



21 世纪大学本科
计算机专业系列教材

齐德昱 编著

数据结构与算法

<http://www.tup.com.cn>

- 根据教育部高教司主持评审的《中国计算机科学与技术学科教程 2002》组织编写
- 与美国 ACM 和 IEEE/CS 《Computing Curricula 2001》同步



清华大学出版社

21世纪大学本科计算机专业系列教材

数据结构与算法

齐德昇 编著

清华大学出版社
北京

内 容 简 介

本书包括数据结构和算法设计方法两部分内容。数据结构部分重点介绍计算机程序设计中所涉及的表、栈、队列、树、图等基本数据对象的面向对象抽象与实现;算法设计方法部分介绍基本的算法设计策略与方法,包括逐步求精法、穷举法、迭代法、递推法、递归法、分治法、回溯法、分支限界法、动态规划法、贪心法等。

本书的数据结构部分将数据抽象与面向对象化作为重点,是对传统的“数据结构”课程的更新与扩充,以抽象观点和类库观点,对基本数据结构赋予新的内涵、新的处理方式,使其上升为面向对象数据结构,这与目前用C++描述数据结构的教材不同。

本书内容丰富,涵盖了“数据结构与算法”课程的国内外最新教学大纲——ACM 和 IEEE CS CC2001和《中国计算机科学与技术学科教程 2002》规定内容,并形成鲜明的特色,适合作为计算机专业本科生或非计算机专业的研究生的“数据结构与算法”教材,也可供软件设计师和程序员用作继续学习面向对象程序设计的教材。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

数据结构与算法/齐德昱编著 —北京:清华大学出版社,2003

(21世纪大学本科计算机专业系列教材)

ISBN 7-302-06866-6

I. 数… II. 齐… 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(2003)第 053934 号

出 版 者:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

地 址:北京清华大学学研大厦

邮 编:100084

客户服务:010-62776969

责任编辑:张瑞庆 王冰飞

封面设计:孟繁聪

版式设计:肖 米

印 装 者:清华大学印刷厂

发 行 者:新华书店总店北京发行所

开 本:185×230 印张:27 字数:556千字

版 次:2003年10月第1版 2003年10月第1次印刷

书 号:ISBN 7-302-06866-6/TP·5093

印 数:1~6000

定 价:33.00元

21 世纪大学本科计算机专业系列教材编委会

名誉主任：陈火旺

主任：李晓明

副主任：钱德沛 焦金生

委员：（按姓氏笔画为序）

马殿富	王志英	王晓东	宁 洪	刘 辰
孙茂松	李大友	李仲麟	吴朝晖	何炎祥
宋方敏	张大方	张长海	周兴社	侯文永
袁开榜	钱乐秋	黄国兴	蒋宗礼	曾 明
廖明宏	樊孝忠			

秘书：张瑞庆

本书责任编辑：蒋宗礼

21 世纪是知识经济的时代,是人才竞争的时代。随着 21 世纪的到来,人类已步入信息社会,信息产业正成为全球经济的主导产业。计算机科学与技术信息产业中占据了最重要的地位,这就对培养 21 世纪高素质创新型计算机专业人才提出了迫切的要求。

为了培养高素质创新型人才,必须建立高水平的教学计划和课程体系。在 20 多年跟踪分析 ACM 和 IEEE 计算机课程体系的基础上,紧跟计算机科学与技术的发展潮流,及时制定并修正教学计划和课程体系是尤其重要的。计算机科学与技术的发展对高水平人才的要求,需要我们从总体上优化课程结构,精炼教学内容,拓宽专业基础,加强教学实践,特别注重综合素质的培养,形成“基础课程精深,专业课程宽新”的格局。

为了适应计算机科学与技术学科发展和计算机教学计划的需要,要采取多种措施鼓励长期从事计算机教学和科技前沿研究的专家教授积极参与计算机专业教材的编著和更新,在教材中及时反映学科前沿的研究成果与发展趋势,以高水平的科研促进教材建设。同时适当引进国外先进的原版教材。

