



教育部高等学校电子信息类专业教学指导委员会规划教材  
高等学校电子信息类专业系列教材

计算机科学与技术

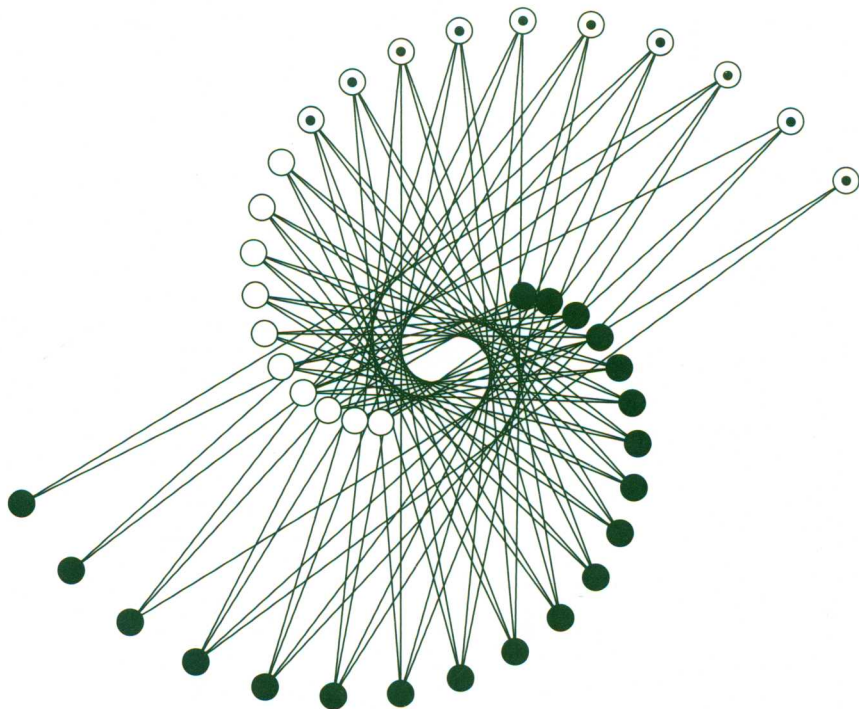
**D**ata Structure  
Description Based on C Programming Language

# 数据结构

——基于C语言的描述

彭波 主编

Peng Bo



清华大学出版社





教育部高等学校电子信息类专业教学指导委员会规划教材  
高等学校电子信息类专业系列教材

Data Structure

Description Based on C Programming Language

# 数据结构

——基于C语言的描述

彭波 主编

Peng Bo

清华大学出版社

北京

## 内 容 简 介

本书系统地介绍数据结构基础理论知识及算法设计,第1~7章从抽象数据类型的角度讨论各种基本类型的数据结构及其应用,主要包括线性表、栈和队列、串、数组和广义表、树和二叉树及图;第8章和第9章主要讨论查找和排序的各种实现方法及其综合比较;第10章介绍不同类型文件的基本操作方法;第11章介绍数据结构课程实验的目的、步骤及内容;附录给出全书习题的参考答案。全书采用类C语言作为数据结构和算法的描述语言,随书配备电子教案,以及第11章实验的源代码。

本书在内容选取上符合人才培养目标的要求及教学规律和认知规律,在组织编排上体现“先理论、后应用、理论与应用相结合”的原则,并兼顾学科的广度和深度,力求适用面广。本书具有结构严谨、层次清楚、概念准确、深入浅出、描述清晰等特点。

本书可作为计算机类专业和信息类相关专业的本科或专科教材,也可供从事计算机工程与应用工作的科技工作者参考。

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:基于C语言的描述/彭波主编. —北京:清华大学出版社,2019

(高等学校电子信息类专业系列教材)

ISBN 978-7-302-53115-9

I. ①数… II. ①彭… 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 ②TP312.8

中国版本图书馆CIP数据核字(2019)第104435号

责任编辑:盛东亮

封面设计:李召霞

责任校对:李建庄

责任印制:宋 林

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](mailto:c-service@tup.tsinghua.edu.cn)

