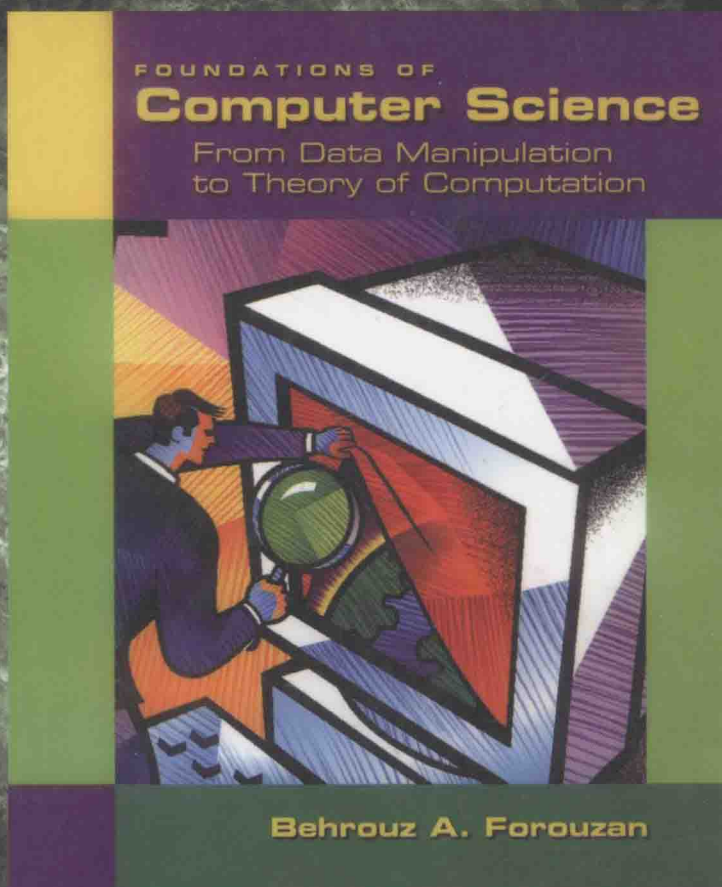




计 算 机 科 学 丛 书

计算机科学导论

(美) Behrouz A. Forouzan 著 刘艺 段立 钟维亚 等译



Foundations of Computer Science
From Data Manipulation to Theory of Computation

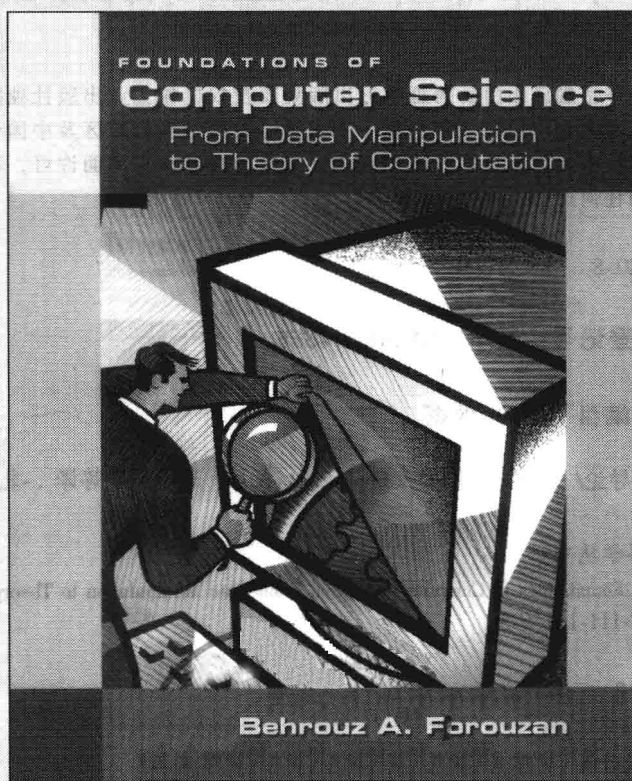


机械工业出版社
China Machine Press

计 算) 丛 书

计算机科学导论

(美) Behrouz A. Forouzan 著 刘艺 段立 钟维亚 等译



Foundations of Computer Science
From Data Manipulation to Theory of Computation



机械工业出版社
China Machine Press

本书是大学计算机相关专业的基础课教材，涉及到计算机科学的各个方面。本书着重讲解基本概念而不是数学模型和技术细节，通过大量的图表和演示范例讲解计算机科学的基础知识；每章后面的关键术语、小结和练习有助于读者掌握和复习知识要点。

