本书全面、深入地探讨了编译器设计方面的重要主题，包括词法分析、语法分析、语法制导定义和语法制导翻译、运行时刻环境、目标代码生成、代码优化技术、并行性检测以及过程间分析技术，并在相关章节中给出大量的实例。与上一版相比，本书进行了全面修订，涵盖了编译器开发方面最新进展。每章中都提供了大量的实例及参考文献。

本书是编译原理课程方面的经典教材，内容丰富，适合作为高等院校计算机及相关专业本科生及研究生的编译原理课程的教材，也是广大技术人员的极佳参考读物。

本书版权登记号：图字：01-2006-6521

图书在版编目（CIP）数据

编译原理　第2版/（美）阿霍（Alfred, V. A.）等著；赵建华等译. —北京：机械工业出版社，2009.1

（计算机科学丛书）

书名原文：Compilers: Principles, Techniques and Tools, Second Edition

ISBN 978-7-111-25121-7

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

中国版本图书馆 CIP 数据核字（2008）第 173435 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：朱　劼

北京京北印刷有限公司印刷

2009 年 2 月第 2 版第 2 次印刷

184mm × 260mm · 40.5 印张

标准书号：ISBN 978-7-111-25121-7

定价：89.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

本社购书热线：（010）68326294

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，还揭示了学术的源变，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和发展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到积极的推动作用，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分社较早意识到“出版要为教育服务”。自1998年开始，华章分社就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与Pearson，McGraw-Hill，Elsevier，MIT，John Wiley & Sons，Cengage等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出Andrew S. Tanenbaum，Bjarne Stroustrup，Brain W. Kernighan，Dennis Ritchie，Jim Gray，Afred V. Aho，John E. Hopcroft，Jeffrey D. Ullman，Abraham Silberschatz，William Stallings，Donald E. Knuth，John L. Hennessy，Larry L. Peterson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、研究及珍藏。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，有的还专程为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，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。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

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。随着计算机科学与技术专业学科建设的不断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将步入一个新的阶段，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。华章分社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：

华章网站：www.hzbook.com
电子邮件：hzjsj@hzbook.com
联系电话：(010) 88379604
联系地址：北京市西城区百万庄南街1号
邮政编码：100037

华章教育

译者序

绝大部分软件是使用高级程序设计语言来编写的。用这些语言编写的软件必须经过编译器的编译，才能转换为可以在计算机上运行的机器代码。编译器所生成代码的正确性和质量会直接影响成千上万个软件。因此，编译器构造原理和技术是计算机科学技术领域中的一个非常重要的组成部分。不仅如此，编译技术在当前已经广泛应用于编译器构造之外的其他领域，比如程序分析/验证、模型转换、语言处理等领域。因此，虽然大部分读者不会参与设计商用编译器，但拥有编译的相关知识仍然会对他们的研究开发生涯产生有益的影响。

A. V. Aho 等人撰写的《Compilers: Principles, Techniques, and Tools》被誉为编译教科书中的“龙书”。这说明这本书具有很高的权威性。我们有幸受机械工业出版社的委托，翻译龙书的第2版。

本书不仅包含了词法分析、语法分析、语义分析、代码生成等传统、经典的编译知识，还详细介绍了一些最新研究成果，比如过程间指针分析的最新进展。这使得本书不仅适用于编译原理的初学者，还可以作为研究人员的参考书目。本书不仅介绍编译器构造的基本原理和技术，还详细介绍一些有用的编译器构造工具，比如对 Lex 和 Yacc 的介绍使得读者可以了解这些工具的工作原理和使用方法。除此之外，读者还可以看到很多其他领域的概念在编译器构造中的应用。比如在第9章，读者可以看到群论中的抽象概念“格”被完美地应用于数据流分析算法的设计。而在第11章，线性规划和整数规划技术被成功地应用于程序并行化技术。这些内容对拓展读者的视野和思路有很大的好处。

