计算机系列教材

编译原理与技术
（第2版）

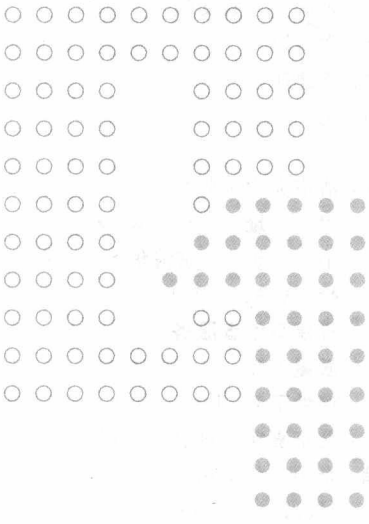
李文生 编著

清华大学出版社

计算机系列教材

李文生 编著

编译原理与技术
(第2版)



清华大学出版社
北京

内 容 简 介

本书系统地介绍了编译程序的设计原理和基本实现技术。主要内容包括词法分析、语法分析、语义分析、中间代码生成、代码生成和代码优化等，还重点介绍了用于实现语义分析和中间代码生成的语法制导翻译技术，以及程序运行时存储空间的组织与管理。

本书在介绍基本理论和方法的同时，也注重实际应用，介绍了 LEX 和 YACC 的使用方法及原理，剖析了 PL/0 语言的编译程序，介绍了 GCC 编译程序的基本结构。配合理论教学，给出了一些实践题目，旨在培养学生分析和解决问题的能力。

本书内容充实、图文并茂、各章节内容循序渐进，并注重理论与实践的结合。

本书可作为高等学校计算机科学与技术专业的本科生教材或参考书，也可供其他专业的学生或从事计算机工作的工程技术人员阅读参考。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。
版权所有，侵权必究。侵权举报电话：010-62782989　13701121933

图书在版编目(CIP)数据

编译原理与技术/李文生编著. --2 版. --北京：清华大学出版社，2016
计算机系列教材
ISBN 978-7-302-44141-0

Ⅰ. ①编…　Ⅱ. ①李…　Ⅲ. ①编译程序－程序设计－教材　Ⅳ. ①TP314

中国版本图书馆 CIP 数据核字(2016)第 139085 号

责任编辑：张瑞庆　柴文强
封面设计：常雪影
责任校对：李建庄
责任印制：沈　露

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http://www.tup.com.cn，010-62795954
印 装 者：北京嘉实印刷有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：27　　**字　　数**：659 千字
版　　次：2009 年 1 月第 1 版　2016 年 10 月第 2 版　　**印　　次**：2016 年 10 月第1 次印刷
印　　数：1～1000
定　　价：49.00 元

产品编号：068058-01

《编译原理与技术（第2版）》前言

根据作者在教学实践中的心得和体会，以及在第1版教材使用过程中得到的积极反馈和建议，结合编译技术的发展和应用，本书对第1版教材的部分内容进行了修订。

本书仍以Pascal语言作为教学语言来讲解编译原理和实现技术，但其中的算法描述采用类C语言代码。作此选择的主要原因有两个：一是Pascal语言和C语言是体现了结构化程序设计思想的代表性语言，但二者有明显的区别：Pascal语言更注重抽象，注重简明严格，是程序设计理论专家的研究成果，最初也是为了教学和编写系统软件而设计的，并且这两个目标都已达到；C语言更注重实用，注重表达式的简捷高效，在很多场合受到了软件开发人员的欢迎，逐渐成为软件开发的主流语言。二是Pascal语言规模适中，其全部语法规则采用形式化的语法描述形式BNF和语法图，作为计算机科学与技术专业相关课程的教学语言仍非常具有代表性，Pascal语言编译程序所用原理和技术对C语言同样适用。

本书继承和发扬了第1版教材理论充实、注重实践的特色，并对原教材需要改善的内容进行了修订，主要体现在：

（1）第4章，对LR分析程序进行了修改，主要涉及分析栈的组织和相应的分析动作，修改后的算法更便于理解和编码实现。

（2）第5章，根据第4章的修改，对相应内容进行了修正，增加5.5节，介绍通用的语法制导翻译函数的构造方法。

（3）第6章，内容组织做了调整，统一了示例语言文法，将声明部分的翻译方案和各语法成分的类型检查方案的内容组织在一节中，使得内容上下衔接更自然，便于理解。

（4）第8章，增加了访问记录中域的翻译方案，重新组织了8.5节的内容，从实现的角度细化了8.6节的内容，根据第6章的文法，重写了8.7节过程调用语句的翻译。

（5）第9章，修正了所用目标机器的指令格式，更便于熟悉Intel汇编语言的读者理解；给出了代码生成算法的伪码表示。

（6）第10章，重新组织了10.3节的内容，并给出了基本块DAG构造算法的伪码表示。

（7）第11章，这是新增加的一章，以C++语言为例介绍面向对象的编译方法。主要

介绍了面向对象语言的核心概念类和对象，及其主要特征继承性、封装性和多态性，介绍了方法的编译和继承的编译方法，以及面向对象语言程序运行时环境。

限于学识水平和时间，书中难免存在不妥或错误之处，欢迎大家批评指正。读者对本书有任何意见或建议，请发送电子邮件至：wenshli@bupt.edu.cn。

作 者

2016年8月

FOREWORD

试读结束：需要全本请在线购买： www.ertongbook.com

《编译原理与技术（第1版）》前言

“编译原理与技术”是计算机科学与技术专业的专业基础课程，通过本课程的学习，不仅可以提高编程技巧，掌握软件设计的一般方法，而且对计算机系统软件有一个比较清楚的认识和理解，为进一步的学习和研究打下良好的基础。

