# 实验一 线性表

## 一 实验内容

根据给定语料，构建一个词云。建议步骤如下：

1.读取分词文件（5分）

2.统计每个词及词频（30分）

2.1 构建词及词频的存储结构WordFreq

2.2 构建元素类型为WordFreq的线性表

2.3 将分词文件中的词及词频存入线性表

2.3.1 如何将fileContent按照词切分

提示：利用String 类的split方法

2.3.2 如何统计词的频次

提示：

如果一个词word不在线性表中，则插入（word,1）；

  否则在线性表中找到这个词，将它的频次加一；

  这里涉及三个方法：insert，find，set

3. 构建停用词表（20分）

3.1 停用词，即对生成词云无用的词，比如“的”、“地”、“得”、“。”等

3.2 可以构建一个停用词文件，然后通过读取文件构建停用词表

3.3 将在停用词表中的词从词频统计线性表中删除

4.输出词频统计线性表信息到屏幕或者文件中（5分）

5.生成词云（10分）

5.1利用https://www.sohu.com/a/162276896\_165070生成词云

5.2利用github中开源的代码生成词云（选作；加分项；）

## 二 实验要求

1. 撰写实验报告（30分；从上机开始，每拖延一周减5分，截止日期后0分）
   1. 实验内容
   2. 设计
      1. 类设计（类图，类说明）
      2. 类结构关系
      3. 核心算法流程
   3. 程序执行过程截图及说明
   4. 实验总结
   5. 附源代码
2. 检查程序（从上机开始，每拖延一周减5分，截止日期后0分）
   1. 演示功能
   2. 口述实现过程
   3. 回答问题