



高等学校计算机精品课程系列教材

数据结构与算法

王昆仑 李红 主编



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

高等学校计算机精品课程系列教材

数据结构与算法

王昆仑 李 红 主编

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

内 容 简 介

本书是以高等工科院校本科计算机类专业、信息技术类专业和相关专业“应用型”人才培养为目标编写的一本实用性教材。以学习软件设计开发中涉及到的各种数据结构及常用算法和解决基本应用问题的实际应用需求为基本点,深入介绍了各种数据结构的定义(逻辑结构、存储结构和基本算法)和基本应用等方面的知识。本书以“数据结构”的逻辑结构作为引线,突出以实例和应用为特色,把数据结构与算法问题同应用问题结合起来,通过介绍大量的应用问题,缩短了理论知识与应用问题之间的距离。算法学习由基本算法和应用问题算法两部分组成,这里注重介绍算法的设计过程和算法分析。为便于教学,每章都有教学目标和教学提示,每章最后都配备有满足教学要求的各类习题。本书所有算法均在 Microsoft Visual C++ 6.0 环境中通过测试,并在附录 A 中给出了源程序以供读者选用。

本书可作为理工科高等院校计算机工程类、软件工程类和信息技术类等相关专业的教材,也可供从事相关工作的科技工作者参考。

图书在版编目(CIP)数据

数据结构与算法/王昆仑主编. —北京:中国铁道出版社, 2007. 6

(高等学校计算机精品课程系列教材)

ISBN 978-7-113-07628-3

I. 数… II. 王… III. ①数据结构—高等学校—教材
②算法分析—高等学校—教材 IV. TP311.12

中国版本图书馆 CIP 数据核字 (2007) 第 081164 号

书 名: 数据结构与算法

作 者: 王昆仑 李 红 许 强

出版发行: 中国铁道出版社 (100054, 北京市宣武区右安门西街 8 号)

策划编辑: 严晓舟 秦绪好

责任编辑: 杨 勇 包 宁

特邀编辑: 薛秋沛

封面设计: 付 巍

封面制作: 白 雪

责任校对: 贾 星

印 刷: 三河市宏达印刷有限公司

开 本: 787×1092 1/16 印张: 27 字数: 628 千

版 本: 2007 年 6 月第 1 版 2007 年 6 月第 1 次印刷

书 号: ISBN 978-7-113-07628-3/TP·2201

定 价: 35.00 元

版权所有 侵权必究

本书封面贴有中国铁道出版社激光防伪标签, 无标签者不得销售

凡购买铁道版的图书, 如有缺页、倒页、脱页者, 请与本社计算机图书批销部调换。

如何合理地组织数据、高效率地处理数据是扩大计算机应用领域、提高软件效率的关键。在软件开发过程中要求“高效地”组织数据和设计出“好的”算法，并使算法用程序来实现，通过调试而成为软件，必须具备数据结构领域和算法设计领域的专门知识。

“数据结构与算法”课程主要学习在软件开发中涉及到的各种常用数据结构及其常用算法，在此基础上，学习如何利用数据结构和算法解决一些基本的应用问题。通过学习，使读者基本掌握相关领域的基础知识和基本应用。

本教材为达到高等工科院校“应用型”人才的培养目标，在吸收了国、内外教材的知识体系结构的基础上，参考了众多的应用资料并根据主编多年在高校讲授《数据结构》课程的体会而编写。教材有以下几个特色。

(1) 学习一种数据结构必须掌握该数据结构的定义（逻辑结构、存储结构和基本算法）和基本应用两个方面的知识。所以，本教材以“结构”为特色，每一个数据结构的學習都围绕着该数据结构的定义，通过数据结构的逻辑结构、存储结构、基本算法和相关应用问题来介绍其基本知识和应用知识。数据的存储结构是学习和掌握算法的基础，本书在“算法的数据类型”部分中加以突出介绍。学生掌握的是数据结构的完整知识并能学为所用。