本书既适合当作大专院校的计算机基础课教材，也可作为一般的计算机基础入门读物。

Behrouz A. Forouzan: Foundations of Computer Science: From Data Manipulation to Theory of Computation (ISBN: 0 534-37968-0)

Copyright © 2003 by Brooks/Cole, a division of Thomson Learning.

Original language published by Thomson Learning (a division of Thomson Learning Asia Pte Ltd). All rights reserved. 本书原由汤姆森学习集团出版。版权所有，盗印必究。

China Machine Press is authorized by Thomson Learning to publish and distribute exclusively this simplified Chinese edition. This edition is authorized for sale in the People's Republic of China only (excluding Hong Kong, Macao SAR and Taiwan). Unauthorized export of this edition is a violation of the Copyright Act. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

本书中文简体字翻译版由汤姆森学习出版集团授权机械工业出版社独家出版发行。此版本仅限在中华人民共和国境内（不包括中国香港、澳门特别行政区及中国台湾）销售。未经授权的本书出口将被视为违反版权法的行为。未经出版者预先书面许可，不得以任何方式复制或发行本书的任何部分。

981-254-191-8

本书版权登记号：图字：01-2003-4655

图书在版编目（CIP）数据

计算机科学导论/（美）佛罗赞（Forouzan, B.A.）著；刘艺等译.-北京：机械工业出版社，2004.1

（计算机科学丛书）

书名原文：Foundations of Computer Science: From Data Manipulation to Theory of Computation
ISBN 978-7-111-13159-5

I. 计… II. ①佛…②刘… III. 计算机科学 IV. TP3

中国版本图书馆 CIP 数据核字（2003）第 092786 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：吴怡

北京瑞德印刷有限公司印刷·新华书店北京发行所发行

2008 年 6 月第 1 版第 16 次印刷

184mm×260mm·22.25 印张

定价：30.00 元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换
本社购书热线电话：（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中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、研究及皮藏。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
石教英
张立昂
邵维忠
周克定
郑国梁
高传善
裘宗燕

王 珊
吕 建
李伟琴
陆丽娜
周傲英
施伯乐
梅 宏
戴 葵

冯博琴
孙玉芳
李师贤
陆鑫达
孟小峰
钟玉琢
程 旭

史忠植
吴世忠
李建中
陈向群
岳丽华
唐世渭
程时端

史美林
吴时霖
杨冬青
周伯生
范 明
袁崇义
谢希仁

译者序

《计算机科学导论》是一本国外 IT 相关专业大学生的基础课教材。通常供大学低年级学生使用。该教材是一本百科全书式的计算机专业基础入门读物，涉及到计算机科学的方方面面。虽然读者对象是计算机专业的学生，但这本书却深入浅出，引人入胜，绝无一般教科书的枯燥和晦涩。作者的用意显然是为了轻松勾画出计算机科学体系的框架，为有志于 IT 行业的学生奠定计算机科学知识的基础，铺设进一步深入专业理论学习的桥梁。所以作者在这本书中强调的是概念而不是数学模型和技术细节。并通过大量的图片、图表和演示增强读者对内容的理解和知识的掌握；通过范例讲解概念和相关的数学模型；通过关键术语、小结和练习帮助读者学习和复习。

这本书既适合国内的大专院校用作计算机基础课教材，也可以供有意在计算机方面发展的非计算机专业读者作入门参考。该书作为一本成熟的教材已经被国外许多大学选用，引进翻译显然对国内的计算机教学有着重要的参考价值。