质量反馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:23.5

字 数:569千字

版 次:2019年10月第1版

印 次:2019年10月第1次印刷

定 价:69.00元

产品编号:082634-01

# 高等学校电子信息类专业系列教材

## 一 顾问委员会

谈振辉	北京交通大学 (教指委高级顾问)	郁道银	天津大学 (教指委高级顾问)
廖延彪	清华大学 (特约高级顾问)	胡广书	清华大学 (特约高级顾问)
华成英	清华大学 (国家级教学名师)	于洪珍	中国矿业大学 (国家级教学名师)
彭启琮	电子科技大学 (国家级教学名师)	孙肖子	西安电子科技大学 (国家级教学名师)
邹逢兴	国防科技大学 (国家级教学名师)	严国萍	华中科技大学 (国家级教学名师)

## 一 编审委员会

主任	吕志伟	哈尔滨工业大学		
副主任	刘旭	浙江大学	王志军	北京大学
	隆克平	北京科技大学	葛宝臻	天津大学
	秦石乔	国防科技大学	何伟明	哈尔滨工业大学
	刘向东	浙江大学		
委员	王志华	清华大学	宋梅	北京邮电大学
	韩焱	中北大学	张雪英	太原理工大学
	殷福亮	大连理工大学	赵晓晖	吉林大学
	张朝柱	哈尔滨工程大学	刘兴钊	上海交通大学
	洪伟	东南大学	陈鹤鸣	南京邮电大学
	杨明武	合肥工业大学	袁东风	山东大学
	王忠勇	郑州大学	程文青	华中科技大学
	曾云	湖南大学	李思敏	桂林电子科技大学
	陈前斌	重庆邮电大学	张怀武	电子科技大学
	谢泉	贵州大学	卞树檀	火箭军工程大学
	吴瑛	解放军信息工程大学	刘纯亮	西安交通大学
	金伟其	北京理工大学	毕卫红	燕山大学
	胡秀珍	内蒙古工业大学	付跃刚	长春理工大学
	贾宏志	上海理工大学	顾济华	苏州大学
	李振华	南京理工大学	韩正甫	中国科学技术大学
	李晖	福建师范大学	何兴道	南昌航空大学
	何平安	武汉大学	张新亮	华中科技大学
	郭永彩	重庆大学	曹益平	四川大学
	刘缠牢	西安工业大学	李儒新	中国科学院上海光学精密机械研究所
	赵尚弘	空军工程大学	董友梅	京东方科技集团股份有限公司
	蒋晓瑜	陆军装甲兵学院	蔡毅	中国兵器科学研究院
	仲顺安	北京理工大学	冯其波	北京交通大学
	黄翊东	清华大学	张有光	北京航空航天大学
	李勇朝	西安电子科技大学	江毅	北京理工大学
	章毓晋	清华大学	张伟刚	南开大学
	刘铁根	天津大学	宋峰	南开大学
	王艳芬	中国矿业大学	靳伟	香港理工大学
	苑立波	哈尔滨工程大学		
丛书责任编辑	盛东亮	清华大学出版社		

# 序

## FOREWORD

我国电子信息产业销售收入总规模在 2013 年已经突破 12 万亿元,行业收入占工业总体比重已经超过 9%。电子信息产业在工业经济中的支撑作用凸显,更加促进了信息化和工业化的高层次深度融合。随着移动互联网、云计算、物联网、大数据和石墨烯等新兴产业的爆发式增长,电子信息产业的发展呈现了新的特点,电子信息产业的人才培养面临着新的挑战。

(1) 随着控制、通信、人机交互和网络互联等新兴电子信息技术不断发展,传统工业设备融合了大量最新的电子信息技术,它们一起构成了庞大而复杂的系统,派生出大量新兴的电子信息技术应用需求。这些“系统级”的应用需求,迫切要求具有系统级设计能力的电子信息技术人才。

(2) 电子信息系统设备的功能越来越复杂,系统的集成度越来越高。因此,要求未来的设计者应该具备更扎实的理论基础知识和更宽广的专业视野。未来电子信息系统的设计越来越要求软件和硬件的协同规划、协同设计和协同调试。

(3) 新兴电子信息技术的发展依赖于半导体产业的不断推动,半导体厂商为设计者提供了越来越丰富的生态资源,系统集成厂商的全方位配合又加速了这种生态资源的进一步完善。半导体厂商和系统集成厂商所建立的这种生态系统,为未来的设计者提供了更加便捷却又必须依赖的设计资源。

