

编译原理实验题目

教学定位与基本安排

编译原理的课程实验既是培养学生理论联系实际的重要手段也是检验我们教学理念的贯彻与实现、检查教学效果、让学生展示能力的重要途径。

为了达到上述目的，作为编译原理课程的上机实验，最好是能完成一个小型语言的基本编译系统，但这需要有较好的实验条件特别是较多的实验学时（理想的情形是专门配置一个两到三周的课程设计）的支撑。对于一个只有 56 学时的编译原理课程的上机实验（仅为 12 学时），构造一个小型语言的完整的编译系统显然是不现实的。为最大程度达到教学目的，将本课程实验分解为 3 个部分：

- （1）词法分析器设计与实现；
- （2）语法分析器的设计与实现；
- （3）语法制导的三地址代码生成器的设计与实现。

虽然这 3 个部分可以分离（分阶段）实现，我们要求大家将其当作将一个简单语言翻译成等价的三地址码的编译系统的设计与实现的 3 个相关模块，在设计和实现中统筹考虑，并且在验收中作为一个系统进行验收。

通过这个系统的设计与实现，首先是引导学生走过构造一个微型编译系统的基本框架的主要工作流程，具体体验如何在理论指导下开展实践活动——理论是如何指导实践的，在实践中如何加深对理论的理解，培养学生按照编译理论和编译系统设计实现的基本方法，实现一个系统（理论指导实践）；其次是培养学生系统能力（系统的设、分析与实现）；第三是培养学生的软件系统实现能力（程序设计与实现）；第四是通过分小组，培养学生的团队合作精神与能力；第五是培养学生查阅资料，获取适当工具、使用适当工具，了解编译技术进展的能力；第六是培养学生表达（书面语口头）能力。

要求学生在理论学习的过程中就逐步考虑实现的问题，开始进行适当的设计与初步实现，在实验学时执行阶段，按照要求进行设计与实现工作，以此来解决 12 个学时的学时量不够的问题。在 12 个学时中，前 10 个学时留给各小组各自集中开

展研究、设计与实现工作，并准备验收报告（要求制作 PPT，准备好系统演示），两个学时用于“验收报告”（口头报告、系统演示），各组可以安排一人主持，保证全体成员参与报告。除了口头的报告外，每个小组还需要撰写一份实验报告，对系统的设计与实现进行描述。

按照要求设计实现这样一个系统，是本专业教学中的一个相对复杂的系统，能够体现对学生进行“复杂工程问题”处理能力的培养。首先，这是一个既有设计又有实现的工程问题；其次这个工程问题的解决必须运用编译系统设计的基本原理，包括使用一些基本的处理方法；第三，必须根据基本原理和系统实现的需要，通过抽象，实现对问题及其处理过程的分析和符号（形式化）描述，包括建立语言相关的描述模型；第四，至少包括 3 个主干子系统和一些辅助系统，整个系统具有一定的规模；第五，该系统的构建，不仅涉及到编译的基本原理，还涉及程序设计方法、算法设计、系统集成，以及相互间协调，按照要求，还涉及到问题的书面语和口头表达等。这些特征表明，该系统的设计与实现确实满足华盛顿协议等对“复杂工程问题”的界定，作为课程中的一项实验，能够体现对学生进行“复杂工程问题”处理能力的培养。

考虑到这是一次教学锻炼，可以根据教学班级的情况，采用随机方式安排分组，每组 3-4 人，各组随机安排组长一名，负责组织组内的分工协调等工作。提醒各组长重视并利用好这次机会，好好锻炼一下系统的设计、任务的恰当划分、人员的分工协调等工作，注意各个子系统之间的关联关系以及各系统负责同学之间的沟通。总体设计的角色不限于组长，本组内的任何成员都是总体设计的候选人。

实验一：词法分析程序的设计与实现

一. 实验目的

基本掌握计算机语言的词法分析程序的开发方法。

二. 实验内容

选择适当的方法（自行设计、使用 Lex 等自动生成工具），设计实现一个能够分析三种整数、标识符、主要运算符和主要关键字的词法分析程序。

三. 实验要求

1. 根据以下的正规式，编制正规文法，画出状态图；

标识符 $\langle \text{字母} \rangle (\langle \text{字母} \rangle | \langle \text{数字字符} \rangle)^*$

十进制整数 $0 | (1|2|3|4|5|6|7|8|9)(0|1|2|3|4|5|6|7|8|9)^*$

八进制整数 $0(1|2|3|4|5|6|7)(0|1|2|3|4|5|6|7)^*$

十六进制整数 $0x(0|1|2|3|4|5|6|7|8|9|a|b|c|d|e|f)(0|1|2|3|4|5|6|7|8|9|a|b|c|d|e|f)^*$

运算符和分隔符 $+ \quad - \quad * \quad / \quad > \quad < \quad = \quad (\quad) \quad ;$

