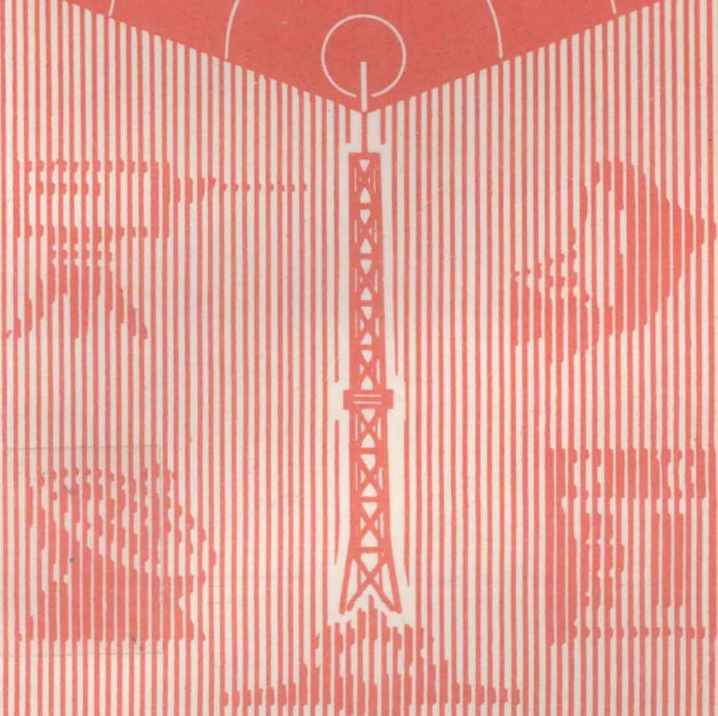


高等学校教材

程序设计语言编译方法

肖军模 编著



大连理工大学出版社

程序设计语言编译方法

肖军模 编著

大连理工大学出版社

(辽)新登字 16 号

内 容 提 要

本书以类 PASCAL 的 SIMPLE 语言为背景,介绍程序设计语言的各种编译技术。本书前三章重点介绍编译技术所涉及的概念、形式语言理论和程序语言的结构与定义方法;第四章至第九章分别介绍各种实用的词法分析、语法分析和语义翻译、存贮管理、局部优化、目标代码生成等技术;第十章介绍连接装配程序的设计方法。书中还按软件工程的观点介绍了 SIMPLE 语言编译器的设计方法。全书内容通俗易懂,理论性、系统性与实用性兼融。

本书可作为计算机专业的大专生、本科生及研究生的教材,也可以作为计算机科技工作者的参考书。本书学时:课内 60 小时,上机实验 20 小时。

程序设计语言编译方法

Chengxusheji Yuyan Bianyi Fangfa

肖军模 编著

* * *

大连理工大学出版社出版发行

(邮政编码: 116024)

大连业发印刷厂印刷

* *

开本: 787×1092 1/32 印张: 15 1/2 字数: 320 千字

1988 年 12 月第 1 版

1995 年 6 月第 2 版

1995 年 12 月第 5 次印刷

印数: 19 001—27 000 册

* * *

责任编辑: 王启太

责任校对: 宋玉珠 赵 静

封面设计: 孙宝福

* * *

ISBN 7-5611-0099-X

TP · 4

定价: 12.00 元

出版说明

根据国务院关于高等学校教材工作分工的规定,我部承担了全国高等学校、中等专业学校工科电子类专业教材的编审、出版的组织工作。由于各有关院校及参与编审工作的广大教师共同努力,有关出版社的紧密配合,从1978年至1985年,已编审、出版了两轮教材,正在陆续供给高等学校和中等专业学校教学使用。

为了使工科电子类专业教材能更好地适应“三个面向”的需要,贯彻“努力提高教材质量,逐步实现教材多样化,增加不同品种、不同层次、不同学术观点、不同风格、不同改革试验的教材”的精神,我部所属的七个高等学校教材编审委员会和两个中等专业学校教材编审委员会,在总结前两轮教材工作的基础上,结合教育形势的发展和教学改革的需要,制订了1986~1990年的“七五”(第三轮)教材编审出版规划。列入规划的教材、实验教材、教学参考书等近400种选题。这批教材的评选推荐和编写工作由各编委会直接组织进行。