(2) 本书的知识结构以“数据结构”的逻辑结构作为引线。第1章作为本书的学习基础和预备知识。其后内容分为4个部分，从数据的逻辑结构来看，第一部分学习的是逻辑结构为“线性”的数据结构，包括第2章~第6章；第二部分学习的是逻辑结构为“树形”的数据结构，包括第7章~第8章；第三部分学习的是逻辑结构为“集合型”的数据元素在散列存储方法下的数据结构，包括第9章；第四部分学习的是逻辑结构为“图形”的数据结构，包括第10章。第11章重点讨论了算法和程序性能分析以及算法设计方法的基本问题。这样划分教材章节的原因是：使采用本书作为教材的读者能够掌握什么是数据结构，如何设计算法以及能解决什么问题。以突出“数据结构”主题。

(3) 教材中注重工科“应用型”人才培养的需求和学习方法。吸收理工科教材的特色，在介绍新的知识点时，没有大段的文字描述，而是尽可能地采用具体的例题来加强其学习效果。

(4) 本教材中介绍了大量的应用问题。把数据结构与算法问题同应用问题结合起来，缩短了理论知识与应用问题之间的距离，在附录A中，给出了第2章~第10章中重要的基本算法和应用问题的源程序，适合工科院校相关专业的学生参考使用。对课程设计环节有很好的辅助作用、指导作用。

(5) 教材中注意了算法设计能力的培养。学习和培养算法设计能力是本课程的主要教学目的之一，本教材中注重介绍算法的设计过程和算法分析；在上机实验环节中，将引领学生编写十余个从简单到有一定难度的算法；在习题中，安排了一定量的基础题和适量的算法设计题，供教师和学生在学习中参考使用。

本书共分 11 章。第 1 章“数据结构和算法”作为全书的导引，主要包括有关数据、数据类型、数据结构、算法、算法实现（算法描述工具——C 语言）、C 语言使用中的相关问题和算法分析等基本概念和相关知识。其中重点是数据、数据类型、数据结构和算法等概念；对于本教材使用的算法描述工具——C 语言，则介绍了指针、结构变量、函数、递归、动态存储分配、文件操作、程序测试与调试问题等内容，以方便本课程与 C 语言课程的衔接，便于教学。

第 2 章～第 6 章是逻辑结构为“线性”的数据结构及其应用知识内容。

第 2 章“顺序表及其应用”主要介绍的是线性逻辑结构的数据在顺序存储方法下的数据结构顺序表（包括顺序串）的概念、数据类型、数据结构、基本运算及相关应用问题。其中重点一是顺序表的定义、数据类型、数据结构、基本算法和性能分析等概念和相关知识；二是顺序表的应用，包括查找问题（简单顺序查找、二分查找、分块查找）、排序问题（直接插入排序、希尔排序、冒泡排序、快速排序、直接选择排序、归并排序）、字符处理问题（模式匹配）等内容。第 3 章“链表及其应用”主要介绍的是线性逻辑结构的数据在链接存储方法下的数据结构链表的相关知识。主要是单链表、循环链表的数据类型描述、数据结构、基本运算及其实现以及链表的相关应用问题，在此基础上介绍了链串的相关知识。在应用方面有多项式的相加问题、归并问题、箱子排序问题和链表在字符处理方面的应用问题等。第 4 章“堆栈及其应用”介绍在两种不同的存储结构下设计的堆栈，即顺序栈和链栈的相关知识，了解堆栈的相关应用，掌握应用堆栈来解决实际问题的思想及方法。第 5 章“队列及其应用”主要介绍顺序存储和链接存储方法下的两种队列、顺序（循环）队列和链队列的数据结构、基本运算及其性能分析以及应用。第 6 章“特殊矩阵、广义表及其应用”将学习数组、稀疏矩阵和广义表的基本概念，几种特殊矩阵的存储结构及基本运算，在此基础上学习特殊矩阵的计算算法与广义表应用等相关问题。本章的重点是相关数据结构的存储结构及基本运算算法。

