**编译原理课程实验报告**

**实验1：词法分析**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 姓名 | 靳贺霖 | | 院系 | | 计算机学院 | | | 学号 | | | 1160300312 | | |
| 任课教师 | | 辛明影 | | | | 指导教师 | 辛明影 | | | | | | |
| 实验地点 | | 格物208 | | | | 实验时间 | 2019-04-14 | | | | | | |
| 实验课表现 | | 出勤、表现得分 | |  | | 实验报告  得分 |  | | 实验总分 | | |  | |
| 操作结果得分 | |  | |
| **一、需求分析** | | | | | | | | | | 得分 | | |  |
| 要求：阐述词法分析系统所要完成的功能  词法分析是编译的第一阶段。词法分析器的主要任务是读入源程序的输入字符、将它们组成词素，生成并输出一个词法单元序列，每个词法单源对应于一个词素。这个词法单元序列被输出到语法分析器进行语法分析。词法分析器通常还要和符号表进行交互。当词法分析器发现了一个标识符的词素时，它要将这个词素添加到符号表中。在某些情况下，词法分析会从符号表中读取有关标识符种类的信息，以确定向语法分析器传送哪个词法单元。  词法分析器在编译器中负责读取源程序，因此还会完成一些识别词素之外的其他任务，任务一是过滤掉源程序中的注释和空白，另一个任务是将编译器生成的错误消息与源程序的位置联系起来。 | | | | | | | | | | | | | |
| **二、文法设计** | | | | | | | | | | 得分 | | |  |
| 要求：对如下内容展开描述   1. 给出各类单词的词法规则描述（正则文法或正则表达式）   各类单词的正则表达式如下所示：  **常数**：  **标识符**：  **注释**：  **关键字**: keywordstruct|if|else|for|do|while|return|int|float|char|double|boolean|void  |true|false|include|string  **运算符**：  **界符**：   1. 各类单词的转换图   **识别常数的DFA ：**    **识别标识符的DFA：**    **识别注释的DFA**    **关键字，运算符，界符的通过简单的字符匹配就可以完成，设计DFA简单但繁琐，不在这里表示。** | | | | | | | | | | | | | |
| **三、系统设计** | | | | | | | | | | 得分 | | |  |
| 要求：分为系统概要设计和系统详细设计。  （1）系统概要设计：给出必要的系统宏观层面设计图，如系统框架图、数据流图、功能模块图等以及相应的文字说明。  **程序设计的方法引用UML图如下所示：**    其中，State和Snap是我们设计的数据结构，其中没有相关的方法可以调用，Lex是我们继承了Application类的图形界面程序的启动类。其他类的关系如所示，MyController类作为总控类，调用的Judger类和DFA类中相应的方法来对整体输入进行识别，输出和移近等。DFA类设计模拟了DFA的识别，在识别各种字符中用了Judger中对不同字符的判别。Judger类作为一个方法提供的类，没有对其他类的方法进行引用。  （2）系统详细设计：对如下工作进行展开描述   * **核心数据结构的设计**   **Token类**：Token类中定义了要输出的三元组，第一维是原字符，第二维是这个字符对应的系统存储的字段，第三个字符存储的是这个字段对应的值；  **StateSnap类**：SteteSnap类中定义了当前数据流的读取状态，其中定义的String类型的input，int类型的current和forward。Input表示此次分析的输入，current表示当前正在识别的字符下标，forward表示前向搜索的字符下标；  **Judger类**：Judger类中定义了一系列表，用于关键字，运算符，界符的判别。同时还在Judger类中定义了每个判别对应的种编码。   * **主要功能函数说明**   **isKeyWord(), isOperator(), isBoundary(), isDigit(), isChar()**是定义在Judger类中的用来判断一个字符(串)是否为关键字，运算符，界符，数字，字母的判别函数，主要就是通过判别输入的字符(串)是否和要求的相等或者存在于已经定义好的关键字，运算符，界符表中。