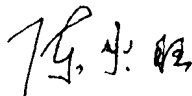
为了提高教学质量,需要不断改革教学方法与手段,倡导因材施教,强调知识的总结、梳理、推演和挖掘,通过加快教案的不断更新,使学生掌握教材中未及时反映的学科发展新动向,进一步拓宽视野。教学与科研相结合是培养学生实践能力的有效途径。高水平的科研可以为教学提供最先进的高新技术平台和创造性的工作环境,使学生得以接触最先进的计算机理论、技术和环境。高水平的科研还可以为高水平人才的素质教育提供良好的物质基础。学生在课题研究中不但能了解科学研究的艰辛和科研工作者的奉献精神,而且能熏陶和培养良好的科研作风,锻炼和培养攻关能力和协作精神。

进入 21 世纪,我国高等教育进入了前所未有的大发展时期,时代的进步与发展对高等教育质量提出了更高、更新的要求。2001 年 8 月,教育部颁发了《关于加强高等学校本科教学工作,提高教学质量的若干意见》。文件指出,本科教育是高等教育的主体和基础,抓好本科教学是提高整个高等教育质量的重点和关键。随着高等教育的普及和高等学校的扩招,在校大学本科计算机专业学生的人数将大量上升,对适合 21 世纪大学本科计算机科学与技术学科课程体系要求的,并且适合中国学生学习的计算机专业教材的需求量

也将急剧增加。为此,中国计算机学会和清华大学出版社共同规划了面向全国高等院校计算机专业本科生的“21 世纪大学本科计算机专业系列教材”。本系列教材借鉴美国 ACM 和 IEEE/CS 最新制定的《Computing Curricula 2001》(简称 CC2001)课程体系,反映当代计算机科学与技术学科水平和计算机科学技术的新发展、新技术,并且结合中国计算机教育改革成果和中国国情。

中国计算机学会教育专业委员会和全国高等学校计算机教育研究会,在清华大学出版社的大力支持下,跟踪分析 CC2001,并结合中国计算机科学与技术学科的发展现状和计算机教育的改革成果,研究出了《中国计算机科学与技术学科教程 2002》(China Computing Curricula 2002,简称 CCC2002),该项研究成果对中国高等学校计算机科学与技术学科教育的改革和发展具有重要的参考价值和积极的推动作用。

“21 世纪大学本科计算机专业系列教材”正是借鉴美国 ACM 和 IEEE/CS CC2001 课程体系,依据 CCC2002 基本要求组织编写的计算机专业教材。相信通过这套教材的编写和出版,能够在内容和形式上显著地提高我国计算机专业教材的整体水平,继而提高我国大学本科计算机专业的教学质量,培养出符合时代发展要求的具有较强国际竞争力的高素质创新型计算机人才。



中国工程院院士

国防科学技术大学教授

21 世纪大学本科计算机专业系列教材编委会名誉主任

2002 年 7 月

前言

FOREWORD

本书内容分两部分:数据结构(第1章,第3章至第10章)和算法设计方法(第2章、第11章至第12章)。本书形式上仍然以传统数据结构的主要内容为主线,但在内容上对传统的“数据结构”课程进行更新与扩充,对基本数据结构赋予新的内涵、新的处理方式,使其上升为面向对象数据结构。本教材的数据结构部分与目前国内外用C++描述数据结构的教材不同,有下列主要特色。

基本目的方面

- 使学生掌握基本程序设计中所需的基本数据设置方法,即以数据为中心的程序设计技术。
- 使学生建立程序设计模块化、抽象化的思想,掌握相关技术。
- 使学生进一步建立面向对象思想,掌握其方法,并将其融入数据结构的处理中。
- 培养计算机专业学生建立现代软件开发中广为使用的类库的基本能力。

基本特征方面

- **对象观点。**将传统数据结构中的每种逻辑结构中的各种成分都视为基本对象,给出它们的抽象访问接口、存储方式及对象实现。面向对象技术实质上是数据结构概念的自然扩展与继续,对象包含数据结构中的主要因素——数据成分与操作。因此,数据结构的“面向对象化”是必然的发展趋势。

- **类库观点。**类库是面向对象的产物,它的出现,使软件IC成为现实。目前世界上流行的二次开发工具/系统/支撑环境,大多是以类库形式出现的。因此,类库的创建,是计算机专业学生所必须掌握的技术。本教材从基本讲起,以传统的数据结构为蓝本,将各种逻辑结构整体分别视为类库,依次讲述其对象描述、抽象接口设置、对象实现。

