

教育部人才培养模式改革和开放教育试点教材

数据结构

殷人昆 编著



清华大学出版社

<http://www.tup.tsinghua.edu.cn>

数据结构

殷人昆 编著

清华大学出版社

(京)新登字 158 号

内 容 简 介

“数据结构”是计算机专业的核心课程,是从事计算机软件开发和应用的人员必备的专业基础。随着计算机的日益普及,“数据结构”课程也在不断发展。

本书按照“数据结构”课程教学大纲的要求,从面向对象的概念、对象类设计的风格和数据结构的层次开始。从线性结构到非线性结构,从简单到复杂,循序渐进,逐步深入地讨论了各种数据结构的内在的逻辑联系及它们在计算机中的实现方式和使用。此外,对常用的迭代、递推、递归、回溯等算法设计技术、搜索和排序算法等,都做了详尽的描述,并引入了简单的算法分析。

全书采用了面向对象的观点讨论数据结构技术,并以兼有面向过程和面向对象双重特色的 C++ 语言作为算法和数据结构的描述工具,强化基本知识与基本技能的双基训练。全书条理清晰,通俗易懂,图文并茂,适于自学。

本书适用于计算机专业的本科学生使用,也可作为教师和相关科研人员的参考书。与本书配套的《数据结构例题分析及习题解答》一书将由清华大学出版社出版。此外,如果有需要用 PowerPoint 制作的教学幻灯片的教师,可与清华大学出版社联系。

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。

书 名: 数据结构

作 者: 殷人昆 编著

出版者: 清华大学出版社(北京清华大学学研大厦,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 北京市人民文学印刷厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 印张: 22 字数: 506 千字

版 次: 2001 年 3 月第 1 版 2001 年 3 月第 2 次印刷

书 号: ISBN 7-302-04271-3/TP·2511

印 数: 8001~16000

定 价: 27.60 元

前 言

自 1978 年美籍华裔学者冀中田教授在国内首开“数据结构”课程以来,经过 20 余年的发展,这门课程已成为各大学计算机专业本科的主干课程,也成为非计算机类学生和研究生学习计算机的必修课程。

“数据结构”课程脱胎于“离散数学结构”,它涉及各种离散结构(如表、向量、集合、树、图、代数方程和多项式等)在计算机上如何存储和处理,这是从事计算机硬、软件科研与开发的人们必须了解的。特别是引入抽象数据类型的概念和面向对象模式,更能够适应软件开发人员的需要。

从 1998 年开始,中央广播电视大学与清华大学合作,开设计算机科学与技术专业本科生课程。依据教学大纲的要求,我们编写了这本《数据结构》教材。教材的编写原则体现基本知识的学习与基本技能的培养并重,采用支持传统的面向过程的开发模式和先进的面向对象开发模式的 C++ 语言作为描述语言,有利于学生在编程方式和习惯方面的转换。教材中给出的算法不再用伪码,直接用可执行的 C++ 代码描述,读者直接输入计算机,简单整理即可运行。

学好“数据结构”课程,将为后续的“数据库系统”、“操作系统”、“编译原理”等专业课程,打下良好的知识基础,而且还为软件开发和程序设计提供了必要的技能训练。

本课程要求学员对计算机的组成有基本的了解,对程序设计方法有一定基础,对 C 语言编程,最好是 C++ 语言编程有一定的经验。对涉及离散数学的知识,如表、树、图和集合等,有初步的了解。事实上,教材中将会简单地补充这些方面的内容。

本教材内容丰富,条理清晰,叙述易懂,算法简洁,可读性强。对于重点或要点,在书中用黑体字加以强调。如果讲课时间不够,可根据时间和能力,适当对全书需要学习的内容加以剪裁。本着少讲多练的原则,可以对每种结构,只介绍类定义和关键操作的实现,其它内容可以让同学自学。通过上机练习,加深理解。

与本教材配套的用 PowerPoint 制作的教学幻灯片已经完成并投入使用,如有需要者可与清华大学出版社联系。在教学过程中习题的安排在“附录 B 教学进度与习题安排参考”中写明。习题的参考答案在与本教材配套的《数据结构实验及习题解答》(清华大学出版社)一书中给出。