返回均为布尔类型的数据；  **getBoudaryName()**是定义在Judger类中用于获得界符对应的种编码方法；  **operatorJudge()**是定义在Judger类中用于判断操作的类型以及获取其对应的种编码的方法；  **digitDFA(), commentDFA(), identifierDFA()**是定义在DFA类中，用于从文件中读取DFA表，并对输入的字符进行判别的函数。对于DFA的识别，首先将状态设置到状态0（开始状态），然后读取当前状态对应的状态表的一行（读取能够转移的状态），如果当前识别的字符能和这一行的某个表示匹配，就转义到这个状态并将识别的标识current后移一位，将forward设置为current。当这个DFA无法匹配即识别结束时，判断当前状态，如果是终结状态就将正确识别的结果写入到保存结果Token的List中，并且如果遇到识别异常，遇到无法识别的字符或者在非终结状态结束，会记录错误信息并通过恐慌模式来识别下一个能够识别的字符前缀；  **browseFile()**是定义在MyController类中的用与执行点击按钮的事件和存放词法分析主控程序的部分，它调用了其他识别DFA的方法，并将结果输出到图形界面中。首先，实例化一个StateSnap类，将要分析的文件对应对于当前要分析的字符，如果是数字，就转入digitDFA中，如果要分析的字符以及前向搜索的一个字符为“/\*”时就进入commentDFA的识别，如果要分析的字符为下划线\_或者字母是，进入到identifierDFA的识别。如果都不是，再进行界符，运算符的匹配识别，如果均无法识别，提示错误进行下一个识别即可；  **GUI.fxml**中定义了javafx图形界面对应的xml，用于图形界面的生成。   * **程序核心部分的程序流程图**   G:\浏览器下载\未命名文件 (7).png | | | | | | | | | | | | | |
| **四、系统实现及结果分析** | | | | | | | | | | 得分 | | |  |
| 要求：对如下内容展开描述。   1. 系统实现过程中遇到的问题；   实验过程中一直无法理解词法分析的目的是什么，只识别出token序列对语义分析作用不大，如果要有实在的作用就要填写符号表，但是符号表的填写是在语义分析完成的，这就意味着只考虑词法分析无法填完整的符号表，在语义分析中需要再扫描一次源程序填写符号表。于是我就退而求其次，将符号表的位置先空出来，将标识符和常数等先提取出，在语义分析中再填写符号表。   1. 针对某测试程序输出其词法分析结果；   **测试使用的样例程序输入如下所示：**   1. **struct** stu{ 2. **char**[5] name; 3. **int** age; 4. **float** grade; 5. } student; 7. **int**[10] a; 8. **float** \_a = 3.1514; 9. **float** \_b\_s = 3.15e+6; 10. **float** c = 3.15e 11. boolean d = **true**; 13. /\*comment\*/ 14. **if**(\_a == 0){ 15. \_a++;  ^ 16. } **else** { 17. \_a--; 18. } 20. **for**(**int** i=0;i<10;i++){ 21. a[i]=0; 22. /\***this** is a wrong comment 23. }   **其中得到的词法分析结果如下图所示：**     1. 输出针对此测试程序对应的词法错误报告；   **提示的错误报告如下：**  **An error happened during identifying digit at line 10**  **Unrecognized symbol at line 15**  **An error happened during handling comments at line 22**  然后，**程序界面整体运行结果如下所示：**     1. 对实验结果进行分析。   提供的示例中，包括了不同类型的标识符（字母，下划线开头，剩余部分有字母，数字，下划线构成），以及不同类型的关键字（整型，浮点型，布尔型，结构体，以及分支结构和循环结构），还有运算符，界符，整数，浮点数常数等。然后在结果中均有体现，对于不同的正确识别的符号均查询表后被转换成种编码打印在右边，对于识别出的错误的符号，均打印出错误信息和位置，如这次测试提供的程序中在第10行有错误的浮点数定义，在第15行有无法识别的字符，在第22行有错误形式的注释，均在报告中显示。 | | | | | | | | | | | | | |
| 指导教师评语：  日期： | | | | | | | | | | | | | |