**华南农业大学数学与信息学院**

**综合性实验报告**

|  |  |
| --- | --- |
| **课程:** | **面向对象程序设计实验** |
| **实验题目:** | **Java源代码注释关键字分析程序** |
| **学号:** | **202014210104** |
| **姓名:** | **古卓恒** |
| **专业:** | **网络工程** |
| **班级:** | **2020级1班** |
| **指导教师:** | **宋鸿陟** |
| **提交日期:** | **2021年12月3日** |

1. 题目介绍与分析

1.1 题目介绍

编写一个Java应用程序，对单个Java源程序文件及某个目录中的所有Java源程序文件（包括子目录）进行分析。

(1)Java源程序文件个数，对目录分析进行分析时需要。

(2)Java源程序中的字符个数，对目录分析时是其中所有源程序文件的字符个数总和。

(3)Java源程序文件中的注释的个数，即源程序文件中共有多少个注释，包括：单行注释和多行注释。对目录分析时是其中所有源程序文件的总和。

(4)Java源程序文件中的注释的字符个数，即源程序文件中所有注释的字符数之和。对目录分析时是其中所有源程序文件的总和。

(5)Java源程序文件中关键字使用情况，即源程序文件各个关键字使用了多少次。对目录分析时是其中所有源程序文件的总和。

1.2 实验要求分析

(1) 程序运行时，首先要有一个主菜单。通过主菜单可以进入不同的子菜单进行操作或选择退出。

(2) 对于源程序文件（包括子目录）进行分析。

该程序首先提示并等待用户输入目标目录的路径。

1. 如果文件或目录不存在，提示用户该名称不是合法的目录名称，要求用户重新输入目录或退出当前界面或退出程序。
2. 如果文件或目录存在，则首先判断该路径是否为文件夹目录。

1)如果不是文件夹，则

2)如果是文件夹，

1. 分析完成后，提示用户“[目标路径文件] 分析完成......”。
2. 最后将文件的信息输出到data的文件夹中，并保存到【文件名+.txt】文件中，详细的说明在（3）中说明。

(3) 对于将目录的详细信息输出到data文件夹中的详细功能分析。

由于将文件打包成jar包，而在jar包中不可以像读取文件夹的方式一样直接读取到jar包中的data文件夹。所以解决的方法是通过在运行jar包的目录下，寻找运行文件夹的目录下有无data目录，若无，则创建一个data文件夹，并把从源目录的分析结果保存到data文件夹下的【文件名+.txt】文件中去，同时，对于文件名也会进行特殊处理，将所有 . \ 替换成 \_ . 以确保文件名不会与路径混淆和直观的看出分析文件位置。如果存在data文件夹，则直接获取data文件夹的实例化对象，将从源目录的分析结果保存到data文件夹下的【文件名+.txt】文件中去。如果已经存在同名的文件，则会更新文件的内容。

(4) 用户输入数字自动打开对应目录文件的分析结果。

自动打开文件，会因系统的不同而导致打开的方法不同。在此主要研究在windows系统下的打开方式。通过Runtime.getRuntime().exec("rundll32 url.dll FileProtocolHandler " + txt\_name)来打开txt文件，若无法打开，则会指引用户手动打开。

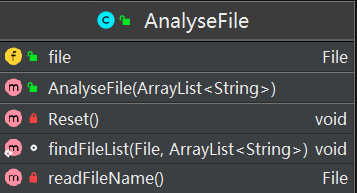
2. 程序静态结构

Main:程序主入口。在这里只有一行代码，即创建Mainproject实例化对象，以达到简洁代码的目的。

Mainproject:程序主任务。这里包含一些界面指令的处理和一些异常处理。

Interface:封装的程序可视化界面，及可视化界面各种操作方法的类。这个类中包含各种情况可能发生的状况而要弹出的界面，将所有界面和关于界面的操作集成要该类下，有几个优点：1.使程序界面整体风格更加统一；2.使代码管理更加方便；3.其他类可更专注于其专业性的方面操作。

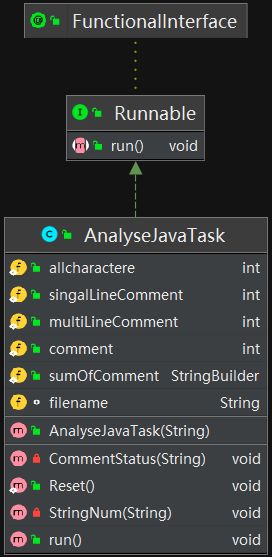
AnalyseFile: 读目录的类。读入路径，若路径最后的后缀名为.java，则直接存入list中，否则，遍历目录，提取java文件到list，若又发现子目录，则递归调用。最终得到的list列表记录了所有Java源文件绝对路径，最终得到所有java文件地址。