由于本书涉及面广，技术内容新，有一定的翻译难度。为此我们不敢疏忽，字斟句酌，力求翻译准确。在翻译过程中我们还对于原书的一些疏漏进行了更正。

由于我们的水平有限，如果读者在阅读中发现我们的工作有不足之处，敬请指正。

另外，除封面署名人员外，参加翻译和校对工作的还有冯坤、刘文颖、尹迪、李启元、宋胜峰、李亚楠、金亮、张显峰、米晨、郭琛、胡星、陈春、卢江涛、王继涛等人，在此表示感谢。

译者

2003年8月8日

E-mail: my_reader@sina.com

前 言

欢迎学习计算机科学！你将开始探索一个美妙和激动人心的世界，这里提供了许多充满挑战和令人兴奋的职业机会。计算机在我们的日常生活中扮演了一个重要的角色而且在未来也将一样。

计算机科学是一个充满了挑战和发展机遇的年轻学科。计算机网络将地球上每一个角落的人们连接在一起。虚拟现实创造了眩目的三维图像。宇宙空间探险的成功也部分归功于计算机的发展。计算机建立的特技效果改变了电影工业。计算机在遗传学研究中也扮演了一个重要的角色。

本书的结构

这本书是为 CS0 课程设计的。它包含了计算机科学的所有领域。全书分为五个部分：计算机和数据、计算机硬件、计算机软件、数据组织和高级主题。

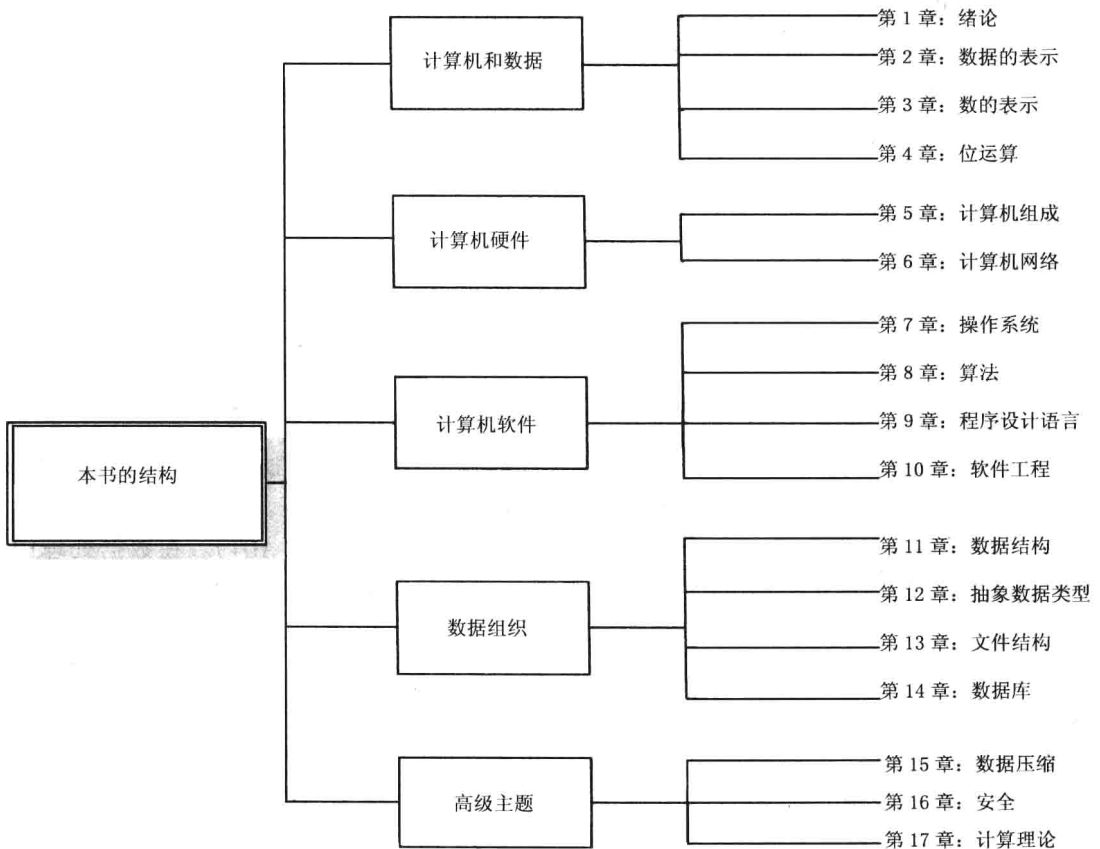


图 P.1 本书结构

第一部分：计算机和数据

在第一部分，我们讨论计算机和它所处理的数据。这部分包含4章。

第1章：绪论 在这一章，我们把计算机当作一个处理数据的实体，我们将介绍冯·诺伊曼概念和一台计算机的大体组成部件。我们将计算机组成的细节讨论放到第5章。

第2章：数据的表示 在这一章，我们讨论文本、图像、音频、视频作为位模式是如何表现的。有关数的表示方面的内容我们将在第3章中介绍。

第3章：数的表示 本章讨论数的表示。我们讲解整型数和浮点型数是如何在计算机里存储的。

第4章：位运算 在这一章，我们讨论位模式的有关运算，包括算术和逻辑的运算。

第二部分：计算机硬件

在第二部分，我们主要介绍计算机硬件。这部分分2章。

第5章：计算机组成 在这一章，我们将计算机作为一个独立的整体，描述其硬件的各个组成部件以及它们是如何工作的。

