**编译原理课程设计**

# 0前言

**1、问题背景**

编译原理课程设计 与 实用编译器的比较

要点：实用编译器比编译原理课程设计要完成的复杂的多

**2、目的**

通过实践（设计与开发），对编译程序的设计原理及技术进一步的掌握和了解，培养学生计算思维能力及系统能力的提升。

**3、问题分析**

编译原理课程设计做什么？

设计并实现一个简单Pascal语言的编译器，主要包括几个模块：

词法分析模块

语法分析模块

语义分析模块

遵循软件的开发过程：分析、设计、编码、集成和测试

**4. Simple-PASCAL语言语法**

**源程序设计语言 G[<语句表>]**

参考编译原理教材179 上机实习题

**<语句表>→<语句> | <语句>;<语句表>**

**<语句>→<赋值语句>|<条件语句>|<WHILE语句>|<复合语句>**

**<赋值语句>→<变量>:=<算术表达式>**

**<条件语句>→IF<关系表达式>THEN<语句>ELSE<语句>**

**<WHILE语句>→WHILE<关系表达式>DO<语句>**

**<复合语句>→BEGIN<语句表>END**

**<算术表达式>→<项>|<算术表达式>+<项>|<算术表达式>-<项>**

**<项>→<因式>|<项>\*<因式>|<项>/<因式>**

**<因式>→<变量>|<常数>|(<算术表达式>)**

**<关系表达式>→<算术表达式><关系符><算术表达式>**

**<变量>→<标识符>**

**<标识符>→<标识符><字母>|<标识符><数字>|<字母>**

**<常数>→<整数>**

**<整数>→0|<非零数字><泛整数>**

**<泛整数>→<数字>|<数字><泛整数>|ε**

**<关系符>→<|<=|==|>|>=|<>**

**<字母>**

**→A|B|C|D|E|F|G|H|I|J|K|L|M|N|O|P|Q|R|S|T|U|V|W|X|Y|Z**

**<非零数字>→1|2|3|4|5|6|7|8|9**

**<数字>→<非零数字>|0**

**注释：**从接连出现的/\*到下一次接连出现的\*/之间的任何文字都是注释。

从某行接连出现的//到该行的结尾的任何文字都是注释。

**白空格：**两个单词之间的任何空格，制表符，回车，换行等都是白空格，除了用来分隔单词以外，没有意义。

**限制条件：**标识符的长度不大于8，整数不大于65535。

**5. 编写Simple-PASCAL语言源程序**

（1）什么是源程序？

需要被处理的程序叫做源程序。

转换程序的输入时源程序。

在这里，源程序就是使用语句表文法编写的程序

（2）编写符合上述语法的3个不同源程序，要求基本覆盖所有产生式，包含两种注释， 包含若干白空格， 满足限制条件。

## 第一部分：词法分析

### 1 词法分析任务

主要设计内容（二选一）：

（1）利用给出的类PASCAL文法，设计相应单词的有限自动机并进行化简；根据简化的有限自动机设计并实现词法分析器，将写好的源程序作为输入，进行词法分析并给出分析结果。

（2）利用给出的类PASCAL文法，以Flex、bison及C语言等工具编写的可被正确编译运行的源程序，将写好的源程序作为输入，进行词法分析并给出分析结果。（参考教材93页）

### 2词法分析

**1、词法分析的“处理”做什么事情？**

1）处理文件

读取源文件的内容

标准套路，（参考C语言的文件操作函数，fopen函数等）

2）把源文件的内容分解为词法单元

词法单元就是单词。

单词有类别（本次课程设计，请设计好单词的分类，参考教材79页）：

标识符为1类：如ABC，XYZ，C

整数设为2类：如123，900

IF为3类

THEN为4类

ELSE, WHILE, DO, BEGIN, END…….

运算符+， -， \*， /,>, <, =, >= , <=, <>, :=为101/102……

分隔符类： ；，（，）为201/201…….

