# 编译原理实验1报告--石卓凡120L021011

## 程序实现了哪些功能？简要说明如何实现这些功能。

**1） 识别词法错误（错误类型 A）**

**2） 识别语法错误（错误类型 B）。**

**3） 识别单行注释“//”和多行注释“/\*…\*/”形式的注释**

**4） 构造与打印语法分析树**

### 识别单行注释与多行注释

**识别单行注释：**

如果识别到”//”,则一直while循环c=input(),直到输入的c=’\n’代表本次注释结束

**识别多行注释：**

标记在刚识别到 “/\*”状态flag=0

如果flag==0且此后识别到”\*”，状态转为flag=1

如果flag==1且此后识别到”/”,状态flag=2并完成识别

如果flag==1但此后识别到的不是”/”,则状态flag = 0

如果一直没匹配成功且结束时flag!=2,则识别为error B

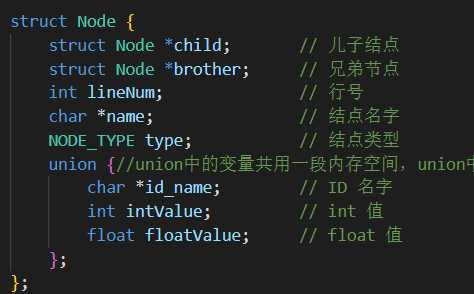
如果是直接识别到”\*/”，则识别为error B

### 语法分析树的构造与输出

单独编写tree.h,tree.c文件来负责语法分析树部分

**基础数据结构：**

* 1. 采用的是“孩子—兄弟”结点，链式存储



图：Node数据结构

* 1. 需要识别每个结点的类型，因此定义了结点类型为 int，并且规定了4种节点类型：，

|  |  |
| --- | --- |
| NONTERMINAL | 0 |
| INT\_TYPE | 1 |
| FLOAT\_TYPE | 2 |
| STRING\_TYPE | 3 |

**构造语法树的实现方案：**

* 1. struct Node \*createNode(char \*name, int line, NODE\_TYPE type)创建一个新的结点，指定节点名name，指定行号line，指定结点类型type，并返回新节点的指针
  2. struct Node \*insertNode(struct Node \*node, char \*name, int line, NODE\_TYPE type)向语法分析树中新创建一个结点，指定节点名name，指定行号line，指定结点类型type，然后设置其子女为我们给定的结点node，最后返回父结点所在的指针

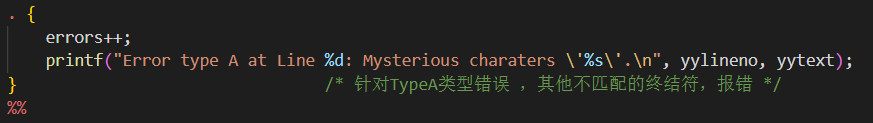
**输出语法树的实现方案：**

* 1. printNode（）根据结点的类型打印输出，就相当于打印语法树的一行信息
     1. 判断node->type类型==NONTERMINAL，且没有孩子的节点直接return，不再打印
     2. 否则，需要打印, 打印语法书前面的空白符号，根据depth深度决定空白符个数
        1. 判断node->type类型==STRING\_TYPE
           1. 如果是ID或者TYPE，打印node->name和node->id\_name
           2. 否则其他类型，只打印node->name
        2. 判断node->type类型==INT\_TYPE，输出 INT 和结点值
        3. 判断node->type类型==FLOAT\_TYPE，输出FLOAT和结点值
        4. 否则输出节点名字和行号
  2. printTree（）采用深度优先的方式（DFS）遍历语法树，方式采用的是函数递归。先孩子后兄弟

### 词法分析

**原理：**

首先根据c—词法，编写出对应的正则表达式，将所有需要匹配的就考虑完毕。那么剩下的未识别未定义的则都会被”.”匹配到，代表着识别到错误类型A，词法错误



利用此处代码，可以对其他不匹配的终结符进行TypeA报错

**如何输出行号：**

**利用**Flex内部提供的变量yylineno，记录着行号

**如何输出错误charaters：**

利用Flex内部提供的yytext，直接提供错误的charaters

### 语法分析

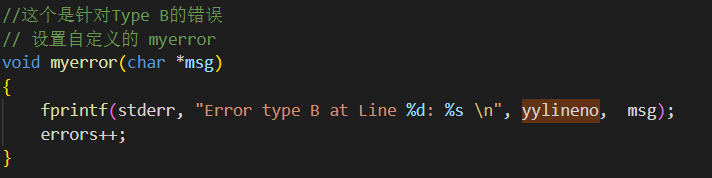
**原理：**

1. 根据给出的c—文法定义，我们可以定义所有的产生式，来匹配定义需要的所有情况。
2. 词法分析程序从yylex()获得了某词法单元
   1. 如果正常匹配了则继续执行
   2. 如果发现当前状态没有针对这个词法单元的动作，证明发生了语法错误。将error+1，并且准备进行错误恢复
3. 编写的带有error的产生式，用来匹配不正确的词法单元，并且输出相应的错误信息再实现再同步，使得这个错误之后的语法分析能够正常进行。将error尽可能放在终结符前面，少数情况也放到一些词法单元前面

**如何打印type B错误：**

重载并注释了yyerror函数，使得该默认函数失效。yyerror 函数是一个错误处理函数，重新编写了一个myerror函数来代替原有的yyerror函数

Myerror函数：接收错误提示信息msg，然后可以根据输出格式要求进行输出错误消息



## b）程序应该如何被编译？可以使用脚本、makefile或逐条输入命令进行编译，请详细说明应该如何编译你的程序。

若不报错，将输出到result.txt文件，若报错则直接输出到控制台

**编译命令为**

1. flex szf.l
2. bison -d syntax.y
3. gcc syntax.tab.c -lfl -ly -o parser tree.c

**运行测试样例命令**

1. chmod +x parser
2. ./parser test01.cmm（以test01.cmm为例）