关键字 `if then else while do`

2. 根据状态图，设计词法分析函数 `int scan()`，完成以下功能：

- 1) 从键盘读入数据，分析出一个单词。
- 2) 返回单词种别（用整数表示），
- 3) 返回单词属性（不同的属性可以放在不同的全局变量中）。

3. 编写测试程序，反复调用函数 `scan()`，输出单词种别和属性。

四. 实验环境

PC

五. 实验步骤

1. 根据状态图，设计词法分析算法
2. 采用适当语言，设计函数 `scan()`，实现该算法

3. 编制测试程序（主函数 main）。
4. 调试程序：输入一组单词，检查输出结果。

六. 测试数据示例

输入数据：0 92+data> 0x3f 00 while

输出结果：这些单词的单词种别及其属性

INT10	0	>	_
INT10	92	INT16	63
+	_	INT8	0
IDN	data	WHILE	_

七. 思考题

1. 词法分析能否采用空格来区分单词？
2. 程序设计中哪些环节影响词法分析的效率？如何提高效率？

实验二、三：语法制导的三地址代码生成程序设计与实现

一. 实验目的

掌握计算机语言的语法分析程序设计与属性文法应用的实现方法。

二. 实验内容

采用适当的方法（递归子程序法、LL 分析法、LR 分析法，使用 Yacc 等自动生成工具），实现一个能够进行语法分析并生成三地址代码的微型编译程序。

三. 实验要求

1. 考虑下述语法制导定义中文法，构造语法分析程序；
2. 参考下述语法制导定义，构造三地址代码生成程序。

产生式	语义规则
$P \rightarrow L$	$P.code = L.code$
$P \rightarrow LP_1$	$P.code = L.code \parallel P_1.code$
$L \rightarrow S ;$	$L.code = S.code$
$S \rightarrow id = E$	$S.code = E.code \parallel gen(id.place := 'E.place)$

$S \rightarrow \text{if } C \text{ then } S_1$	$C.\text{true} = \text{newlabel}; \quad C.\text{false} = S.\text{next};$ $S_1.\text{next} = S.\text{next};$ $S.\text{code} = C.\text{code} \parallel \text{gen}(E.\text{true}':') \parallel S_1.\text{code}$
$S \rightarrow \text{while } C \text{ do } S_1$	$S.\text{begin} = \text{newlabel}; \quad C.\text{true} = \text{newlabel};$ $C.\text{false} = S.\text{next}; \quad S_1.\text{next} = S.\text{begin};$ $S.\text{code} = \text{gen}(S.\text{begin}':') \parallel C.\text{code} \parallel$ $\text{gen}(E.\text{true}':') \parallel S_1.\text{code} \parallel \text{gen}(\text{'goto' } S.\text{begin});$
$S \rightarrow \text{if } C \text{ then } S_1$ $\text{else } S_2$	
$S \rightarrow \{ P \}$	$S.\text{code} = P.\text{code}$
$C \rightarrow E_1 > E_2$	$C.\text{code} = E_1.\text{code} \parallel E_2.\text{code} \parallel$ $\text{gen}(\text{'if' } E_1.\text{place}' > E_2.\text{place}' \text{goto' } C.\text{true}) \parallel$ $\text{gen}(\text{'goto' } C.\text{false})$
$C \rightarrow E_1 < E_2$	$C.\text{code} = E_1.\text{code} \parallel E_2.\text{code} \parallel$ $\text{gen}(\text{'if' } E_1.\text{place}' < E_2.\text{place}' \text{goto' } C.\text{true}) \parallel$ $\text{gen}(\text{'goto' } C.\text{false})$
$C \rightarrow E_1 = E_2$	$C.\text{code} = E_1.\text{code} \parallel E_2.\text{code} \parallel$ $\text{gen}(\text{'if' } E_1.\text{place}' = E_2.\text{place}' \text{goto' } C.\text{true}) \parallel$ $\text{gen}(\text{'goto' } C.\text{false})$
$E \rightarrow E_1 + T$	$E.\text{place} = \text{newtemp};$ $E.\text{code} = E_1.\text{code} \parallel T.\text{code} \parallel$ $\text{gen}(E.\text{place}': = E_1.\text{place}' + T.\text{place})$
$E \rightarrow E_1 - T$	$E.\text{place} = \text{newtemp}; \quad E.\text{code} = E_1.\text{code} \parallel T.\text{code} \parallel$ $\text{gen}(E.\text{place}': = E_1.\text{place}' - T.\text{place})$
$E \rightarrow T$	$E.\text{place} = T.\text{place}; \quad E.\text{code} = T.\text{code}$
$T \rightarrow F$	$T.\text{place} = F.\text{place}; \quad T.\text{code} = F.\text{code}$
$T \rightarrow T_1 * F$	$T.\text{place} = \text{newtemp};$ $T.\text{code} = T_1.\text{code} \parallel F.\text{code} \parallel$ $\text{gen}(T.\text{place}': = T_1.\text{place}' * F.\text{place})$
$T \rightarrow T_1 / F$	$T.\text{place} = \text{newtemp}; \quad T.\text{code} = T_1.\text{code} \parallel F.\text{code} \parallel$ $\text{gen}(T.\text{place}': = T_1.\text{place}' / F.\text{place})$
$F \rightarrow (E)$	$F.\text{place} = E.\text{place}; \quad F.\text{code} = E.\text{code}$
$F \rightarrow \text{id}$	$F.\text{place} = \text{id.name}; \quad F.\text{code} = ' '$
$F \rightarrow \text{int8}$	$F.\text{place} = \text{int8.value}; \quad F.\text{code} = ' '$
$F \rightarrow \text{int10}$	$F.\text{place} = \text{int10.value}; \quad F.\text{code} = ' '$
$F \rightarrow \text{int16}$	$F.\text{place} = \text{int16.value}; \quad F.\text{code} = ' '$

