编译原理

实

验

报

告

**实验名称：实验一：词法分析**

**姓名： 方澳阳**

**学号： 180110115**

**学院： 计算机科学与技术**

**专业： 计算机类**

1. **实验目的与方法**

（1）通过本实验加深对词法分析程序的功能及实现方法的理解

（2） 设计并编程实现一个词法分析程序，对C语言源程序段进行词法分析，加深对高级语言的认识

（3）对C语言的文法描述有更深的认识，体会有穷自动机，编码表和符号表在编译的整个过程中的应用

1. **实验总体流程与函数功能描述**

**流程：**

1. **将代码串读取进来**
2. **根据编码表初始化c实验的种别码**
3. **对读进来的代码串进行扫描，循环判断读进来的字符串，将其转为token。**
4. **遍历所有token串，将标识符存到符号表中，并且初始化其他字段。**

**read\_program(): 从txt文件中读取代码,存到program**

**initMap(): 初始化C语言的种别码**

**getNext(); 使pos指针加1,使ch为下一个字符**

**scanner(); 扫描器,对ch进行判断,存入对应的**

1. **实验内容**

**编码表：**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| auto | 1 | switch | 26 | / = | 51 |
| break | 2 | typedef | 27 | : | 52 |
| case | 3 | union | 28 | ; | 53 |
| char | 4 | unsigned | 29 | ? | 54 |
| const | 5 | void | 30 | [ | 55 |
| continue | 6 | volatile | 31 | ] | 56 |
| default | 7 | while | 32 | ^ | 57 |
| do | 8 | - | 33 | ^ | 58 |
| double | 9 | = | 34 | { | 59 |
| else | 10 | -= | 35 | | | 60 |
| enum | 11 | -> | 36 | || | 61 |
| extern | 12 | ! | 37 | |= | 62 |
| float | 13 | != | 38 | } | 63 |
| for | 14 | % | 39 | ~ | 64 |
| goto | 15 | %= | 40 | + | 65 |
| if | 16 | & | 41 | ++ | 66 |
| int | 17 | && | 42 | += | 67 |
| long | 18 | &= | 43 | < | 68 |
| register | 19 | ( | 44 | << | 69 |
| return | 20 | ) | 45 | <<= | 70 |
| short | 21 | \* | 46 | <= | 71 |
| signed | 22 | \* | 47 | = | 72 |
| sizeof | 23 | , | 48 | == | 73 |
| static | 24 | . | 49 | > | 74 |
| struct | 25 | / | 50 | >= | 75 |
| >> | 76 | \ | 78 | # | 80 |
| >>= | 77 | \= | 79 | /\*注释\*/ | 81 |
| 常数 | 82 | 标识符 | 83 |  |  |

## Token串和符号表的逻辑结构及存贮结构

**Token串的存储使用一个动态数组存储，定义为vector<pair<int,string>>，数组中的每一个元素为一个pair，pair.first代表种类，pair.second表示具体的串。当pair.first为标识符时，可以通过pair.second去符号表中索引到对应的信息。**

**符号表使用Map作为存储结构。通过设置对应的string串作为key，可以在Map中得到其对应的value。符号表定义为map<string,SYSTABLE> sysMap,其中，SYSTABLE为一个类，里面存储了对应标识符的type和address。**

## 算法描述

1. **实验结果与分析**

**对实验的输入输出结果进行展示与分析**

1. **实验中遇到的困难与解决办法**

**描述实验中遇到的困难与解决办法，对实验的意见与建议或所取得的收获。**