

重点大学计算机专业系列教材

数据结构基础教程（C语言版） 学习指导书

陈 瑛 叶小平 编著



清华大学出版社

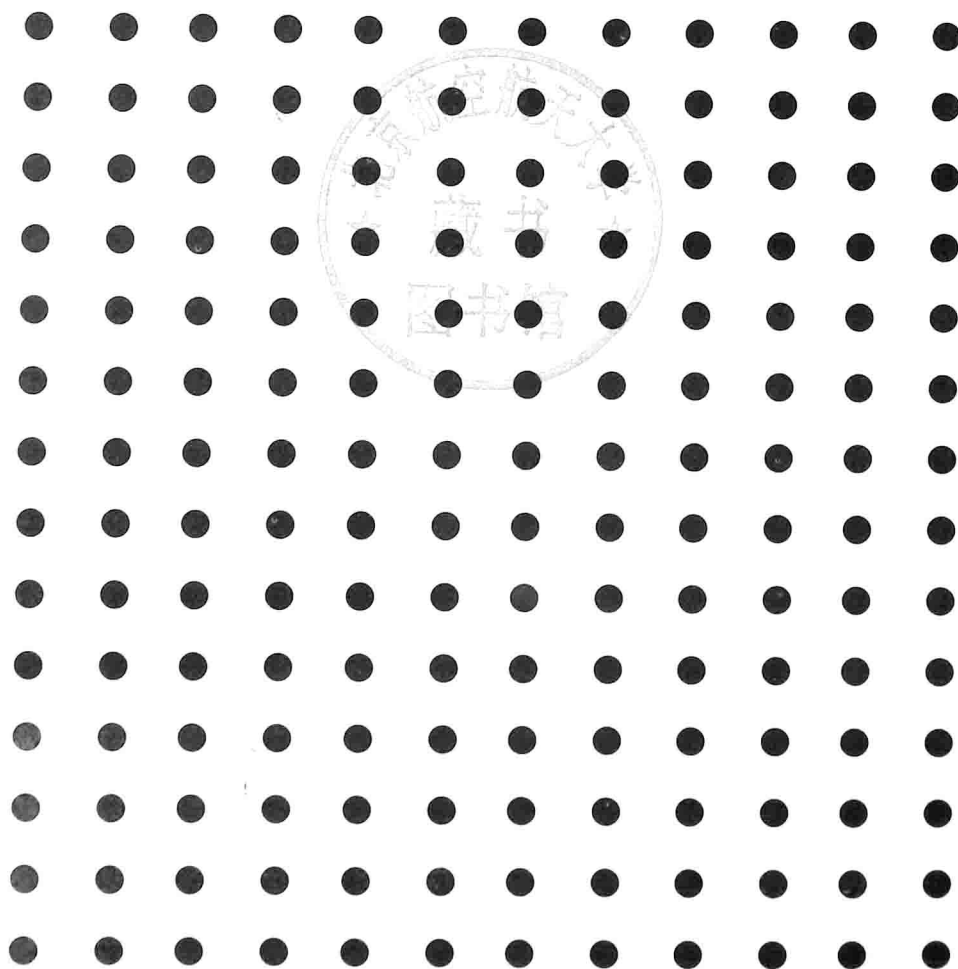
014057526

TP312C-42
221

重点大学计算机专业系列教材

数据结构基础教程 (C语言版) 学习指导书

陈 璞 叶小平 编著



清华大学出版社



北航

C1745945

TP312C-42
221

内 容 简 介

本书是结合两位作者多年讲授“数据结构”课程的教学经验,以及在理论答疑、上机实践、课外辅导过程中所遇到的实际问题编写而成的。本书虽然是《数据结构基础教程》立体化教材建设的配套用书,但也可以独立使用。本书由10章组成,每章由5节组成。第1节学习要点,给出课程内容的主要知识点,突出重点,指导学习者理顺课程内容,完成知识构建;第2节重点难点分析,给出课程学习要求和线索,对重点难点问题进行分析解答,突破学习者学习的难点;第3节例题,提供一些具体的、有代表意义的典型例题,并对答案进行解析;第4节习题,题目覆盖面广,形式多变,指导学生将知识点转化为解决实际问题的能力;第5节实验指导,包含设计型、综合型题目并提供程序分析,帮助学习者将教材知识点落到上机编程的实处,激发学习热情。

