**2021-2022学年第二学期《互联网编程》编程测验**

**姓名： 郑雨婷 学号： 2021150122**

**专业： 计算机科学与技术 学院： 计算机软件与学院**

**题目：单词频次统计-流的运用**

**一、实验目的与内容**

**1. 实验目的**

掌握Java流的主要方法及运用，自学文本、字符串处理方法。

**2.  实验内容**

扫描指定文件夹里面的所有文件，给出各个文件的属性信息。另外，针对每个文件（如A.txt）统计里面文章中单词的出现频次并输出统计结果到文件Sta\_A.txt。Sta\_A.txt文件包含A.txt文件里面每个单词出现的频次。

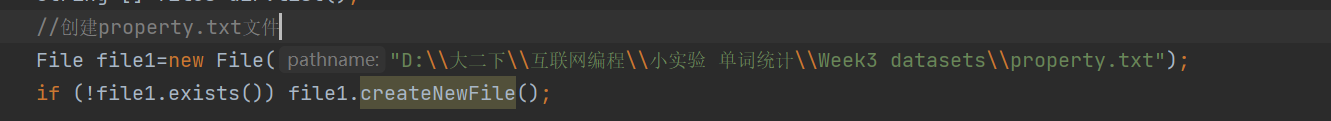
**3. 实验要求：**

1. 指定文件夹里面所有的文件的属性都得列出，输出到property.txt。
2. 指定文件夹里面每个文件的单词频次都得统计，并保存到对应输出文件。输出文件中单词按频次降序排列。
3. 统计程序运行时间，包括总体运行时间和各个文件单词频次统计所花时间。
4. 尽可能快的完成所有文件的单词频次统计。优化代码！