教育部 2012 年颁布了新版《高等学校本科专业目录》,将电子信息类专业进行了整合,为各高校建立系统化的人才培养体系,培养具有扎实理论基础和宽广专业技能的、兼顾“基础”和“系统”的高层次电子信息人才给出了指引。

传统的电子信息学科专业课程体系呈现“自底向上”的特点,这种课程体系偏重对底层元器件的分析与设计,较少涉及系统级的集成与设计。近年来,国内很多高校对电子信息类专业课程体系进行了大力度的改革,这些改革顺应时代潮流,从系统集成的角度,更加科学合理地构建了课程体系。

为了进一步提高普通高校电子信息类专业教育与教学质量,贯彻落实《国家中长期教育改革和发展规划纲要(2010—2020 年)》和《教育部关于全面提高高等教育质量若干意见》(教高【2012】4 号)的精神,教育部高等学校电子信息类专业教学指导委员会开展了“高等学校电子信息类专业课程体系”的立项研究工作,并于 2014 年 5 月启动了《高等学校电子信息类专业系列教材》(教育部高等学校电子信息类专业教学指导委员会规划教材)的建设工作。其目的是为推进高等教育内涵式发展,提高教学水平,满足高等学校对电子信息类专业人才培养、教学改革与课程改革的需要。

本系列教材定位于高等学校电子信息类专业的专业课程,适用于电子信息类的电子信

息工程、电子科学与技术、通信工程、微电子科学与工程、光电信息科学与工程、信息工程及其相近专业。经过编审委员会与众多高校多次沟通,初步拟定分批次(2014—2017年)建设约100门课程教材。本系列教材将力求在保证基础的前提下,突出技术的先进性和科学的前沿性,体现创新教学和工程实践教学;将重视系统集成思想在教学中的体现,鼓励推陈出新,采用“自顶向下”的方法编写教材;将注重反映优秀的教学改革成果,推广优秀的教学经验与理念。

为了保证本系列教材的科学性、系统性及编写质量,本系列教材设立顾问委员会及编审委员会。顾问委员会由教指委高级顾问、特约高级顾问和国家级教学名师担任,编审委员会由教育部高等学校电子信息类专业教学指导委员会委员和一线教学名师组成。同时,清华大学出版社为本系列教材配置优秀的编辑团队,力求高水准出版。本系列教材的建设,不仅有众多高校教师参与,也有大量知名的电子信息类企业支持。在此,谨向参与本系列教材策划、组织、编写与出版的广大教师、企业代表及出版人员致以诚挚的感谢,并殷切希望本系列教材在我国高等学校电子信息类专业人才培养与课程体系建设中发挥切实的作用。

吕志伟 教授

# 前言

## PREFACE

“数据结构”课程是计算机、电子信息类及相关专业的专业基础。它在整个课程体系中处于承上启下的核心地位：一方面扩展和深化在离散数学、程序设计语言等课程学到的基本技术和方法；另一方面为进一步学习操作系统、编译原理、数据库等专业知识奠定坚实的理论与实践基础。本课程在教给学生数据结构设计和算法设计的同时，培养学生的抽象思维能力、逻辑推理能力和形式化思维方法，增强分析问题、解决问题和总结问题的能力，更重要的是培养专业兴趣，树立创新意识。本书在内容选取上符合人才培养目标的要求及教学规律和认知规律，在组织编排上体现“先理论、后应用、理论与应用相结合”的原则，并兼顾学科的广度和深度，力求适用面广。

全书共分 11 章。第 1 章综述数据、数据结构和抽象数据类型等基本概念及算法描述与分析方法；第 2~7 章主要从抽象数据类型的角度分别讨论线性表、栈和队列、串、数组和广义表、树和二叉树、图等基本类型的数据结构及其应用；第 8 章和第 9 章讨论查找和排序的各种方法，着重从时间性能、应用场合及使用范围方面进行分析和比较；第 10 章主要介绍顺序文件、索引文件、索引顺序文件、哈希文件、多关键字文件的基本操作方法；第 11 章介绍数据结构课程实验的目的、步骤及内容。本书对数据结构众多知识点的来龙去脉做了详细解释和说明；每章后面配有难度各异的适量习题，并在附录中给出习题的参考答案，供读者理解知识及复习提高之用。随书配备电子教案，以及第 11 章实验的源代码。

