C-语言词法分析器说明

**1.文件说明**

提交的压缩包内共包含10项文件：



|  |  |  |
| --- | --- | --- |
| 序号 | 名称 | 说明 |
| 1 | custom\_build\_rules文件夹 | Flex安装包内自带文件，无需对这些文件进行任何操作 |
| 2 | data文件夹 |
| 3 | FlexLexer.h文件 |
| 4 | win\_flex.exe |
| 5 | data.txt | 样例程序。在进行代码测试时，应对data.txt中的样例代码进行修改 |
| 6 | lex.l | Lex源程序 |
| 7 | lex.yy.c | flex生成的C-语言词法分析器程序 |
| 8 | lex.exe | 可执行的词法分析器 |
| 9 | 文档说明.docx | 本篇说明文档 |
| 10 | 样例识别结果.png | 可执行词法分析器对data.txt中的样例程序分析后的结果截图 |

**2.运行环境说明**

操作系统：Windows 10

工具：

|  |  |  |
| --- | --- | --- |
| 名称 | 下载地址 | 功能说明 |
| Flex | <https://sourceforge.net/projects/winflexbison/> | 将.l文件转换为lex.yy.c |
| MinGW | <https://sourceforge.net/projects/mingw/> | 提供gcc, gdb等功能环境 |
| MinGW安装教程 | <http://c.biancheng.net/view/8077.html> | MinGW安装教程 |

**3.测试方法**

3.1直接测试样例程序

在解压后的文件夹内，启动Windows自带的power shell或cmd。输入以下命令：.\lex.exe

3.2测试自定义程序

在解压后的文件夹内，修改data.txt中的程序内容后，启动Windows自带的power shell或cmd。输入以下命令：.\lex.exe

**4.测试结果说明**

下面以压缩包内自带程序data.txt作为演示进行说明。

/\*this is a annotation\*/

import pack;

var name = 'this is a name';

var num = 10;

var f\_num = -10.01;

var arr[2] = {0, 1, 2};

while(num <= 15){

    num++;

    print(name);

}

if(num != false){

    f\_num = num+10/(0x10);

}

else{

    num = 0x11>>1;

}

var str = input();

str = str + name;

switch(1):

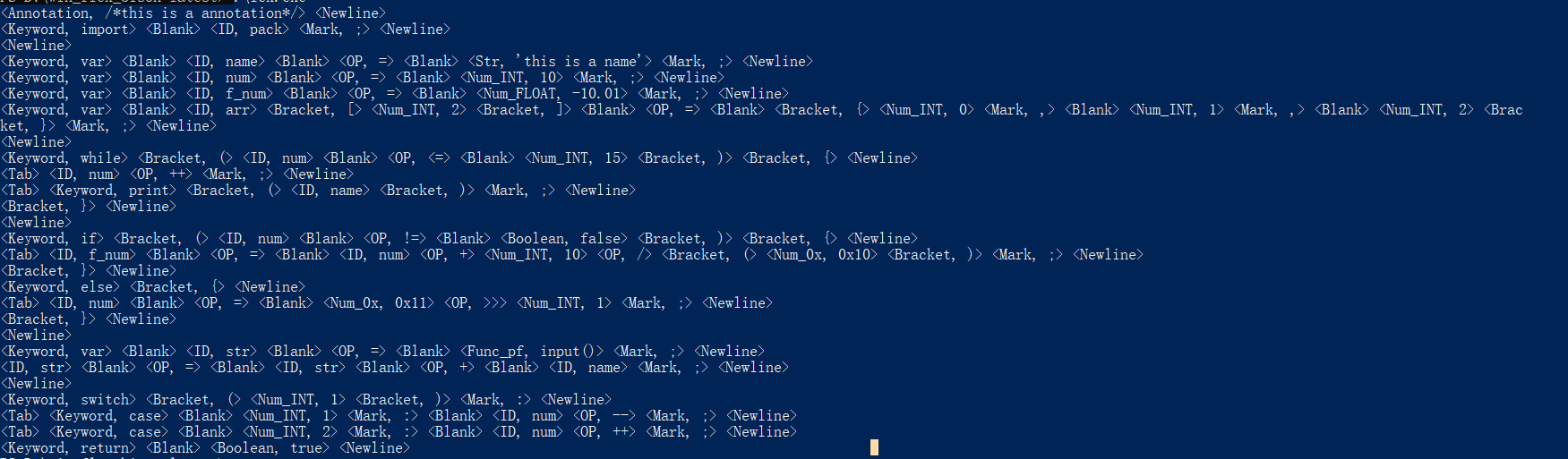
    case 1: num--;

    case 2: num++;

return true

这段程序包含了lex.l中定义的所有类型。lex.l中定义的类型及说明见下表：

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 类型名称 | 说明 | 举例 |
| 1 | **Annotation** | 注释 | /\*this is a annotation\*/ |
| 2 | **Keyword** | 关键字 | import, var, while, if, return |
| 3 | **ID** | 变量 | name, num, arr, f\_num |
| 4 | **OP** | 运算符 | =, +, -, \*, /, >>, <=, ++, -- |
| 5 | **Mark** | 标点 | , . : ; |
| 6 | **Boolean** | 布尔值 | true, false |
| 7 | **Bracket** | 括号 | ( ), [ ], { } |
| 8 | **Blank** | 空格 | ‘ ’ |
| 9 | **Newline** | 换行符 | /n |
| 10 | **Tab** | 缩进 | Tab |
| 11 | **Num\_0x** | 二进制数字 | 0x0101 |
| 12 | **Num\_INT** | 有符号整数 | 100, -100 |
| 13 | **Num\_FLOAT** | 有符号浮点数 | 1.10, -1.50 |
| 14 | **Str** | 字符串 | ‘this is a string’ |
| 15 | **Func\_pf** | 无参函数 | input() |
| 16 | **ERROR** | 无法匹配的内容 | 直接打出内容，而不是以<type, value>形式print |

运行结果展示及说明：  


结果中以**<type, value>**形式来展现每一个token的**类型(type)**和**对应的值(value)**，结果中的每一行词法分析结果对应样例程序的每一行源程序。

**5.出错解决方法**

|  |  |
| --- | --- |
| 出错提示 | 解决办法 |
| 找不到data.txt路径 | 将lex.exe和data.txt放在同一文件夹内重新编译 |
| gcc命令无效 | 安装MinGW, 并确认MinGW已经配置在系统环境变量中 |
| .\lex.exe无法识别 | 重新解压安装包。在解压后安装包文件夹内打开power shell 或 cmd并按照顺序输入以下指令：  step1: .\win\_flex --nounistd lex.l  step2: gcc -o lex lex.yy.c  step3: .\lex.exe |
| 其他任何问题 | 请联系：  QQ：842315974  Email：[imingjun.ma@outlook.com](mailto:imingjun.ma@outlook.com) |