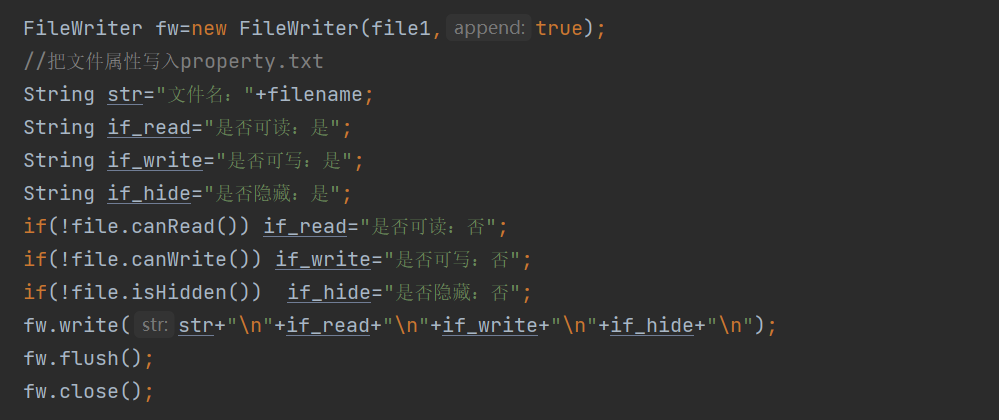
注意：需要现场验收程序。（返校后）

**二、实验报告**

贴出关键代码及注释、说明，并且附上测试性能结果截图，如运行时间等。

1. 创建property.txt文件

遍历所有文件，访问每个文件的属性，利用FileWriter写入属性。

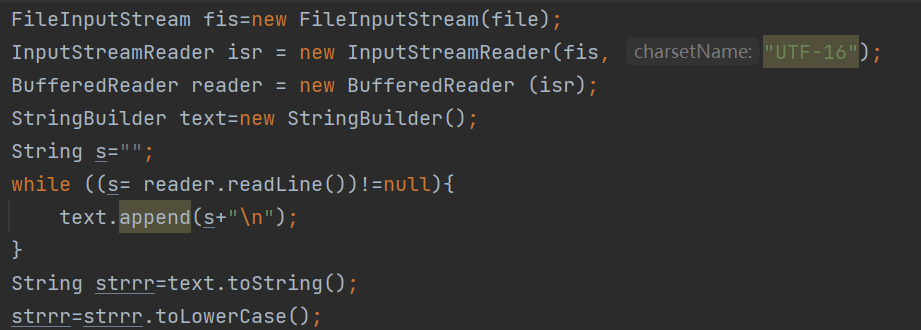


1. 统计单词频次。

创建Sta\_A.txt文件

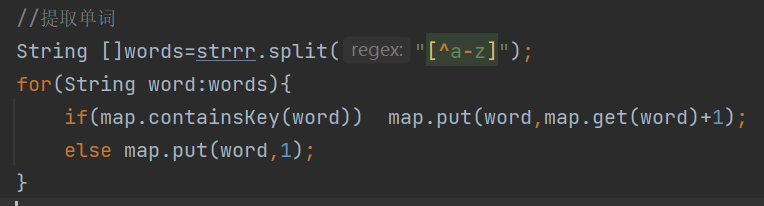


读取文件，利用输入流，将读取到的文件返回为StringBuilder,再转为String类型，并且全部转换为小写字母，这样统计频次时可以不用考虑字母大小写。

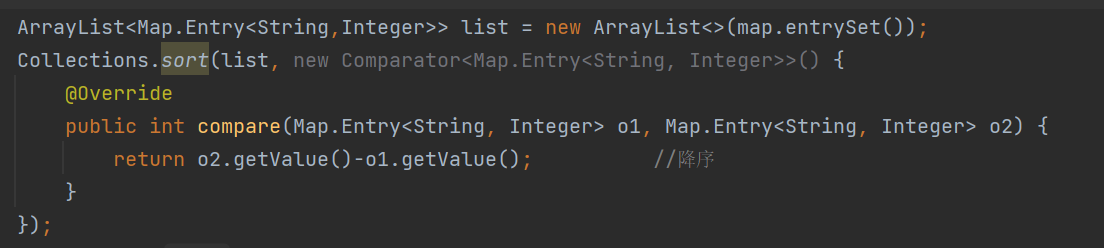


利用HashMap没有重复元素的特性，来存放单词，若重复，则value++;

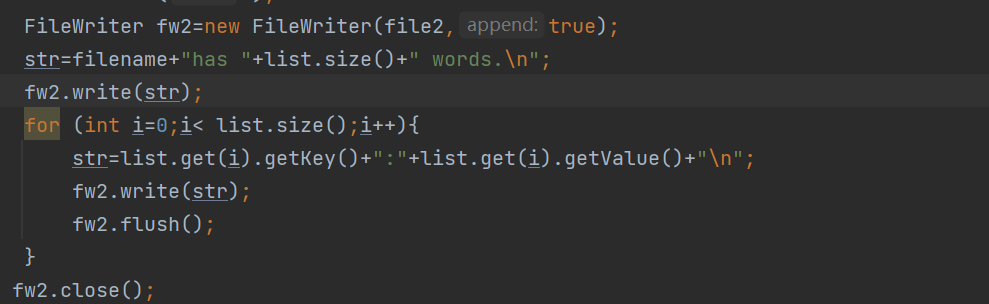
这里分割单词的时候用到正则表达式，^代表除了，之前已经全部转换为小写，所以除了a-z的地方都进行分割，就可以把单词提取出来。



按单词出现次数降序排序。



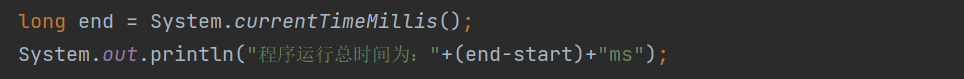
用FileWriter按照排好的顺序写入文件Sta\_A.txt



统计运行时间

用System.currentTimeMillis()返回开始和结束时的时间，二者相减就是运行时间。





运行结果如下：

