# 实验 1：PL/0 语言编译器分析实验

### 一、实验目的

通过阅读与解析一个实际编译器（PL/0语言编译器）的源代码，

加深对编译阶段（包括词法分析、语法分析、语义分析、中间代

码生成等）和编译系统软件结构的理解，并达到提高学生学习兴

趣的目的。

### 二、实验要求

(1)要求掌握基本的程序设计技巧（C语言）和阅读较大规模程序

源代码的能力；

(2)理解并掌握编译过程的逻辑阶段及各逻辑阶段的功能；

(3)要求能把握整个系统（PL/0语言编译器）的体系结构，各功能

模块的功能，各模块之间的接口；

(4)要求能总结出实现编译过程各逻辑阶段功能采用的具体算法

与技术。

三．实验步骤  
 (1)根据PL/0语言的语法图，理解PL/0语言各级语法单位的结构，掌握PL/0语言合法程序的结构；

(2)从总体上分析整个系统的体系结构、各功能模块的功能、各模块之间的调用关系、各模块之间的接口；

(3)详细分析各子程序和函数的代码结构、程序流程、采用的主要算法及实现的功能；

(4)撰写分析报告，主要内容包括系统结构框图、模块接口、主要算法、各模块程序流程图等

##### 四．分析报告

①从总体上分析整个系统的体系结构、各功能模块的功能、各模块之间的调用关系

1. 系统整体结构：  
   文字描述：  
    编译过程采用一遍扫描方式，系统以语法、语义分析程序为核心，词法分析和目标代码生成作为辅助过程调用：在语法分析需要读单词的时候就调用词法分析函数；遇到标识符就调用表格管理程序进行查找；在语法、语义分析过程中遇到差错时，调用出错处理程序；当语法和语义分析完成之后就调用目标代码生成函数。
2. **各个功能模块的功能：**

1.模块功能：  
 i）词法分析模块：

识别单词，包括关键字、标识符等。

ii）语法、语义分析模块：

对源代码的语法检查和语义分析。

iii）目标代码生成模块：

对语法语义正确的语句进行机器指令的生成， 生成的代码顺序放在数组 code 中，code 为一维数组，数组元素为记录型数据。每一个记录就是一条目标指令。

iV）出错处理模块：

对词法和语法、语义分析遇到的错误给出在源程序中出错的位置和与错误 性质有关的编号，并进行错误恢复。

V）表格管理模块：

实现变量，常量和过程标识符的信息的登录与查找。

1. 主要函数的功能：

|  |  |
| --- | --- |
| 函数名 | 功能 |
| getsym | 词法分析，读取一个单词 |
| getch | 漏掉空格，读取一个字符 |
| gen | 生成目标代码，并送入目标程序区 |
| test | 测试当前单词符号是否合法 |
| block | 分程序分析处理过程 |
| enter | 登录符号表 |
| position | 查找标识符在符号表中的位置 |
| constdeclaration | 常量定义处理 |
| vardeclaration | 变量定义处理 |
| listcode | 列出目标代码清单 |
| statement | 语句部分处理 |
| expression | 表达式处理 |
| term | 项处理 |
| factor | 因子处理 |
| condition | 条件处理 |
| interpret | 对目标代码的解释执行程序 |

②详细分析各子程序和函数的代码结构、程序流程、采用的主要算法及实现的功能

（1）词法分析模块

读源程序（getch)、滤空格、识别保留字、识别标识符、拼数、识别单字符单词、拼双字符单词，得到一个合法的单词。

1. 语法分析模块

1.文字描述：

采用递归子程序法：对于每个非终结符号，编写一个子程序，由该子程序负责识别该语法单位是否正确。从读入第一个单词开始，由非终结符<程序>（即开始符）出发，沿语法描述图箭头所指出的方向进行分析。

当遇到非终结符时，则调用相应的处理过程（需要Select集），从语法描述图看，也就进入了一个语法单元，再沿当前所进入的语法单元所指箭头方向继续进行分析。

当遇到描述图中的终结符时，则判断当前读入的单词是否与图中的终结符相匹配，若匹配，再读取下一个单词继续分析。

遇到分支点时，将当前的单词与分支点上多个终结符逐个相比较，若都不匹配时可能是进入下一个非终结符语法单位或是出错。

1. 语义分析模块  
    文字描述：

对语句进行语法分析后，当遇到标识符的引用时就调用 POSITION 函数查 TABLE 表，看是否有过正确定义，若已有，则从表中取相应的有关信息，供代码的生成使用。若无定义则错。当语法语义正确时，就生成相应语句功能的目标代码。

1. 目标代码生成模块：  
    文字描述：

当语法语义都正确时，根据特定的语句，根据相应的机器指令格式生成目标代码，生成的代码顺序放在数组 code 中，code 为一维数组，数组元素为记录型数据。每一个记录就是一条目标指令。

1. 表格管理模块：  
    文字描述：

对于登录表格信息：根据单词的属性在表格中记录相应的信息。对于查找表格信息：根据标识符查找对应记录，返回标识符的相关信息。

1. 出错处理模块：

对于易于校正的错误，如丢了逗号，分号等，指出出错位置，加以校正，继续进行分析。对于难以校正的错误，给出错误的位置与性质，跳过后面的一些单词，直到下一个可以进行正常语法分析的语法单位。  
 具体为：在进入某个语法单位时，调用test，检查当前符号是否属于该语法单位的开始符号集合。若不属于，则滤去开始符号和后跟符号集合之外的所有符号。在语法单位分析结束时，调用TEST，检查当前符号是否属于调用该语法单位时应有的后跟符号集合。若不属于，则滤去后跟符号和开始符号集合之外的所有符号。