

**编译原理实验报告**

题 目 语义分析

学 院 计算学部

专 业 软件工程

学 号 120L022109

学 生 李世轩

任 课 教 师 陈殷

哈尔滨工业大学计算机科学与技术学院

2023.4

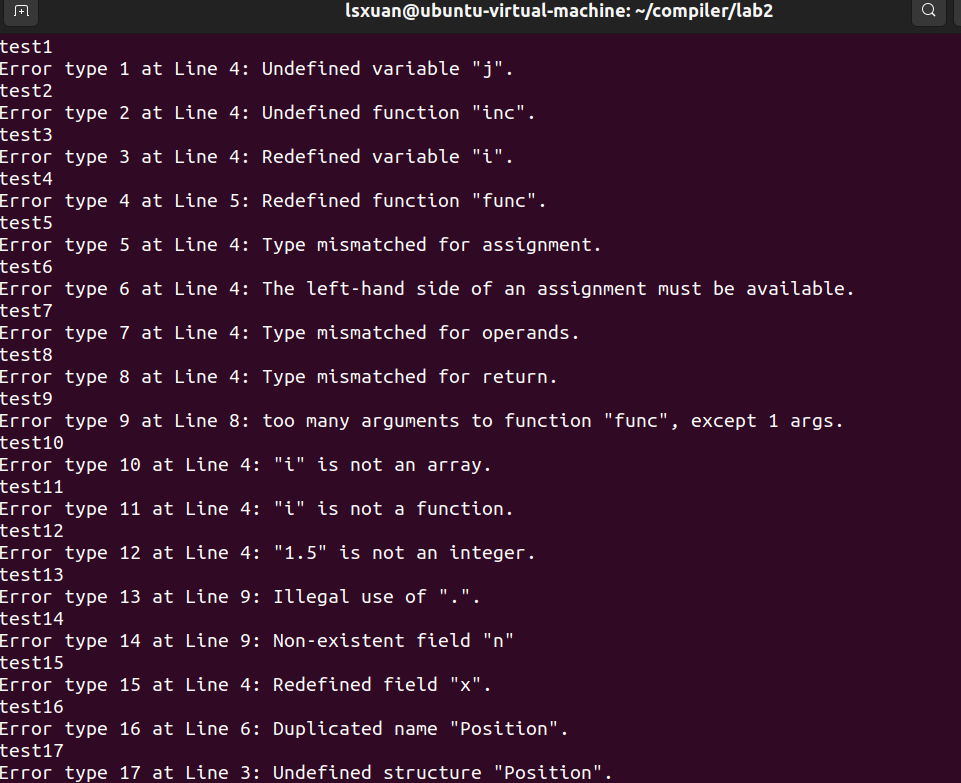
a）程序实现了哪些功能？简要说明如何实现这些功能。清晰的说明有

助于助教对你的程序所实现的功能进行合理的测试。

接受一个c--文件，对于那些没有语义错误的输 入文件，你的程序不输出任何内容。对于那些存在语义错误的输入文件，程序输出相应的错误信息，这些信息包括错误类型、出错的行号以及说明文字，其格式为：

Error type【错误类型】at Line【行号]：【说明文字】.

如图，即为对应指导书中的17个样例的输出。除了用例9，为了避免不断重复拼接字符串，其他输出说明文字格式均与指导书相同。



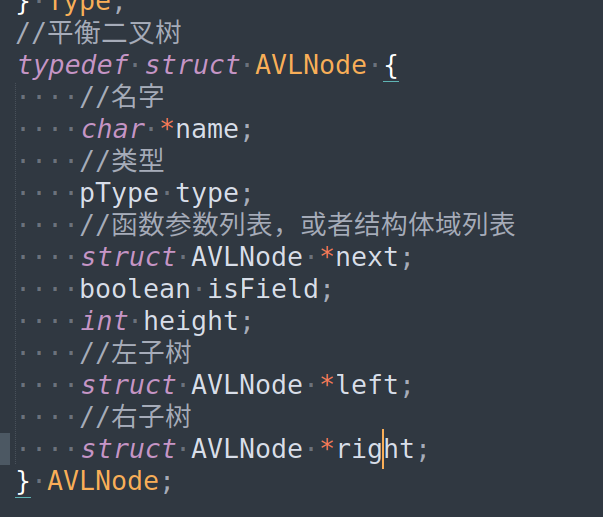
程序实现：

首先介绍以下定义的数据结构。

首先是关于符号表的定义，这里使用一个平衡二叉树作为符号表，name代表符号的名字，即变量名，函数名，结构体名等等。Type代表节点类型，其定义在 旁边。Next指针只想函数的参数或者结构体的域（这个指针的作用可以看作链表），heigh是当前二叉树节点的高度，用于计算二叉树的平衡因子。Left，right为二叉树的左右子树。isField用来标记这个符号是否为结构体的域，如果是，那么只能通过结构体变量来访问。

Type表示符号表中一项的类型，Kind是一个枚举类型，BASIC, ARRAY, STRUCTURE, FUNCTION分别代表基础数据类型（int，float），数组，结构体，函数。

共用体u中的BasicType代表int或float。array代表数组的一些信息，elem为数组中每个元素的类型，size表示数组的尺寸。Structure代表结构体名称（该值只有在声明结构体变量时才会被使用，在符号表中表示一个结构体定义时，这里不会有值。）function中argc代表参数数量，returnType代表返回类型。

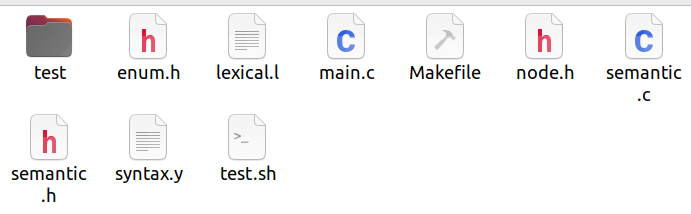


程序的算法中心思想就是，遍历整个语法树，当遇到变量声明，结构体定义，函数声明等等语句时，检查该符号是否已经出现在符号表中，若不冲突，向符号表插入语句。当在遇到变量被使用时，判断其是否被定义或操作类型是否正确的检查。

b）程序应该如何被编译？可以使用脚本、makefile或逐条输人命令进行

编译， 请详细说明应该如何编译你的程序。

程序目录如下



编译使用makefile，即切换到源程序目录下，使用make指令。

另外这里提供一个测试使用的脚本test.sh，其中每条执行指令的格式都如./parser ./test/test1.cmm，所以如果要使用该测试脚本，目录结构必须一样（test目录下存放所有的17个测试用例）。