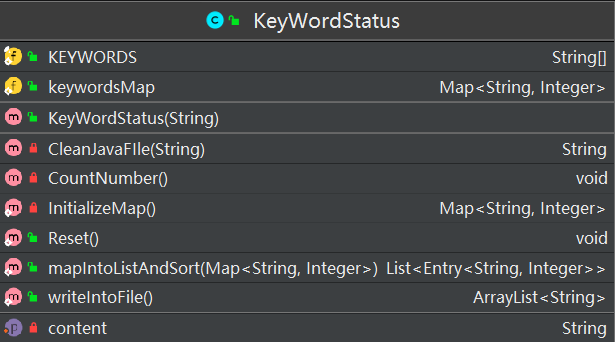
AnalyseJava: 对于Analysefile处理过后得到的list进行多线程操作，建立线程池。

TaskPool:封装了建立线程池操作的细节。

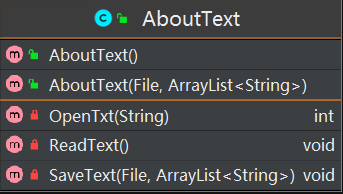
AnalyseJavaTask:对于单个java文件进行统计分析，得到该文件的字符个数、关键字使用情况、注释个数、注释字符数。

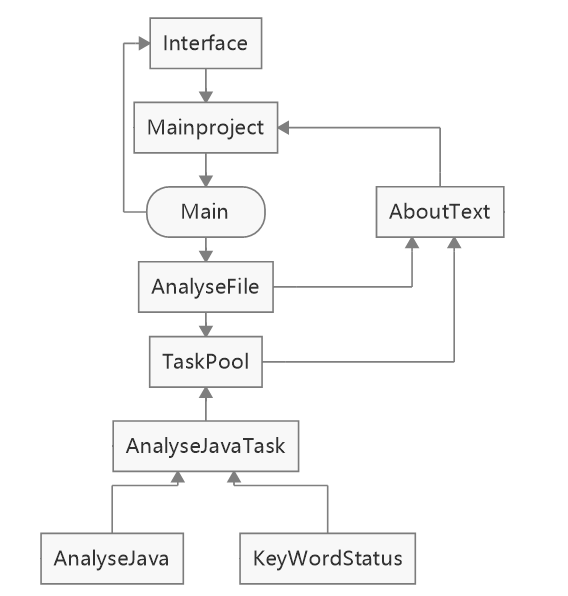


KeyWordStatus: 统计关键词使用情况。

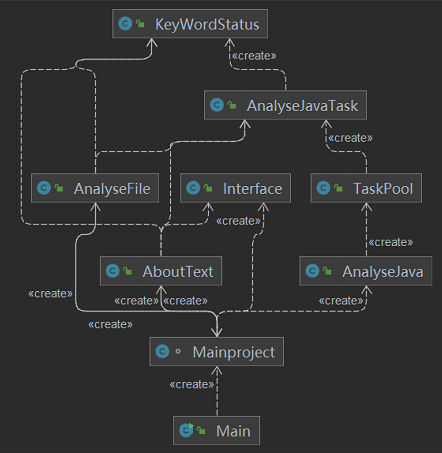


AboutText:封装文件写入相关。



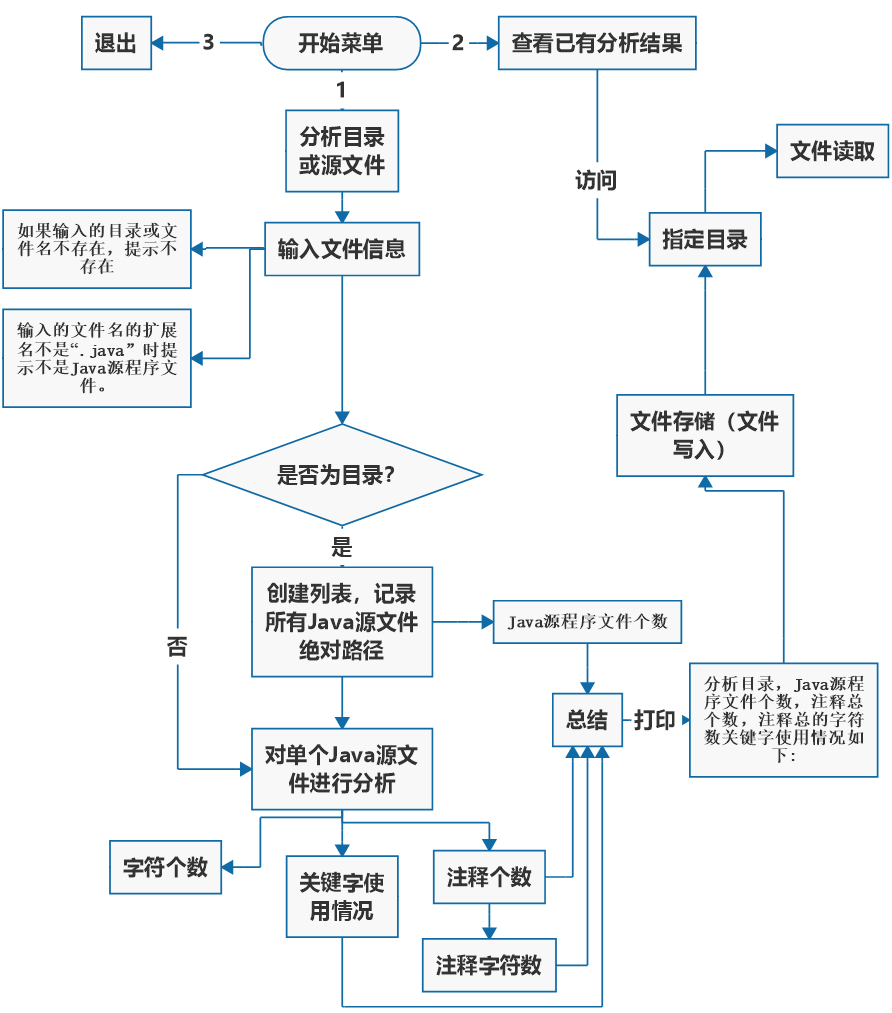


图表 1初始规划类之间的关系



图表 2具体实现

3. 程序主要功能

1. 分析目录或源文件程序。输入文件名后将分析结果存储到指定目录中。
2. 在目录中筛选所有Java文件，并将绝对地址存入列表中。通过equals方法比较。
3. 统计字符个数。逐行遍历，通过计算每行的长度可以得到总字符个数
4. 注释个数。通过正则表达式Pattern.compile("((?<!:)[\\/\\/.\*)|(\\/\\\*(\\s|.)\*?\\\*\\/)](file:///\\/\\/.*)|(\\/\\*(\\s|.)*%3f\\*\\/))")和matcher.group()来判断是单行注释还是多行注释。