在本教材的成书过程中,得到清华大学计算机系和中央广播电视大学有关领导的关心,以及清华大学出版社的支持。中央电大徐孝凯副教授参与了全书的编写过程,对教材内容、教材体系安排、习题选择等都提出了宝贵的意见。清华大学计算机系聂宇、王映和魏小亮同学对书中的许多程序做了调试。此外,在教材的大纲和考核要求的制订过程中,还得到了北京航空航天大学计算机系唐发根教授、石油化工大学周树杰教授的帮助,最后

由北京大学计算机系孙家骕教授,北京理工大学计算机系江涛教授,北京工商大学计算机系刘杰副教授对本书进行了终审,在此表示衷心的感谢。

由于作者的水平和经验所限,教材中还可能存在着错误和不足,恳请读者给予批评指正。

作 者

2000年11月于北京

序

我们正处在跨越世纪的门槛上,人类社会在一股股变革性力量的推动下发生着根本性的变化。知识经济时代的到来向我们显示,一个国家最重要的资源已经不再是土地、劳动力或资本,而是其国民的知识和创造力;国与国的竞争虽然常常表现为政治、经济或军事事实的较量,但归根到底已是一场教育和科技的竞争。换言之,国家的综合实力将主要由其国民的教育水平来决定。一时间,世界各国的校长们、跨国企业的巨头们乃至许多的政府首脑们都在纷纷议论 21 世纪的教育,以迎接知识经济的挑战。我们中华民族有着几千年的文明,为在世界民族之林重振雄风,再展辉煌,发出了时代的特强音:实施科教兴国,提高全民素质。从中央领导到广大群众,都对教育提出了更高的要求,寄予了更大的希望,同时也给予了更多的支持。人们在这方面的思想观念和实践探索正在以空前的速度发展着。

中国的高等教育已经走完了一个世纪的路程。已经过去的 20 世纪正是它从无到有、从小到大、由产生到发展的一段百年历史。中国人民在短短的数十年时间里构筑了资本主义国家好几百年才形成的高等教育体系,涌现出一批高水平的学校,培养了一大批高层次优秀人才,取得了辉煌的成就。但是在新时期,教育不适应现代化建设需要的矛盾不断显露,我国劳动者受教育水平普遍较低的现象无法面对新世纪的机遇和挑战,我国高等教育的发展现状也难以满足广大人民群众空前强烈的受教育愿望。一代伟人邓小平早在 10 年前就一针见血地指出,我们的最大失误是教育,一是放松了对青少年的思想道德教育,二是教育规模发展不够快。现在看来,这两个问题依然是症结所在。一个 12 亿人口的泱泱大国,高等学校的毛入学率仅 10% 左右,实在很不相称。我国的高等教育已经面临着大力发展、高速发展、从根本上改变落后状态的紧迫问题。

令人欣慰和鼓舞的是中国有一所全世界最大的大学——中国广播电视大学,上百万的学生遍布在 960 万 km^2 的辽阔土地上。它突破传统教育在空间上的限制,不断减弱时间上的束缚,以覆盖面广、全方位为各类社会成员提供教育服务的优势,成为中国高等教育体系中的一个重要组成部分。二十多年来,它为实现高等教育大众化,为提高我国劳动者的整体素质,为变巨大的人口包袱为巨大的人力资源,以形成浩浩荡荡的高水平建设大军,发挥了不可磨灭的作用。最近,中央电大又有重大改革举措,进一步面向社会开展了“开放教育”等项试点工作,在教育思想、招生对象、培养模式、管理机制方面进行新的探索。尤其引人注目的是中央电大与国内的一些重点高校形成了紧密的合作关系,携手为我国现代远程教育开拓新路。重点高校有学科和教学上的优势,它们的加盟有利于电大提高教学质量、办出特色;而中央电大有很丰富的教育资源,有完整的办学系统,有一支富有经验的教学与管理队伍,特别是有较强的社会服务意识和人才市场意识,这对于需要进一步向社会开放的普通高校而言,又有许多值得学习和借鉴之处。我们完全有理由相信,中央电大和重点高校的结合,不仅可以在现阶段实现优势互补、资源共享,而且有可能成