第6章：计算机网络 在本章中，我们讨论计算机是如何相互连接成局域网和互联网的。

第三部分：计算机软件

在第三部分中，我们主要关注计算机软件各个方面。

第7章：操作系统 本章讨论计算机软件里的核心部分——操作系统，具体介绍操作系统的任务及其如何演化，同时讨论部分现代的操作系统的。

第8章：算法 算法在计算机科学领域扮演了重要角色，甚至有些人认为计算机科学也就是关于算法的学问。我们将给出算法的定义，同时利用一些工具来描述算法。关于这些工具的详细介绍见附录C、D和E。

第9章：程序设计语言 本章首先从总体上介绍计算机语言，然后以当前比较流行的C语言为例具体讨论。

第10章：软件工程 在本章中，我们讨论计算机课程体系中很重要的一门学科：软件工程。

第四部分：数据组织

在第四部分里，我们重新回到了数据上，但我们主要是从使用者的角度来介绍。

第11章：数据结构 在本章中，我们讨论数据结构。数据在更抽象的层次上被组织成结构。我们将介绍现在常用的一些数据结构，如数组、记录、链表等。

第12章：抽象数据类型 在本章中，我们讨论抽象数据类型（ADT）。在数据处理中，需要把数据定义成数据包，其中包括了为该数据包定义的有关运算。我们将描述列表、栈、队列、树和图，以便学生能从抽象角度去理解数据。

第13章：文件结构 本章讨论文件结构，我们揭示了文件是如何在逻辑上组织起来的，我们讨论了顺序访问和随机访问文件。学生在学习程序设计课程之前需要懂得这些概念。

第14章：数据库 一个组织的数据很少会是单独地或孤立地存储着，通常它们是存储在一个叫数据库的实体中。我们将介绍关系数据库，同时涉及一些数据库语言（SQL）的知识，使用这种语言可以在数据库中查询信息。

第五部分：高级主题

在这部分里，我们讨论3个在计算机科学中愈来愈重要的高级主题：数据压缩、安全和

计算理论。如果课时不够，或者学生缺乏相关背景知识，那么这些主题可以跳过。

第 15 章：数据压缩 在本章中，我们介绍两种数据压缩：无损压缩和有损压缩。我们讨论游程长度编码、霍夫曼编码，并以 Lempel Ziv 算法为例介绍无损压缩。同时还以 JPEG 和 MPEG 为例介绍有损压缩。

