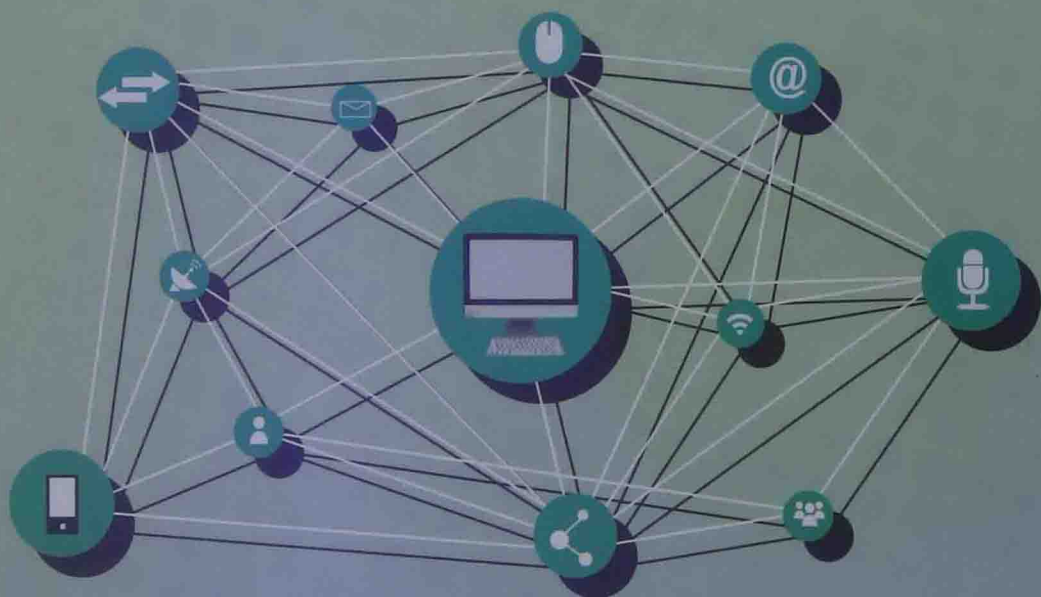




普通高等教育“十二五”应用型本科规划教材·计算机系列
山西省高等学校省级教学成果奖配套教材



数据结构

SHUJU JIEGOU

中国高等教育学会 组织编写

冯俊 编著

教育部“十二五”普通高等教育本科国家级规划教材
教育部“十二五”普通高等教育本科国家级规划教材



数据结构

第2版

清华大学出版社

ISBN 7-302-47777-7

清华大学出版社



普通高等教育“十二五”应用型本科规划教材·计算机系列
山西省高等学校省级教学成果奖配套教材

数据结构

中国高等教育学会 组织编写

冯 俊 编著

中国人民大学出版社
· 北京 ·

图书在版编目 (CIP) 数据

数据结构/冯俊编著. —北京: 中国人民大学出版社, 2014. 7
普通高等教育“十二五”应用型本科规划教材. 计算机系列
ISBN 978-7-300-19538-4

I. ①数… II. ①冯… III. ①数据结构-高等学校-教材 IV. ①TP311.12

中国版本图书馆 CIP 数据核字 (2014) 第 151666 号

普通高等教育“十二五”应用型本科规划教材·计算机系列
山西省高等学校省级教学成果奖配套教材

数据结构

中国高等教育学会 组织编写

冯 俊 编著

Shuju Jiegou

出版发行 中国人民大学出版社

社 址 北京中关村大街 31 号

电 话 010-62511242 (总编室)

010-82501766 (邮购部)

010-62515195 (发行公司)

网 址 <http://www.crup.com.cn>

<http://www.ttrnet.com> (人大教研网)

经 销 新华书店

印 刷 北京七色印务有限公司

规 格 185mm×260mm 16 开本

印 张 22

字 数 510 000

邮政编码 100080

010-62511770 (质管部)

010-62514148 (门市部)

010-62515275 (盗版举报)