这批教材的书稿,是从通过教学实践、师生反映较好的讲义中经院校推荐,由编审委员会(小组)评选择优产生出来的。

广大编审者、各编审委员会和有关出版社为保证教材的出版和提高教材的质量,作出了不懈的努力。

限于水平和经验,这批教材的编审、出版工作还会有缺点和不足之处,希望使用教材的单位,广大教师和同学积极提出批评建议,共同为不断提高工科电子类专业教材的质量而努力。

电子工业部教材办公室

前 言

本教材系按电子工业部制定的工科电子类专业教材 1986~1990 年编审出版规划,由计算机和自动控制教材编审委员会,计算机教材编审小组组织征稿、评选、推荐出版的。

本教材由南京通信工程学院肖军模同志编著,哈尔滨工业大学郭福顺同志担任主审。

本课程的课内教学时数为 60 学时,主要介绍程序设计语言的各种编译技术。全书共分为十章:第一章(引言)主要介绍编译技术的基本概念术语;第二章(编译基础)主要介绍与编译技术有关的形式语言理论;第三章(程序语言的结构与定义)介绍语言的组成特点、定义方法;第四章(词法分析)介绍扫描器的各种设计方法;第五章(语法分析)介绍算符优先、递归下降、LL(1)及 LR 等常用的语法分析技术;第六章(语法制导翻译技术)重点介绍自底向上的语义翻译方法,还介绍递归下降、LL(1)和属性翻译等语法制导翻译技术;第七章(运行时的存贮分配)主要介绍栈式管理技术,对静态与堆式管理技术也做适当介绍;第八章(代码优化)介绍利用树结构和 DAG 对基本块进行局部优化的方法;第九章(目标代码生成)介绍目标代码的生成方法和寄存器分配技术;第十章介绍各种运行准备工作。在四、六、八、九章中,分别介绍用软件工程观点设计 SIMPLE 编译器各部分的方法。考虑到不少院校都广泛采用陈火旺等同志主编的《程序设计语言编译原理》一书

作为编译教材,所以本书尽量采用《程序设计语言编译原理》中的名词术语,以使得教员在使用本书时感到方便。

使用本教材时,应要求学员耐心阅读书中的程序或算法,并适当多做一些习题或补充题。学好本书最好的方法是结合教学进度完成书中的上机习题(上机实验 20 小时),编写并调试一个完整的 SIMPLE 语言的小编译程序,或至少完成其中的词法分析与语法制导翻译部分。书中带“*”的章节只供选用。

参加本教材审阅的同志还有蒋宗礼同志,他为本书提出了许多宝贵意见。另外庄雯同志几乎抄写了全部书稿并做了一些文字加工,加快了书稿的完成时间。在这里向他们表示诚挚的感谢。由于编者水平有限,书中难免还存在一些缺点和错误,殷切希望广大读者批评指正。

编 者

1988 年 1 月 25 日

再 版 前 言

首先感谢广大读者对本书的厚爱,以及大连理工大学出版社的大力支持,使得本书有了一次再版的机会,有了一次增补与修改的机会。

本教材在我院教学实施过程中,同学们的反应是良好的。他们感到本书以形式语言原理为基础介绍各种编译方法与技术,学习起来逻辑性强,容易理解。同时,在教学过程中也发现了不少错误,其中有的是印刷错误、文字错误,有的则是内容错误。这次修定对已发现的错误做了相应的修改,但还不能说已全部修改完。希望广大读者发现问题后,给我来信,以便今后有机会再修改。

在这次修订中增加了词法分析器和语法分析器的自动生成原理。读者可以从这些内容中了解到任何问题只要形式化了,就可以由计算机自动生成。这一观念对目前计算机界正在深入研究的计算机辅助软件工程(CASE)工具有很重要意义。

