

## 实验名称：C 语言子集编译器

### 一、实验目的

通过设计编制调试一个具体的 C 语言子集编译器，加深对编译原理的理解。并掌握在对程序设计语言源程序进行编译过程中各个步骤的具体功能作用及它们之间的衔接。

### 二、实验要求与内容

设计编制调试一个具体的 C 语言子集编译器，主要包括以下步骤：

#### 1. 词法分析

设计、编制并调试该 C 语言子集的词法分析程序，要求结果以二元组形式保存在文件中。

#### 2. 语法分析

编制一个语法分析程序，分析方法不限，实现对词法分析程序所提供的单词序列进行语法检查和结构分析。

#### 3. 语义分析

采用语法制导翻译原理，将语法分析所识别的语法成分变换成四元式的中间代码序列，结果以文件保存。

### 三、实验预习提示

#### 1、C 语言子集编译器的功能和输出格式

C 语言子集编译器的功能是输入源程序，输出中间代码。要求实现最简单的 C 语言程序的编译工作：C 语言程序无子程序、无控制语句、包含四则运算，输入输出语句、变量常量定义。可根据个人能力增加功能。

#### 2、C 语言子集的文法

可参考课本 PL/0 的文法表示，要求用产生式的形式描述，并在实验报告中给出。

#### 3、模块结构、算法设计参照教材第二章和附录中的代码对 PL/0 语言的编译器的设计。

### 四、实验过程和指导：

#### 1、准备

- a. 阅读课本有关章节；
- b. 设计算法和流程；
- c. 初步编制好程序；
- d. 准备多组测试数据。

#### 2、为了能设计好程序，注意以下事情：

- a. 模块设计：将程序分成合理的多个模块（函数），每个模块做具体的同一事情。
- b. 写出（画出）设计方案：模块关系简图、流程图、全局变量、函数接口等。
- c. 编程时注意编程风格：空行的使用、注释的使用、缩进的使用等。

### 五、评定标准

#### 1、算法设计

- 1) 对 C 语言子集文法表示
- 2) 词法分析过程的设计
- 3) 语法分析过程的设计
- 4) 语义分析过程的设计

### 5) 中间代码的表示与优化过程的设计

## 2、整体运行

- 1) 系统整体算法的可行性、可读性
- 2) 系统整体算法的正确性、健壮性

## 3、实验报告

- 1) 实验报告各项的完整性：实验目的、实验要求与内容、实验过程、实验结果、分析体会（总结）；
- 2) 实验报告对实验完成过程的陈述：算法设计的具体情况、文法表示、各步骤地设计、部件间的衔接、模块划分、模块关系简图、流程图、全局变量、函数接口等。
- 3) 对实验的正确总结。
- 4) 要求在实验报告各项后面具有附录：

附录 A：编译系统代码

附录 B：系统使用说明

附录 C：参考书目

## 六、相关问题提示

### 1. 待分析的 C 语言的子集文法

- (1)  $\langle \text{程序} \rangle \rightarrow \text{main}() \langle \text{语句块} \rangle$
- (2)  $\langle \text{语句块} \rangle \rightarrow \{ \langle \text{语句串} \rangle \}$
- (3)  $\langle \text{语句串} \rangle \rightarrow \epsilon | \langle \text{语句} \rangle ; \langle \text{语句串} \rangle$
- (4)  $\langle \text{语句} \rangle \rightarrow \langle \text{赋值语句} \rangle | \langle \text{条件语句} \rangle | \langle \text{循环语句} \rangle$
- (5)  $\langle \text{赋值语句} \rangle \rightarrow \text{ID} = \langle \text{表达式} \rangle$
- (6)  $\langle \text{条件语句} \rangle \rightarrow \text{if} (\langle \text{条件} \rangle) \langle \text{语句块} \rangle$
- (7)  $\langle \text{循环语句} \rangle \rightarrow \text{while} (\langle \text{条件} \rangle) \langle \text{语句块} \rangle$
- (8)  $\langle \text{条件} \rangle \rightarrow \langle \text{表达式} \rangle \langle \text{关系运算符} \rangle \langle \text{表达式} \rangle$
- (9)  $\langle \text{表达式} \rangle \rightarrow \langle \text{项} \rangle | \langle \text{表达式} \rangle + \langle \text{项} \rangle | \langle \text{表达式} \rangle - \langle \text{项} \rangle$
- (10)  $\langle \text{项} \rangle \rightarrow \langle \text{因子} \rangle | \langle \text{项} \rangle * \langle \text{因子} \rangle | \langle \text{项} \rangle / \langle \text{因子} \rangle$
- (11)  $\langle \text{因子} \rangle \rightarrow \text{ID} | \text{NUM} | (\langle \text{表达式} \rangle)$
- (12)  $\langle \text{关系运算符} \rangle \rightarrow < | <= | > | >= | == | !=$

### 2. 待分析的 C 语言的子集词法

#### a. 关键字

main if else int return void while

关键字必须小写

#### b. 运算符和界符

= + - \* / < <= > >= == != ; : , { } [ ] ( )

#### c. 标志符 ID 和常数 NUM

通过以下正规式定义其他标记：

ID→letter(letter|digit)\*

NUM→digit digit\*

letter→a|...|z|A...|Z

digit→0|...|9

d. 空格由空白、制表符和换行符组成

空格一般用来分隔 ID、NUM、专用符号和关键字，词法分析阶段常常被忽略。

e. 各种单词符号对应的种别码

单词符号	种别码	单词符号	种别码
main	1	[	28
int	2	]	29
char	3	{	30
if	4	}	31
else	5	,	32
for	6	:	33
while	7	;	34
ID	10	>	35
NUM	20	<	36
=	21	>=	37
+	22	<=	38
-	23	!=	39
*	24	==	40
/	25	'\0'	1000
(	26	ERROR	-1
)	27		

f. 输出的四元式举例

对于语句串：

i=2\*3+4;

if (i>10) j=3;

while (j>0) k=1;

输出的四元式序列如下：

(1) (\*, 2, 3, T1)

(2) (+, T1, 4, T2)

(3) (=, T2, , i)

(4) (j>, i, 10, 6)

(5) (j, , , 7)

(6) (=, 3, , j)

(7) (j>, j, 0, 9)

(8) (j, , , 11)

(9)  $(=, 1, , k)$

(10)  $(j, , , 7)$

(11)...

## 附录 1：实验报告格式

### 实验报告

#### 一、实验项目名称

#### 二、实验目的

目的要明确，要抓住重点，可以从理论和实践两个方面考虑。在理论上，验证定理、公式、算法，并使实验者获得深刻和系统的理解；在实践上，掌握使用实验设备的技能技巧和程序的调试方法。一般需说明是验证型实验还是设计型实验，是创新型实验还是综合型实验。

#### 三、实验环境

实验用的软硬件环境。

#### 四、实验内容

说明本次实验的具体内容，要求实验内容饱满，难度适中。

#### 五、实验步骤

要求说明依据何种原理、定律、算法或操作方法进行实验，写清楚程序实现过程。

#### 六、源程序清单、测试数据、结果

实验中编写的源程序和测试数据以及程序运行结果。

#### 七、实验小结和思考

针对实验过程中出现的问题以及实验结果进行详细的分析，并写出本次实验的体会。