

《编译原理》实践指导书

适用专业：计算机类各专业

杭州电子科技大学信息工程学院

2017 年 3 月

前言

“编译原理”是计算机类专业的一门重要专业基础课。设置该课程的目的在于系统地向学生讲述编译系统的结构、工作流程及编译程序各组成部分的设计原理和实现技术,使学生通过学习既掌握编译理论和方法方面的基本知识,也具有设计、实现、分析和维护编译程序的初步能力。

“编译原理”是一门理论性和实践性都很强的课程。进行上机实验的目的是使学生通过完成上机实验题目加深对课堂教学内容的理解,同时培养学生实际动手能力。针对我校学生的实际情况,本实验指导书设计了 4 个实验项目,在实验内容方面难易适中,在实验要求方面分成不同的层次,循序渐进,力争使学生经过一定的努力,都能够完成相应题目,收获成功的喜悦,从而激发起他们学习的兴趣和积极性。实现项目中实验 1、3 为基于 Parser Generator 软件实现词法分析和语法分析;实验 2、4 为基于高级程序设计语言的词法分析和语法分析实现。

实验 1：基于 LEX 的词法分析实验

一、实验目的

熟悉 LEX 基本语法，掌握 Parser Generator 软件的使用；通过设计、开发通用高级语言一个单词种类的词法分析程序，加深对课堂教学内容（包括正规文法、正规表达式、有限自动机、NFA 到 DFA 的转换、DFA 的最小化）的理解，提高词法分析方法的实践能力。

二、实验要求

根据附录 A 中 S 语言的语法表示，选择或改造一种单词种类的语法表示，编写一个 LEX 程序，使之生成可识别该类单词符号的高级语言程序。

三、实验原理

LEX 是词法分析器的生成程序，LEX 源程序是用一种面向问题的语言写成的，核心是正规表达式，描述输入串的词法结构。LEX 可以将后缀为 1 的文件编译生成目标语言程序（如 C，C++ 及 JAVA 程序）。Parser Generator 是 Windows 环境下的 LEX 和 YACC 集成环境，它包括了一个图形用户界面，LEX 和 YACC 的两个版本 ALEX 和 AYACC。

LEX 编译器的工作原理是：LEX 编译器将 LEX 程序中的正规式经过若干步骤的转换（正规式->NFA->DFA->最小状态 DFA），最终转换成相应的等价确定有限状态自动机，并将其动作插入到 LEX 输出程序的适当地方。控制流是由确定有限状态自动机的解释器掌握，解释器是 LEX 的构成部分，对不同的输入源程序来说解释器是相同

的。

四、实验步骤

- (1)熟悉 Parser Generator 的基本使用方法和 LEX 编程方法，并确定一种比较熟悉的语言作为生成后的源语言；
- (2)使用正规表达式正确表示有关单词符号；
- (3)编写 LEX 程序生成该类单词符号的识别程序；
- (4)分析生成的高级语言程序结构。
- (5)撰写实验报告。

实验 2：基于高级语言的词法分析实验

一、实验目的

熟悉有限自动机的基本原理，利用熟悉的高级语言构建用于识别单词符号的词法分析器。通过设计、开发一个高级语言的词法分析程序，加深对课堂教学内容（包括正规文法、正规表达式、有限自动机、NFA 到 DFA 的转换、DFA 的最小化）的理解，提高词法分析方法的实践能力。

二、实验要求

- (1) 深入理解、掌握有限自动机及其应用；
- (2) 掌握根据语言的词法规则构造识别其单词的有限自动机的方法；
- (3) 掌握 NFA 到 DFA 的等价变换方法、DFA 最小化的方法；
- (4) 掌握设计、编码、调试词法分析程序的技术与方法，
- (5) 根据附录 A 中 S 语言的语法表示，选择或改造一种单词种类的语法表示具体实现 S 语言（见附录 A）的词法分析程序。

三、实验原理

词法分析是编译过程的第一个阶段。它的任务是对输入的字符串形式的源程序按顺序进行扫描，根据源程序的词法规则识别具有独立意义的单词符号，并输出单词的序列。

有限自动机是描述程序设计语言单词结构的工具，而状态转换图是有限自动机的比较直观的描述方法。根据程序设计语言的词法规则构造描述该语言单词结构的有限自动机，获取识别各类单词的形式模