- **面向对象技术。**本教材在以基本数据结构为对象介绍类库的构造的同时,还介绍了面向对象程序设计中的多方面的基本技术,如异常处理、可变类型处理、继承设计等,是面向对象程序设计课程的重要补充与发展。

本书的算法设计部分则抛开了面向对象思想,将重点放到了问题的解决过程的思路和方法上,力求清晰、严密地归纳各种基本的算法设计策略和方法及其应用,并给出严密

的数学推导和证明。

本书主要面向“数据结构与算法”课程,是根据最新教学大纲与国内外最新教材动向,结合作者自己的教学研究成果编写而成。该课程是计算机类专业的专业基础课,为必修课中的主干课。学习本课程,需要具备高级语言程序设计的基础和C++基础。如果没有学过C++,教师可在讲授前,用2~4个学时简要介绍C++中类和对象的概念和定义方法,而对于C++中其他问题,可在讲授正文中穿插介绍。本书内容较多,全部讲授至少需要72学时。如果学时不够,建议略去带“*”的部分。

本教材使用的C++编译器是C++ Builder 6.0中自带的C++编译器(非Windows环境,即Console环境)。

为了阅读方便,书中部分内容用特殊形式标明,说明如下:

: “强调”,表示对应的内容比较重要,强调指出。

: “课外资料”,介绍相关背景或基础性的内容,旨在扩大知识和辅助理解。

: “提醒”,用于指出一些易于出现错误的情况,起到提醒作用。

北京大学李晓明教授、北京工业大学蒋宗礼教授认真审阅了全书,并提出了许多宝贵意见。华南理工大学李仲麟教授、肖南峰教授、张齐副教授、陈俊风副教授以及张见威、李缨、杨清洪等老师,都对本书的编写给予了大力支持和帮助。值此出版之际,作者一并表示衷心感谢!

限于水平与时间,书稿虽几经修改,仍难免存在缺点或错误,恳请广大读者惠予批评指正。

作 者

qideyu@scut.edu.cn

2003年2月



CONTENTS

第 1 章 概述	1
1.1 数据结构的兴起与发展	1
1.2 数据结构的研究对象	2
1.3 数据结构的概念	3
1.4 数据结构的图示	4
1.5 数据结构的分类	5
1.5.1 集合	5
1.5.2 线性结构	5
1.5.3 树形结构	5
1.5.4 图状结构	6
1.6 数据结构的存储	6
1.6.1 存储器表示	6
1.6.2 存储映像	7
1.6.3 基本存储方法	7
1.7 数据结构的访问接口	9
1.7.1 访问接口与逻辑结构	9
1.7.2 基本操作的种类	10
1.7.3 基本操作的实现	10
1.8 面向对象方法	10
1.8.1 对象与类	10
1.8.2 面向对象方法要素	11
1.8.3 面向对象方法的若干述评*	13
1.8.4 面向对象程序设计语言*	14
1.9 面向对象与数据结构	18
1.9.1 面向对象与数据结构的关系	18

1.9.2 面向对象数据结构	18
1.9.3 数据结构的对象模型	19
本章小结	20
习题	20
第2章 程序设计基本策略与方法	21
2.1 算法	21
2.1.1 算法的概念	21
2.1.2 算法的时间复杂度与空间复杂度	22
2.1.3 算法时间复杂度的度量	24
2.2 穷举法	26
2.3 递推法与迭代法	26
2.3.1 递推法	26
2.3.2 迭代法	27
2.4 递归法	29
2.4.1 递归与递归程序的概念	29
2.4.2 递归程序设计要点	31
2.4.3 递归程序执行机理	31
2.4.4 Hanoi 塔问题与运行图	33
2.5 逐步求精法	35
2.5.1 基本思想	35
2.5.2 应用示例	35
2.6 分治法	37
2.6.1 基本思想	37
2.6.2 平面分治法示例——顺序统计	38
2.6.3 迭代分治法示例——循环赛赛程安排*	40
本章小结	42
习题	43
第3章 线性表	45
3.1 线性表的逻辑结构	45
3.1.1 基本概念	45
3.1.2 线性表抽象模型	45
3.2 线性表的顺序存储结构	49