由于本书覆盖的范围非常广，不可能在一个学期内讲完本书的全部内容。因此我建议采用本书作为本科生教材的老师只选择讲授其中的基础部分，即第1章到第9章中的大部分内容。第2章是对后面各章内容的介绍，可以在讲授相应内容之前指导学生预习。

最后感谢机械工业出版社的温莉芳女士以及姚蕾和朱劼两位编辑在本书的翻译过程中给予我们的有力帮助，也感谢其他给予我们支持的同事。由于水平有限，翻译中的错漏之处在所难免，欢迎读者批评指正。

译者

2008年6月于南京

前言

从本书的1986版出版到现在，编译器设计领域已经发生了很大的改变。随着程序设计语言的发展，提出了新的编译问题。计算机体系结构提供了多种多样的资源，而编译器设计者必须能够充分利用这些资源。最有意思的事情可能是，古老的代码优化技术已经在编译器之外找到了新的应用。现在，有些工具利用这些技术来寻找软件中的缺陷，以及(最重要的是)寻找现有代码中的安全漏洞。而且，很多"前端"技术——文法、正则表达式、语法分析器以及语法制导翻译器等——仍然被广泛应用。

因此，本书先前的版本所体现的我们的价值观一直没有改变。我们知道，只有很少的读者将会去构建甚至维护一个主流程序设计语言的编译器。但是，和编译器相关的模型、理论和算法可以被应用到软件设计和开发中出现的各种各样的问题上。因此，我们会关注那些在设计一个语言处理器时常常会碰到的问题，而不考虑具体的源语言和目标机器究竟是什么。

使用本书

要学完本书的全部或大部分内容至少需要两个学季(quarter)，甚至两个学期(semester)[⊖]。通常会在一门本科课程中讲授本书的前半部分内容；而本书的后半部分(强调代码优化)会在研究生层面或另一门小范围的课程中讲授。下面是各章的概要介绍：

- 第1章给出一些关于学习动机的资料，同时也将给出一些关于计算机体系结构和程序设计语言原则的背景知识。
- 第2章会开发一个小型的编译器，并介绍很多重要概念。这些概念将在后面的各章中深入介绍。这个编译器本身将在附录中给出。
- 第3章将讨论词法分析、正则表达式、有穷状态自动机和词法分析器的生成器工具。这些内容是各种文本处理的基础。
- 第4章将讨论主流的语法分析方法，包括自顶向下方法(递归下降法、LL技术)和自底向上方法(LR技术和它的变体)。
- 第5章将介绍语法制导定义和语法制导翻译的基本思想。
- 第6章将使用第5章中的理论，并说明如何使用这些理论为一个典型的程序设计语言生成中间代码。
- 第7章将讨论运行时刻环境，特别是运行时刻栈的管理和垃圾回收机制。
- 第8章将主要讨论目标代码生成技术。该章会讨论基本块的构造，从表达式和基本块生成代码的方法，以及寄存器分配技术。
- 第9章将介绍代码优化技术，包括流图、数据流分析框架以及求解这些框架的迭代算法。

⊖ 美国大学的学制大致可以分为两种：quarter(学季)和semester(学期)。前者是把一年分为4个quarter，每个quarter 3个月，原则上是上3个quarter修一个quarter的假；而后者则类似我国国内的寒暑假制的大学学制，大概4个月一个semester。——编辑注

- 第 10 章将讨论指令级优化。该章的重点是从小段指令代码中抽取并行性，并在那些可以同时做多件事情的单处理器上调度这些指令。
- 第 11 章将介绍大规模并行性的检测和利用。这里的重点是数值计算代码。这些代码具有对多维数组进行遍历的紧致循环。
- 第 12 章将介绍过程间分析技术。它将讨论指针分析、别名和数据流分析。这些分析都考虑了到达代码中某个给定点时的过程调用序列。