型，进而通过编程模拟该形式模型的运行，可实现词法分析程序。

四、实验步骤

- (1)根据 S 语言的语法规则，总结 S 语言的单词种类与各类单词的结构特征；
- (2)编写或改写 S 语言一类单词的正规表达式；
- (3)利用熟悉的高级程序设计语言编写有限自动机；
- (4)撰写实验报告。

实验 3：利用 YACC 生成语法分析器

一、实验目的

熟悉 YACC 基本语法，掌握 Parser Generator 软件的使用；通过设计、开发高级语言一个语法结构的语法分析程序。了解 YACC 处理二义性的方法。

二、实验要求

选用典型的二义性文法（如表达式文法，悬空二义性文法等）作为处理对象，编写 YACC 程序生成该文法的语法分析程序（注：语法分析前的词法分析可以简化，可以用单个字母表示一个单词）。

例如：生成如下文法表示的表达式对应的计算器：

$$\begin{aligned} \text{exp} \rightarrow & \text{exp} + \text{exp} \mid \text{exp} - \text{exp} \\ & \mid \text{exp} * \text{exp} \mid \text{exp} / \text{exp} \\ & \mid \text{exp} ^ \text{exp} \mid - \text{exp} \\ & \mid (\text{exp}) \mid \text{NUM} \end{aligned}$$

对于输入的中缀表达式，要给出计算结果。如 $3+(4*5)$ 结果应为 23。要求能连续处理若干个数学表达式，直到输入结果或文件结束。NUM 可以使用实验 1 中的整数或浮点数词法分析的分析结果。

三、实验原理

YACC 是一个编译程序的生成程序，用户提供的语言的语法描述规格说明它基于 LALR 语法分析原理，YACC 输入将自动构造一个该语言的语法分析器；同时，它还能根据规格说明中给出的语义子程序建立规定的翻译。

一般来说,二义文法有两种解决方法,一种是直接修改文法,使该文法不存在二义性;另一种方法是保留文法的二义性,当语法分析出现冲突的时候规定一种解决办法。YACC 在解决二义性上采用了比较简单实用的方法:

1、产生“归约-归约”冲突时,按照规则说明中产生式的排列顺序,选择排在前边的产生式进行归约;

2、当产生“移进-归约”冲突时,选择执行移进动作。并且可以再说明部分选用`%left`、`%right`以及该符号所形成语句的次序给终结符赋予优先级和结合性。

四、实验步骤

(1)熟悉 YACC 的基本使用方法,并确定一种比较熟悉的语言作为生成后的源语言。

(2)熟悉 LEX 与 YACC 程序之间的数据交换,掌握语义动作符号的使用。

(3)选择一种上下文无关文法作为处理对象编写 YACC 程序,生成词法和语法分析程序。

(4)修改 YACC 程序,测试 YACC 在解决二义性问题上的处理方式。

(5)撰写实验报告。

实验 4：基于高级语言的语法分析实验

一、实验目的

熟悉语法分析常用方法（LL (1), LR (0) 等）。通过设计高级语言一个语法结构的下推自动机，实现语法分析程序。掌握编译程序语法分析部分的构建。

二、实验要求

选用一种典型的上下文无关文法作为处理对象，如表达式文法，条件语句文法等（文法也可以是对 S 语言中的一个语法结构改造而来，不限）。利用熟悉的高级语言编写一个下推自动机，并实现对处理语言的句子分析。

三、实验原理

语法分析程序是对上下文无关语言中的句子的分析程序，分析结果是形成分析树或者是推导序列、归约序列。不同的分析方法会形成不同的分析结果。下推自动机是语法分析程序常用的分析模型，不同的分析方法采用不同的细化结构和分析方法。

如 LL (1) 自顶向下分析方法，首先需要对文法分析并构造预测分析表，只有没有多重入口的预测分析表才能进行确定的自顶向下分析。然后通过编程实现下推自动机，置入预测分析表模拟自顶向下分析上下文无关语言的句子，给出推导序列。