5. 关键字使用情况统计。先用正则表达式除去所有注释、所有自符串直接量，再逐个单词进行比对，将关键字使用情况存在字典里。
6. 分析结果存储到指定目录中。通过字符串数组和BufferedWriter逐行写入。
7. 查看已有的分析结果。加载目录下结果文件名，通过数字选择打开。
8. 加载目录。通过dir.list()存入文件名，并打印文件。
9. 打开文件。通过Runtime.getRuntime().exec("rundll32 url.dll FileProtocolHandler " + txt\_name)打开。

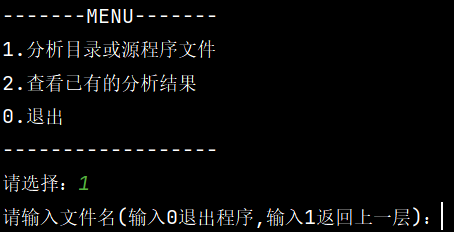
程序整体流程图

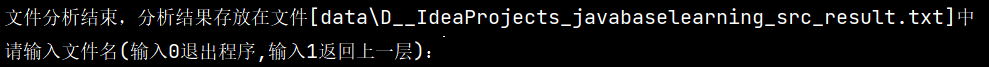
4.程序测试

注意：通过cmd运行jar包时，需要添加参数 -Dfile.encoding=utf-8

**在cmd 中输入命令行 java -jar -Dfile.encoding=utf-8 Conservativeness.jar**

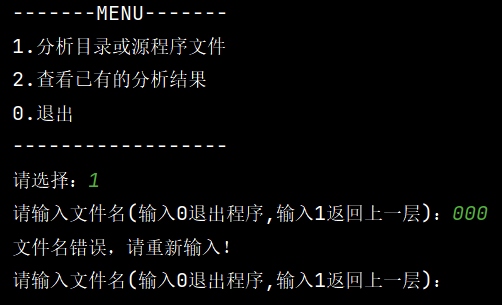
(1)以分析D:\IdeaProjects\javabaselearning\src为例，输入1后，再输入文件目录的绝对路径，可以以极快的速度得到结果并存在D\_\_IdeaProjects\_javabaselearning\_src\_result.txt中，若检测到目录中已存在同名txt文件，则对该txt文件执行重新写入操作，以达到分析结果的更新。最后结果如图。