版 次 2014 年 7 月第 1 版

印 次 2014 年 7 月第 1 次印刷

定 价 42.80 元

版权所有 侵权必究

印装差错 负责调换

教师信息反馈表

为了更好地为您服务，提高教学质量，中国人民大学出版社愿意为您提供全面的教学支持，期望与您建立更广泛的合作关系。请您填好下表后以电子邮件或信件的形式反馈给我们。

您使用过或正在使用的我社教材名称		版次	
您希望获得哪些相关教学资料			
您对本书的建议（可附页）			
您的姓名			
您所在的学校、院系			
您所讲授的课程名称			
学生人数			
您的联系地址			
邮政编码		联系电话	
电子邮件（必填）			
您是否为人大社教研网会员	☐ 是，会员卡号：_____ ☐ 不是，现在申请		
您在相关专业是否有主编或参编教材意向	☐ 是 ☐ 否 ☐ 不一定		
您所希望参编或主编的教材的基本情况（包括内容、框架结构、特色等，可附页）			

我们的联系方式：北京市西城区马连道南街 12 号
 中国人民大学出版社应用技术分社
 邮政编码：100055
 电话：010-63311862
 网址：<http://www.crup.com.cn>
 E-mail：smooth.wind@163.com



前

言

为贯彻落实《国家中长期教育改革和发展规划纲要（2010—2020 年）》，结合“高等学校本科教学质量与教学改革工程”万种新教材建设项目，充分发挥教材在提高人才培养质量中的基础性作用，积极推进高等院校教学改革和教材建设，中国高等教育学会启动了《普通高等教育“十二五”应用型本科规划教材》立项项目。

本书是普通高等教育“十二五”应用型本科规划教材，是山西省高等学校省级教学成果奖配套教材。教材充分吸收学科新知识、新技术、新成果，使其具有基础性、理论性、指导性、概括性，注重应用性、实践性，以增强读者计算思维能力与运用计算机实际问题能力。

2005 年 6 月，美国总统信息技术咨询委员会（PITAC）给美国总统提交报告《计算科学：确保美国竞争力》（Computational Science: Ensuring America's Competitiveness）。报告中写道：虽然计算本身也是一门学科，但是其具有促进其他学科发展的作用。21 世纪科学上最重要的、经济上最有前途的研究前沿都有可能通过熟练掌握先进的计算技术和运用计算科学而得到解决。报告建议，将计算科学长期置于国家科学与技术领域的中心领导地位。

2006 年 3 月，美国卡内基·梅隆大学计算机科学系主任周以真（Jeannette M. Wing）教授在美国计算机权威杂志、ACM 会刊《Communications of the ACM》杂志上给出计算思维（Computational Thinking, CT）的定义，计算思维是运用计算机科学的基础概念进行问题求解、系统设计以及人类行为理解等涵盖计算机科学之广度的一系列思维活动。

20 世纪 80 年代，计算技术开始渗透到大多数学科领域。1985 年春，（美国）计算机协会（Association for Computing Machinery, ACM）与（美国）电气和电子工程师学会计算机分会（Institute of Electrical and Electronics Engineers—Computer Society, IEEE-CS）联手组成攻关组，经过近 4 年工作，提交了《计算作为一门学科》（Computing as a discipline）的报告。1990 年，攻关组在该报告的基础上提交了计算学科教学计划的 Computing Curricula 1991（简称 CC1991）报告。1998 年秋，ACM 与 IEEE-CS 再次联手组成

任务组,经过3年多的努力工作,于2001年12月提交了计算学科教学计划的 Computing Curricula 2001 (简称 CC2001) 报告。该报告将计算学科划分为14个领域,每个领域分成若干知识单元,每个知识单元又包括若干主题。CC2013 又将计算学科扩展到18个领域。CC2001 将这些领域的大多数知识单元规定为计算机及相关学科本科学生必须掌握的核心知识。数据结构主要涉及到程序设计基础 (PF)、算法与复杂性 (AL) 和程序设计语言 (PL) 等领域。掌握这些领域的知识对于利用计算机资源,开发高效的程序系统是必不可少的。本书内容涵盖了上述3个领域中的多个知识单元,主要包括 PF3 (基本数据结构)、PF4 (递归技术)、AL1 (算法分析基本方法)、AL3 (基本计算算法)、PL4 (数据类型)、PL6 (面向对象程序设计) 和 PL9 (数据抽象) 等。

