**词频统计的结构化分析、设计、实现**

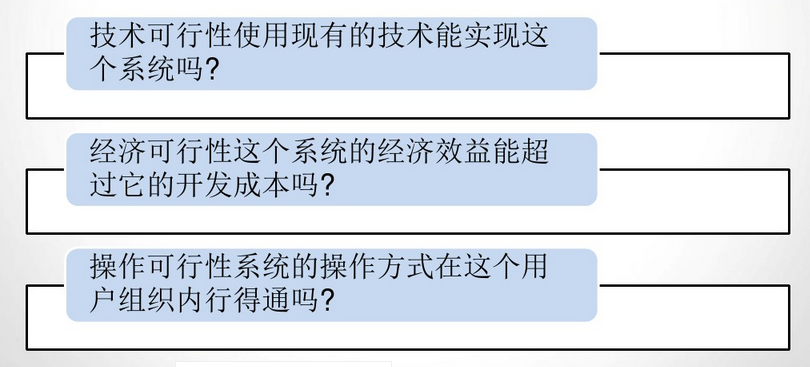
1. **结构化分析**

**1、可行性研究**

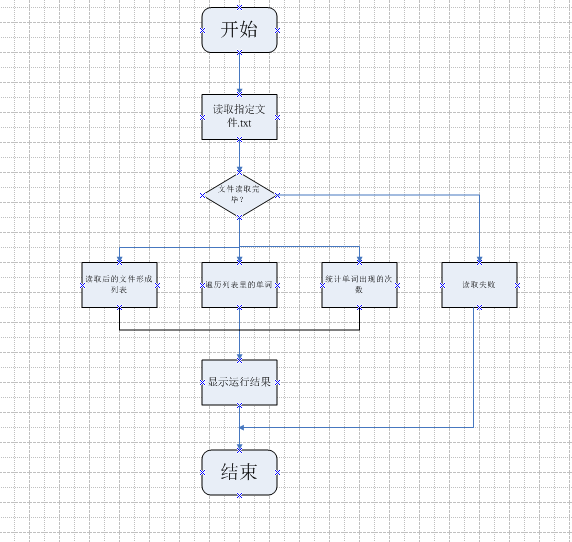
操作可行性系统操作简单，且界面效果加入人性化设计理念，使得用户能够 快速熟悉操作，简单的选择文本文件即可

**1.1可行性研究的任务**

可行性研究的目的不是解决问题，而是确定问题是否值得被解决。可从下面 三方面说明可行性：



**1.2大致流程图**



1. **需求分析**

**2.1功能需求**

1. 设计一个图形化窗口，包含文件选择框，选择文件按钮，统计按钮， 结果显示文本框
2. 、对存储在文本文件下的英语文本进行读取，形成列表，遍历列表中 的单词
3. 、统计文本单词及词频，并输出结果
4. 、在图形界面显示文件，通过点击文件显示文件具体统计结果

具体要求如下：

1. 在GUI界面选择文本文件
2. 选择的文件进行读取得出结果

**2.2非功能需求**

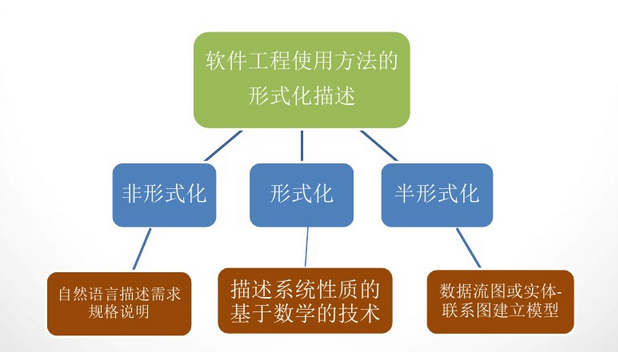
**2.2.1性能需求**

1. 正确性：选择正确格式的文本文件，输入其他格式的文件将输出错误
2. 高效性：选择读取的文件，无论该文本文件里面包含多少单词文本，都可以读取并输出最后的结果

**2.2.2输入输出需求**

1. 在可视化界面选择文件
2. 在内部进行读取输出
3. **形式化说明**

形式化技术有效提高了开发的效率、改进了质量，容易在软件的规约上取得一致性。



**3.1 概述**

**3.1.1非形式化方法的缺点**

用自然语言书写的系统规格说明书，可能存在矛盾二义性、含糊性、不完整性及抽象层次混乱等问题。

**3.1.2形式化方法的优点**

为了客服非形式化的缺点，把数学引入软件开发过程。

**3.2词频统计方法**

读取文本文件中的所有单词，然后将单词形成列表，在从列表中的单词存放在列表中，读取列表中的单词，如果列表下标的单词是和之前的单词有冲突则+1

1. **词频统计的设计与实现**
2. **词频统计的设计背景**

词频统计是为了统计一个文档中出现的单词的数量以及次数，根据单词出现的次数可以让人们更加清晰的看出该篇文章中单词的重要性排序

1. **词频统计逻辑分析**

（1）、程序需要读取一个文本里面的单词，可以是任意英文文本

（2）、至少统计100个词以上

（3）、用户可以同时查询多个单词出现的次数

（4）、单词出现的次数进行降序显示

（5）、统计文本单词及词频，并输出结果

（6）、在图形界面显示文件，通过点击文件显示文件具体统计结果

**C、词频设计**

1. 设计一个图形化窗口，包含文件选择框，选择文件按钮，统计按钮，结果显示文本框
2. 对存储在文本文件下的英语文本进行读取，形成列表，遍历列表中的单词
3. 统计单词出现的次数，并在文本框中显示最后的统计结果
4. **词频的实现**

具体代码如下：



