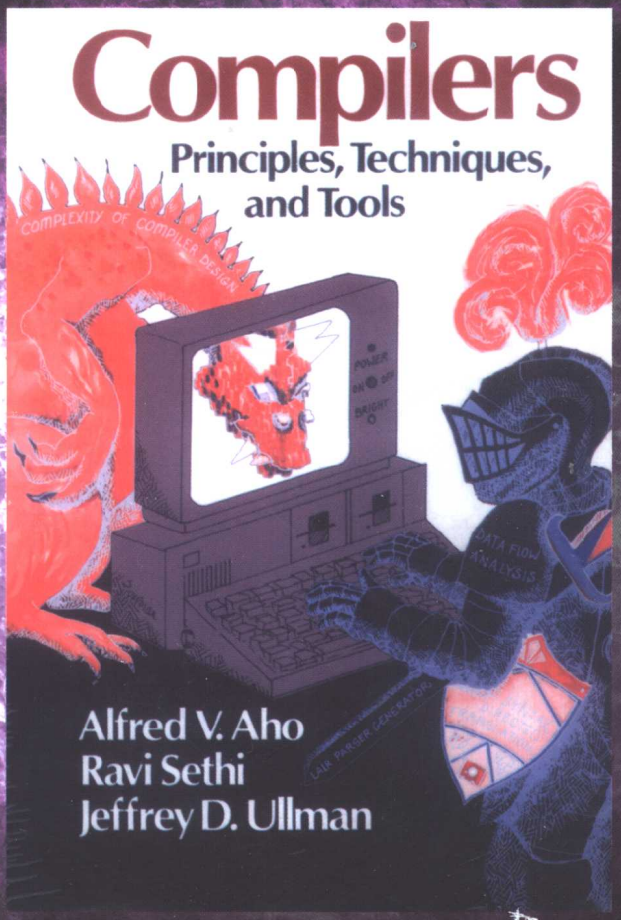
计 算 机 科 学 丛 书

编译原理

(美) Alfred V. Aho 贝尔实验室 Ravi Sethi Avaya实验室 Jeffrey D. Ullman 斯坦福大学 著 李建中 姜守旭 译



Compilers
Principles, Techniques, and Tools

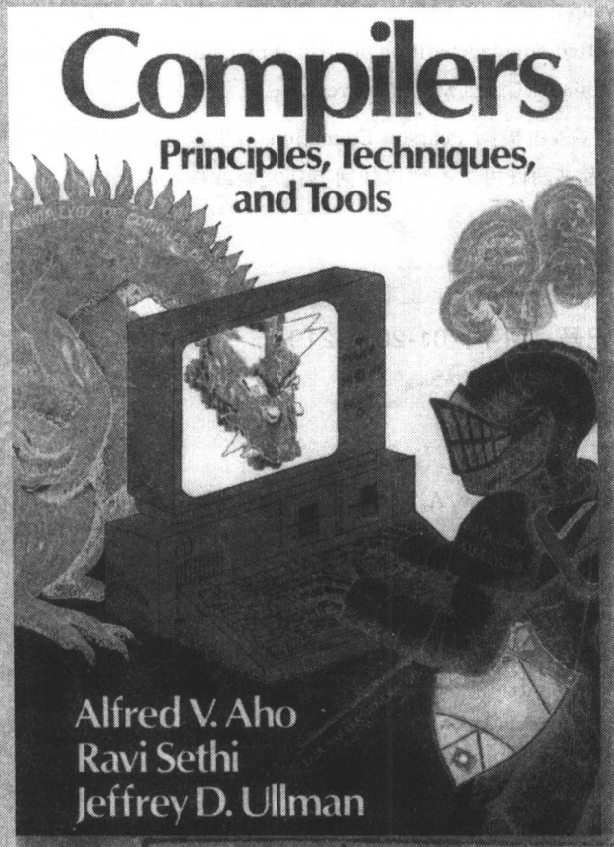
机械工业出版社
China Machine Press

PEARSON
Addison Wesley

计 算 机 科 学 丛 书

编译原理

(美) Alfred V. Aho 贝尔实验室 Ravi Sethi Avaya实验室 Jeffrey D. Ullman 斯坦福大学 著 李建中 姜守旭 译



Compilers
Principles, Techniques, and Tools

机械工业出版社
China Machine Press

本书深入讨论了编译器设计的重要主题，包括词法分析、语法分析、语法制导分析、类型检查、运行环境、中间代码生成、代码生成、代码优化等，并在最后两章中讨论了实现编译器的一些编程问题和几个编译器实例，每章都提供了大量的练习和参考文献。本书从介绍编译的原理性概念开始，然后通过构建一个简单的一遍编译器来逐一解释这些概念。

本书是编译原理课程的经典教材，作者曾多次使用本书的内容在贝尔实验室、哥伦比亚大学、普林斯顿大学和斯坦福大学向本科生和研究生讲授初等及高等编译课程。

本书版权登记号：图字：01-2001-2194

图书在版编目（CIP）数据

编译原理 /（美）阿霍（Aho, A. V.）等著；李建中等译. -北京：机械工业出版社，2003.8
（计算机科学丛书）
书名原文：Compilers: Principles, Techniques, and Tools
ISBN 7-111-12349-2

Ⅰ. 编… Ⅱ. ① 阿… ② 李… Ⅲ. 编译程序 - 程序设计 - 教材 Ⅳ. TP314

中国版本图书馆CIP数据核字（2003）第050118号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）
责任编辑：杨海玲
北京中加印刷有限公司印刷·新华书店北京发行所发行
2003年8月第1版第1次印刷
787mm × 1092mm 1/16 · 34印张
印数：0 001-5 000册
定价：55.00元

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall，Addison-Wesley，McGraw-Hill，Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum，Stroustrup，Kernighan，Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及庋藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师们服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T.