白空格不是单词

注释以及注释标记不是单词

**练习3** 手工把练习1写好的源程序分解为单词类别和单词词文组成的二元式流

2、**怎样描述词法单元：正规文法，正规表达式**

上面给出的语句表文法实际上是该语言的BNF，其中包了单词描述和文法描述。需要从中提取出描述单词的正规文法或者正规表达式。

采用正规文法:

<保留字>→IF|THEN|……

<整数>→ 0|1<泛整数>|2<泛整数>

采用正规表达式:

(A|…|Z)(A|…|Z|0|…|9)\*|(0|…|9)|(1|…|9)(0|…|9)\*|IF|THEN|………….

**练习4** 写出描述语句表语言的词法单元的正规文法或者正规表达式（参考教材80页）

**3、怎样分解词法单元：有穷自动机，递归分析程序**

**怎样编写词法分析程序：**

1）预处理：把源文件一个字符一个字符的读入词法分析程序设置的输入字符结构体数组中（输入缓冲区），读入过程要删除注释，删除多余的白空格.

2）输入缓冲区要记录每个符号的ASCII码，以及该符号在源程序中的坐标值（行号和列号）

输入符号结构体数组：

struct inchar

{

char input；

int x，y；

}INCHAR[]；

**练习5** 编写预处理函数，从源文件得到输入符号流，要求按照上面的形式保留每个符号的坐标

3）从源程序字符数组中获得单词, 编码为二元式.

二元式采用结构体数组存储, 把单词类型和词元记录下来

例如:

struct dual

{ int dual\_type;

union {

char lexeme\_text[]；

int lexeme\_num[];

} lexeme

} DUAL[]

4）分解单词的方法：

1. 利用case多路转换语句根据单词的特点直接编写。
2. 通过描述单词的正规文法得到相应的有穷自动机，通过case多路转换语句完成有穷自动机的处理流程。参考课本P78的例子
3. 将有穷自动机转换为状态矩阵，构造状态矩阵驱动程序完成有穷自动机的流程。参考课本p55例子。（此处的课本是蒋立源主编的教材）

**4、编写词法分析程序要注意的问题**：

1)检查词法是否有错误

检查是否有非法字符：如a, aBC, @, &, !

检查是否有非法单词：如12A, BC\_1,

检查标志符和数字是否满足限制条件

检查注释符号是否配对

2)界符分隔单词

能够区分两个单词的符号为界符

有些界符不是单词：如白空格

有些界符仅仅用来分隔：如；

有些界符本身还是源程序不可缺少的单词，如(, ), +, /, 等等

有些界符包含两个字符：如<>, >=等等

3）输出词法错误

如果没有错误，返回0

如果有错误，需要报告词法错误在源程序中的位置。并且，要能够越过错误，分解下一个单词，直到源程序结束。

4) 输出的二元式流保存在二元式结构体数组中，供语法分析使用。

为了方便接下来的语法分析能够定位语法错误在源程序中的位置，需要在二元式结构体数组中保留每个单词在源程序中的坐标（单词的首字母所在的行号列号）。

实际的二元式结构体数组应该是：

struct dual

{ int dual\_type;

union {

char lexeme\_text[]；

int lexeme\_num[];

} lexeme

int x, y;

} DUAL[]

我们还是把它叫做二元式

**练习6**  编写词法分析程序，从输入符号流得到二元式流（参考教材81页算法）

词法出错提示格式：

Error type [错误类型] at line[行号]：说明文字

错误类型：A（词法错误）

**5、词法分析程序的总结**

1）预处理函打开源文件，读取源程序内容到设置好的字符数组（输入缓冲区），这个过程要删除注释，删除多余的白空格，所有的白空格最好都转为单个空格。预处理得到了字符流。