第 7 章和第 8 章是逻辑结构为“树形”的数据结构及其应用知识内容。

第 7 章“二叉树及其应用”的知识结构主要是：非线性数据结构二叉树的定义、性质、逻辑结构、存储结构及其各种基本运算算法，包括二叉树的建立、遍历、线索化和表达式求值等算法。在此基础上，介绍二叉树的一些应用问题，包括哈夫曼编码问题、（平衡）二叉排序树问题和堆排序问题等。第 8 章“树和森林及其应用”介绍树和森林的数据结构、基本算法及其性能分析，树和森林与二叉树之间的转换算法等，在此基础上介绍树的应用——B 树，应用 B 树来实现数据元素的动态查找。

第 9 章“散列结构及其应用”是逻辑结构为“集合型”的数据元素在散列存储方法下的数据结构及其应用知识内容。主要介绍散列结构的概念、散列存储结构——散列表、散列函数和散列表中解决冲突的处理方法——开放定址法、链地址法以及散列表的基本算法及其性能分析，在散列结构的应用方面介绍散列结构的查找问题、LZW 压缩/解压缩问题和直接存取文件问题。

第 10 章“图及其应用”是逻辑结构为“图形”的数据结构及其应用知识内容，主要介绍图的定义和基础知识，图的 4 种存储结构，图的基本算法以及图的典型应用问题（最小生成树、最短路径、拓扑排序和关键路径等）。

第 11 章“算法性能分析和算法设计方法简介”主要对算法和程序性能分析中的目的、时间复杂性和空间复杂性、复杂性要素和分析方法、时间复杂性上（下）限值、算法性能测量等问题进行讨论，并结合货箱装船、0/1 背包和迷宫老鼠等问题介绍优化问题、分而治之、贪婪算法和回溯算法等基本的算法设计方法的基本知识，介绍 NP-复杂问题和 NP-完全问题。通过对本章的学习，使读者初步了解算法设计的常用方法，知道什么是“优质”算法和程序以及如何测量、评价算法的知识。

每章都有教学目标和教学提示，每章后面都配备有一定量的填空题、判断题、选择题、简答题和算法设计题，供读者选用。习题的参考答案或者提示可以在网站 <http://edu.tqbooks.net> 上下载。

本教材适合理工科高等院校本科计算机工程类专业、信息技术类专业和软件工程类相关专业使用。减少部分教学内容，也可以作为专科教学使用。同时也是相关专业的读者了解和學習数据结构与算法的一本很好的入门教材。

需要特别说明的是，本书中较细致地介绍了较多的应用问题及其算法，教学过程中由于受到教学课时的限制或者根据本校培养目标要求以及学生的实际情况，可以选讲本书中打“*”号的章节内容，也可安排有兴趣的学生选修。

本教材由王昆仑、李红主编，编写工作由王昆仑、李红和许强完成，其中王昆仑编写了第 1、5、6、9、11 章，李红编写了第 2、3、4、7 章，许强编写了第 8、10 章。教材中的全部算法由项响琴在 Microsoft Visual C++ 6.0 环境中进行了调试，董靖完成了部分章节的绘图工作，屠菁、黄小杰、林晓燕、彭晓舟等为本书的编写工作给予了很大的支持。另外，本书的编写工作得到了省、校两级精品课程建设项目基金的资助，还得到了计算机教育界同行的关心和帮助，在此一并致谢！