哥伦比亚大学、哈佛大学、斯坦福大学已经开设了讲授本书内容的课程。哥伦比亚大学定期开设一门关于程序设计语言和翻译器的课程，使用了本书前 8 章的内容。该课程常年面向高年级本科生/一年级研究生讲授，这门课程的亮点是一个长达一个学期的课程实践项目。在该项目中，学生分成小组，创建并实现一个他们自己设计的小型语言。学生创建的语言涉及多个应用领域，包括量子计算、音乐合成、计算机图形学、游戏、矩阵运算和很多其他领域。在构建他们自己的编译器时，学生们使用了很多种可以生成编译器组件的工具，比如 ANTLR、Lex 和 Yacc；他们还使用了第 2 章和第 5 章中讨论的语法制导翻译技术。后续的研究生课程的重点是本书第 9 章到第 12 章的内容，着重强调适用于当代计算机(包括网络处理器和多处理器体系结构)的代码生成和优化技术。

斯坦福大学开设了一门历时一个学季的入门课程，大致涵盖了本书第 1 章到第 8 章的内容，同时还会简介本书第 9 章中全局代码优化的相关内容。第二门编译器课程包括本书第 9 章到第 12 章的内容，另外还包括第 7 章中更为深入的有关垃圾收集的内容。学生使用一个该校开发的、基于 Java 的系统 Joeq 来实现数据流分析算法。

预备知识

学习本书的读者应该拥有一些“计算机科学的综合知识”，至少学过两门程序设计课程，以及数据结构和离散数学的课程。具备多种程序设计语言的知识对学习本书会有所帮助。

练习

本书包含内容广泛的练习，几乎每一节都有一些练习。我们用感叹号来表示较难的练习或练习中的一部分。难度最大的练习有两个感叹号。

万维网上的支持

在本书的主页(http://dragonbook.stanford.edu)[⊖]上可以找到本书已知错误的勘误表以及一些支持性资料。我们希望将我们讲授的每一门与编译器相关的课程的可用讲义(包括家庭作业、答案和练习等)都提供出来。我们也计划公布由一些重要编译器的作者撰写的关于这些编译器的描述。

致谢

本书封面由 Strange Tonic Productions 的 S. D. Ullman 设计。

Jon Bentley 针对本书的初稿中的多章内容与我们进行了广泛深入的讨论。我们收到了来自下列人员的有帮助的评价和勘误：Domenico Bianculli、Peter Bosch、Marcio Buss、Marc Eaddy、

⊖ 读者可以从 http://infolab.stanford.edu/~ullman/dragon/errata.html 找到本书的勘误表，并可以从 http://infolab.stanford.edu/~ullman/dragon.html 处找到本书的一些支持资料。

Stephen Edwards、Vibhav Garg、Kim Hazelwood、Gaurav Kc、Wei Li、Mike Smith、Art Stamness、Krysta Svore、Olivier Tardieu 和 Jia Zeng。我们衷心感谢这些人的帮助。当然，书中还可能有错漏之处，希望得到指正和反馈。

另外，Monica 希望能够向她在 SUIF 编译器团队的同事表示感谢，感谢他们在 18 年的时间里给予她的支持和帮助，他们是：Gerald Aigner、Dzintars Avots、Saman Amarasinghe、Jennifer Anderson、Michael Carbin、Gerald Cheong、Amer Diwan、Robert French、Anwar Ghuloum、Mary Hall、John Hennessy、David Heine、Shih-Wei Liao、Amy Lim、Benjamin Livshits、Michael Martin、Dror Maydan、Todd Mowry、Brian Murphy、Jeffrey Oplinger、Karen Pieper、Martin Rinard、Olatunji Ruwase、Constantine Sapuntzakis、Patrick Sathyanathan、Michael Smith、Steven Tjiang、Chau-Wen Tseng、Christopher Unkel、John Whaley、Robert Wilson、Christopher Wilson 和 Michael Wolf。

A. V. A. , Chatham NJ
M. S. L. , Menlo Park CA
R. S. , Far Hills NJ
J. D. U. , Stanford CA
2006 年 6 月

目 录