长出一种符合我国国情发展教育的最具潜力的新型教育模式。

现在摆在我们面前的这套中央广播电视大学本科(专科起点)“计算机科学与技术”专业教材,就是中央电大和清华大学合作的产物。在开放教育试点启动之际,在计算机及其网络技术日新月异、其爆炸式发展和神话般应用使人们眼花缭乱、不知所措之时,在我国至少缺乏数十万计算机软件及网络技术人才的当口,这套教材像雪里送炭,像清风送爽,终于在人们的企盼和惊喜中问世了。它确实及时和解渴。教材的编者是清华大学计算机系一批学术水平高、教学经验丰富的教授,他们以知识、能力和素质的全面训练为目标,将教材的先进性、实用性和可读性融为一体。教材纲目清楚,重点突出,深入浅出,便于自学。书中每章有小结,章章有习题,有的还配有实验指导和习题解答,不仅对计算机专业学生适用,其他专业的学生也可以此入门。清华大学的老师们还准备为这套教材制作多媒体导读光盘和网络辅导教材,指明教学基本要求,区分应该熟练掌握和只需一般了解的内容,并进行重点难点分析和讲解。这全套的教材称得上是难得的好书。

对于中国广播电视大学我是颇有感情的,不只是因为它过去的功绩和带给人们未来的曙光,还因为我本人二十年前也曾参与过中央电大“电子技术基础”课程的教学工作。那时我收到许多电大学生热情洋溢的来信,强烈感受到他们对知识与教育的渴求,感受到他们学习的艰辛和坚韧不拔的毅力,同时也感受到了广大学生对我的信任和鼓励。当年的电大学生如今多数已成为我国经济建设和社会发展中的骨干,一些人后来获得了博士学位,有的已成为我国重点大学的教授。中央电大的成功实践已在社会上赢得了很好的声誉,而当前扩大教育规模、构建终身学习体系的社会呼唤又给电大今后的发展提供了新的难得的机遇。近年来,信息网络与多媒体技术突飞猛进,也使电大的远程教育形式跃上了现代化的新台阶。这次中央电大和清华大学合作,共同在计算机专业开放教育改革试点中付出了辛勤的劳动,播下了希望的种子。我期待着中央电大有更多的创新,更大的发展,更加充满活力。我也殷切希望电大的学生们为中华民族的强盛而自强不息,学有所成。

努力吧,中国广播电视大学一定能成为中国教育界一颗璀璨的明珠。

清华大学副校长、教授 胡东成

2000年8月于北京

目 录

前言	(I)
第 1 章 绪论	(1)
本章要点	(1)
1.1 数据结构的概念及分类	(1)
1.1.1 数据与数据结构	(1)
1.1.2 数据结构的分类	(2)
1.2 抽象数据类型及面向对象概念	(4)
1.2.1 数据类型	(4)
1.2.2 数据抽象与抽象数据类型	(5)
1.2.3 面向对象的概念	(6)
1.2.4 用于描述数据结构的语言	(7)
1.3 算法定义	(7)
1.4 算法性能分析与度量	(9)
1.4.1 算法的性能标准	(9)
1.4.2 算法的后期测试	(10)
1.4.3 算法的事前估计	(11)
1.4.4 渐进的时间复杂度	(16)
1.4.5 渐进的空间复杂度	(19)
小结	(19)
习题	(20)
第 2 章 数组	(23)
本章要点	(23)
2.1 作为抽象数据类型的数组	(23)
2.1.1 在 C++ 中数组的定义和初始化	(23)
2.1.2 作为抽象数据类型的数组	(24)
2.1.3 数组的顺序存储方式	(26)
2.2 顺序表	(28)
2.2.1 线性表的概念	(28)
2.2.2 顺序表的定义和特点	(29)
2.2.3 顺序表的搜索、插入和删除	(32)
2.2.4 作为抽象数据类型,使用顺序表的事例	(33)

2.3	稀疏矩阵	(34)
2.3.1	稀疏矩阵的抽象数据类型	(34)
2.3.2	稀疏矩阵的压缩表示	(35)
2.4	字符串	(36)
2.4.1	字符串抽象数据类型和类定义	(36)
2.4.2	字符串操作的实现	(37)
	小结	(40)
	习题	(41)
第3章	链表	(45)
	本章要点	(45)
3.1	单链表	(45)
3.1.1	单链表的概念	(45)
3.1.2	单链表的类定义	(46)
3.1.3	单链表中的插入与删除	(47)
3.1.4	带表头结点的单链表	(49)
3.1.5	单链表的模板类	(50)
3.1.6	静态链表	(54)
3.2	循环链表	(55)
3.3	多项式及其相加	(56)
3.3.1	多项式抽象数据类型与 * this 指针	(56)
3.3.2	多项式的表示	(57)
3.3.3	多项式的加法	(59)
3.4	双向链表	(61)
3.4.1	双向链表的概念	(61)
3.4.2	带表头结点的双向循环链表	(62)
3.4.3	双向循环链表的搜索、插入和删除算法	(63)
3.5	稀疏矩阵	(65)
	小结	(67)
	习题	(68)
第4章	栈和队列	(71)
	本章要点	(71)
4.1	栈	(71)
4.1.1	栈的定义	(71)
4.1.2	顺序栈——栈的数组存储表示	(71)
4.1.3	链式栈——栈的链接存储表示	(74)
4.2	表达式的计算	(76)