第 16 章：安全 在本章中，我们讨论数据安全方面的 4 个内容：保密、验证、完整和不可否认性。我们将演示如何使用加密/解密和数字签名来创建安全系统。

第 17 章：计算理论 在本章中，我们简明地浏览一下计算理论。我们阐明在解决问题方面没有哪种语言比其他语言更优越。同时也说明有些问题无论用何种计算机语言来编写程序都是无法解决的。

鸟瞰

读者必须记住本书没有深入讨论计算机科学的各个主题，要那样的话需要更多的篇幅，本书试图覆盖计算机科学的相关主题。我们的经验表明，学生掌握了数据的表示和处理，就能够更好地使用低级和高级计算机语言进行程序设计，而掌握有关计算机科学的一般知识，学生可以更容易学好网络和互联网课程。本书只是计算机科学的一个鸟瞰式的纵览。

本书的特点

下面的一些特色不仅使得本书独树一帜，同时也特别适合刚入门的学生学习。

概念

贯穿本书，我们始终强调概念要比数学模型更重要，我们认为对概念的理解必然左右对模型的理解。

图文并茂

本书图文并茂，全书有将近 300 幅图片。虽然这样一来增加了书的篇幅，但图片增进了对文字的理解。

例子

我们尽可能地运用例子来表述概念和数学模型。

每章结尾的材料

每一章结尾的材料包括三个部分：关键术语、小结、练习。

关键术语 列出本章介绍过的重要术语，每一个关键术语在术语表中都有定义。

小结 包括对本章中所有关键内容的简明概括。

练习 每个练习包括三部分内容：复习题、多项选择题和课后练习题。

- 复习题：测试本章中所有的要点和概念。
- 多项选择题：测试对材料的理解。
- 课后练习题：测试学生能否运用概念和公式。

附录

本书有 7 篇附录，是各章中需要速查的表格和材料，这些附录是：

- ASCII 码
- Unicode
- 流程图

- 伪代码
- 结构图
- 离散余弦变换
- 缩略语

术语表

本书最后的术语表收录了所有关键术语。

练习题的答案

所有编号为奇数的复习题、多项选择题和课后练习题的答案都可以在以下网址上找到：
www.brookscole.com/compsci。

教辅材料

复习题、多项选择题和课后练习题答案中的所有图和要点的幻灯片可以在以下网址找到：
www.brookscole.com/compsci。

致谢

此类教材的出版离不开众人的支持，尤其是这本书。

我们要感谢 De Anza 学院教职员工一直以来的大力支持和他们的建议。我们特别要感谢 Scott DeMouthe，他帮我们阅读手稿并解决问题。

对于那些没有参加编写工作的人，我们同样感谢他们专心评阅本书。写书很容易“一叶蔽目，不见泰山”。那些站在幕后评读全稿的审稿人的重要指导是无法估价的。我们特别要感谢那些审稿人的贡献，他们是：北卡罗来纳州大学夏洛特分校的 Essam El-Kwae、Fairleigh Dickinson 大学的 Norman J. Landis、加州大学洛杉矶分校的 John A. Rohr、波士顿学院的 Robert Signorile、新泽西理工大学的 Robert Statica。

我们同样感谢我们的出版人 Bill Stenquist，组稿编辑 Kallie Swanson，编辑助理 Carla Vera，产品编辑 Kelsey McGee。我们同样感谢 Matrix 公司的 Merrill Peterson，文字编辑 Frank Hubert，校对 Amy Dorr。

最后，最应该感谢的是我们的家人和朋友。曾经有位作家把写作描写为“把自己锁在房间里”。在我们忍受写作之苦的过程中，家人和朋友同样忍受着我们的缺席。我们只能希望当他们看到最后的作品时，他们感到自己的牺牲是值得的。