另外,这次修订还附上了部分较难习题的答案,为读者自学提供方便。但一个习题往往可有多种解法,所附答案仅供参

考。答案大部分是王智学副教授做的。他用本书讲授编译课多年,提出了不少修改意见,在这里特别表示诚挚的谢意。

最后,再次谢谢读者,谢谢所有关心这本书和为这本书出版作出贡献的人们。

编 者

于南京通信工程学院

1995 年 4 月 12 日

目 录

第一章 引 言	1
1.1 计算机的语言层次与翻译	1
1.2 编译的阶段(phase)	3
1.3 编译的遍(pass)	7
1.4 SIMPLE 编译的结构设计	8
第二章 编译基础	10
2.1 字母表、串和语言	10
2.2 文法的引例.....	11
2.3 文法的形式定义.....	13
2.4 文法的分类.....	15
2.5 正规文法和有穷自动机.....	17
2.5.1 正规集与正规文法.....	17
2.5.2 有穷自动机.....	19
2.5.3 正规文法与有穷自动机.....	24
2.5.4 正规文法和正规式.....	26
2.6* 正规式、NFA 和 DFA 之间的等价变换	28
2.6.1 从正规式构造 NFA	29
2.6.2 NFA 的确定化	31
2.6.3 DFA 的最小化	34
2.7 上下文无关文法与下推自动机.....	39
2.7.1 上下文无关文法.....	39

2.7.2	上下文无关文法的变换	42
2.7.3	语法树与文法的二义性	45
2.7.4	下推自动机	47
习 题		52
第三章 程序设计语言的构造基础与定义		56
3.1	程序设计语言的构造基础	56
3.1.1	对象的名字、属性与汇聚	56
3.1.2	数据型对象及其存贮表示	59
3.1.3	运行时的存贮管理	64
3.1.4	运算及运算的复合	65
3.1.5	运算的顺序控制	66
3.1.6	参数传递	67
3.2	程序设计语言的定义	72
3.2.1	程序语言的语法定义	73
3.2.2	程序语言的语义定义	83
3.3	SIMPLE 语言的主要成分	86
习 题		86
附录 SIMPLE 语言的定义		89
第四章 词法分析		93
4.1	单词的种类与机内表示法	93
4.1.1	单词的分类	93
4.1.2	单词的机内表示法	94
4.2	扫描器的任务与设计考虑	97
4.3	扫描器的设计	99
4.3.1	扫描器的界面与数据结构	99
4.3.2	扫描器的数据流图	101

4.3.3 扫描器的详细设计	104
4.4 扫描器的其它设计技术	112
4.4.1 有穷自动机技术	112
4.4.2 正规式技术	116
4.5 某些高级语言的词法分析问题	117
4.6 词法分析阶段的错误处理	119
4.7 扫描器的自动生成原理	121
上机作业: 实现 SIMPLE 扫描器	126
习 题	126
第五章 语法分析	129
5.1 自顶向下分析法	130
5.1.1 自顶向下分析法中的问题研究	130
5.1.2 左递归与回溯问题	132
5.1.3 LL(1)文法的定义	135
5.1.4 LL(1)文法的判别	139
5.1.5 LL(1)文法的讨论	142
5.2 递归下降法	143
5.3 LL(1)分析法	148
5.4 自底向上分析法	153
5.4.1 自底向上分析法中的问题研究	153
5.4.2 句柄的定义与识别方法	156
5.5 算符优先分析技术	161
5.5.1 算符优先文法	162
5.5.2 算符优先矩阵	166
5.5.3 算符优先分析法的实现问题	171
5.5.4 算符优先分析的数据流图与	

模块结构图	175
5.5.5 算符优先分析技术的改进	179
5.6 LR 分析技术	182
5.6.1 概述	182
5.6.2 LR(0)分析表的构造方法	184
5.6.3 SLR(1)分析表的构造方法	191
5.6.4 LR 分析器总控程序	195
5.7 语法分析中的错误处理	198
5.8 各种语法分析技术的比较	200
5.9 语法分析器自动生成的原理	203
5.9.1 文法定义语言的功能	203
5.9.2 其它文法类分析表的自动生成问题	207
习 题	208
第六章 语法制导翻译技术	215
6.1 语法制导翻译概述	215
6.2 中间代码	217
6.3 自底向上的语法制导翻译	219
6.3.1 自底向上翻译的特点	219
6.3.2 说明语句的翻译	222
6.3.3 简单赋值句的翻译	225
6.3.4 布尔表达式的翻译	228
6.3.5 IF 语句的翻译	236
6.3.6 REPEAT 语句的翻译	239
6.3.7 FOR 循环语句的翻译	241
6.4 自顶向下的语法制导翻译	243
6.4.1 递归下降的语法制导翻译	243