全书采用类 C 语言描述数据结构和操作算法。类 C 语言是 C 语言的一个精选子集，同时又采用了 C++ 对 C 非面向对象的增强功能，使本书对各种抽象数据类型的定义和与数据结构相关的操作算法的描述更加简明清晰，可读性更好，既不拘泥于 C 语言的细节，又容易转换成能够上机执行的 C 程序或 C++ 程序。

从课程性质上讲，“数据结构”是高等院校计算机科学、电子信息科学及相关专业考试计划中的一门专业基础课；其教学要求是学会分析研究计算机加工的数据结构的特性，以便为应用涉及的数据选择适当的逻辑结构、存储结构及其相应的算法，并初步掌握算法的时空分析技术。从课程学习上讲，“数据结构”的学习是复杂程序设计的训练过程；其教学目的是着眼于原理与应用的结合，在深化理解和灵活掌握教学内容的基础上，学会把知识用于解决实际问题，书写出符合软件工程规范的文件，编写出结构清晰及正确易读的程序代码。可以说，“数据结构”比“高级程序设计语言”等课程有着更高的要求，它更注重培养学生分析抽象数据的能力。

在本书的构思与编写过程中，得到了孙一林、邱李华等多位教授，以及多位研究生的帮助，在此表示感谢。本书可作为计算机类专业和电子信息类相关专业的本科或专科教材，也

可供从事计算机工程与应用工作的科技工作者参考。本书结构严谨、层次清楚、概念准确、深入浅出、通俗易懂、便于自学。

由于编者水平有限,书中不当之处敬请读者提出批评和建议。订购本书作为教材的教师可联系编者获取第 11 章实验的源代码(见清华大学出版社官方网站本书页面)。

编 者

2019 年 8 月

# 目录

## CONTENTS

<b>第 1 章 绪论</b> .....	1
1.1 数据结构的范畴 .....	1
1.1.1 计算机处理问题的分类 .....	1
1.1.2 非数值性问题的求解 .....	2
1.2 数据结构发展的概况 .....	3
1.3 数据结构相关的概念 .....	4
1.3.1 数据的概念 .....	4
1.3.2 结构的概念 .....	5
1.3.3 类型的概念 .....	7
1.4 算法描述与算法分析 .....	9
1.4.1 算法的概念 .....	9
1.4.2 算法描述 .....	11
1.4.3 算法分析 .....	13
习题 .....	17
<b>第 2 章 线性表</b> .....	21
2.1 线性表的类型定义 .....	21
2.1.1 线性表的定义 .....	21
2.1.2 线性表的抽象数据类型 .....	22
2.2 线性表的顺序表示及操作实现 .....	23
2.2.1 顺序表的定义 .....	23
2.2.2 顺序表的操作实现 .....	24
2.3 线性表的链式表示及操作实现 .....	34
2.3.1 单链表的定义 .....	34
2.3.2 单链表的操作实现 .....	35
2.3.3 循环链表 .....	46
2.3.4 双向链表 .....	47
2.3.5 静态链表 .....	50
2.4 线性表两种存储表示的比较 .....	52
2.4.1 基于空间的比较 .....	52
2.4.2 基于时间的比较 .....	52
习题 .....	53
<b>第 3 章 栈和队列</b> .....	55
3.1 栈 .....	55