目 录

出版者的话
专家指导委员会
译者序
前言

第一部分 计算机和数据

第 1 章 绪论	1	2.2.2 位模式	12
1.1 计算机作为黑盒	1	2.2.3 字节	12
1.1.1 数据处理器	1	2.3 表示数据	12
1.1.2 可编程数据处理器	1	2.3.1 文本	12
1.2 冯·诺伊曼模型	3	2.3.2 数	14
1.2.1 4 个子系统	3	2.3.3 图像	14
1.2.2 存储的程序概念	3	2.3.4 音频	16
1.2.3 指令的顺序执行	4	2.3.5 视频	16
1.3 计算机硬件	4	2.4 十六进制表示法	16
1.4 数据	4	2.5 八进制表示法	18
1.4.1 存储数据	4	2.6 关键术语	19
1.4.2 组织数据	4	2.7 小结	19
1.5 计算机软件	5	2.8 练习	20
1.5.1 必须存储程序	5	第 3 章 数的表示	22
1.5.2 指令的序列	5	3.1 十进制和二进制	22
1.5.3 算法	5	3.1.1 十进制系统	22
1.5.4 语言	5	3.1.2 二进制系统	22
1.5.5 软件工程	6	3.2 转换	23
1.5.6 操作系统	6	3.2.1 二进制数向十进制数转换	23
1.6 历史	6	3.2.2 十进制数向二进制数转换	23
1.6.1 机械计算机器 (1930 年以前)	6	3.3 整数的表示法	24
1.6.2 电子计算机的诞生 (1930—1950)	7	3.3.1 无符号整数格式	25
1.6.3 计算机的诞生 (1950—现在)	7	3.3.2 符号加绝对值格式	26
1.7 关键术语	8	3.3.3 二进制反码格式	28
1.8 小结	8	3.3.4 二进制补码格式	30
1.9 练习	9	3.3.5 整数表示法小结	32
第 2 章 数据的表示	11	3.4 EXCESS 系统	32
2.1 数据的类型	11	3.5 浮点表示法	33
2.2 计算机内部的数据	11	3.5.1 转换成二进制	33
2.2.1 位	12	3.5.2 规范化	34
		3.5.3 符号、幂和尾数	34
		3.5.4 IEEE 标准	35
		3.6 十六进制表示法	36
		3.7 关键术语	36
		3.8 小结	37
		3.9 练习	37
		第 4 章 位运算	42

7.7 练习	113	9.3 程序的执行	141
第 8 章 算法	117	9.4 语言的分类	142
8.1 概念	117	9.4.1 过程化(强制性)语言	142
8.1.1 非正式定义	117	9.4.2 面向对象的语言	143
8.1.2 示例	117	9.4.3 函数型语言	145
8.1.3 定义动作	118	9.4.4 说明性(逻辑)语言	146
8.1.4 精化	118	9.4.5 专用语言	147
8.1.5 泛化	120	9.5 过程化语言: C	148
8.2 三种结构	120	9.5.1 标识符	148
8.2.1 顺序	121	9.5.2 数据类型	149
8.2.2 判断	121	9.5.3 变量	149
8.2.3 循环	121	9.5.4 常量	150
8.3 算法的表示	121	9.5.5 输入和输出	151
8.3.1 流程图	121	9.5.6 表达式	151
8.3.2 伪代码	121	9.5.7 语句	152
8.4 更正式的定义	124	9.5.8 函数	153
8.4.1 有序集合	124	9.5.9 选择	155
8.4.2 明确步骤	124	9.5.10 循环	155
8.4.3 产生结果	124	9.5.11 派生数据类型	157
8.4.4 在有限的时间内终止	125	9.5.12 递归	157
8.5 子算法	125	9.6 关键术语	157
8.6 基本算法	126	9.7 小结	158
8.6.1 求和	126	9.8 练习	159
8.6.2 乘积	126	第 10 章 软件工程	163
8.6.3 最大和最小	127	10.1 软件生命周期	163
8.6.4 排序	127	10.1.1 分析阶段	163
8.6.5 查找	130	10.1.2 设计阶段	164
8.7 递归	132	10.1.3 实现阶段	164
8.7.1 迭代的定义	132	10.1.4 测试阶段	165
8.7.2 递归的定义	132	10.2 开发过程模型	165
8.8 关键术语	134	10.2.1 瀑布模型	165
8.9 小结	134	10.2.2 增量模型	166
8.10 练习	135	10.3 模块化	166
第 9 章 程序设计语言	138	10.3.1 工具	166
9.1 演化	138	10.3.2 耦合	167
9.1.1 机器语言	138	10.3.3 内聚	168
9.1.2 符号语言	139	10.4 质量	169
9.1.3 高级语言	139	10.4.1 质量定义	169
9.1.4 自然语言	140	10.4.2 质量因素	169
9.2 构建程序	140	10.4.3 质量周期	171
9.2.1 编写和编辑程序	140	10.5 文档	172
9.2.2 编译程序	141	10.5.1 用户文档	172
9.2.3 链接程序	141	10.5.2 系统文档	172

