

JAVA 编程进阶上机报告



学 院 智能与计算学部

专 业 软件工程

班 级 二班

学 号 3018216084

姓 名 田家硕

一、实验要求

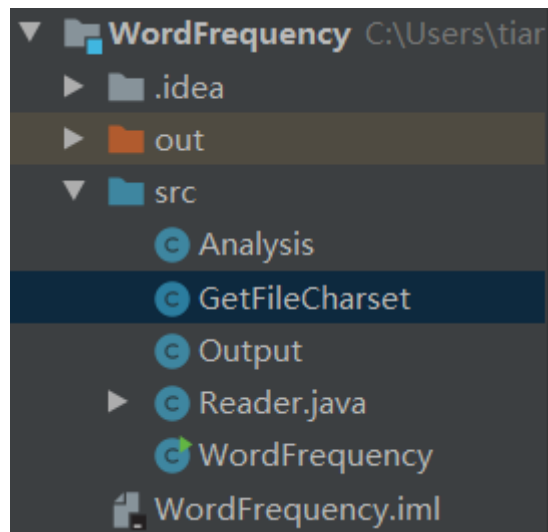
输入:了不起的盖茨比（英文版）.txt

输出:创建一个 output.txt 。输出格式为单词+空格+频次，结果按照单词的频次倒序排列。

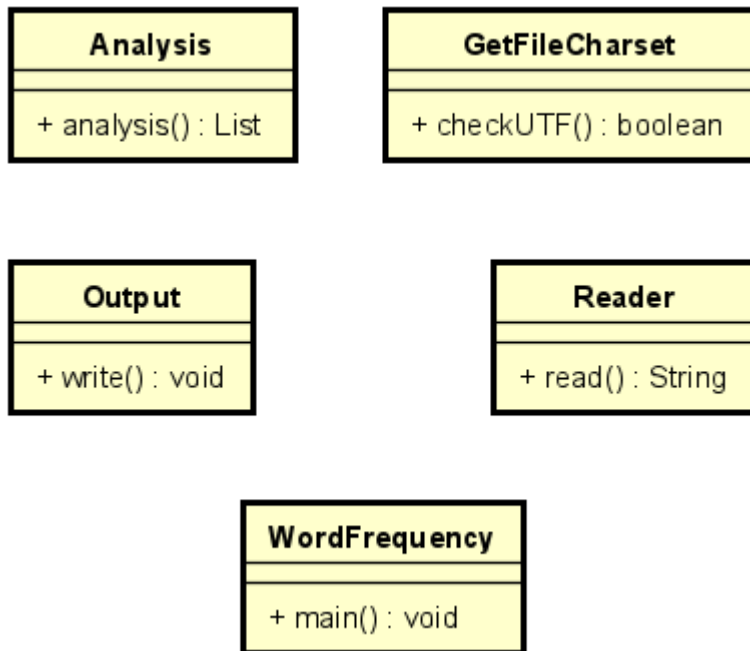
异常处理：例如文件不存在，文件没有读权限，文件编码错误等

二、设计思路和UML

- 项目目录：



- 分为Reader Analysis Output 三个环节来实现字频统计。其中Reader类中做了文件不存在，文件没有读权限，文件编码错误的异常处理。GetFileCharset类中封装了判断文件编码是否为UTF-8的方法。



三、源代码

WorldFrequency

```

import java.io.*;
import java.util.*;

public class WordFrequency {
    public static void main(String[] args) throws IOException {
        //读入
        String content = Reader.read("了不起的盖茨比英文.txt");
        List<Map.Entry<String, Integer>> list =
Analysis.analysis(content);
        Output output = new Output();
        output.write(list);
    }
}
  
```

Output

```

import java.io.File;
import java.io.FileWriter;
import java.io.IOException;
import java.io.Writer;
  
```

```

import java.util.List;
import java.util.Map;

public class Output {
    public static void write(List<Map.Entry<String, Integer>> list) throws
IOException {
        //写入output.txt
        File resultFile = new File("D:" + File.separator + "output.txt");
        Writer wrt = new FileWriter(resultFile);
        for(Map.Entry<String,Integer> mapping:list){
            char[] c = (mapping.getKey()+"
"+mapping.getValue()+"\n").toCharArray();
            wrt.write(c);
        }
        wrt.flush();
    }
}

```

Analysis

```

import java.util.*;

public class Analysis {
    public static List analysis(String content) {
        String[] contents = content.split("\\s+");//按空格+分割
        Map<String, Integer> map = new HashMap<>();
        //存入map
        for (int i = 0; i < contents.length; i++) {
            if (map.containsKey(contents[i])) {
                map.put(contents[i], map.get(contents[i]) + 1);
            } else {
                map.put(contents[i], 1);
            }
        }
        //排序
        List<Map.Entry<String, Integer>> list = new ArrayList<>
(map.entrySet());
        //将map.entrySet()转换成list
    }
}

```

```

//然后通过比较器来实现排序
Collections.sort(list, new Comparator<Map.Entry<String, Integer>>
() {
    //降序排序
    public int compare(Map.Entry<String, Integer> o1,
                        Map.Entry<String, Integer> o2) {
        return o2.getValue().compareTo(o1.getValue());
    }
});
return list;
}
}

```

Reader

```

import java.io.*;

class AddException extends Exception{
    public AddException(String msg){
        super(msg);
    }
}

public class Reader {
    public static String read(String name) throws IOException {
        File file = new File("D:" + File.separator + name);
        //判断是否存在
        if (!file.exists()) {
            try {
                throw new AddException("文件不存在");
            }
            catch (AddException e){
                e.printStackTrace();
            }
        }
        //判断是否可读
        if (!file.canRead()){
            try {
                throw new AddException("没有读权限");
            }

```

```

    }
    catch (AddException e){
        e.printStackTrace();
    }
}
//读入
//使用文件输入流实例化BufferedReader类对象
String str = null;    //接收输入数据
String content = "";
    BufferedReader buf = new BufferedReader(new FileReader(file));
    while ((str = buf.readLine()) != null) {    //判断是否存在
        content += " " + str;
    }
    buf.close();

//判断编码
if(!GetFileCharset.checkUTF(content)){
    try {
        throw new AddException("文件编码不正确");
    }
    catch (AddException e){
        e.printStackTrace();
    }
}

return content;
}
}

```

GetFileCharset


```

public class GetFileCharset {
    public static boolean checkUTF(String str) {
        boolean flag = java.nio.charset.Charset.forName("UTF-8").newEncoder().canEncode(str);
        return flag;
    }
}

```

四、实验结果

输入到D:/output.txt

 output.txt - 记事本

文件(E) 编辑(E) 格式(O) 查看(V) 帮助(H)

the 1887

I 1270

and 1267

a 1208

of 1008

to 999

in 677

was 671

he 546

that 503