4.2.1	表达式	(76)
4.2.2	应用后缀表示计算表达式的值	(77)
4.2.3	中缀表示与其他表示之间的转换	(80)
4.3	队列	(83)
4.3.1	队列的定义	(83)
4.3.2	循环队列——队列的顺序存储表示	(83)
4.3.3	链式队列——队列的链接存储表示	(86)
4.3.4	队列的应用举例——打印二项展开式 $(a+b)^i$ 的系数	(87)
4.4	优先级队列	(89)
4.4.1	优先级队列的定义	(89)
4.4.2	优先级队列的存储表示和实现	(90)
	小结	(91)
	习题	(92)
第5章	递归	(94)
	本章要点	(94)
5.1	递归的概念	(94)
5.2	递归过程与递归工作栈	(98)
5.2.1	递归工作栈	(98)
5.2.2	用栈实现递归过程的非递归算法	(99)
5.2.3	用迭代法实现递归过程	(101)
5.3	用回溯法求解迷宫问题	(103)
5.4	广义表	(107)
5.4.1	广义表的概念	(108)
5.4.2	广义表的表示及操作	(109)
5.4.3	广义表存储结构的实现	(110)
5.4.4	广义表的递归算法	(114)
	小结	(121)
	习题	(121)
第6章	树与森林	(124)
	本章要点	(124)
6.1	树和森林的概念	(124)
6.1.1	树的定义	(124)
6.1.2	树的术语	(125)
6.1.3	树的抽象数据类型	(126)
6.2	二叉树	(126)
6.2.1	二叉树的定义	(126)

6.2.2	二叉树的性质	(127)
6.2.3	二叉树的抽象数据类型	(128)
6.2.4	二叉树的表示	(129)
6.3	遍历二叉树	(134)
6.3.1	遍历二叉树的递归算法	(135)
6.3.2	应用遍历二叉树的事例	(136)
6.3.3	遍历二叉树的非递归算法	(138)
6.3.4	二叉树的计数	(141)
6.4	线索化二叉树	(144)
6.4.1	线索	(144)
6.4.2	中序线索化二叉树	(144)
6.5	堆	(150)
6.5.1	堆的定义	(150)
6.5.2	堆的建立	(151)
6.5.3	堆的插入与删除	(153)
6.6	树与森林	(154)
6.6.1	树的存储表示	(155)
6.6.2	森林与二叉树的转换	(157)
6.6.3	树的遍历	(158)
6.6.4	森林的遍历	(160)
6.7	霍夫曼树	(161)
6.7.1	路径长度	(161)
6.7.2	霍夫曼树	(162)
6.7.3	霍夫曼编码	(164)
	小结	(165)
	习题	(166)
第7章	集合与搜索	(169)
	本章要点	(169)
7.1	集合及其表示	(169)
7.1.1	集合基本概念	(169)
7.1.2	用位向量实现集合抽象数据类型	(169)
7.1.3	用有序链表实现集合的抽象数据类型	(172)
7.1.4	并查集	(177)
7.2	静态搜索表	(180)
7.2.1	搜索的概念	(180)
7.2.2	静态搜索结构	(181)
7.2.3	顺序搜索	(183)