3.1.1	栈的类型定义 .....	55
3.1.2	栈的存储表示及操作实现 .....	56
3.1.3	栈与递归问题 .....	62
3.2	队列 .....	66
3.2.1	队列的类型定义 .....	66
3.2.2	队列的存储表示及操作实现 .....	68
习题	.....	78
<b>第4章</b>	<b>串</b> .....	<b>80</b>
4.1	串的类型定义 .....	80
4.1.1	串的定义 .....	80
4.1.2	串的抽象数据类型 .....	81
4.2	串的存储表示及操作实现 .....	83
4.2.1	定长顺序存储表示 .....	83
4.2.2	堆分配存储表示 .....	87
4.2.3	串的块链存储表示 .....	91
4.3	串的模式匹配 .....	96
4.3.1	简单的模式匹配方法——BF 算法 .....	96
4.3.2	改进的模式匹配方法——KMP 算法 .....	98
习题	.....	102
<b>第5章</b>	<b>数组和广义表</b> .....	<b>104</b>
5.1	数组 .....	104
5.1.1	数组的类型定义 .....	104
5.1.2	数组的顺序表示及操作实现 .....	106
5.2	矩阵的压缩存储 .....	109
5.2.1	特殊矩阵的压缩存储 .....	109
5.2.2	稀疏矩阵的压缩存储 .....	113
5.3	广义表 .....	122
5.3.1	广义表的类型定义 .....	122
5.3.2	广义表的链式表示及操作实现 .....	125
习题	.....	134
<b>第6章</b>	<b>树和二叉树</b> .....	<b>136</b>
6.1	树 .....	136
6.1.1	树的类型定义 .....	137
6.1.2	树的存储表示及操作实现 .....	142
6.2	二叉树 .....	148
6.2.1	二叉树的类型定义 .....	149
6.2.2	二叉树的重要性质 .....	152
6.2.3	二叉树的存储表示及操作实现 .....	155
6.2.4	线索二叉树 .....	160
6.3	树和森林与二叉树的转换 .....	164
6.3.1	树与二叉树的转换 .....	165
6.3.2	森林与二叉树的转换 .....	167

6.4	哈夫曼树及其应用 .....	169
6.4.1	哈夫曼树 .....	169
6.4.2	哈夫曼编码 .....	174
	习题 .....	178
<b>第7章</b>	<b>图 .....</b>	<b>181</b>
7.1	图的类型定义 .....	181
7.1.1	图的定义 .....	181
7.1.2	图的抽象数据类型 .....	186
7.1.3	图的遍历 .....	187
7.2	图的存储表示与操作实现 .....	188
7.2.1	邻接矩阵 .....	189
7.2.2	邻接表 .....	190
7.2.3	十字链表 .....	192
7.2.4	邻接多重表 .....	193
7.2.5	图的操作实现 .....	194
7.3	图的连通性及其应用 .....	198
7.3.1	无向图的连通分量 .....	198
7.3.2	生成树和生成森林 .....	198
7.3.3	最小生成树 .....	200
7.4	有向无环图及其应用 .....	205
7.4.1	拓扑排序 .....	206
7.4.2	关键路径 .....	209
7.5	最短路径 .....	214
7.5.1	单源最短路径 .....	214
7.5.2	其他最短路径 .....	217
	习题 .....	218
<b>第8章</b>	<b>查找 .....</b>	<b>221</b>
8.1	查找的基本概念 .....	221
8.2	静态查找表 .....	223
8.2.1	静态查找表的类型定义 .....	223
8.2.2	顺序表的查找 .....	223
8.2.3	有序表的查找 .....	224
8.2.4	索引顺序表的查找 .....	227
8.3	动态查找表 .....	229
8.3.1	动态查找表的类型定义 .....	230
8.3.2	二叉排序树和平衡二叉树 .....	230
8.3.3	B <sub>-</sub> 树、B <sup>+</sup> 树和键树 .....	245
8.4	哈希表 .....	255
8.4.1	哈希表的定义 .....	255
8.4.2	哈希函数的构造 .....	257
8.4.3	处理冲突的方法 .....	259
8.4.4	哈希表上的查找 .....	260
	习题 .....	264

<b>第 9 章 排序</b> .....	267
9.1 排序的基本概念 .....	267
9.2 插入排序 .....	269
9.2.1 直接插入排序 .....	269
9.2.2 希尔排序 .....	271
9.3 交换排序 .....	273
9.3.1 冒泡排序 .....	273
9.3.2 快速排序 .....	274
9.4 选择排序 .....	278
9.4.1 简单选择排序 .....	278
9.4.2 堆排序 .....	279
9.5 归并排序 .....	284
9.5.1 2-路归并排序 .....	285
9.5.2 归并排序 .....	286
9.6 基数排序 .....	287
9.6.1 多关键字排序 .....	287
9.6.2 链式基数排序 .....	288
9.7 排序方法比较 .....	292
习题 .....	295
<b>第 10 章 文件</b> .....	298
10.1 文件的基本概念 .....	298
10.2 顺序文件 .....	300
10.2.1 顺序文件的查找 .....	300
10.2.2 顺序文件的修改 .....	301
10.2.3 顺序文件的特点 .....	301
10.3 索引文件 .....	301
10.3.1 索引文件的分类 .....	302
10.3.2 索引文件的存储 .....	302
10.3.3 索引文件的操作 .....	302
10.3.4 利用查找表建立多级索引 .....	303
10.4 索引顺序文件 .....	304
10.4.1 ISAM 文件 .....	304
10.4.2 VSAM 文件 .....	306
10.5 哈希文件 .....	308
10.5.1 哈希文件的操作 .....	309
10.5.2 哈希文件的特点 .....	309
10.6 多关键字文件 .....	309
10.6.1 多重表文件 .....	310
10.6.2 倒排文件 .....	311
10.7 文件综合举例 .....	312
习题 .....	314