3.2.1 基本存储方法	49
3.2.2 面向对象描述	50
3.3 异常处理与下标选择器*	59
3.3.1 异常处理	59
3.3.2 下标选择器	62
3.4 线性表的链式存储——线性链表	66
3.4.1 链式存储方法	66
3.4.2 线性链表的面向对象描述	66
3.4.3 线性链表的面向对象实现	69
3.5 几种特殊线性链表	79
3.5.1 带头结点的链表	79
3.5.2 循环链表	80
3.5.3 双向链表	81
3.6 线性表应用示例	82
3.6.1 集合运算*	82
3.6.2 一元多项式相加	83
3.6.3 一元多项式的乘法*	92
本章小结	92
习题	93
第4章 特殊线性表——栈、队列、串	94
4.1 栈	94
4.1.1 栈的逻辑结构	94
4.1.2 栈的顺序存储结构	96
4.1.3 栈的链式存储结构	100
4.2 队列	104
4.2.1 队列的逻辑结构	104
4.2.2 队列的顺序存储结构	105
4.2.3 队列的链式存储结构	110
4.3 串	112
4.3.1 串的逻辑结构	112
4.3.2 串的存储结构	116
本章小结	119
习题	120

第 5 章 数组与十字链表 121

5.1 数组 121

5.1.1 数组的定义与运算 121

5.1.2 数组的存储结构与寻址问题 122

5.1.3 一维数组的存储与寻址 122

5.1.4 二维数组的存储与寻址 123

5.1.5 多维数组的存储与寻址 124

5.1.6 寻址公式的计算 125

5.2 特殊数组 126

5.2.1 对称矩阵 126

5.2.2 下/上三角矩阵 127

5.3 稀疏矩阵 127

5.3.1 稀疏矩阵的逻辑表示 127

5.3.2 三元组表存储法 128

5.3.3 三元组表的操作 130

5.3.4 转置操作 132

5.4 十字链表 138

5.4.1 存储方式 138

5.4.2 十字链表对象 140

5.4.3 基本操作的实现 142

本章小结 146

习题 146

第 6 章 树形结构 148

6.1 树形结构的基本概念 148

6.1.1 树形结构的定义 148

6.1.2 基本术语 151

6.2 二叉树 152

6.2.1 二叉树的基本概念 152

6.2.2 几种特殊二叉树 152

6.2.3 二叉树的基本性质 153

6.2.4 二叉树的遍历 155

6.3 二叉树的存储结构 157

6.3.1 顺序存储结构 157

6.3.2 链式存储结构	158
6.4 二叉树对象模型	160
6.4.1 二叉树结点对象	160
6.4.2 二叉树对象	161
6.5 二叉树的遍历操作	165
6.5.1 前序遍历操作	165
6.5.2 中序遍历操作	169
6.5.3 后序遍历操作	170
6.6 二叉树的解析表示与存储结构之间的转化	173
6.6.1 双遍历结果转化为树	173
6.6.2 根据广义表表示创建树	175
6.6.3 根据存储结构创建广义表*	178
6.6.4 根据前序扩展序列创建树*	179
6.7 二叉树的线索化	181
6.7.1 线索化的概念	181
6.7.2 线索化算法	182
6.8 树与森林	184
6.8.1 树与森林的遍历	184
6.8.2 树、森林与二叉树之间的转化	186
6.8.3 树的存储结构	187
6.9 树对象模型*	190
6.9.1 树结点对象	190
6.9.2 树类	191
6.10 树的应用示例——哈夫曼树	194
6.10.1 哈夫曼树的基本概念	194
6.10.2 哈夫曼树构造算法	195
6.10.3 哈夫曼树构造算法的实现	196
6.10.4 哈夫曼判定树	199
6.10.5 哈夫曼编码与数据压缩	200
本章小结	201
习题	202
第7章 图结构	204
7.1 图的基本概念	204