10.5.3 进行过程中的文档化	173	12.7.1 术语	201
10.6 关键术语	173	12.7.2 图的操作	202
10.7 小结	174	12.7.3 图的实现	204
10.8 练习	174	12.7.4 图的应用	204
第四部分 数据组织			
第 11 章 数据结构	177	12.8 关键术语	205
11.1 数组	177	12.9 小结	206
11.1.1 数组的应用	179	12.10 练习	208
11.1.2 二维数组	180	第 13 章 文件结构	213
11.2 记录	181	13.1 存取方法	213
11.3 链表	182	13.1.1 顺序存取	213
11.3.1 节点	182	13.1.2 随机存取	213
11.3.2 链表的指针	182	13.2 顺序文件	214
11.3.3 链表的操作	183	13.3 索引文件	216
11.4 关键术语	184	13.4 哈希文件	217
11.5 小结	184	13.4.1 哈希方法	217
11.6 练习	185	13.4.2 冲突	218
第 12 章 抽象数据类型	188	13.5 文本与二进制	220
12.1 背景	188	13.5.1 文本文件	220
12.1.1 定义	188	13.5.2 二进制文件	220
12.1.2 抽象数据类型模型	189	13.6 关键术语	220
12.1.3 抽象数据类型的操作	189	13.7 小结	221
12.2 线性列表	189	13.8 练习	221
12.2.1 线性列表的操作	190	第 14 章 数据库	225
12.2.2 广义线性列表的实现	192	14.1 数据库管理系统	225
12.2.3 线性列表的应用	192	14.2 体系结构	226
12.3 栈	192	14.2.1 内层	226
12.3.1 栈的操作	192	14.2.2 概念层	226
12.3.2 栈的实现	193	14.2.3 外层	226
12.3.3 栈的应用	194	14.3 数据库模型	226
12.4 队列	194	14.3.1 层次模型	226
12.4.1 队列的操作	195	14.3.2 网络模型	227
12.4.2 队列的实现	195	14.3.3 关系模型	227
12.4.3 队列的应用	195	14.4 关系模型	227
12.5 树	196	14.5 关系的操作	228
12.5.1 树的基本概念	196	14.5.1 插入	228
12.5.2 树的操作	197	14.5.2 删除	229
12.6 二叉树	198	14.5.3 更新	229
12.6.1 二叉树的操作	199	14.5.4 选择	229
12.6.2 二叉树的实现	200	14.5.5 投影	230
12.6.3 二叉树的应用	201	14.5.6 连接	230
12.7 图	201	14.5.7 并	231
		14.5.8 交	231
		14.5.9 差	231