本书注重突出关键细节,并强化实例讲解。最后配套考试模拟题以及模拟题、习题、实验的参考答案,所有算法均经过VC平台上调试运行。

本书可供高等院校计算机信息科学与技术及其相关专业本科教学使用,同时也适合非计算机专业人员自学使用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据结构基础教程(C语言版)学习指导书/陈瑛,叶小平编著.--北京:清华大学出版社,2014

重点大学计算机专业系列教材

ISBN 978-7-302-36145-9

I. ①数… II. ①陈… ②叶… III. ①数据结构—高等学校—教材 ②C语言—程序设计—高等学校—教材 IV. ①TP311.12 ②TP312

中国版本图书馆CIP数据核字(2014)第072651号

责任编辑:刘向威 薛 阳

封面设计:常雪影

责任校对:李建庄

责任印制:王静怡

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦A座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

课 件 下 载: <http://www.tup.com.cn>, 010-62795954

印 刷 者:北京市人民文学印刷厂

装 订 者:三河市漂源装订厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:14.5

字 数:365千字

版 次:2014年8月第1版

印 次:2014年8月第1次印刷

印 数:1~2000

定 价:29.00元

出版说明

随着国家信息化步伐的加快和高等教育规模的扩大,社会对计算机专业人才的需求不仅体现在数量的增加上,而且体现在质量要求的提高上,培养具有研究和实践能力的高层次的计算机专业人才已成为许多重点大学计算机专业教育的主要目标。目前,我国共有 16 个国家重点学科、20 个博士点一级学科、28 个博士点二级学科集中在教育部部属重点大学,这些高校在计算机教学和科研方面具有一定优势,并且大多以国际著名大学计算机教育为参照系,具有系统完善的教学课程体系、教学实验体系、教学质量保证体系和人才培养评估体系等综合体系,形成了培养一流人才的教学和科研环境。

重点大学计算机学科的教学与科研氛围是培养一流计算机人才的基础,其中专业教材的使用和建设则是这种氛围的重要组成部分,一批具有学科方向特色优势的计算机专业教材作为各重点大学的重点建设项目成果得到肯定。为了展示和发扬各重点大学在计算机专业教育上的优势,特别是专业教材建设上的优势,同时配合各重点大学的计算机学科建设和专业课程教学需要,在教育部相关教学指导委员会专家的建议和各重点大学的大力支持下,清华大学出版社规划并出版本系列教材。本系列教材的建设旨在“汇聚学科精英、引领学科建设、培育专业英才”,同时以教材示范各重点大学的优秀教学理念、教学方法、教学手段和教学内容等。

本系列教材在规划过程中体现了如下一些基本组织原则和特点。

1. 面向学科发展的前沿,适应当前社会对计算机专业高级人才的培养需求。教材内容以基本理论为基础,反映基本理论和原理的综合应用,重视实践和应用环节。

2. 反映教学需要,促进教学发展。教材要能适应多样化的教学需要,正确把握教学内容和课程体系的改革方向。在选择教材内容和编写体系时注意体现素质教育、创新能力与实践能力的培养,为学生知识、能力、素质协调发展创造条件。

3. 实施精品战略,突出重点,保证质量。规划教材建设的重点依然是专业基础课和专业主干课;特别注意选择并安排了一部分原来基础比较好的优秀教材或讲义修订再版,逐步形成精品教材;提倡并鼓励编写体现重点大学

计算机专业教学内容和课程体系改革成果的教材。

4. 主张一纲多本,合理配套。专业基础课和专业主干课教材要配套,同一门课程可以有多本具有不同内容特点的教材。处理好教材统一性与多样化的关系;基本教材与辅助教材以及教学参考书的关系;文字教材与软件教材的关系,实现教材系列资源配套。

5. 依靠专家,择优落实。在制订教材规划时要依靠各课程专家在调查研究本课程教材建设现状的基础上提出规划选题。在落实主编人选时,要引入竞争机制,通过申报、评审确定主编。书稿完成后要认真实行审稿程序,确保出书质量。

繁荣教材出版事业,提高教材质量的关键是教师。建立一支高水平的以老带新的教材编写队伍才能保证教材的编写质量,希望有志于教材建设的教师能够加入到我们的编写队伍中来。

教材编委会

前言

计算机所加工处理的对象是数据,数据具有一定的组织结构,而数据结构正是研究如何组织、存储数据和对数据进行操作的一门科学。数据结构课程是计算机专业的一门核心基础课程,并且逐步发展成为其他理工科专业的选修课,因此具有相当重要的地位和作用。

