

Java 集合使用实验报告

一、实验目的

- 1.1 掌握 Java 中抽象的数据集合框架
- 1.2 掌握各种集合接口和集合类所代表的数据结构
- 1.3 熟练运用常用的几个集合类
- 1.4 掌握泛型的概念和使用
- 1.5 掌握文件读写和集合排序

二、相关知识

- 2.1 基本的数据集合框架类有 LinkedList, ArrayList, Hashtable 等, 我们常用的数据集合有 ArrayList, HashMap, Vector, TreeSet, Stack 等
- 2.2 泛型的本质是参数化类型, 也就是所操作的数据类型被指定为一个参数。参数类型可以在类、接口和方法的创建中, 分别称为泛型类、泛型接口、泛型方法。

三、实验内容

- 3.1 设计一个统计文本文件中词频最高的前 K 个单词的程序

四、实现思路

4.1 程序设计

4.1.1 代码实现思路

1. 整体思路封装为两个函数: fileRead 和 print_topK, 第一个函数完成文件的读取并将文本进行分割存入并统计词频, 存入 HashMap 中; 第二个函数将 HashMap 转为 List, 使用 comparator 进行自定义排序, 输出词频最高的前 K 个单词。
2. FileRead(String from): 使用 FileReader 和 BufferedReader 完成文件的读入, 使用 readLine() 每次读取一行, 使用 replace 替换一些常见的标点符号为空格, 根据空格使用 split 分割文本, 存入 HashMap 并统计词频, 返回一个 HashMap.
3. Print_topK(HashMap count, int k): 使用 list 数据结构将 HashMap 的 entrySet 保存, 使用 list.sort() 进行自定义排序, 重载 compare 进行频率的比较, 最后根据 K 的要求输出词频最高的前 K 个单


词。

4. 1. 2 代码截选

图 1: fileRead

图 2: print_TopK

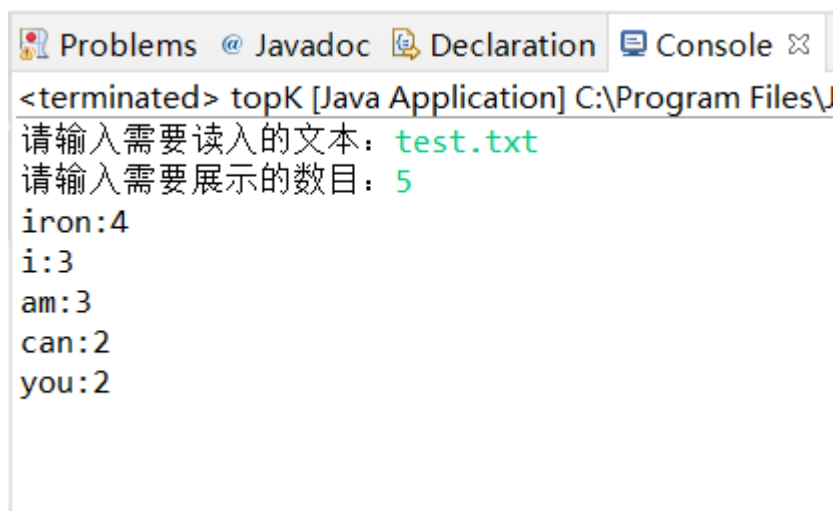
五、 结果截图

 test - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
i am iron man am am iron iron iron,i can swim and dance.can you?  
i love you three thousand times  
live or die,is a question.
```

测试文本



测试结果正确

六、 实验收获