，Stanford，U.C. Berkeley，C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程，而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下，读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证，但我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

电子邮件：hzedu@hzbook.com
联系电话：（010）68995264
联系地址：北京市西城区百万庄南街1号
邮政编码：100037

专家指导委员会

（按姓氏笔画顺序）

尤晋元	王　珊	冯博琴	史忠植	史美林
石教英	吕　建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周克定	周傲英	孟小峰	岳丽华	范　明
郑国梁	施伯乐	钟玉琢	唐世渭	袁崇义
高传善	梅　宏	程　旭	程时端	谢希仁
裘宗燕	戴　葵			

译 者 序

编译器产生于20世纪60年代，在计算机科学技术的发展历史中发挥了巨大作用，是开发计算机应用系统不可缺少的重要工具。编译器的原理和技术具有十分普遍的意义。在每一个计算机科学技术工作者的职业生涯中，这些原理和技术都被反复用到。编译器的编写涉及到程序设计语言、计算机体系结构、语言理论、算法和软件工程等学科，是计算机科学技术的重要基础。作为计算机科学技术学科的专业基础课，编译器原理和技术是计算机科学技术专业学生的必修课程。本书是一部优秀的编译器原理和技术教材。

本书是 Alfred V. Aho 和 Jeffrey D. Ullman 所著的《*Principles of Compiler Design*》一书的后裔版。本书作者 Alfred V. Aho 是 AT&T 贝尔实验室计算机原理研究部负责人，Jeffrey D. Ullman 是斯坦福大学计算机科学系教授，Ravi Sethi 是 AT&T 贝尔实验室研究人员。Alfred V. Aho 和 Jeffrey D. Ullman 是世界著名的计算机科学家，他们在计算机科学理论、数据库等很多领域都做出了杰出贡献。他们的很多著作都被国际公认为是权威性著作，深受读者的喜爱。本书的英文版出版于20世纪80年代，是一部著名的编译器原理与技术方面的教材，一直被国际著名高等院校特别是美国著名大学作为编译器原理与技术的教科书。这部著作对我国计算机界也具有重大影响。机械出版社独具慧眼，决定将这部著作翻译成中文在国内出版，这必将对我国计算机科学技术的编译原理教学工作产生积极的推动作用。有幸承担该书的翻译工作，我们感到十分荣幸。