“数据结构”具有理论抽象程度较高和应用实践性极强的特点,要求在“原理”与“实践”之间自由转换,但作为一种普遍现象,“原理与实践”的冲突在“数据结构”课程学习中表现明显,作为教师,有责任对学习者的困惑和难点做出指导。

本书是结合两位作者多年讲授“数据结构”课程的教学经验,以及在理论答疑、上机实践、课外辅导过程中所遇到的实际问题编写而成的。注重整合课程涉及的各种知识点与方法,注重重点难点问题的解析,注重算法思想到程序实现的转换过程,突出关键细节,强化基本算法的具体设计与实现,同时通过归纳、对比、图表等各种手段直观演示,提供学生独立思考与创新的思路,注意促使学生从编程实现的“终极”层面上掌握所学习的重要原理和基本算法。使得学生能够在适当的理论背景与实践框架下完成对所学知识的整合构建。

本书作为《数据结构基础教程》立体化教材建设的配套用书设计编写,但脱离《数据结构基础教程》一书,也可以独立使用。本书为学习者提供了学习指导,作为教材的有力补充,可满足课后学习的需要。本书由10章组成,每章由5节组成。第1节学习要点,给出课程内容的主要知识点,突出重点,指导学习者理顺课程内容,完成知识构建;第2节重点难点分析,给出课程学习要求和线索,对重点难点问题进行分析解答,突破学习者学习的难点;第3节例题,提供一些具体的、有代表意义的典型例题,并对答案进行解析;第4节习题,题目覆盖面广,形式多变,指导学生将知识点转化为解决实际问题的能力;第5节实验指导,包含设计型、综合型题目并提供程序分析,帮助学习者将教材知识点落到上机编程的实处,激发学习热情。

最后本书还配套考试模拟题以及模拟题、习题、实验的参考答案,让学习者在课后学习中及时检查,发现问题,保持学习动机的持续性。本书中所有

算法均经过 VC 平台上调试运行。

清华大学出版社给予本书编写以很大的支持与细心指导,在此表示衷心的感谢!

在教材编写与课程实践中,本书参考了国内外众多数据结构与算法方面的优秀教材,其中大多列举在书后的参考文献当中。在此,对这些教材的编著者表示衷心感谢。限于作者水平和其他条件约束,书中疏漏和不足之处在所难免,诚恳地希望专家学者、选用本书的教师和同学能不吝赐教,以便本书能进一步得以改进和完善。

作 者

2014年5月

目录

第 1 章 绪论	1
1.1 学习要点	1
1.1.1 数据相关概念	1
1.1.2 数据模型与数据结构	2
1.1.3 数据操作与算法分析	3
1.1.4 数据结构地位与内容体系	3
1.2 重点难点分析	4
1.2.1 学习要求与线索	4
1.2.2 重点难点解析	4
1.3 例题	5
1.4 习题	6
1.5 实验指导	7
第 2 章 线性表	9
2.1 学习要点	9
2.1.1 线性表的逻辑结构	9
2.1.2 线性表的顺序存储结构	9
2.1.3 线性表的链式存储结构	10
2.1.4 线性表存储结构比较	14
2.2 重点难点分析	15
2.2.1 学习要求与线索	15
2.2.2 重点难点解析	15
2.3 例题	18
2.4 习题	20
2.5 实验指导	24
2.5.1 数组基本操作	24
2.5.2 纸牌游戏	25

2.5.3	链表基本操作	26
2.5.4	有序链表的合并	28
2.5.5	一元多项式相减	29
2.5.6	Joseph 环	31
第 3 章	栈与队列	33
3.1	学习要点	33
3.1.1	栈	33
3.1.2	队列	35
3.2	重点难点分析	38
3.2.1	学习要求与线索	38
3.2.2	重点难点解析	39
3.3	例题	41
3.4	习题	42
3.5	实验指导	45
3.5.1	栈的应用之单括号匹配	45
3.5.2	栈的应用之多级括号匹配	46
3.5.3	十进制整数转化为 R 进制整数	48
3.5.4	队列基本操作	49
3.5.5	跳舞配对问题 1	51
3.5.6	跳舞配对问题 2	52
第 4 章	数组、串和广义表	54
4.1	学习要点	54
4.1.1	数组	54
4.1.2	串	58
4.1.3	广义表	61
4.2	重点难点分析	62
4.2.1	学习要求与线索	62
4.2.2	重点难点解析	62
4.3	例题	63
4.4	习题	64
4.5	实验指导	67
4.5.1	对称矩阵的压缩存储与输出	67
4.5.2	三元组快速转置	68
4.5.3	文章编辑	70
第 5 章	二叉树及应用	72
5.1	学习要点	72

