华南理工大学

《编译原理》课程实验报告

实验题目： 设计SAMPLE语言的词法分析器

姓名： 张博 学号：

班级： 计科2班 组别：

合作者： 无

指导教师： 董敏

|  |
| --- |
| **实验概述** |
| 【实验目的及要求】  实验目的：  设计一个词法分析器，能生成一系列token二元组(种别，单词的值)并且发现一些具体错误。  实验要求：   1. 启动程序后，先输出作者姓名、班级、学号（可用汉语、英语或拼音）； 2. 请求输入测试程序名，键入程序名后自动开始词法分析并输出结果； 3. 输出结果为单词的二元式序列（样式见样板输出1和2）； 4. 要求能发现下列词法错误和指出错误性质和位置：   字符常数缺右边的单引号；  注释部分缺右边的界符\*/（注释要求左右边分别用/\*和\*/界定）。  【实验环境】  操作系统：Windows win 10 |
| **实验内容** |
| 【实验过程】   1. 实验主要过程： 2. 熟悉实验要求以及sample语言的保留字和值。这部分通过观看实验文档了解清楚。 3. 了解通过NFA、DFA构造词法分析程序的方法。这部分我通过课本以及PPT进行学习分析。 4. 对任务目的、实现方式了解后开始写代码，通过c++来实现词法分析程序。我用的软件是VS2019。   4. 源码保存在complier-word.cpp中，标识符表、样例保存在test.txt、example.txt、example1.txt中。  二、实验结果  1. 能对给出的样例进行词法分析，生成一系列token二元式，内容为单词的种别码以及对应的值。    2. 能发现字符常量定义错误，即字符常数右边缺少单引号。     1. 能发现注释错误，即/\*不能和\*/配对的情况。     三、实验代码及原理分析：  1. 将词法分析的保留字表读入程序，这部分我把各个保留字按顺序放到指定文件test.txt中。在分析时便能直接按各个保留字所在的位置直接给保留字赋值。  2. 通过哈希表is\_reserved、word\_value、value\_word储存各个保留字以及对应的值。  在init()中实现对保留字的记录，is\_reserved用来判断一个字符串是否为保留字，word\_value用来记录保留字和对应保留字的值(因为在文件读入时我的顺序已经安排好了，所以用cnt来记录位置，同时这个位置的值就是对应保留字的值)。value\_word用来记录一个值对应的保留字，方便后面操作时的比较分析。  3. 对字符常量、注释以及符号做特别判断。  对字符常量的判断，先判断读入的字符是否为单引号’，是的话继续往下读，直到遇到匹配的另一个单引号’，如果输入字符串后续没有另一个单引号，则输出字符常量错误，否则就返回该字符常量并存储到token中。    对字符常量的判断，先判断读入的字符是否为/\*，是的话继续往下读，直到遇到匹配的\*/，如果输入字符串后续没有对应的符号\*/，则输出注释错误，否则略过这段注释部分，cur位置继续向下读取字符。  对特殊符号的判断，判断读入的符号是否为.或者\_或者:或者各种关系符号比如>,<,>=,<=等，这部分判断的是分隔符和逻辑运算符。     1. 对字母、数字以及标识符进行判断   首先用一个字符串tmp来储存遇到分隔符之前的字符串值，如果tmp是保留字则直接存储到token中并移动cur位置，否则调用word函数对tmp进行判断，判断其为字符常量、数字或者是标识符(如果数字非法，即数字中出现字母的情况，word会返回-1开头的pair，此时会报错，并提示整数格式错误)。  word函数用来判断传入字符串为标识符、字符常量或者数字，首先判断传入字符是否为空，是的话返回-1开头的pair表示错误，如果第一个字符为字母，则把该字符串定义为标识符并返回，否则则定义为整数，然后往下读字符，如果遇到字母或者其他非数字的情况，则返回-1开头的pair表示错误，并输出整数格式错误。   1. 将结果保存在二元式向量token中并打印结果   直接将读取过程中存储的pair一个一个打印出来。 |
| **小结** |
| 词法分析程序是后面语法分析、语义分析的基础，相关的内容知识也比较重要，我通过查阅课本、教学PPT等方式，了解到多种实现词法分析程序的方法，最后终于完成这个有趣的实验任务，这个实验对我的编码能力、代码工程能力是个很好的锻炼，让我受益匪浅。尤其是让我熟悉了对字符串的处理以及哈希表的运用，这是一个很好的实践例子，也让我能更加灵活高效的调用STL的各个函数，对C++11的特性理解也更加丰富。 |
| **指导教师评语及成绩** |
| 评语：  成绩：           指导教师签名：                                                 批阅日期： |