本书是编译器原理与技术的基本教程，旨在介绍编译器的一般原理，解决人们在编译器设计中遇到的普遍问题。本书在系统地介绍编译器的一般原理的同时，特别注重编译原理和技术的实际应用，给出了许多启示，并配置了大量的例题和习题。本书的内容适用于所有源语言和目标机器。本书介绍的概念和技术不仅适用于编译器的设计，也适用于一般的软件设计。显然，本书在目前只有少数人涉及编译器的构造和维护的情况下仍然具有重要的意义和价值。

本书共有十二章和一个附录。第1章介绍编译器的基本结构；第2章描述了一个变中缀表达式为后缀表达式的翻译器；第3章介绍了词法分析器、正规表达式、有穷自动机以及词法分析器的自动生成工具；第4章详述常用的语法分析技术；第5章阐述了语法制导翻译的基本概念；第6章介绍了实现静态语义检查的基本思想；第7章讨论了程序运行环境的存储组织问题；第8章首先介绍了中间语言的概念，然后讨论如何把一般的程序设计语言翻译成中间代码的问题；第9章介绍目标代码生成技术和代码生成器的自动生成方法；第10章全面介绍了代码优化方法；第11章讨论了实现编译器的一些编程问题；第12章提供了几个编译器实例；附录A描述了一种简单的语言，学生可以把它作为源语言，构造一个编译器。

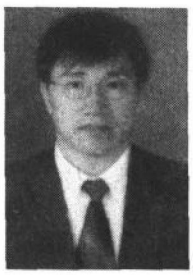
本书可以作为高等院校计算机专业本科生和研究生编译原理与技术课程的教材，也可以作为计算机软件工作者的技术参考书。

限于译者水平，译文中疏漏和错误难免，欢迎批评指正。

李建中，姜守旭

2003年7月1日

译者简介



李建中，哈尔滨工业大学教授，博士生导师，国家杰出青年基金获得者，中国计算机学会理事，中国计算机学会数据库专业委员会副主任。从事计算机科学技术的教学、研究、开发工作二十余年。主要研究领域为数据库系统与并行计算，主持完成研究项目20余项，在统计与科学数据库、并行数据库、数据仓库、数据挖掘等方面取得了一系列研究成果，在IEEE Transactions on Knowledge and Data Engineering、VLDB、ACM SIGMOD等国内外重要学术刊物和学术会议发表学术论文180余篇，出版学术专著和教材4部，获得各类科学技术奖励多项。

姜守旭，哈尔滨工业大学副教授，硕士生导师。1990年毕业于哈尔滨工业大学计算机及应用专业并留校任教，1995年获得计算机软件硕士学位。多年来一直从事计算机科学技术的教学与科研工作，主要研究方向为操作系统与对等计算。曾获部科技进步三等奖一项，出版教材一部。

前　言

本书是 Alfred V.Aho和Jeffrey D.Ullman 所著的《*Principles of Compiler Design*》一书的后续版本。与后者类似，本书也是编译器设计基础课程的教材。本书的重点是解决人们在设计语言翻译器时遇到的普遍问题，而不论源和目标机器是什么。

本书介绍的概念和技术不仅适用于编译器的设计，也适用于一般的软件设计。例如，建立词法分析器的串匹配技术已用于文本编辑器、信息检索系统和模式识别器；上下文无关文法和语法制导定义等概念已用于设计许多诸如本书产生的排版、绘图系统这样的小语言；代码优化技术已用于程序验证器和从非结构化程序产生结构化程序的程序检验器之中。显然，本书在目前只有少数人涉及编译器的构造和维护的情况下仍然具有重要的意义和价值。

