填入  $1 \times 19$  的表格;
- (2) 附件 3、附件 4 的结果:将碎片序号按复原后顺序填入  $11 \times 19$  的表格;
- (3) 附件 5 的结果:将碎片序号按复原后顺序填入两个  $11 \times 19$  的表格;
- (4) 不能确定复原位置的碎片,可不填入上述表格,单独列表。