7.1.1	图的概念	204
7.1.2	图的基本操作	207
7.2	图的对象抽象模型	207
7.2.1	图结点抽象模型	208
7.2.2	图的边对象抽象模型	208
7.2.3	图抽象对象模型	209
7.3	图的存储结构	211
7.3.1	邻接矩阵法	212
7.3.2	邻接表	214
7.3.3	十字链表*	218
7.3.4	邻接多重表*	220
7.4	图的遍历	222
7.4.1	概述	222
7.4.2	深度优先遍历	222
7.4.3	深度优先遍历的性质	227
7.4.4	广度优先遍历	228
7.4.5	广度优先遍历的性质	232
7.5	拓扑排序	232
7.5.1	拓扑序列与 AOV 网	233
7.5.2	拓扑排序算法与实现	234
7.6	AOE 网与关键路径	238
7.6.1	AOE 网与关键路径的概念	238
7.6.2	关键路径的识别	241
	本章小结	242
	习题	242
第 8 章	广义表	244
8.1	广义表的逻辑结构	244
8.1.1	基本概念	244
8.1.2	广义表逻辑图	246
8.1.3	广义表的遍历	246
8.1.4	基本特性	247
8.1.5	基本操作	247
8.2	广义表的存储结构	247

8.2.1 基本存储方法	247
8.2.2 链式结构的高级语言描述	249
8.3 广义表对象模型*	250
8.3.1 广义表元素接口	250
8.3.2 广义表接口	251
8.4 广义表的分支单链表对象*	254
8.4.1 结点对象	254
8.4.2 分支单链表对象	255
8.5 广义表操作的实现*	256
8.5.1 一般问题	256
8.5.2 遍历操作	257
8.5.3 广义表统计计数	258
8.5.4 广义表的串行化与逆串行化	259
8.5.5 广义表的复制与求尾	261
8.6 广义表结构的应用	262
8.6.1 多元多项式的表示	262
8.6.2 层次结构的表示	263
本章小结	264
习题	264

第9章 检索结构 266

9.1 概述	266
9.1.1 检索的概念	266
9.1.2 检索结构	267
9.1.3 检索算法的时间与空间复杂度分析	267
9.1.4 检索算法的判定树	268
9.2 线性结构的检索	270
9.2.1 顺序检索	270
9.2.2 折半检索	271
9.2.3 斐波那契检索*	275
9.2.4 插值检索*	280
9.3 线性索引结构	280
9.3.1 概述	280
9.3.2 稠密索引	281

9.3.3	分块索引	281
9.4	树形索引结构与二叉排序树	283
9.4.1	树形索引结构概述	283
9.4.2	二叉排序树的概念	283
9.4.3	二叉排序树的检索	284
9.4.4	二叉排序树的插入	286
9.4.5	二叉排序树的删除	287
9.4.6	二叉排序树的分析与最优二叉排序树	292
9.5	平衡二叉排序树	293
9.5.1	基本概念	293
9.5.2	若干性质	293
9.5.3	局部平衡调整算法	294
9.6	B 树	299
9.6.1	B 树的概念	299
9.6.2	B 树的存储结构	300
9.6.3	B 树的基本操作	301
9.6.4	B 树的检索方法	301
9.6.5	B 树的插入	301
9.6.6	B 树的删除	304
9.6.7	B ⁺ 树	306
9.6.8	B 树对象模型	307
9.7	散列结构	307
9.7.1	概念	307
9.7.2	散列技术中的主要问题	308
9.7.3	散列过程	308
9.7.4	散列函数的设计	308
9.7.5	冲突解决	310
	本章小结	311
	习题	312
第 10 章	外存与文件组织	313
10.1	外存结构	313
10.1.1	外存简介	313
10.1.2	磁带结构	313

10.1.3	磁盘结构	314
10.2	文件	315
10.2.1	文件的概念	315
10.2.2	文件操作与存取方式	316
10.2.3	文件的物理组织	316
10.2.4	缓冲技术	318
10.3	顺序文件	318
10.4	索引文件	319
10.5	ISAM*	320
10.5.1	ISAM 的概念	320
10.5.2	ISAM 结构的操作	321
10.6	VSAM	322
10.6.1	VSAM 的概念	322
10.6.2	VSAM 结构的操作	323
10.7	散列方式	324
10.8	多索引文件	325
10.8.1	多重表文件	325
10.8.2	倒排文件	326
	本章小结	327
	习题	327
第 11 章	排序算法	329
11.1	概述	329
11.2	插入排序	330
11.2.1	直接插入排序	331
11.2.2	其他插入排序算法	332
11.3	交换排序	332
11.3.1	冒泡排序	333
11.3.2	冒泡算法的改进	334
11.3.3	快速排序*	335
11.4	选择排序	336
11.4.1	直接选择排序	336
11.4.2	堆排序	337
11.5	归并排序	345