第 11 章 课程实验 .....	316
11.1 实验概述 .....	316
11.1.1 教学目的 .....	316
11.1.2 实验步骤 .....	317
11.2 实验内容 .....	318
11.2.1 线性表 .....	318
11.2.2 栈和队列 .....	320
11.2.3 串 .....	321
11.2.4 数组和广义表 .....	323
11.2.5 树和二叉树 .....	325
11.2.6 图 .....	326
11.2.7 查找 .....	328
11.2.8 排序 .....	330
附录 习题参考答案 .....	332

### 主要知识点

- 数据结构的范畴及相关的概念。
- 算法的基本概念及描述方法。
- 算法的时间复杂度及空间复杂度的分析方法。

使用计算机求解任何问题都离不开程序设计,而程序设计的实质就是数据表示和数据处理。数据要能被计算机处理,首先必须能够存储在计算机的内存中,这项任务称为数据表示,数据表示的核心任务是数据结构的设计;一个实际问题的求解必须满足各项处理的要求,这项任务称为数据处理,数据处理的核心任务是算法设计。因此,数据结构主要讨论数据表示和数据处理的基本问题。

## 1.1 数据结构的范畴

数据结构起源于程序设计。随着计算机应用领域的扩大和软件与硬件技术的飞速发展,计算机的应用已远远超出了科学和工程计算的范围,并广泛地应用于情报检索、信息管理、系统工程,乃至人类社会活动的一切领域。与此同时,计算机的处理对象也从简单的纯数值性数据发展到非数值性和具有一定结构的数据,例如文本、图形、图像、音频、视频及动画等,处理的数据量也越来越大,这就给程序设计带来一个问题:应该如何组织待处理的数据以及数据之间的关系(结构)。

### 1.1.1 计算机处理问题的分类

计算机处理的问题可以分为数值性问题和非数值性问题。

#### 1. 数值性问题

众所周知,20世纪40年代,电子计算机问世的直接原因是解决弹道计算问题。早期的电子计算机的应用范围只限于科学和工程计算,处理对象是纯数值性数据,通常人们把这类问题称为数值性问题。例如,线性方程求解问题涉及的运算对象是简单的整型、实型或布尔型数据。程序设计者的主要精力集中于程序设计的技巧,不需要重视数据结构。

#### 2. 非数值性问题

根据统计,当今处理非数值性问题占用了90%以上的机器时间,这类问题涉及的数据结构更为复杂,数据元素之间的相互关系一般无法用数学方程式描述。因此,解决此类问题

的关键已经不再是数学分析和计算方法,而是需要建立问题的数学模型和设计相应的算法,才能有效地解决问题。

### 1.1.2 非数值性问题的求解

1968年,美国唐纳德·克努特(Donald E. Knuth)教授开创了数据结构的最初体系,他所著的《计算机程序设计艺术(第I卷)·基本算法》是第一本比较系统地阐述数据逻辑结构和存储结构及其操作的著作。1976年,瑞士计算机科学家尼古拉斯·沃斯(Niklaus Wirth)教授(图灵奖获得者)提出:数据结构+算法=程序。斗转星移至今,尽管新的技术方法不断涌现,这句名言依然焕发着无限的生命力,它借助面向对象知识的普及,使数据结构技术更加完善和易于使用。由此,也说明了数据结构在计算机学科中的地位和不可替代的独特作用。