本书的使用方法

本书深入地讨论了编译器设计的重要主题。

第1章介绍编译器的基本结构，是阅读本书其余部分的基础。

第2章描述了一个变中缀表达式为后缀表达式的翻译器。这个翻译器使用本书介绍的一些基本技术构建。后面的一些章节逐渐地扩展了第2章介绍的内容。

第3章介绍了词法分析器、正规表达式、有穷状态机以及词法分析器的生成器工具。本章的内容已经被广泛地用于文本处理。

第4章深入介绍了常用的语法分析技术。本书讨论的语法分析技术比较广泛，从适用于手工实现的递归下降技术到用于语法分析器生成器的计算更密集的LR技术。

第5章介绍了语法制导翻译的主要概念。本章的内容将被用于本书中说明和实现翻译的其余各章。

第6章介绍了实现静态语义检查的主要思想，详尽讨论了类型检查与合一问题。

第7章讨论了用于支持程序运行环境的存储组织问题。

第8章首先介绍了中间语言的概念，然后讨论如何把一般的程序设计语言结构翻译成中间代码的问题。

第9章介绍目标代码生成技术，包括简单的代码生成方法以及产生表达式代码的优化方法。本章也讨论了窥孔优化方法和代码生成器的生成器。

第10章全面介绍了代码优化，详细讨论了各种数据流分析方法和几种主要的全局优化方法。

第11章讨论实现编译器的一些编程问题。软件工程和软件测试在构造编译器的过程中是非常重要的。

第12章提供几个编译器实例。这些编译器都使用了本书介绍的技术。

附录A描述了一种简单的语言。这种语言是Pascal语言的“子集”，它可以用做实现项目的基础。

本书作者曾使用本书的内容多次在贝尔实验室、哥伦比亚大学、普林斯顿大学和斯坦福

大学为本科生和研究生讲授初等和高等编译课程。

初等编译课程可以由本书以下章节构成：

简介	第1章、2.1~2.5节
词法分析	2.6节、3.1~3.4节
符号表	2.7节、7.6节
语法分析	2.4节、4.1~4.4节
语法制导翻译	2.5节、5.1~5.5节
类型检查	6.1~6.2节
运行环境的组织	7.1~7.3节
中间代码的生成	8.1~8.3节
代码生成	9.1~9.4节
代码优化	10.1~10.2节

第2章描述了编程项目所需的信息。

以编译器构造工具为核心的课程可以包括3.5节的词法分析器的生成器、4.8节和4.9节的语法分析器的生成器、9.12节的代码生成器的生成器以及第11章中有关编译器构造技术的内容。

高等编译课程可以由本书的以下内容组成：第3章和第4章介绍的词法分析器的生成器和语法分析器的生成器中使用的算法、第6章介绍的有关类型等价、重载、多态和合一的内容、第7章介绍的程序运行环境的存储组织、第9章介绍的模式制导的代码生成方法、第10章介绍的代码优化。

习题

我们用星号来标记习题的难度。没有星号的习题检验学生对基本定义的理解；具有一个星号的习题用于高等编译课程；具有两个星号的习题是值得深思熟虑的难题。

致谢

在本书的编写过程中，很多人都对手稿给予了有价值的建议。在此，我们对以下人员表示衷心的感谢，他们是：Bill Appelbe、Nelson Beebe、Jon Bentley、Lois Bogess、Rodney Farrow、Stu Feldman、Charles Fischer、Chris Fraser、Art Gittelman、Eric Grosse、Dave Hanson、Fritz Henglein、Robert Henry、Gerard Holzmann、Steve Johnson、Brian Kernighan、Ken Kubota、Daniel Lehmann、Dave MacQueen、Dianne Maki、Alan Martin、Doug Mcllroy、Charles McLaughlin、John Mitchell、Elliott Organick、Robert Paige、Phil Pfeiffer、Rob Pike、Kari-Jouko Räihä、Dennis Ritchie、Sriram Sankar、Paul Stoecker、Bjarne Stroustrup、Tom Szymanski、Kim Track、Peter Weinberger、Jennifer Widom 和 Reinhard Wilhelm.

作者特别感谢 Patricia Solomon 对本书图片的加工制作。她的工作热情和专业水准是非常令人钦佩的。在本书的编写过程中，J.D.Ullman 得到了以色列艺术和科学学会的爱因斯坦奖学金的资助。最后，我们要特别感谢贝尔实验室在本书准备过程中对我们的支持。

A. V. A., R. S., J. D. U.

目　录