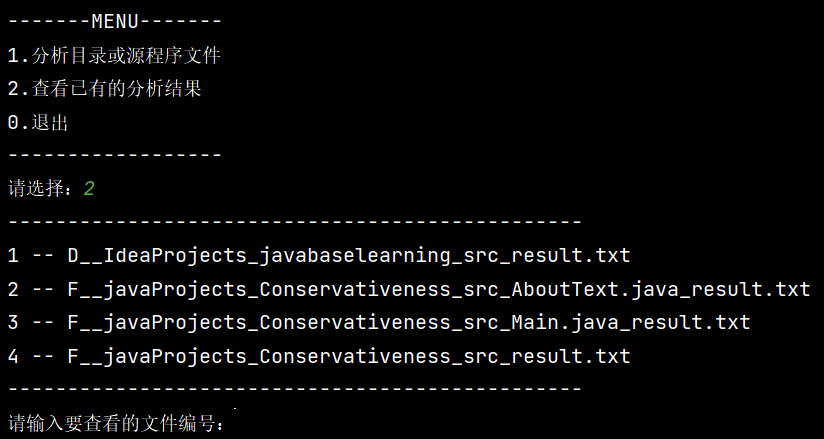
此时，可选择继续分析或退到上一层。

若输入文件名错误或无法访问，则会要求重新输入。



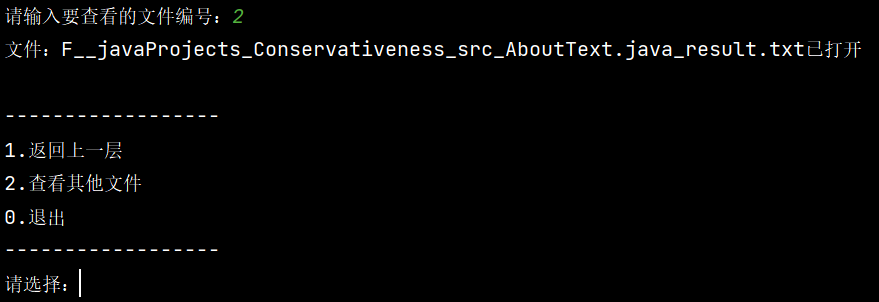
(2)以分析F:\javaProjects\Conservativeness\src\AboutText.java为例，结果类似(1)。