7.2.4	基于有序顺序表的折半搜索	(185)
7.3	二叉搜索树	(188)
7.3.1	定义	(188)
7.3.2	二叉搜索树上的搜索	(189)
7.3.3	二叉搜索树的插入	(191)
7.3.4	二叉搜索树的删除	(192)
7.3.5	二叉搜索树的搜索效率	(194)
7.4	AVL 树	(196)
7.4.1	AVL 树的定义	(196)
7.4.2	平衡化旋转	(197)
7.4.3	AVL 树的插入和删除	(200)
7.4.4	AVL 树的高度	(203)
小结		(203)
习题		(204)
第 8 章	图	(207)
本章要点		(207)
8.1	图的基本概念	(207)
8.1.1	图的基本概念	(207)
8.1.2	图的抽象数据类型	(210)
8.2	图的存储表示	(210)
8.2.1	邻接矩阵	(211)
8.2.2	邻接表	(213)
8.2.3	邻接多重表	(217)
8.3	图的遍历与连通性	(219)
8.3.1	深度优先搜索	(220)
8.3.2	广度优先搜索	(221)
8.3.3	连通分量	(222)
8.3.4	重连通分量	(224)
8.3.5	图的遍历举例:欧拉回路问题	(225)
8.4	最小生成树	(227)
8.4.1	克鲁斯卡尔算法	(228)
8.4.2	普里姆算法	(230)
8.5	单源最短路径问题	(232)
8.6	活动网络(activity network)	(235)
8.6.1	用顶点表示活动的网络	(235)
8.6.2	用边表示活动的网络	(239)
小结		(243)

习题	(244)
第9章 排序	(247)
本章要点	(247)
9.1 概述	(247)
9.2 插入排序	(249)
9.2.1 直接插入排序	(249)
9.2.2 折半插入排序	(251)
9.2.3 链表插入排序	(251)
9.2.4 希尔排序	(253)
9.3 交换排序	(255)
9.3.1 起泡排序	(255)
9.3.2 快速排序	(256)
9.4 选择排序	(259)
9.4.1 直接选择排序	(260)
9.4.2 锦标赛排序	(261)
9.4.3 堆排序	(261)
9.5 归并排序	(264)
9.5.1 归并	(264)
9.5.2 迭代的归并排序算法	(265)
9.5.3 递归的链表归并排序	(267)
9.6 基数排序	(268)
9.6.1 多排序码排序	(269)
9.6.2 链式基数排序	(270)
9.7 外排序	(272)
9.7.1 外排序的基本过程	(272)
9.7.2 k 路平衡归并与败者树	(274)
9.7.3 初始归并段的生成	(277)
9.7.4 最佳归并树	(279)
小结	(281)
习题	(282)
第10章 索引与散列	(285)
本章要点	(285)
10.1 静态索引结构	(285)
10.1.1 线性索引	(285)
10.1.2 倒排表	(287)
10.1.3 m 路静态搜索树	(288)

10.2	动态索引结构	(289)
10.2.1	动态的 m 路搜索树	(289)
10.2.2	B 树	(291)
10.2.3	B 树的插入	(293)
10.2.4	B 树的删除	(295)
10.2.5	B+ 树	(297)
10.3	散列	(300)
10.3.1	词典的抽象数据类型	(300)
10.3.2	散列表与散列方法	(301)
10.3.3	散列函数	(302)
10.3.4	处理冲突的闭散列方法	(305)
10.3.5	处理冲突的开散列方法——链地址法	(310)
10.3.6	散列表分析	(312)
10.4	可扩充散列	(313)
10.4.1	二叉 Trie 树	(313)
10.4.2	将二叉 Trie 树转换为目录表	(314)
10.4.3	目录表扩充与收缩	(316)
10.4.4	性能分析	(317)
	小结	(318)
	习题	(318)
附录 A	用 C++ 描述面向对象程序	(322)
A.1	用模板定义 C++ 中的类	(322)
A.2	类中成员函数的实现	(323)
A.3	函数名重载和操作符重载	(327)
A.4	C++ 中的主函数	(328)
附录 B	教学进度与习题安排参考	(329)
附录 C	词汇索引	(330)
参考文献		(337)