由于数据结构与算法的应用发展迅速，加之编者水平有限，书中疏漏和不妥之处恳请读者批评指正，电子邮箱：kunlunwang@126.com。

编 者

2007 年 2 月

目录

CONTENTS

第 1 章 数据结构和算法.....	1
1.1 数据和数据类型.....	1
1.1.1 数据和数据元素.....	1
1.1.2 数据类型.....	2
1.1.3 抽象数据类型.....	4
1.1.4 抽象数据类型程序应用实例	5
1.1.5 数据对象.....	6
1.2 数据结构	6
1.2.1 数据的逻辑结构.....	6
1.2.2 数据元素的存储结构.....	8
1.2.3 常用的数据运算.....	10
1.3* 算法描述工具——C 语言	12
1.3.1 指针类型与指针变量.....	13
1.3.2 结构类型与结构变量.....	15
1.3.3 函数与参数.....	17
1.3.4 递归定义和递归函数.....	18
1.3.5 动态存储分配.....	19
1.3.6 文件操作.....	21
1.3.7 程序测试与测试集.....	23
1.3.8 测试数据的设计.....	23
1.3.9 程序调试问题.....	25
1.4 算法和算法评价.....	25
1.4.1 算法的概念.....	25
1.4.2 算法的性质.....	27
1.4.3 算法的评价标准.....	28
1.5 算法性能分析.....	29
1.5.1 算法的时间性能分析	29
1.5.2 算法的空间性能分析	33
小结.....	33
习题.....	34
第 2 章 顺序表及其应用.....	38
2.1 顺序表的基本概念.....	38
2.1.1 顺序表的定义.....	38
2.1.2 顺序表的数据结构分析.....	38

2.1.3 顺序表的数据类型描述	39
2.2 顺序表基本算法	40
2.3 顺序表基本算法性能分析	43
2.3.1 时间性能分析	43
2.3.2 空间性能分析	44
2.4 顺序表的应用 1——查找问题	44
2.4.1 查找的概念	44
2.4.2 简单顺序查找算法	45
2.4.3 有序表的二分查找算法	47
2.4.4 分块查找算法	51
2.4.5 3 种查找算法的性能比较	52
2.5 顺序表的应用 2——排序问题	53
2.5.1 排序的概念	53
2.5.2 顺序表的数据类型	54
2.5.3 插入排序——直接插入排序算法	54
2.5.4 插入排序——希尔排序算法	57
2.5.5 交换排序——冒泡排序算法	59
2.5.6 交换排序——快速排序算法	61
2.5.7 选择排序——直接选择排序算法	65
2.5.8 归并排序算法	67
2.5.9 排序算法的性能分析与比较	71
2.6 顺序表的应用 3——字符处理问题	71
2.6.1 串和顺序串的定义及相关概念	71
2.6.2 顺序串的数据结构分析	72
2.6.3 顺序串的基本运算	72
2.6.4 顺序串的数据类型定义	74
2.6.5 顺序串的基本运算算法	74
2.6.6 串的模式匹配算法	77
小结	78
习题	78
第 3 章 链表及其应用	83
3.1 链表的基本概念	83
3.1.1 链表的定义	83
3.1.2 链表的逻辑结构	83
3.1.3 链表的存储结构	84
3.1.4 静态链表和动态链表	85
3.1.5 链表基本运算	86
3.2 单链表的数据结构	86

3.2.1	单链表的逻辑结构.....	86
3.2.2	单链表的存储结构.....	86
3.3	单链表基本算法.....	88
3.3.1	单链表的基本算法.....	88
3.3.2	单链表基本算法性能分析.....	93
3.4	循环链表.....	94
3.4.1	单循环链表.....	94
3.4.2	双向循环链表.....	95
3.4.3	双向循环链表的结点插入算法.....	96
3.4.4	双向循环链表的结点删除算法.....	97
3.5	链表的应用.....	98
3.5.1	多项式相加问题.....	98
3.5.2	两个链表的归并问题.....	101
3.5.3	箱子排序问题.....	103
3.5.4	链表在字符处理方面的应用——链串.....	105
小结		107
习题		107
第 4 章	堆栈及其应用.....	112
4.1	堆栈的基本概念.....	112
4.1.1	堆栈的定义.....	112
4.1.2	堆栈的逻辑结构.....	112
4.1.3	堆栈的基本算法.....	113
4.2	顺序栈及其基本算法.....	113
4.2.1	顺序栈的概念及数据类型.....	113
4.2.2	顺序栈的基本算法.....	114
4.2.3	顺序栈基本算法性能分析.....	115
4.3	链栈及其基本算法.....	116
4.3.1	链栈的概念及数据类型.....	116
4.3.2	链栈的基本算法.....	117
4.3.3	链栈基本算法性能分析.....	118
4.4	堆栈的应用.....	118
4.4.1	数制转换问题.....	118
4.4.2	简单的文字编辑器问题.....	119
4.4.3	表达式计算问题.....	120
4.4.4	括号匹配问题.....	126
4.4.5*	汉诺塔问题.....	128
4.4.6*	火车车厢重排问题.....	132
4.4.7*	开关盒布线问题.....	136