14.6 结构化查询语言	232	17.1 简单语言	263
14.7 其他数据库模型	234	17.1.1 递增语句	263
14.7.1 分布式数据库	235	17.1.2 递减语句	263
14.7.2 面向对象数据库	235	17.1.3 循环语句	263
14.8 关键术语	235	17.1.4 简单语言的威力	264
14.9 小结	236	17.1.5 结论	266
14.10 练习	236	17.2 图灵机	266
第五部分 高级主题			
第 15 章 数据压缩	241	17.2.1 图灵机组成部件	266
15.1 无损压缩	241	17.2.2 对简单语言的模拟	268
15.1.1 游程长度编码	241	17.2.3 结论	270
15.1.2 霍夫曼编码	242	17.3 歌德尔数	270
15.1.3 Lempel Ziv 编码	244	17.3.1 描述程序	271
15.2 有损压缩方法	247	17.3.2 翻译数	271
15.2.1 图像压缩: JPEG	247	17.4 停机问题	272
15.2.2 视频压缩: MPEG	250	17.5 可解的和不可解的问题	273
15.3 关键术语	251	17.5.1 不可解问题	273
15.4 小结	251	17.5.2 可解问题	274
15.5 练习	252	17.6 关键术语	275
第 16 章 安全	254	17.7 小结	275
16.1 保密	254	17.8 练习	275
16.1.1 加密/解密	254	附 录	
16.1.2 使用组合的保密	257	附录 A ASCII 码	279
16.2 数字签名	258	附录 B Unicode	283
16.2.1 签署整个文档	258	附录 C 流程图	286
16.2.2 签署摘要	258	附录 D 伪代码	290
16.3 关键术语	260	附录 E 结构图	293
16.4 小结	260	附录 F 离散余弦变换	297
16.5 练习	260	附录 G 缩略语和术语表	298
第 17 章 计算理论	263	索引	323

第一部分 计算机和数据

第1章 绪 论

今天**计算机科学**一词是一个非常广泛的概念。尽管如此，在本书里，我们将其定义为“和计算机相关的问题”。本章首先阐述什么是计算机，接着探索和计算机直接相关的一些问题。首先我们将计算机看成一个**黑盒**去猜测它的行为，然后通过深入这个箱子来揭示每个计算机的共同之处。这就形成了被人们广泛认同的计算机基础的模型——**冯·诺伊曼模型**。然后，我们简要地讨论冯·诺伊曼模型的分支和如何接受它。到这里，我们涉及到的章节和章节里的内容都属于这些问题。最后我们介绍计算机（这一改变文明的装置）的简明历史。

2

1.1 计算机作为黑盒

如果不关心计算机的内部结构，可以简单地认为计算机是一个黑盒。但是，仍然需要定义计算机所完成的工作来区别它和其他黑盒之间的差异。这里我们提供两个常见的计算机模型。

1.1.1 数据处理器

可以认为计算机是一个**数据处理器**。依照这种定义，可以认为计算机是一个接受输入数据，处理数据，产生输出数据的黑盒（如图 1-1 所示）。尽管这个模型能够体现现代计算机的功能，但是它的定义还是太广泛。按照这种定义，可以认为便携式计算器也是计算机（按照字面意思，它也符合定义模型）。

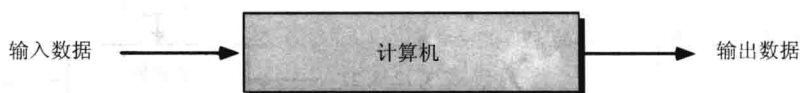


图 1-1 数据处理器模型

另一个问题是这个模型并没有说明它所处理的类型以及是否可以处理一种以上的类型。换句话说，它并没有清楚地说明基于这个模型的机器能够完成的操作类型和数量。它是专用机器还是通用机器呢？

这种模型可以表示为一种用来完成特定任务的专用计算机（或者处理器），比如用来控制建筑物温度或汽车油料使用。尽管如此，如今所说的计算机是一种通用的机器，它可以完成各种不同的工作。这表明我们需要改变对计算机定义模型来反映当今的计算机。

1.1.2 可编程数据处理器

一个相对较好的具有通用性的计算机模型如图 1-2 所示。图中给计算机添加了一个额外