(3)查看已有分析结果。在主界面输入2，可查看已有分析结果。



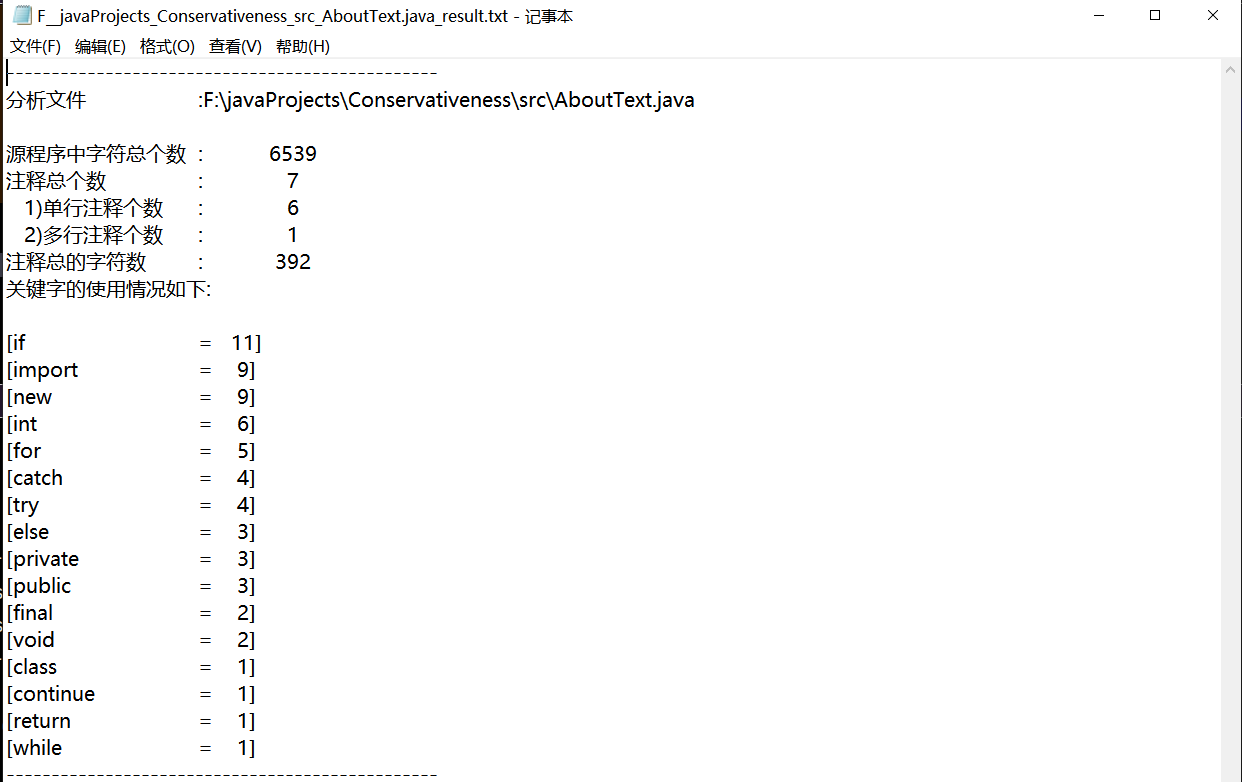
输入想要查看的结果最前面的文件编号，即可自动打开该txt文件。

同时，输出框也会有跟进动作，提示文件已自动打开，同时输入1可返回上一层，输入2可继续查看其他文件，0则可退出。

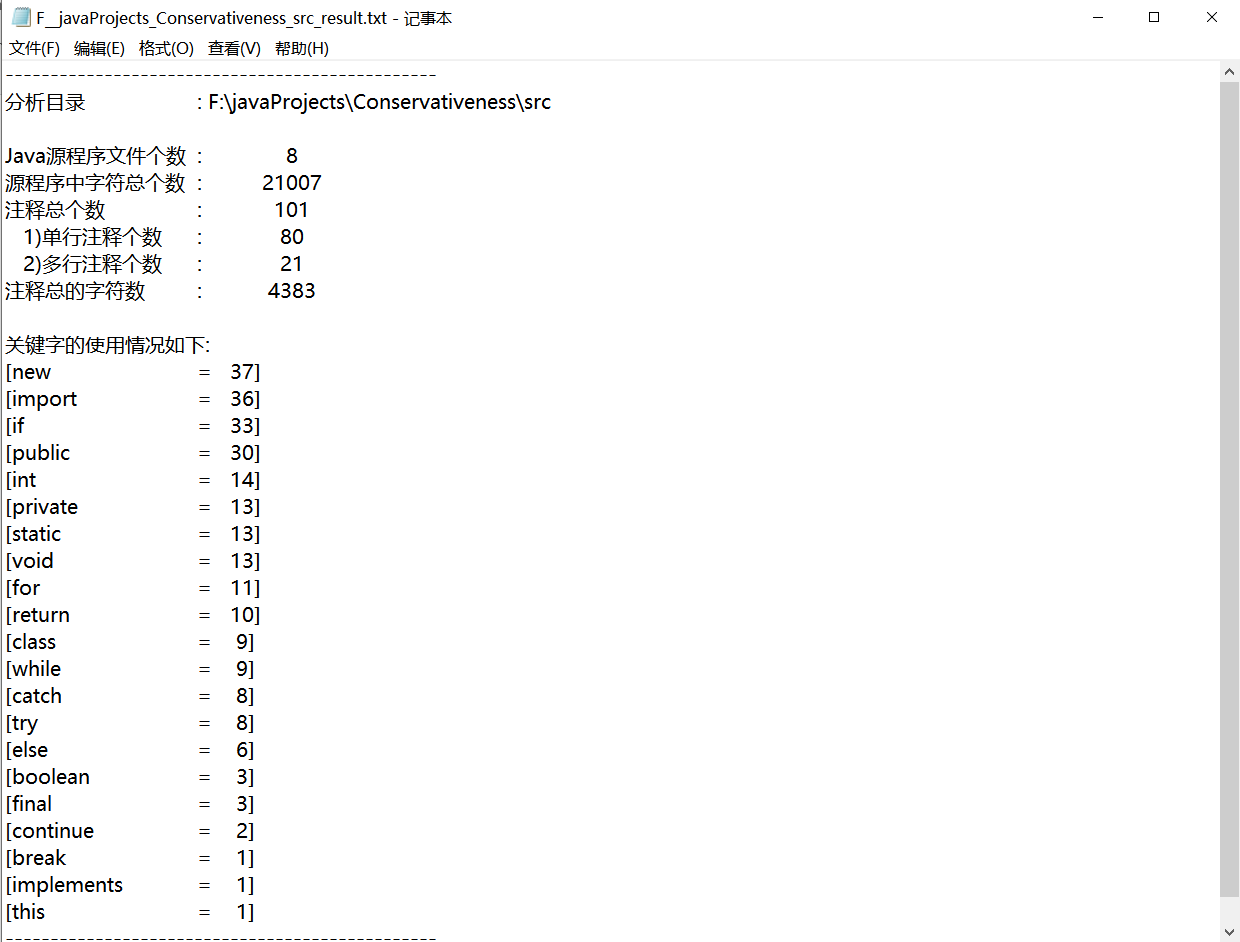


输入数字后弹出对应txt文件

输入数字后自动打开，即弹出对应txt文件。

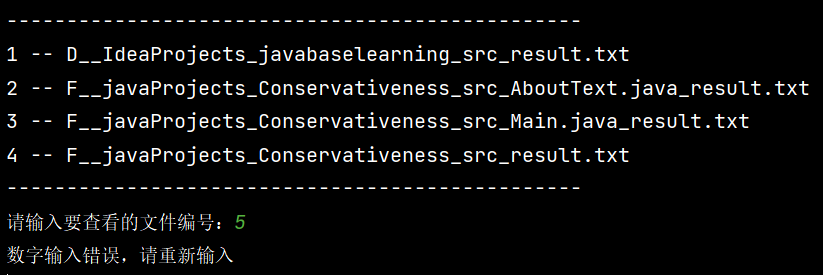


分析单个java源程序



分析目录

若输入错误数字，则会提示重新输入。



5. 实验总结

在这一次综合性实验中，我个人的Java技能得到了实际的应用，对于Java的各类知识和概念有了更深层次的认识。多功能程序的设计、面对对象的思想要求我们有抽象观、全局观。除此之外,为了满足实验要求，我还自学了一些知识，例如Java正则表达式、封装、聚合、组合、继承、IO等知识。在这一过程中拓展了面对对象程序设计的知识图谱。

在写代码的过程中，我有充分的认识，就是每个类都是流水线上面的个体，我要做的就是让每个类只操心自己的工作，而不需要关注数据从哪里来、是什么数据结构，我也非常清楚每个类的功能，每个类需要去实现的具体的功能，需要将数据处理到什么样的效果。