四、实验步骤

(1)根据 S 语言 BNF 形式的语法规则（见附件 A），写出 S 语言一种语法结构的上下文无关文法；或者选择一种典型的上下文无关文法

作为处理对象。

(2) 构建相应文法的预测分析表。

(3) 使用熟悉的高级语言编写一个下推自动机 (LL 或者 LR, 不限)。运行为输入的句子输出推导或归约产生式序列。

(4) 撰写实验报告。

附录 A: S 语言语法的 BNF 表示

- (1) $\langle \text{程序} \rangle \rightarrow [\langle \text{常量说明} \rangle][\langle \text{变量说明} \rangle]\langle \text{语句} \rangle$
- (2) $\langle \text{常量说明} \rangle \rightarrow \text{Const} \langle \text{常量定义} \rangle \{, \langle \text{常量定义} \rangle\};$
- (3) $\langle \text{常量定义} \rangle \rightarrow \langle \text{标识符} \rangle = \langle \text{无符号整数} \rangle$
- (4) $\langle \text{无符号整数} \rangle \rightarrow \langle \text{数字} \rangle \{ \langle \text{数字} \rangle \}$
- (5) $\langle \text{字母} \rangle \rightarrow a|b|c| \dots |z$
- (6) $\langle \text{数字} \rangle \rightarrow 0|1|2| \dots |9$
- (7) $\langle \text{标识符} \rangle \rightarrow \langle \text{字母} \rangle \{ \langle \text{字母} \rangle | \langle \text{数字} \rangle \}$
- (8) $\langle \text{变量说明} \rangle \rightarrow \text{Var} \langle \text{标识符} \rangle \{, \langle \text{标识符} \rangle\};$
- (9) $\langle \text{语句} \rangle \rightarrow \langle \text{赋值语句} \rangle | \langle \text{条件语句} \rangle | \langle \text{当循环语句} \rangle | \langle \text{复合语句} \rangle | \epsilon$
- (10) $\langle \text{赋值语句} \rangle \rightarrow \langle \text{标识符} \rangle = \langle \text{表达式} \rangle;$
- (11) $\langle \text{表达式} \rangle \rightarrow [+|-]\langle \text{项} \rangle \{ \langle \text{加法运算符} \rangle \langle \text{项} \rangle \}$
- (12) $\langle \text{项} \rangle \rightarrow \langle \text{因子} \rangle \{ \langle \text{乘法运算符} \rangle \langle \text{因子} \rangle \}$
- (13) $\langle \text{因子} \rangle \rightarrow \langle \text{标识符} \rangle | \langle \text{无符号整数} \rangle | '(\langle \text{表达式} \rangle)'$
- (14) $\langle \text{加法运算符} \rangle \rightarrow +|-$
- (15) $\langle \text{乘法运算符} \rangle \rightarrow *|/$
- (16) $\langle \text{条件语句} \rangle \rightarrow \text{if} \langle \text{条件} \rangle \text{ then} \langle \text{语句} \rangle | \text{if} \langle \text{条件} \rangle \text{ then} \langle \text{语句} \rangle \text{ else} \langle \text{语句} \rangle$
- (17) $\langle \text{条件} \rangle \rightarrow \langle \text{表达式} \rangle \langle \text{关系运算符} \rangle \langle \text{表达式} \rangle$
- (18) $\langle \text{关系运算符} \rangle \rightarrow ==|<=|<|>|>=|<>$
- (19) $\langle \text{当循环语句} \rangle \rightarrow \text{while} \langle \text{条件} \rangle \text{ do} \langle \text{语句} \rangle$
- (20) $\langle \text{复合语句} \rangle \rightarrow \text{begin} \langle \text{语句} \rangle \{; \langle \text{语句} \rangle\} \text{end}$

注：产生式中 \langle 、 \rangle 括起的部分表示一个非终结符号， $[$ 、 $]$ 括起的部分表示可选项， $\{$ 、 $\}$ 括起的部分表示可重复，符号 $|$ 表示“或”。

附录 B：实验报告基本格式

编译原理实验报告

实验 1	
实验 2	
实验 3	
实验 4	

姓 名：
学 号：
班 级：

实验一 XXXXXXXXXX

一、 实验目的

.....

二、 实验内容

.....

三、 实验设计

.....

四、 实验实现及结果

.....

五、 实验心得

.....