**例 1-1** 学生情况管理问题。

用计算机来完成学生管理就是用计算机程序来处理学生情况登记表,实现增加、删除、修好、查找等功能。图 1-1 是一张简单学生情况登记表。在学籍管理问题中,计算机的操作对象是每个学生的情况信息(档案表项),各档案表项之间的关系可以用称为线性表的数学模型来描述。在该数学模型中,计算机处理的数据之间存在的是“一个对一个”的线性关系,这类数学模型可称为线性的数据结构。

姓名	性别	年龄	籍贯	班别	成绩			
					数学	物理	化学	外语
孙臣	男	18	北京	6003	95	90	92	96
钱晓	女	20	上海	6002	90	95	85	80
常依	女	19	长沙	6004	85	90	80	75
吴伟	男	19	湖北	6001	85	80	75	90
...	...	...	...	...	...	...	...	...

图 1-1 学生情况登记表

**例 1-2** 人机对弈问题。

计算机之所以能够和人对弈,是因为对弈的策略实现已经存入计算机内。在对弈问题中,计算机的操作对象是对弈过程中可能出现的棋盘状态(格局),而格局之间的关系是由对弈规则决定的。因为从一个格局可以派生出多个格局,所以这种关系通常不是线性的。如图 1-2(a)所示为井字棋(又称三子连珠,由两个人对弈,棋盘为  $3 \times 3$  的方格,当一方的三个棋子占同一行、同一列或同一对角线时便为胜方)的一个格局,从该格局出发可以派生出五个新格局,从新格局出发,还可以再派生出新格局,如图 1-2(b)所示。格局之间的关系可以用称为树的数学模型来描述。在该数学模型中,计算机处理的数据之间存在“一个对多个”的层次关系,这类数学模型可称为树的数据结构。

**例 1-3** 城市最小造价通信网问题。

用计算机求解城市最小造价通信网问题就是由计算机程序在已知某些城市之间直接的通信线路预算造价的情况下,求出  $n$  个城市中任意两个城市之间直接或间接的通信线路,使网络造价最低。图 1-3(a)所示为 7 个城市预算造价通信网,图 1-3(b)所示为 7 个城市最小

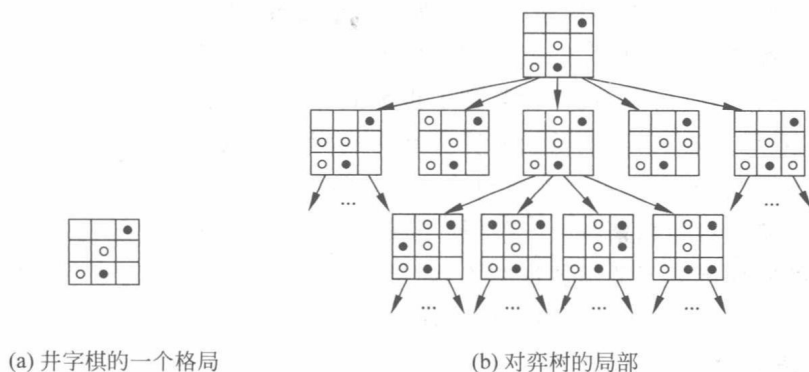


图 1-2 对弈问题中格局之间的关系

造价通信网。在城市最小造价通信网问题中,如果用圆圈表示一个城市,两个圆圈之间的连线表示对应城市之间的通信线路,连线上的数值表示该通信线路的造价,则城市之间的通信关系可以用称为图的数学模型来描述。在该数学模型中,计算机处理的数据之间存在的是“多个对多个”的任意关系,这类数学模型可称为图的数据结构。

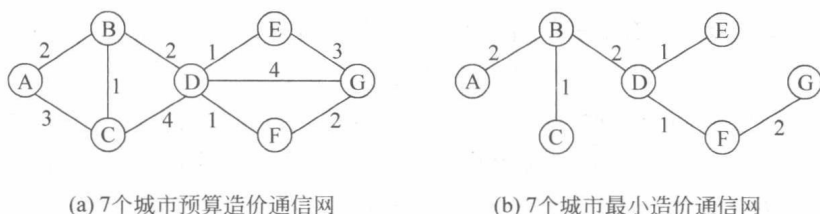


图 1-3 通信网问题中城市之间的关系

由上述三个例子可以看出,描述这些非数值问题的模型已经不再是数学方程,而是线性表、树、图等的数据结构。在抽象出问题的模型之后,数据结构的任务还包括对这个模型进行求解。因此,简单说来,数据结构是一门研究非数值性计算的程序设计问题中计算机的操作对象以及它们之间的关系和操作等的学科。