在程序测试部分，发现bug、解决bug的过程中，需要细心和耐心，还有查阅资料的能力，在这些方面我都得到了经验和提升。

众所周知，实验报告是一个实验成果的结果的集中展示和靓丽门面，也是实验者总结自身、反思自我和提升自己的必由路径。对于产品来说，再好的产品没有介绍说明书，可能就会导致使用门槛的增高；对于一个程序软件来说，没有必要的说明和介绍，使用者也难以顺利的使用；对于科研结果来说，如果没有字面上的、图片上的解释和说明，其价值也很难被发现。我一开始没有认识到这点，我在2021年11月10几号完成代码部分后，便去做了其他事而没有攥写实验爆个鬼，导致最后实验报告过于简单和肤浅，在实验报告中，我没有完全展现出打代码过程中的一些理解和思考，在最后甚至发现了一些hug,以至于我在2021年12月3日凌晨1点还在完善我的实验，而且已经过了deadline几小时了。在此，感谢学委的通融，这也给我上了极为宝贵的一课：程序员不仅要注重代码的质量，还要注重文档的攥写，这不仅帮自己厘清了自己的思路，也帮他人可以更了解自己的产出。

在这一次综合性实验中，我的逻辑思维能力、自学能力、分类管理能力、探究问题能力都有了提升。君子曰：学不可以已”，我也会通过不断的学习，改进自己写代码的质量。完善整个设计的流程，掌握面向对象开发程序的真正内涵。最后，感谢老师的指导。