四. 实验环境

PC

五. 实验步骤——以递归子程序法为例

1. 考虑给定的文法，消除左递归，提取左因子；

2. 编制并化简语法图;
3. 编制递归子程序的算法;
4. 编制各个递归子程序;
5. 连接实验一的词法分析函数 `scan()`, 进行测试;
6. 设计实现三地址代码生成子系统的数据结构和算法;
7. 将各个递归子程序改写为代码生成函数;
8. 编制测试程序 (main 函数);
9. 进行调试: 输入一个语句, 检查输出的三地址代码; 输入一段程序, 检查输出的三地址码;
10. 根据测试情况, 对系统进行修改和优化。

六. 测试数据示例

输入数据 1:

```
while (a3+15)>0xa do if x2 = 07 then while y<z do y = x * y / z; c=b*c+d;
```

输入数据 2:

```
while (a3+15)>0xa do
    if x2 = 07 then
        while y<z do
            y = x * y / z;
        c=b*c+d;
```

正确结果: 等效的三地址代码序列

L1: t1 := a3 + 15	L4: t2 = x * y
if t1 > 10 goto L2	t3 = t2 / z
goto L0	y = t3
L2: if x2 = 7 goto L3	goto L3
goto L1	goto L1
L3: if y < z goto L4	L0: t4=b*c
goto L1	t5=t4+d

c=t5

七. 实验报告要求

(1) 综合报告

① 组内分工与贡献介绍；

② 系统功能概述；

③ 包含实验指导书要求的内容分系统报告；

✓ 词法分析子系统

i. 词法的正规式描述

ii. 变换后的正规文法

iii. 状态图

iv. 词法分析程序的主要数据结构与算法

✓ 如果采用自动生成技术，请描述采用的技术、相应的描述程序（如 LEX 程序）、辅助程序，以及生成的词法分析器的总体结构。

✓ 语法分析子系统

i. 根据选择的语法分析方法进行描述。例如，如果采用递归子程序法，建议包括改写后的产生式集合，化简后的语法图。

ii. 语法分析子系统结构

iii. 语法分析子系统的主要数据结构与算法

✓ 如果采用自动生成技术，请描述采用的技术、相应的描述程序（如 YACC 程序）、辅助程序，以及生成的语法分析器的总体结构。

✓ 三地址代码生成器

i. 语法制导定义。

ii. 算法基本思想

④ 实验体会。

(2) 源程序

① 源程序；

② 实验输入文件典型示例；

③ 对应的实验输出文件典型示例。

(3) 可运行程序包，其中启动文件名为“***学院 20**年第 X 组.exe”

八. 思考题

这些思考题作为实验设计与实现过程中要考虑的问题的示例，希望引导大家在实验过程中考虑相关的类似问题，并不是要求大家在实验报告后面像一个个简答题那样驻各地进行解答。相应的内容可以体现在综合报告的第四部分——实验体会中。

1. 词法分析能否采用空格来区分单词？
2. 程序设计中哪些环节影响词法分析的效率？如何提高效率？
3. 生成的三地址代码可否直接输出（不采用数据结构来实现属性 code）？
4. 如何保证四则运算的优先关系和左结合性？
5. 如何采用代码段相对地址代替三地址代码序列中的标号？
6. 完成一个实验，通常要进行实验设计、实验实现、实验结果分析，你在本次实验中，是否都涉及的这三个类工作，在大学学习的这几年的实验中，是不是都或多或少设计了这三类工作？