《数据结构》课程是介于数学、计算机硬件和计算机软件之间的一门计算学科核心课程,是程序设计、编译原理、操作系统、数据库系统、人工智能等课程的基础。同时,数据结构技术也广泛应用于信息科学、系统工程、应用数学以及各种工程领域。学习本课程的目的是分析研究数据对象特性,掌握如何表示、存储与操作这些数据的技术,为使用计算机处理各种实际应用问题打下良好基础。随着科学技术发展,将会有更多的课程涉及数据结构内容和应用数据结构技术。本课程应为这些课程提供必要的基础知识与应用技术。

本书共分为9章。第1章介绍什么是数据结构,什么是算法,什么是计算思维,阐明学习数据结构的重要性,讨论算法设计的思想与方法。试图引导读者掌握研究、讨论数据结构的方法,提高读者分析问题与解决问题的能力。第2章至第5章重点讨论线性结构。线性表是应用最广泛的数据结构之一。在线性表中,根据结点的不同存储形式可分为顺序表与链表;根据对运算的不限定条件可分为栈与队列。力求使读者了解线性表的各种组织形式与处理方法,掌握数据结构的各种存储技术和应用技巧。通过非递归算法设计与内排序算法设计,使读者掌握算法设计的技术与方法,提高复杂程序设计能力。同时,介绍几种特殊线性表:串、数组、矩阵和广义表,为读者提供一系列处理线性表的基础理论和有效方法。第6章主要研究重要的非线性结构:树和二叉树。它们能很好地描述结构的分支关系与层次特性。第7章讨论更为复杂的非线性结构:图与网络。第8章重点讨论线性表查找、树型结构查找、哈希表及其查找。第9章介绍了文件与外部排序。每章都开辟了“课程设计相关知识与应用示例”一节,介绍算法设计方法与应用示例,旨在培养读者计算思维能力。这是本书的特色之一。

由于 PASCAL 语言具有丰富的数据类型和良好的结构,所以在数据存储结构表示中,拟选用类 PASCAL 语言和 C 语言对照描述;为了着重体现算法设计的思想与算法结构,对算法的描述拟选用结构化流程图 (N—S 图) 作为工具;为了方便读者上机实践,主要算法都采用 C 语言编程实现。这是本书的又一个特色。

本书条理清楚,内容详实。概念表述严谨,逻辑推理严密,语言精练。既注重数据结构原理介绍,又注重算法设计能力培养;既重视理论知识与科学思想介绍,又重视应用技术与动手能力培养。算法结构清晰,构思精巧。本书内容深入浅出,并且配有大量的实例和图示,每章都有丰富的练习题和课程设计,课程设计题目包括问题描述、基本要求、测试数据、实现提示和问题拓展,旨在提高读者分析问题和解决问题的能力。

本课程在教学计划中至少应为6学分,课堂教学在60~80学时之间。《数据结构》是一门技术性很强的课程,为了使學生能真正掌握有关的理论知识和应用技术,在整个教学

过程中至少应安排 6 个以上课程设计，必须保证学生有足够的课下思考作业时间和上机时间。上机时数、课下思考作业时数和课堂讲授时数的比例应不低于 0.5 : 6 : 1。

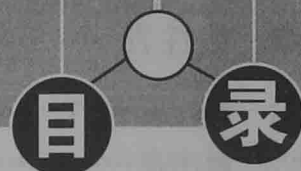
本书作为山西省高等学校省级教学成果奖以及山西财经大学《数据结构》精品课程的组成部分，在教育教学改革研究与课程建设过程中，得到各级领导的支持、专家的指导和同事的帮助，在此表示谢意。

本书凝结了作者 30 多年来的教学科研成果和在讲授《数据结构》等课程中的教学经验。在编写过程中，参考了多种数据结构与算法的优秀教材。在编辑出版过程中，得到了中国人民大学出版社各级领导的支持，负责本书编辑工作的全体同仁付出了辛勤劳动。在此一并表示衷心感谢。