第 1 章 绪 论

本章要点

(1) 理解什么是数据、数据对象、数据元素、数据结构、数据的逻辑结构与物理结构和逻辑结构与物理结构间的关系。

(2) 理解什么是数据类型、抽象数据类型、数据抽象和信息隐蔽原则。了解什么是面向对象。

(3) 理解算法的定义、算法的特性、算法的时间代价和算法的空间代价。

(4) 熟悉用 C++ 语言描述算法的方法,能够使用 C++ 语言编写程序。

1.1 数据结构的概念及分类

1.1.1 数据与数据结构

人们在日常生活中会遇到各种信息,如用语言交流思想,银行与商店的商业交易、战争中用于传递命令的旗语等。这些信息必须转换成数据才能在计算机中进行处理。因此,数据定义是:数据(data)是信息的载体,是描述客观事物的数、字符,以及所有能输入到计算机中并被计算机程序识别和处理的符号的集合。数据大致可分为两类:一类是数值性数据,包括整数、浮点数、复数、双精度数等,主要用于工程和科学计算,以及商业事务处理;另一类是非数值数据,主要包括字符和字符串,以及文字、图形、图像、语音等的

数据。

从传统的观点来看,在解决应用问题时,总把数据按其性质归类到一些称之为数据对象(data object)的集合中。在数据对象中所有数据成员,即数据元素,都具有相同的性质,它们是数据的子集。例如,整数数据对象可以是集合 $N = \{0, \pm 1, \pm 2, \pm 3, \dots\}$ 。英文字母数据对象可以是集合 $LETTER = \{'A', 'B', \dots, 'Z'\}$ 。

学校中的学生是更复杂的数据对象,它的每一个数据元素就是一个学生记录,主要包括学生的学号、姓名、性别、籍贯、出生年月、成绩等数据成员或数据项,以表明学生在某一方面的属性。这些数据项可以分为两种:一种叫做初等项,如学生的性别、籍贯等,这些数据项是在数据处理时不能再分割的最小单位;另一种叫做组合项,如学生的成绩,它可以再划分为物理、化学等课程更小的项。通常,在解决实际应用问题时是把每个学生记录当作一个基本单位进行访问和处理的。

在学生选课系统中,可能以一个班级的学生记录作为学生数据对象,也可能以一个年级的学生记录作为学生数据对象。所以,如何选择数据对象将依据要求而定。

一个数据对象中所有数据成员之间一定存在某种关系。例如,招生考试时把所有考

生按考试成绩从高到低排队,这样,所有考生记录都处在一种有序的序列中。又例如,在 n 个网站之间建立通信网络,要求以最小的代价将 n 个网站连通,如图 1.1(a)所示。这样,在所有网站之间形成一种树形关系;如果要求网络中任一网站出现故障,整个网络仍然保持畅通,则在所有网站之间形成一种网状关系,如图 1.1(b)所示。若综合考虑数据对象及其所有数据成员之间的关系,就可得到数据结构的定义:

数据结构由某一数据对象及该对象中所有数据成员之间的关系组成。记为:

$$\text{Data_Structure} = \{ D, R \}$$

其中, D 是某一数据对象, R 是该对象中所有数据成员之间的关系的有限集合。

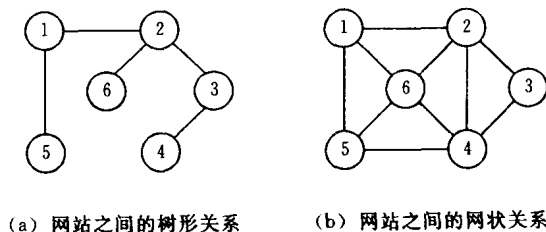


图 1.1 n 个网站之间的连通关系

1.1.2 数据结构的分类

依据数据成员之间关系的不同,数据结构分为两大类:线性结构和非线性结构。

线性结构也称为线性表,在这种结构中所有数据成员(也称为数据元素)都按某种次序排列在一个序列中,如图 1.2 所示。对于线性结构类中每一数据元素,除第一个元素外,其他每一个元素都有一个且仅有一个直接前驱,第一个数据元素没有直接前驱;除最后一个元素外,其他每一个元素都有一个且仅有一个直接后继。

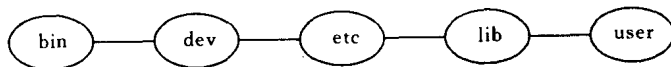


图 1.2 线性结构中各数据成员之间的线性关系

根据对线性结构中数据元素存取方法的不同,又可分为直接存取结构、顺序存取结构和词典(广义索引)结构。对于直接存取结构,可以直接存取某一指定项而不需先访问其前驱。像数组、记录和文件都属于这一类。可以按给定下标直接存取数组中某一数组元素;可以按记录号直接检索记录集合或文件中的某一记录。对于顺序存取结构,只能从序列中第 1 个数据元素起,按序逐个访问,直到指定的元素。例如,在第 2 章和第 3 章中讨论的顺序表和链接表就是这种情况。另外,一些限制存取位置在表的一端或两端的表(如栈、队列和优先级队列等)也是这种情况。词典与数组有类似之处,但数组是通过整数下标进行索引,而词典是通过关键词(key)进行索引。设定数据记录中某一数据项或某一组合数据项为关键词,通过关键词来识别记录。例如,对于学生记录,可设定学生的学号为关键词,用它来识别是哪一位学生的记录。索引表和散列表都属于这一类。