2）预处理后关闭源文件，对输入缓冲区进行词法处理：分解输入字符流，得到的二元式存入二元式结构体。词法处理以后得到二元式流

3）错误处理要完善，特别是坐标准确

**6、词法分析程序的要求**

（1）有正确的单词流输出（类型码、坐标、单词的值）

（2）能够检查单词的拼写错误，并报错

（3）能够识别注释，并去掉注释

（4）整数能够处理越界

（5）标识符识别前8位

## 第二部分：语法分析

### 1 语法分析任务

1、语法分析做什么事情？

语法分析检查源程序是否符合语句表中的语言文法。语法分析程序利用词法分析程序得到的二元式流，将源程序用指定文法进行分析，检查是否存在语法错误。

2、语法分析任务：利用词法分析的结果或者调用词法分析模块进行语法分析程序设计。

主要设计内容（选择其中一种方法）：

（1）递归下降分析法：对给定的文法进行改造，消除左递归，提取左公因子，编写递归下降分析程序；

（2）预测分析法：对给定的文法进行改造，消除左递归，提取左公因子，设计预测分析表，编写预测分析程序；

（3）SLR分析法：首先根据文法构造SLR分析表，然后编写分析表驱动程序，编写语法分析程序；

（4）利用语法分析器自动生成工具Bison编写语法分析程序。

设计要求：语法分析要求检查语法错误，报告语法错误，要给出语法在源程序中的位置，进行语法错误恢复。

### 2语法分析

**1、语法分析要注意的问题**

检查语法错误，类型很多，不一一列举

报告语法错误，要给出语法在源程序中的位置

语法错误恢复，通过语法界符同步，语法界符就是 ；或者 END

**2、递归下降分析法：**对给定的文法进行改造，消除左递归，提取左公因子，编写递归下降分析程序；

递归下降分析法参考教材P116-117的程序，在编写过程中，需要将该程序扩展，编写符合要求的全部语法程序，并给出分析过程，或构造出语法树

**3、预测分析法：**对给定的文法进行改造，消除左递归，提取左公因子，设计预测分析表，编写预测分析程序；

预测分析器的驱动程序的编写参考教材P120，对分析器的流程进行形式化，从而细化为程序。提前构造好相关分析表，教材中的分析表为算术运算，所以需要将该分析表扩展。要求：打印出如表4-3所示的分析过程。

**4、SLR分析法：**首先根据文法构造SLR分析表，然后编写分析表驱动程序，编写语法分析程序；相关的LR类分析器的驱动程序见教材P149的算法，对于指导中给出的语法结构去构造分析表，要求给出如表4-12所示的分析过程。

**5、利用语法分析器自动生成工具Bison编写**语法分析程序，首先熟悉GNU Bison工具，然后书写语法规则。

## 第三部分：语法分析

1、工作：通过语法制导翻译给出中间代码

2、中间语言采用通用的四元式格式，设计中间语言

3、根据语法分析所用的策略编写语法制导翻译文法，

自顶向下分析，L属性文法

自底向上分析，S属性文法

4、根据语法分析采用的策略，编写相应的属性翻译文法

**课程设计要求及成绩**

**1、本次课程设计分组进行，每组由1~3每名同学组成**

**2、每个模块要求先做需求分析，再按照原理进行设计，最后做程序实现。**

**3、考勤**

**4、要求在有效的时间内请指导老师检查程序，并提交课程设计报告**

**5、总评成绩=平时成绩×10﹪+考核成绩\*90﹪**

**其中，考核成绩=软件系统完成及答辩成绩×60﹪+报告成绩×40﹪**

**6、禁止迟到、早退，注意实验室纪律，有事需要提前请假**

**注意：由于时间紧，请大家合理计划任务，充分利用课后时间**

**时间：第14周~第17周**

**课程设计说明书要求及格式（参考）**

**题目：Simple PASCAL编译器的设计与实现**

**摘要**

**目录**

**1概述**

**2课程设计任务及要求**

**3 词法分析程序设计及实现**

**（包括文法、数据结构、分析过程、设计流程图、有限自动机、原理等）**

**4 语法分析程序设计及实现**

**（包括文法、数据结构、分析过程、分析表、设计流程图、原理以及主要算法等）**

**5 语法分析程序设计及实现**

**（包括文法、数据结构、分析过程、设计流程图和原理等）**

**6 系统结果及分析**

**7总结（遇到问题及解决方法、收获及体会等）**

**8 参考文献**

编号：

****

课程设计说明书

题 目： **一个Simple-PASCAL编译器**

**的设计与实现**

学 院： 计算机与信息安全学院

专 业： 计算机科学与技术

学生姓名：

学 号：

指导教师： 王慧娇

职 称： 副教授

√

√

**题目类型：**  理论研究 实验研究 工程设计 工程技术研究 软件开发

2020年6月20日