由于作者水平有限，加之学科理论与技术发展日新月异，书中疏漏之处在所难免，恳请广大读者指正。E-mail: fengj1682000@126.com

编 者

2014 年 4 月



第1章 绪论	1
1.1 数据结构	1
1.1.1 学习数据结构的重要性	1
1.1.2 什么是数据结构	3
1.1.3 数据的逻辑结构	5
1.1.4 数据的存储结构	9
1.1.5 数据的运算	11
1.2 算法	12
1.2.1 什么是算法	12
1.2.2 算法的描述	13
1.2.3 算法设计	15
1.2.4 算法设计要求	17
1.2.5 算法分析	18
1.3 程序、程序设计与程序设计方法学	19
1.3.1 程序	19
1.3.2 程序设计	20
1.3.3 程序设计方法学	20
1.4 计算思维与抽象数据类型	22
1.4.1 计算思维	23
1.4.2 计算机求解问题	24
1.4.3 抽象数据类型	25
1.4.4 数据类型、数据结构与抽象数据类型	27

1.5 课程设计相关知识	27
1.5.1 课程设计目的与内涵	28
1.5.2 课程设计步骤	28
1.5.3 课程设计报告规范	29
课后练习与课程设计	30
第2章 线性表	32
2.1 线性表 ADT	32
2.2 顺序表	34
2.2.1 向量的顺序存储表示	34
2.2.2 向量运算	35
2.2.3 应用举例	41
2.2.4 不等长结点的索引表示——目录表	44
2.3 链表	45
2.3.1 指针与指针对象	45
2.3.2 单链表	48
2.3.3 线性表的两种存储比较	60
2.3.4 可利用空间表及其管理	60
2.4 线性表的其他链接存储表示	62
2.4.1 循环链表	62
2.4.2 双向链表	63
2.4.3 静态链表	65
2.5 课程设计相关知识与应用示例	67
2.5.1 算法设计中的基本方法	67
2.5.2 一元多项式相乘	68
课后练习与课程设计	72
第3章 栈和队列	75
3.1 栈 ADT	75
3.2 顺序栈	77
3.2.1 顺序栈的存储表示	77
3.2.2 顺序栈基本运算的实现	77
3.3 链接栈	81
3.3.1 链接栈的存储表示	81
3.3.2 链接栈基本运算的实现	81
3.4 栈的应用举例	82
3.4.1 数制转换	82

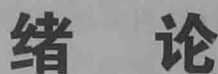
3.4.2 表达式求值	84
3.5 队列 ADT	88
3.6 顺序队列	89
3.6.1 顺序队列的存储表示	89
3.6.2 顺序队列的基本运算的实现	90
3.7 链接队列	92
3.7.1 链接队列的存储表示	92
3.7.2 链接队列基本运算的实现	93
3.8 随机事件模拟	96
3.9 课程设计相关知识与应用示例	99
3.9.1 递归概念	99
3.9.2 递归过程设计	101
3.9.3 递归过程与非递归过程	101
课后练习与课程设计	104
第4章 几种特殊线性表	107
4.1 串	107
4.1.1 串的基本概念	107
4.1.2 串 ADT	109
4.1.3 串的存储表示	110
4.1.4 串运算的实现	112
4.2 数组	113
4.2.1 数组 ADT	113
4.2.2 数组的顺序存储表示	115
4.3 矩阵	116
4.3.1 对称矩阵的压缩存储	116
4.3.2 对角矩阵的压缩存储	117
4.3.3 稀疏矩阵的三元组顺序表存储表示	119
4.3.4 稀疏矩阵的十字链表存储表示	120
4.3.5 稀疏矩阵的转置运算	121
4.3.6 稀疏矩阵的插入运算	122
4.4 广义表	124
4.4.1 广义表的基本概念	124
4.4.2 广义表 ADT	125
4.4.3 广义表的存储表示	126
4.4.4 多元多项式的存储表示	128