4.4.8*	离线等价类问题	139
4.4.9*	迷宫老鼠问题	142
小结		145
习题		146
第5章	队列及其应用	149
5.1	队列的基本概念	149
5.1.1	队列的定义	149
5.1.2	队列的逻辑结构	150
5.1.3	队列的基本算法	150
5.2	顺序队列及其基本算法	150
5.2.1	顺序队列的概念及数据类型	150
5.2.2	顺序循环队列	151
5.2.3	循环队列基本算法	153
5.2.4	循环队列基本算法性能分析	154
5.3	链队列及其基本算法	155
5.3.1	链队列的概念及数据类型	155
5.3.2	链队列基本算法	156
5.3.3	链队列基本算法性能分析	157
5.4	基数排序问题	157
5.4.1	多关键字排序问题	158
5.4.2	链式基数排序算法思想	158
5.4.3	链式基数排序算法实现的技术要点	160
5.4.4	链表基数排序算法及其性能分析	160
5.5*	火车车厢重排问题	162
5.5.1	问题分析及算法思想	162
5.5.2	火车车厢重排算法设计	162
5.5.3	火车车厢重排算法及其性能分析	163
5.6*	电路布线问题	165
5.6.1	电路布线问题分析	165
5.6.2	电路布线问题算法思想	165
5.6.3	电路布线问题算法设计	166
5.6.4	电路布线问题算法及其性能分析	166
5.7*	识别图元问题	168
5.7.1	识别图元问题分析	168
5.7.2	识别图元算法设计	168
5.7.3	识别图元算法及其分析	169
5.8*	工厂仿真问题	170
5.8.1	工厂仿真问题分析	170

5.8.2	工厂仿真问题算法设计	171
5.8.3	工厂仿真问题算法及其性能分析	173
小结		174
习题		174
第 6 章	特殊矩阵、广义表及其应用	178
6.1	数组与矩阵	178
6.1.1	数组与矩阵的概念及其相互关系	178
6.1.2	数组的存储结构	179
6.2	特殊矩阵的压缩存储	180
6.2.1	对称矩阵及其存储结构	181
6.2.2	三角矩阵及其存储结构	181
6.2.3	对角矩阵及其存储结构	182
6.2.4	稀疏矩阵及其存储结构和建表算法	183
6.3	矩阵的应用实例	187
6.3.1	稀疏矩阵的转置问题	187
6.3.2	稀疏矩阵的加法运算问题	189
6.4	广义表	193
6.4.1	广义表的概念	193
6.4.2	广义表的存储结构	194
6.4.3	广义表的应用—— m 元多项式的表示	196
小结		198
习题		198
第 7 章	二叉树及其应用	201
7.1	二叉树的基本概念	201
7.1.1	二叉树的定义	201
7.1.2	二叉树的基本术语	202
7.1.3	两种特殊的二叉树	202
7.1.4	二叉树的性质	203
7.2	二叉树的存储结构	205
7.2.1	二叉树的顺序存储结构	205
7.2.2	二叉树的链接存储结构	205
7.2.3	建立二叉树的算法	206
7.3	二叉树的遍历算法	208
7.3.1	二叉树遍历的概念	208
7.3.2	二叉树遍历递归算法	209
7.3.3	二叉树先序遍历非递归算法	211
7.3.4	二叉树中序遍历非递归算法	212
7.3.5	二叉树后序遍历非递归算法	213