本书的前身是北京邮电大学出版社出版的《编译程序设计原理与技术》，主要介绍编译程序的设计原理和基本实现技术。根据多年的教学实践，对原书的内容进行了调整、补充和完善，并加强了实践环节。本书主要以 Pascal 和 C 语言为背景、就编译原理和技术有关的主要课题进行了系统和深入的讨论。

全书共分 11 章。第 1 章对编译程序的组成、功能及有关的前后处理器等进行了介绍，读者可以从中了解编译程序的概况。第 2 章介绍了有关形式语言与自动机的基本概念，这是学习编译原理必备的基础理论知识。第 3 章引入了词汇、模式等概念，介绍了利用状态转换图手工编写词法分析程序的方法和步骤，并对词法分析程序自动生成工具 LEX 的使用和工作原理作了介绍。第 4 章详细讨论了常用的语法分析技术，如适于手工实现的递归调用预测分析方法、适合利用分析程序自动生成工具实现的完备的 LR 分析技术等，介绍了语法分析程序自动生成工具 YACC 的使用方法等。第 5 章讨论了语法制导翻译技术，介绍了语法制导定义和翻译方案的概念，以及根据语法制导定义和翻译方案设计相应的翻译程序的基本方法，后继的语义分析和中间代码生成就是基于这种技术实现的。第 6 章介绍了语义分析的基本概念和要求，讨论了编译程序所用的重要数据结构——符号表的组织和管理，详细介绍了借助符号表、利用语法制导翻译技术实现类型检查的方法。第 7 章讨论了程序运行时的存储组织与管理问题，介绍程序运行相关的问题及解决方案，有助于读者理解程序设计中的问题，如非局部名字的访问、参数传递机制等。第 8 章介绍了中间语言，讨论了如何利用语法制导翻译技术把一般的程序设计语言结构翻译成中间代码。第 9 章介绍了目标代码生成的思想和一个简单的实现算法。第 10 章简单讨论了常用的代码优化技术。最后一章介绍了编译程序实现的一般方法，剖析了 PL/0 语言的编译程序，介绍了 GCC 编译程序的基本结构，并提供了一个课程设计题目，

按照软件工程的思想，对课程设计提出了基本要求，希望通过实际操作，有助于加深读者对编译原理的理解及对编译技术的掌握。

由于作者水平所限，书中难免存在缺点和不妥之处，真诚地希望得到广大读者和同行、专家的批评指正。

作 者

2008年7月

试读结束：需要全本请在线购买： www.ertongbook.com

第1章 编译概述

编译程序是现代计算机系统中重要的系统软件之一，是高级程序设计语言的支撑软件。众所周知，用高级程序设计语言（比如C/C++）书写的源程序是不能直接在计算机上运行的（起码现有的计算机不支持），要想运行它并得到预期的结果，首先必须把源程序转换成与之等价的目标程序，这个过程就是所谓的编译。编译原理与技术是计算机科学技术中的一个重要分支，现在已经基本形成了一套比较系统的、完整的理论和方法。编译原理与技术是计算机工作者所必须具备的专业基础知识。

编译程序的设计涉及程序设计语言、形式语言与自动机理论、计算机体系结构、数据结构、算法分析与设计、操作系统，以及软件工程等各个方面。本章通过描述编译程序的组成及编译程序的工作环境来介绍编译程序相关的基本概念，以便大家了解本课程的主要内容。

1.1 翻译和解释

1.1.1 程序设计语言

正如语言是人们进行交流的媒介和手段一样，在计算机应用领域，程序设计语言充当了人与问题，以及协助解决问题的计算机之间的通信工具。一种高效的程序设计语言应该能够提高计算机程序的开发效率、具有较好的问题表达能力，并在人们通常的思维方式与计算机执行所要求的精确性之间架起桥梁。程序设计语言同时也是开发人员之间的交流工具，在许多大型软件项目中，项目成功的关键在于程序员是否能读懂别人写的程序代码。

在计算机发展初期，程序员直接用机器语言编写程序。机器语言程序很不直观，难写、难读、难修改，并且对机器硬件的依赖性很强、移植性差。程序设计人员必须受过一定的训练并且熟悉计算机硬件，这在很大程度上限制了计算机的推广应用。

之后，出现了符号语言，即用比较直观的符号来代替纯粹数字表示的机器指令代码，这样使程序便于记忆、阅读和检查，在此基础上又进一步发展为汇编语言。在汇编语言中，除了用直观的助记符代替操作指令以对应一条条的机器指令外，还增加了若干宏指令，每条宏指令对应一组机器指令、完成特定的功能。这些宏指令构成指令码的扩展。汇编语言仍然是依赖于机器的，使用起来还是很不方便，并且程序开发效率也很低。

为进一步解决这些问题，John Backus等人参照数学语言设计了第一个描述算法的语言并于1954年正式对外发布，这就是FORTRAN Ⅰ，1957年第一个FORTRAN编译器在IBM704计算机上实现，并首次成功运行了FORTRAN程序。之后，又相继出现了得到广泛应用的过程性语言，如图灵奖获得者Peter Naur在1960年主编的《算法语言

试读结束：需要全本请在线购买： www.ertongbook.com