## 1.2 数据结构发展的概况

数据结构随着程序设计的发展而发展。程序设计经历了三个阶段:无结构阶段、结构化阶段和面向对象阶段。相应地,数据结构的发展也经历了三个阶段。

### 1. 无结构阶段

20世纪40年代至60年代,计算机的应用主要是针对科学计算,程序设计技术以机器语言及汇编语言为主,程序处理的数据是纯粹的数值,数据之间的关系主要是数学公式或数学模型。在这一阶段,人类的自然语言与计算机编程语言之间存在着巨大的鸿沟,程序设计属于面向计算机的程序设计,设计人员关注的重心是使程序尽可能被计算机接受并按指令正确执行,至于程序能否让人理解并不重要。

### 2. 结构化阶段

20世纪60年代至80年代,计算机开始广泛应用于非数值处理领域,数据表示成为程

程序设计的重要问题,人们认识到程序设计规范化的重要性,提出了程序结构模块化,并开始注意数据表示与操作的结构化。数据结构及抽象数据类型就是在这种背景下形成的。数据结构概念的引入对程序设计的规范化起到了重大的作用。从沃斯提出的著名公式“数据结构+算法=程序”可以看到,数据结构和算法是构成程序的两个重要组成部分,一个软件系统通常是以一个或几个关键数据结构为核心而组成的。

随着软件系统的规模越来越大、复杂性不断增加,人们不得不对结构化技术进行重新评价。软件系统的实现依赖于关键数据结构,如果这些关键数据结构的一个或几个有所改变,则会涉及整个系统,甚至导致整个系统彻底崩溃。

### 3. 面向对象阶段

面向对象技术(首先是面向对象程序设计)开始于20世纪80年代初,是目前最流行的程序设计技术。在面向对象技术中,问题世界的相关实体视为一个对象,对象由属性和方法构成,属性用于描述实体的状态或特征,方法用于改变实体的状态或描述实体的行为。一组具有相同属性和方法的对象的集合抽象为类,而每个具体的对象都是类的一个实例。例如,“学生”是一个类,“张三”“李四”等对象都是“学生”类的实例。

由于对象(类)将密切相关的属性(数据)和方法(操作)定义为一个整体,从而实现了封装和信息隐藏。使用类时,不需要了解其内部的实现细节,如果数据(结构)修改了,只需要修改类内部的局部代码,而软件系统的其余部分不需要修改。

数据结构主要强调两个方面的内容:一个是数据之间的关系;另一个是针对这些关系的基本操作。这两个方面实际上蕴含着面向对象的思想:类重点描述实体的状态与行为,而数据结构重点描述数据之间的关系及其基本操作,数据及其相互关系构成了对实体状态的描述,针对数据元素之间关系的操作构成了对实体行为的描述。由此可见,类与数据结构之间具有对应关系,如图1-4所示。

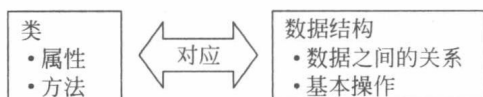


图 1-4 类和数据结构之间的对应关系

值得注意的是,数据结构的发展并未终结。其一,数据结构将继续随着程序设计的发展而发展;其二,面向各个专门领域中特殊问题的数据结构得到研究和发展,例如多维图形数据结构等,各种实用的高级数据结构被研究出来,各种空间数据结构也在探索中;其三,从抽象数据类型的观点来讨论数据结构已经成为一种必然的趋势。

## 1.3 数据结构相关的概念

在本节中将先对以后各章节中反复出现的数据结构相关概念赋以确定的含义,以便在后面的学习中与读者取得“共同的语言”。

### 1.3.1 数据的概念

#### 1. 数据

数据(Data)是信息的载体,在计算机科学中指所有能输入到计算机中并能被计算机程序识别、存储和处理的符号集合。数据是计算机程序加工的“原料”。

例如,一个利用数值分析方法求解代数方程的程序,其处理对象是整数和实数;一个编