7.4	线索二叉树	214
7.4.1	线索二叉树的概念	214
7.4.2	二叉树的线索化	215
7.4.3	线索二叉树的查找算法	216
7.5	二叉树的应用 1——基本算法	217
7.6	二叉树的应用 2——哈夫曼树	220
7.6.1	基本概念	220
7.6.2	哈夫曼树	221
7.6.3	哈夫曼树的构造过程	221
7.6.4	哈夫曼树的存储结构及哈夫曼算法	222
7.6.5	哈夫曼算法	223
7.6.6	哈夫曼树的应用——哈夫曼编码	223
7.7	二叉树的应用 3——二叉排序树	225
7.7.1	二叉排序树及其性质	225
7.7.2	二叉排序树的存储结构	226
7.7.3	二叉排序树的结点查找算法	226
7.7.4	二叉排序树的结点插入算法	227
7.7.5	二叉排序树的生成算法	228
7.7.6	二叉排序树的结点删除算法	229
7.7.7	二叉排序树结点查找算法性能分析	231
7.7.8*	平衡二叉排序树	232
7.8	二叉树的应用 4——堆和堆排序	239
7.8.1	堆和堆排序的概念	239
7.8.2	堆的调整算法	240
7.8.3	建堆	242
7.8.4	堆排序算法及性能分析	243
	小结	244
	习题	244
第 8 章	树和森林及其应用	253
8.1	树和森林的基本概念	253
8.1.1	树和森林的定义	253
8.1.2	树的性质	254
8.1.3	树和森林的遍历	255
8.1.4	树、森林与二叉树的关系	256
8.2	树的存储结构	258
8.2.1	树的顺序存储结构	258
8.2.2	树的链接存储结构	259
8.3	树的基本算法及性能分析	261

8.3.1	树转换为二叉树算法及性能分析	261
8.3.2	森林转换为二叉树算法及性能分析	262
8.3.3	二叉树转换为树算法及性能分析	263
8.3.4	二叉树转换为森林算法及性能分析	265
8.3.5	树的遍历算法	265
8.3.6	森林的遍历算法	266
8.4	树的应用——B 树	267
8.4.1	B 树的概念	267
8.4.2	B 树的数据存储类型	267
8.4.3	B 树的查找算法	268
8.4.4	B 树的插入算法	270
8.4.5	B 树的删除算法及实例	275
	小结	278
	习题	278
第 9 章	散列结构及其应用	282
9.1	散列结构的基本概念	282
9.1.1	散列结构的概念	282
9.1.2	散列存储结构——散列表	283
9.1.3	散列函数	284
9.1.4	冲突处理方法——开放定址法	287
9.1.5	冲突处理方法——链地址法	289
9.2	线性探测散列算法	290
9.2.1	线性探测散列算法的数据类型	290
9.2.2	线性探测散列基本算法	290
9.3	链地址法散列算法	293
9.3.1	链地址散列算法的数据类型	293
9.3.2	链地址散列基本算法	294
9.4	散列结构的查找性能分析	295
9.5*	散列结构应用 1——LZW 压缩问题	296
9.5.1	LZW 压缩问题	296
9.5.2	LZW 压缩算法及其算法性能分析	297
9.5.3	LZW 解压缩问题	298
9.5.4	LZW 解压缩算法及其算法性能分析	299
9.6*	散列结构应用 2——直接存取文件	300
9.6.1	基本概念	300
9.6.2	直接存取文件的存储结构及类型	301
9.6.3	文件操作 1——查找结点算法	301
9.6.4	文件操作 2——增加结点算法	302