第 7 章 图 105

7.1.5	最短路径	110
7.1.6	有向无环网及应用	111
7.2	重点难点分析	111
7.2.1	学习要求与线索	111
7.2.2	重点难点解析	112
7.3	例题	115
7.4	习题	117
7.5	实验指导	121
7.5.1	图的建立与输出	121
7.5.2	图的深度优先遍历	122
第8章	查找	125
8.1	学习要点	125
8.1.1	相关概念	125
8.1.2	基于线性表查找	125
8.1.3	基于二叉树查找	126
8.1.4	基于散列表查找	128
8.2	重点难点分析	129
8.2.1	学习要求与线索	129
8.2.2	重点难点解析	130
8.3	例题	131
8.4	习题	133
8.5	实验指导	136
8.5.1	二分查找	136
8.5.2	建立二叉排序树	137
8.5.3	二叉排序树查找	139
8.5.4	散列查找	141
第9章	排序	143
9.1	学习要点	143
9.1.1	排序相关概念	143
9.1.2	插入排序	143
9.1.3	交换排序	144
9.1.4	选择排序	146
9.1.5	归并排序	146
9.1.6	各种排序方法比较	147
9.2	重点难点分析	147
9.2.1	学习要求与线索	147
9.2.2	重点难点解析	148

9.3 例题	149
9.4 习题	151
9.5 实验指导	153
9.5.1 直接插入排序	153
9.5.2 快速排序	155
9.5.3 简单选择排序	156
9.5.4 学生成绩管理系统	157
第 10 章 文件	160
10.1 学习要点	160
10.1.1 文件基本概念	160
10.1.2 顺序文件	161
10.1.3 索引文件	162
10.1.4 动态索引 B 树	163
10.1.5 散列文件	165
10.1.6 多关键字文件	165
10.2 重点难点分析	166
10.2.1 学习要求与线索	166
10.2.2 重点难点解析	166
10.3 例题	167
10.4 习题	168
10.5 实验指导	170
附录 A 考试模拟题	172
考试模拟题(一)	172
考试模拟题(二)	175
考试模拟题(三)	178
考试模拟题(四)	181
考试模拟题(五)	184
附录 B 考试模拟题参考答案	188
考试模拟题(一)参考答案	188
考试模拟题(二)参考答案	190
考试模拟题(三)参考答案	192
考试模拟题(四)参考答案	194
考试模拟题(五)参考答案	196
附录 C 习题参考答案	200
C.1 绪论参考答案	200

C.2 线性表参考答案	200
C.3 栈与队列参考答案	202
C.4 数组、串和广义表参考答案	203
C.5 二叉树及应用参考答案	205
C.6 树与森林参考答案	208
C.7 图参考答案	209
C.8 查找参考答案	213
C.9 排序参考答案	214
C.10 文件参考答案	216
附录 D 实验报告内容	219
参考文献	220

绪 论

第 1 章

1.1 学习要点

1.1.1 数据相关概念

数据(Data)是对客观事物的描述,是信息的载体,是能够被计算机识别、存储和加工处理的各种形式的总和。

从数据处理角度来看,计算机处理的数据可分为数值型数据和非数值型数据两种。

(1) 数值型数据:主要包括整数、实数和复数等可以“直接”计算的数据,这些数据可以通过数制转换“直接”进入计算机。

(2) 非数值型数据:主要包括字符、文字、图像、音频和视频等,这些数据需要经过各种比较复杂的编码方式才能进入计算机,计算机系统对这些数据主要不是进行“直接计算”,而是着眼于这些数据之间的各种关系进行存储和处理,其主要应用于文字处理、多媒体管理以及更为复杂的逻辑推理等领域中的数据组织、管理和查找等与“直接计算”无关的问题。非数值型数据的基本特征是数据处理方法相对简单,而数据间关系的计算机存储与管理却相当复杂,涉及的数据量也极为庞大(海量数据)。研究这些问题的计算机学科就是即将学习的“数据结构”。数据结构课程的主要任务就是研究程序设计过程中计算机的操作对象(数据)的组织结构以及对这些结构化数据操作的实现,而不是单纯的数据值的具体计算。