4.5 课程设计相关知识与应用示例	129
4.5.1 梵天塔问题	129
4.5.2 再谈递归过程设计	131
4.5.3 求广义表的深度	132
课后练习与课程设计	133
第5章 内排序	137
5.1 排序的基本概念	137
5.2 直接插入排序	138
5.3 直接选择排序	141
5.4 冒泡排序	143
5.5 希尔排序	145
5.6 快速排序	147
5.7 归并排序	152
5.7.1 排序文件的归并	152
5.7.2 2—路归并排序	153
5.8 基数排序	155
5.9 各种排序方法的比较	159
5.10 课程设计相关知识与应用示例	161
5.10.1 自顶向下设计技术与方法	161
5.10.2 逐步求精设计技术与方法	162
5.10.3 积木游戏	162
课后练习与课程设计	165
第6章 树和二叉树	168
6.1 树的基本概念	168
6.2 树的存储结构	171
6.2.1 树的标准形式存储结构	171
6.2.2 树的逆形式存储结构	172
6.2.3 树的扩充标准形式存储结构	173
6.2.4 树的双亲数组存储结构	174
6.2.5 树的孩子链表存储结构	174
6.2.6 树的三重链表存储结构	176
6.3 树的遍历和树的线性表示	177
6.3.1 树的前序遍历	177
6.3.2 树的后序遍历	179
6.3.3 树的层次遍历	180

6.3.4 树的线性表示	182
6.4 二叉树	183
6.4.1 二叉树定义和二叉树 ADT	183
6.4.2 二叉树的性质	185
6.4.3 二叉树与树(或森林)之间的转换	186
6.4.4 二叉树的存储表示	188
6.5 二叉树遍历	192
6.5.1 二叉树遍历	192
6.5.2 由遍历序列构造二叉树	203
6.5.3 基于二叉树遍历操作的算法举例	204
6.6 线索二叉树	206
6.6.1 线索二叉树	206
6.6.2 线索二叉树的利用	208
6.6.3 二叉树的线索化	210
6.7 哈夫曼树及其应用	212
6.7.1 二叉树的路径长度和带权路径长度	212
6.7.2 哈夫曼树和哈夫曼算法	213
6.7.3 哈夫曼树的应用	216
6.8 课程设计相关知识与应用示例	223
6.8.1 结构化方法概述	223
6.8.2 模块化设计技术与方法	223
6.8.3 模块化应用示例	224
课后练习与课程设计	227
第7章 图	230
7.1 图的基本概念	230
7.2 图的存储结构	234
7.2.1 邻接矩阵	234
7.2.2 邻接表	236
7.3 图的遍历	238
7.3.1 深度优先搜索法	239
7.3.2 广度优先搜索法	240
7.3.3 求图的(强)连通分量	242
7.4 生成树与最小生成树	242
7.4.1 生成树概念	242
7.4.2 最小生成树	243
7.5 最短路径	248
7.5.1 单源最短路径	249

7.5.2 每一对顶点间的最短路径	253
7.6 拓扑排序	255
7.6.1 AOV 网	255
7.6.2 拓扑排序	256
7.7 关键路径	258
7.7.1 AOE 网	258
7.7.2 关键路径	258
7.8 课程设计相关知识与应用示例	262
7.8.1 递推算法	262
7.8.2 分治法概述	263
7.8.3 数字旋转方阵	264
课后练习与课程设计	267
第8章 查 找	271
8.1 线性表查找	271
8.1.1 顺序查找	271
8.1.2 折半查找	273
8.1.3 分块查找	275
8.2 树型结构查找	276
8.2.1 二叉排序树	276
8.2.2 平衡查找树	281
8.2.3 B树和B ⁺ 树	288
8.2.4 堆排序	293
8.3 哈希表及其查找	296
8.3.1 哈希表	296
8.3.2 哈希函数构造方法	297
8.3.3 解决冲突方法	299
8.3.4 哈希表的查找	301
8.4 课程设计相关知识与应用示例	304
8.4.1 动态规划法概述	304
8.4.2 贪心法概述	305
8.4.3 背包问题	305
课后练习与课程设计	308
第9章 文件与外部排序	311
9.1 文件概述	311
9.2 文件组织	312
9.2.1 顺序文件	312