6.4.2 LL(1)语法制导翻译	248
6.5 属性文法与属性翻译	253
6.5.1 属性文法与 L_属性文法	253
6.5.2 属性计值规则的改造与属性翻译	257
6.6 SIMPLE 语言的语法制导 翻译程序的设计	265
6.7 某些复杂语言成分的翻译	279
6.7.1 分程序结构的符号表	279
6.7.2 数组变量说明的翻译	283
6.7.3 过程语句的翻译	284
6.7.4 过程调用语句和返回语句的翻译	286
6.7.5 复杂赋值句的翻译	288
6.7.6 CASE 语句的翻译	293
6.7.7 输入输出语句的翻译	295
6.8 P-代码	298
上机作业: 实现 SIMPLE 语法制导翻译程序	301
习 题	301
第七章 运行时的存贮分配	307
7.1 数据区的内容	308
7.2 静态存贮分配	310
7.3 栈式存贮分配	314
7.3.1 栈式存贮管理和过程数据区格式	314
7.3.2 运行时地址的计算	322
7.3.3 过程调用与返回语句的目标代码	323
7.3.4 过程的目标代码结构和运行栈的管理	326
7.3.5 递归过程的调用	332

7.4 堆式存贮分配	334
7.4.1 堆空间的分配管理	336
7.4.2 堆空间的释放与敛集	339
习 题	343
第八章 代码优化	347
8.1 局部优化	347
8.1.1 基本块与程序控制流图的构造	348
8.1.2 适合于优化的中间代码	352
8.2 合并已知量	355
8.3 利用公共子表达式和消除无用赋值	366
8.4 其它优化技术介绍	374
上机作业: 调试 SIMPLE 代码优化程序(选做)	377
习 题	377
第九章 目标代码生成	380
9.1 目标机	380
9.2 各类四元式的翻译方法	383
9.3 寄存器分配与基本块代码的生成	385
9.3.1 引用信息与活跃信息	385
9.3.2 寄存器的分配问题	388
9.3.3 基本块的代码生成算法	389
9.4 DAG 的目标代码生成	392
9.5 SIMPLE 目标代码生成器的设计	399
9.6 目标代码的其它优化技术	403
9.6.1 窥孔优化	403
9.6.2 区域内寄存器的分配	404
上机作业: 实现 SIMPLE 代码生成器	406

习 题	406
第十章 运行准备	409
10.1 装配程序	410
10.1.1 绝对装配程序	410
10.1.2 重定位装配程序	411
10.2 外部访问的辨认	416
10.3 连接程序(LINKER)	417
10.4 库管理	423
10.5 存贮分配	424
习 题	427
部分习题参考答案	429
参考文献	475

第一章 引言

1.1 计算机的语言层次与翻译

在计算机系统中,语言的层次可用图 1.1 来描述。计算机使用的语言可以分为 3 个层次,高级语言层、汇编语言层和机器语言层。将源语言程序转换为另一种语言程序的程序通称

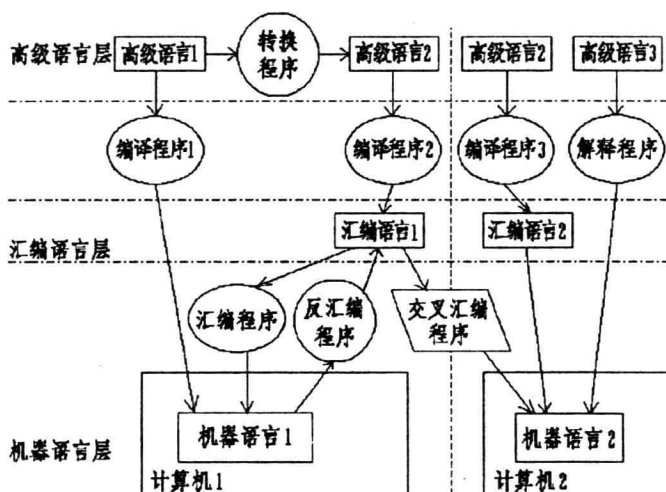


图 1.1 计算机系统语言层次和转换关系