数据项(Data Item):数据语义表示中不可再分割的最小单位,自身具有独立含义。

数据元素(Data Elements):计算机进行技术处理的一个不可再进行分割的最小单位,它在计算机程序中通常作为一个整体考虑。

关于数据项、数据元素、数据的关系,可以用图 1-1 来表示。

数据类型(Data Type)是一组性质相同的值的集合以及定义在这个集合

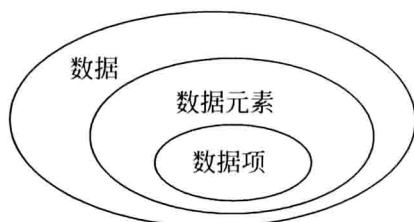


图 1-1 数据项、数据元素、数据之间的关系

上一组操作的总称。

抽象数据类型(Abstract Data Type, ADT)是对一种数据模型和定义于该模型上的一组基本操作集合的总称。

1.1.2 数据模型与数据结构

数据结构(Data Structure)是相互之间存在一种或多种特定关系的数据元素的集合,既反映数据元素的信息,又反映元素间的关系。“数据”是数据元素的有限集合,元素间的关系即数据的逻辑结构。

根据计算机处理的问题不同,处理的数据对象的逻辑结构也不同。例如,在学籍表的管理中,每个数据元素之间存在某种线性关系,如按学号的递增关系;人事管理信息系统中,一个上司与多个下属的领导与被领导关系;公交线路查询中各个公交站点错综复杂的关系等。根据这些逻辑结构的不同,可将数据结构分为 4 种。

(1) 集合结构:数据元素同属一个集合,别无任何关系。每个数据元素都和其他任何元素无关系,孤立存在,如图 1-2(a)所示。

(2) 线性结构:数据元素存在一个对一的线性关系,如前面所提到的学籍表,记录一个连着一个,像一条线,如图 1-2(b)所示。

(3) 树型结构:数据元素存在一个对多个的关系,层次结构很明显,如前面提到的人事系统,像一棵倒立的树,如图 1-2(c)所示。

(4) 图型结构:数据元素之间的关系很任意,错综复杂,存在多个对多个的关系,如前面提到的公交站点,像一个网,所以图型结构也称为网状结构,如图 1-2(d)所示。

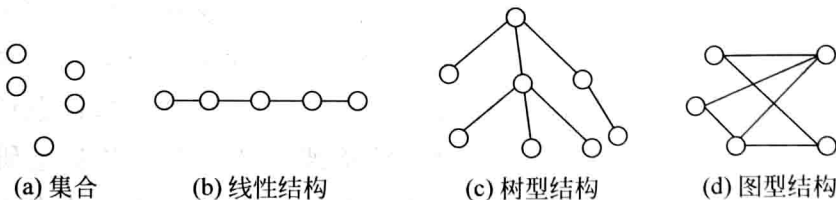


图 1-2 4 种基本数据结构

学习数据结构,要注意以下三个要素。

(1) 数据的逻辑结构——数据元素之间的逻辑关系。

(2) 数据的存储结构——数据元素及其关系在计算机存储器内的表示。

(3) 行为特征——对数据进行操作的特点和规范。

数据存储结构(Storage Structure)就是研究数据及其数据之间关系(逻辑结构)在计算机内存中的存储方式。

数据的存储结构是逻辑结构用计算机语言的实现,即建立计算机的机内表示,在存储空间中建立各结点之间的关联来表示数据元素之间的逻辑关系。4种主要的存储结构有顺序存储结构、链式存储结构、散列存储结构以及索引存储结构,本书主要介绍前面两种。

顺序存储结构中的元素都依次存储在一个连续的区域中,元素的物理地址就体现了元素的逻辑关系。链式存储结构中的结点不仅包含一个数据元素,还包含一个指针,指向该元素逻辑上的直接后继结点。

1.1.3 数据操作与算法分析

数据运算是在数据逻辑结构上对数据进行的操作,可以看做是在数据逻辑结构框架下对数据进行的一种加工和再处理,加工处理的结果需要保持原先的数据逻辑结构不变,即数据运算具有数据逻辑结构的“封闭性”。因此,数据运算不能脱离数据的逻辑结构而独立存在。每种逻辑结构框架下定义的各种运算就构成了相应逻辑结构的一个运算集合,也就是前述 ADT 的核心部件。

数据运算与数据逻辑结构不可分割,但逻辑结构和建立在其上的运算并不能直接在计算机上予以实现。数据运算在计算机上具体实现需要通过算法来完成。算法实际上就是相应数据运算在计算机上的具体实现途径,由此算法与数据的存储结构密切相关,不可分割。

算法是指能够解决一个实际问题的程序模块,表现形式为一系列指令构成的运算序列。一个“合法”的算法至少需要具有输入性、有穷性、确定性、可行性和输出性 5 方面的特征。