9.2.2 索引文件	313
9.2.3 哈希文件	315
9.2.4 多关键字文件	316
9.3 外部排序	318
9.3.1 2—路平衡归并排序法	318
9.3.2 多路平衡归并排序法	320
9.3.3 置换—选择排序	322
9.3.4 最优归并树	326
9.4 课程设计相关知识与应用示例	327
9.4.1 回溯法概述	328
9.4.2 分支限界法概述	329
9.4.3 数据结构课程设计示例	331
课后练习与课程设计	333
参考文献	335



它清楚地揭示了计算机科学中数据结构和算法这两个概念的重要性和统一性。我们不能离开数据结构去抽象地分析求解问题的算法，也不能脱离算法去孤立地研究程序的数据结构。N. 沃思教授还说，不了解施加于数据上的算法，就无法决定如何构造和组织数据；反之，算法的选择常常在很大程度上要依赖于作为基础的数据结构。本章将介绍数据结构、算法和程序等有关概念及其基础知识。

(2) 原材料的加工处理——采用什么样的工艺技术、按照什么样的操作顺序对原材料进行加工处理，最后得到合格产品。

工厂生产模型，如图 1—1 所示。

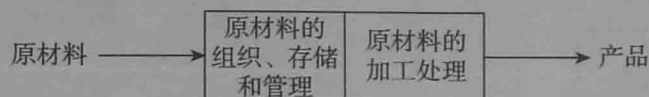


图 1—1 工厂生产模型

由此可见，原材料的组织、存储、管理与原材料的加工处理是工厂进行正常生产的关键。

学习计算机科学的目的是运用计算机来解决实际问题。而计算机的解题过程也可以看作是对原材料进行组织、存储、管理、加工处理，最后得到产品的过程。只不过这里的原材料是数据，产品是处理结果，对数据的加工处理是由算法决定的。

数据是对客观事物采用计算机能够识别、存储和处理的形式所进行的描述。随着计算机的发展和应用领域的扩大，数据量越来越大，数据间的联系也越来越复杂。因此，对数据组织形式的研究也越来越受到重视。

算法是精确定义的一系列规则。它指出怎样从给定的输入数据经过有限步加工处理得到输出信息。输入数据是要处理的数据，输出信息是产生的结果。

计算机解题模型，如图 1—2 所示。

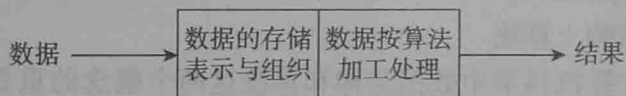


图 1—2 计算机解题模型

由此可见，计算机解题的关键是进行数据的组织和算法的设计。而数据结构就是研究数据组织和算法设计的。

简单地说，数据结构是研究一些数据的集合。就是根据数据的性质、数据元素之间的关系，研究如何表示、存储、操作这些数据的技术。实现数据结构的存储和操作都涉及算法，它们之间存在着本质的联系，失去一方，另一方就没有意义。例如：当讨论一种类型的数据结构时，总离不开讨论对这种类型数据结构需要施加的各种运算，而且只有通过对这些运算的算法进行研究，才能更清楚地理解这种数据结构的意义和作用；反过来，当讨论一个算法时，总是自然地联系到作为该算法的处理对象——数据的组织及其存储表示。因此，在数据结构学习过程中，必须将二者紧密地联系起来。

在计算学科教学计划 2001 (Computing Curricula 2001, CC2001) 的 14 个研究领域，数据结构和算法的基本内容主要涵盖程序设计基础 (Programming Fundamentals, PF)、算法与复杂性 (Algorithms and Complexity, AL) 和程序设计语言 (Programming Language, PL) 3 个领域。CC2001 强调算法和程序设计。由此可见，人们越来越重视数据结构和算法，认为程序设计的实质就是对给定的问题选择一种好的数据结构和设计一个有效率的算法。因此，《数据结构》课程已不仅仅是计算机相关专业教学计划中的核心课程之一，而且是非计算机专业的主要选修课程之一。

《数据结构》课程是计算学科中一门综合性的专业基础课。数据结构的研究不仅涉及