9.6.5 文件操作 3——删除结点算法	303
小结	304
习题	304
第 10 章 图及其应用	307
10.1 图的概念	307
10.1.1 图的定义	307
10.1.2 图的基础知识	308
10.1.3 图的基本运算	311
10.2 图的存储结构及其基本算法	312
10.2.1 邻接矩阵及其数据类型	312
10.2.2 邻接矩阵的基本算法	314
10.2.3 邻接表及其数据类型	315
10.2.4 邻接表的基本算法	317
10.2.5 逆邻接表	318
10.2.6 十字链表及其数据类型	319
10.2.7 十字链表的基本算法	320
10.2.8 邻接多重表及其数据类型	321
10.2.9 邻接多重表的基本算法	322
10.3 图的遍历及算法	323
10.3.1 深度优先搜索遍历	324
10.3.2 深度优先搜索遍历算法	324
10.3.3 广度优先搜索遍历	325
10.3.4 广度优先搜索遍历算法	326
10.4 有向图的连通性和最小生成树	327
10.4.1 有向图的连通性分析	327
10.4.2 连通网的最小生成树问题——通信网络问题	328
10.5 图的(最小)生成树问题	328
10.5.1 图的生成树	328
10.5.2* 最小生成树算法——普利姆算法	329
10.5.3* 最小生成树算法——克鲁斯卡尔算法	331
10.6 非连通图的生成森林算法	332
10.7 最短路径	335
10.7.1 最短路径问题	335
10.7.2* 从某源点到其余各顶点之间的最短路径问题——迪杰斯特拉算法	336
10.7.3* 每一对顶点之间的最短路径问题——弗洛伊德算法	338
10.8 有向无环图及其应用	340
10.8.1 有向无环图及其应用问题实例	340
10.8.2 AOV 网及其特性	342

10.8.3	拓扑排序及其算法	343
10.8.4	AOE 网	345
10.8.5	关键路径及有关概念	345
10.8.6	关键路径求解及其算法思想	346
10.8.7	AOE 网关键路径求解算法的数据类型	347
10.8.8	AOE 网中求解关键路径的算法	347
10.8.9	关键路径求解算法分析	349
小结		350
习题		351
第 11 章*	算法性能分析和算法设计方法简介	357
11.1	算法和程序	357
11.1.1	算法与程序分析	357
11.1.2	空间复杂性分析	358
11.1.3	时间复杂度分析	359
11.2	再谈算法性能问题	360
11.2.1	时间复杂度的上限值	361
11.2.2	定理 11.2 的证明	361
11.2.3	时间复杂度的下限值	362
11.2.4	有关渐进符号的计算	362
11.3	算法性能测量问题	363
11.3.1	算法性能测量问题分析	363
11.3.2	算法性能测量实例	363
11.4	算法设计方法简介	364
11.4.1	分而治之的算法设计方法	365
11.4.2	最优化问题简介	365
11.4.3	贪婪算法设计方法	366
11.4.4	回溯算法设计方法	368
11.5	NP 问题简介	369
11.6	散列表查找性能	370
11.7	讨论题	371
附录 A	本书算法源程序	373
参考文献		414

第 1 章 数据结构和算法

教学目标：本书主要学习如何组织数据和设计算法，学习软件开发中所涉及到的各种常用数据结构。作为全书导引，本章要求掌握数据、数据类型、数据结构、算法及算法分析等基本概念和基础知识。另外，本章还结合课程学习要求，复习和掌握算法描述工具—C 语言中的指针类型与指针变量、结构类型与结构变量、函数与参数、递归定义和递归函数、动态存储分配、文件操作、程序测试与测试集、测试数据的设计和程序调试等问题。

教学提示：如何合理地组织数据、高效率地处理数据是扩大计算机应用领域、提高软件效率的关键。因此，必须完整地讲解和理解数据结构（逻辑结构、存储结构和相关算法）的定义及其实现的方法，算法数据类型的定义本质上是存储结构的实现。算法的时间性能分析是难点，算法的空间性能分析不可忽视。程序调试问题将影响算法的实现，在实践中要注意。