在非线性结构中各个数据成员不再保持在一个线性序列中,每个数据成员可能与零

个或多个其他数据成员发生联系。根据关系的不同,可分为层次结构和群结构。

层次结构是按层次划分的数据元素的集合,指定层次上的元素,可以有零个或多个处于下一个层次上的、直接的所属下层元素。在第6章将重点讨论树形结构,它是典型的层次结构。树中的元素叫做结点。树可以为空,也可以不为空。若树不为空,它有一个叫做根的结点,其他结点都是从它派生出来的。除根以外,每一个结点都有一个处于该结点直接上层的结点。树的结构如图1.3(a)所示。

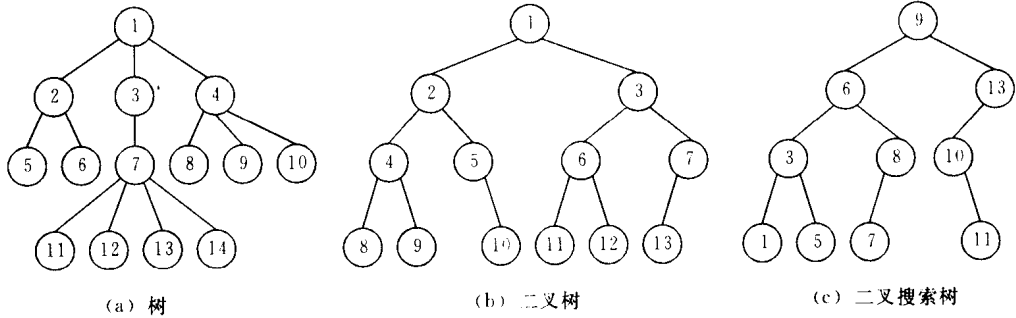


图 1.3 树形结构

一种特殊的树形结构叫做二叉树,如图1.3(b)所示。它的每一个结点最多有两个直接下层结点。二叉树中各结点上的数据经过特殊安排,可成为一棵二叉搜索树,如图1.3(c)所示。这种树形结构可用于构造小型多级索引,通过逐级比较实现对元素的快速访问。

堆是一种特殊的树形结构,如图1.4所示。图1.4(a)是“最大”堆,其根结点中的数据总是堆中所有数据中的最大者;图1.4(b)是“最小”堆,其根结点中的数据总是堆中所有数据中的最小者。优先级队列常常利用堆来实现。

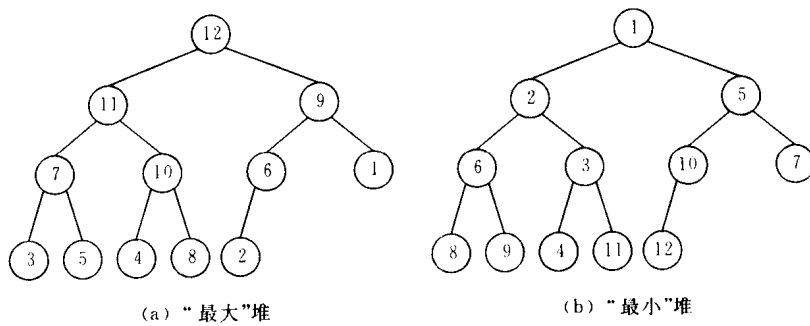


图 1.4 堆结构

群结构中所有元素之间无顺序关系。集合就是一种群结构,将在第7章讨论。在集合中没有重复的元素。另一种群结构就是图结构,如图1.5(a)所示。它是由图的顶点集合和连接顶点的边集合组成。还有一种图的特殊形式,即网络结构。它给每条边赋予一个权值,这个权值指明了在遍历图时经过此边时的耗费。例如,在图1.5(b)中,顶点代表城市,赋予边的权值表示两个城市之间的距离。