一个“好”的“合法”算法通常还需要满足正确性、可读性、健壮性和高效性 4 个方面的要求。

评价一个算法的效率主要看完成整个算法需要耗费多少机器资源,机器资源指的是完成某求解问题规模为 n 的算法所需要的时间和执行算法过程中需要占用的空间,记为 $S(n)$ 和 $T(n)$,分别称为空间复杂度和时间复杂度。

在计算时间复杂度时,只关心完成问题规模 n 的算法时间的数量级,记为 $O(n)$ 。数量级则只关心 n 的最高次幂而忽略其他低次项和系数。一般数据结构中常见的时间复杂度的数量级有 7 个:

$O(1)$ 常数型	$O(n)$ 线性型	$O(n^2)$ 平方型	$O(n^3)$ 立方型
$O(2^n)$ 指数型	$O(\log_2 n)$ 对数型	$O(n \log_2 n)$ 二维型	

当 n 足够大时,时间复杂度的增长趋势可以做如下排序:

$$O(1) < O(\log_2 n) < O(n) < O(n \log_2 n) < O(n^2) < O(n^3) < O(2^n)$$

时间复杂度的分析贯穿整本书的各个章节,要学会在分析了算法流程后,估算出每个算法的时间复杂度并表示出来。

1.1.4 数据结构地位与内容体系

数据结构在整个课程体系中处于承上启下的地位,是计算机科学的核心课程。

“数据结构”与计算机专业其他课程的关系如图 1-3 所示。

本书基本内容的组织体系结构如图 1-4 所示。

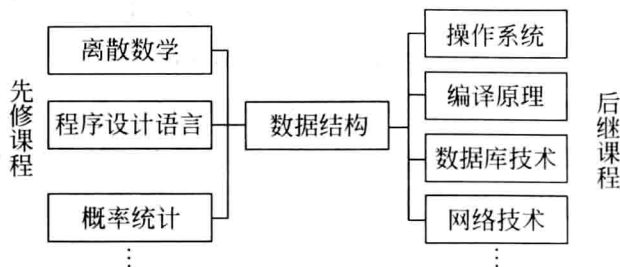


图 1-3 数据结构与其他课程关系图

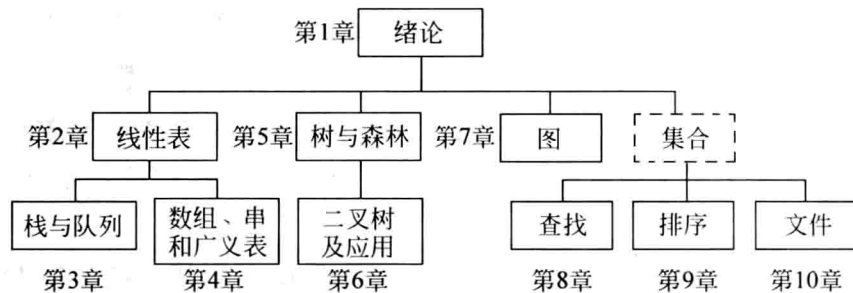


图 1-4 本书内容组织体系

1.2 重点难点分析

1.2.1 学习要求与线索

数据结构在整个计算机科学课程体系中处于承上启下的地位,是计算机科学的核心课程。但内容又比较抽象,专业性比较强。在本章的学习中,有以下几点学习要求。

- (1) 了解数据结构课程的学习目标。
- (2) 掌握各种基本术语的含义、区别与联系。
- (3) 掌握基本的数据结构类型和它们的主要特点,并能举例。
- (4) 掌握计算语句频度和估算算法时间复杂度、空间复杂度的方法。

学习者在学习过程中,可以沿着如下的学习线索进行学习。

- (1) 学习数据结构在计算机知识体系中的地位。
- (2) 学习什么是数据结构及几个相关概念(数据、数据项、数据元素、数据类型、抽象数据类型)。
- (3) 学习 4 种基本的数据结构类型(集合、线性、树、图型)和它们的主要特点。
- (4) 学习 4 种基本的数据存储结构(顺序、链式、索引、散列)和它们的主要特点。
- (5) 学习算法与数据结构的关系,算法的基本特征和要求。
- (6) 学习分析数据结构对算法的影响和算法时间复杂度分析。

1.2.2 重点难点解析

1. 什么是数据结构

计算机所要处理的不是单个数据元素而是具有某种共同特性的数据对象——数据元素