1.1 数据和数据类型

通常，我们将计算机的处理对象称为“数据”。数据的类型可分为数值型和非数值型两大类。本节介绍数据、数据元素、数据项、关键项、关键字、数据类型和抽象数据类型等有关知识。

1.1.1 数据和数据元素

1. 数据

定义 1.1 在计算机科学中，数据是指描述客观事物的数值、字符、相关符号等所有能够输入到计算机中并能被计算机程序处理的符号的总称。

在计算机系统的表示层次，数据以各种数据类型来表示，在计算机系统的物理层次，数据都是以二进制形式表示的。

【例 1-1】在计算机系统中，除数值型数据之外，字符、声音、图像、图形等信息是数据吗？

在计算机高级语言程序设计课程中，我们描述的客观事物通常以数值数据来表示。例如，从客观事物抽象而来的数值，通常用整型、实型、布尔型等基本数据类型数值来表示。

字符、声音、图像、图形等以及数据之间带有更复杂的结构关系的数据就是非数值型的数据，但能够通过编码后以二进制码输入到计算机中存储、处理和输出。所以，数据包括数值型和非数值型两大类。利用非数值型数据处理的问题很多，可以举出很多例子。

在计算机科学与技术专业中，数值型数据的处理方法通常在“计算方法”或者叫“数值计算”课程中学习，非数值型数据的处理方法通常在“数据结构与算法”课程中学习。

2. 数据元素

定义 1.2 数据中具有独立意义的个体称为数据元素。

数据元素是数据的基本单位，在程序设计时通常作为一个整体进行考虑和处理。在有些场合，数据元素也被称为元素或者记录、结点、顶点等。有时，一个数据元素可由一个数据项组成（简单型数据元素），也可由若干个数据项组成（复杂型数据元素）。

定义 1.3 数据项是数据不可分割的最小单位。

定义 1.4 关键项是可以唯一标识一条数据元素的数据项。关键项可以是一个数据项，也可以由多个数据项组合而生成一个关键项。

定义 1.5 关键项中的每一个值就称为所在数据元素的关键字（Key Word 或 Key）。

【例 1-2】为实现图书馆书目的自动检索，将与图书相关的数据做成如表 1-1 所示的表，试分析表中的数据元素（记录）、数据项、关键项、关键字。

表 1-1 图书目录关系表

书 号	书 名	作 者	价 格	...
8420001	计算机原理	张明	17.00	...
8420002	数据结构	陈英	23.00	...
8420003	C 语言	王范	17.60	...
8420004	大学英语	解东红	21.00	...
8420005	大学物理	洪亮	23.50	...
...	...	...	...	...

表中某一本书的相关数据（表中每一行）都是一个数据元素，每一个数据元素都具有独立意义。每一个数据元素由 4 个简单数据项（书号、书名、作者、价格）组成。书号是关键项，8420001 是第一条记录的关键字。用关键项代替所有记录，用关键字代替所在记录。也可以由“书号”和“作者”这两个数据项组合而生成一个关键项。满足“书号=8420003,作者=王范”的关键项的记录是表 1-1 中的第 3 个记录。

表中数据元素之间的关系是自上而下的线性顺序关系，所以表 1-1 也称为线性表。书目的自动检索就是计算机按照某个特定的要求（如给定书名），对某一张表按某种查询方法（如按自上而下的顺序）进行查询。依此类推，学生成绩表、工资表、员工信息表甚至人口普查表等都可以是线性表。

一般来讲，数据元素是相对于所讨论的问题而言的：对二维表来说，每个记录就是它的数据元素；对字符串来说，每个字符就是它的数据元素；对数组来说，每个成分就是它的数据元素等。

1.1.2 数据类型

1. 数据类型的概念和定义

数据类型是和数据结构密切相关的一个概念，在高级程序语言中，用以表示程序的操作对象的特性。根据计算机所处理数据的方式和结果的不同，高级程序语言中定义了几种数据