



数据结构 (C++版)

陈宝平 主编

张巨萍 孙宝军 阿雅娜 副主编



清华大学出版社

21世纪高等学校规划教材 | 计算机科学与技术

数据结构 (C++版)

陈宝平 主编
张巨萍 孙宝军 阿雅娜 副主编

清华大学出版社
北 京

内 容 简 介

数据结构是计算机专业教学计划中的核心课程,也是计算机及相关专业考研和水平等级考试的必考科目。要从事和计算机科学与技术相关的工作,尤其是计算机应用领域的开发和研制工作,必须具备坚实的数据结构基础。本书介绍了学习数据结构所用到的预备知识,叙述了数据结构、算法以及抽象数据类型的概念,介绍了线性表、栈、队列和串、数组和广义表、树和二叉树、图等常用数据结构,讨论了常用的查找、排序和索引技术。

本书内容丰富,层次清晰,讲解深入浅出,可作为计算机及相关专业本专科数据结构课程的教材,也可供从事计算机软件开发和应用的工程技术人员阅读、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(C++版)/陈宝平主编. —北京:清华大学出版社,2012.2

(21世纪高等学校规划教材·计算机科学与技术)

ISBN 978-7-302-27258-8

I. ①数… II. ①陈… 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(2011)第 228821 号

责任编辑:闫红梅 薛 阳

封面设计:傅瑞学,

责任校对:梁 毅

责任印制:何 芊

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

印 装 者:北京市密东印刷有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:16.25 字 数:405千字

版 次:2012年2月第1版 印 次:2012年2月第1次印刷

印 数:1~3000

定 价:26.00元

产品编号:043981-01

编审委员会成员

(按地区排序)

清华大学	周立柱	教授
	覃 征	教授
	王建民	教授
	冯建华	教授
	刘 强	副教授
北京大学	杨冬青	教授
	陈 钟	教授
	陈立军	副教授
北京航空航天大学	马殿富	教授
	吴超英	副教授
中国人民大学	姚淑珍	教授
	王 珊	教授
	孟小峰	教授
	陈 红	教授
	周明全	教授
北京师范大学	阮秋琦	教授
北京交通大学	赵 宏	副教授
北京信息工程学院	孟庆昌	教授
	杨炳儒	教授
北京科技大学	陈 明	教授
石油大学	艾德才	教授
天津大学	吴立德	教授
复旦大学	吴百锋	教授
	杨卫东	副教授
	苗夺谦	教授
同济大学	徐 安	教授
	邵志清	教授
华东理工大学	杨宗源	教授
华东师范大学	应吉康	教授
	乐嘉锦	教授
东华大学	孙 莉	副教授

浙江大学	吴朝晖	教授
	李善平	教授
扬州大学	李 云	教授
南京大学	骆 斌	教授
	黄 强	副教授
南京航空航天大学	黄志球	教授
	秦小麟	教授
南京理工大学	张功萱	教授
南京邮电学院	朱秀昌	教授
苏州大学	王宜怀	教授
	陈建明	副教授
江苏大学	鲍可进	教授
中国矿业大学	张 艳	教授
武汉大学	何炎祥	教授
华中科技大学	刘乐善	教授
中南财经政法大学	刘腾红	教授
华中师范大学	叶俊民	教授
	郑世珏	教授
	陈 利	教授
江汉大学	颜 彬	教授
国防科技大学	赵克佳	教授
	邹北骥	教授
中南大学	刘卫国	教授
湖南大学	林亚平	教授
西安交通大学	沈钧毅	教授
	齐 勇	教授
长安大学	巨永锋	教授
哈尔滨工业大学	郭茂祖	教授
吉林大学	徐一平	教授
	毕 强	教授
山东大学	孟祥旭	教授
	郝兴伟	教授
中山大学	潘小轰	教授
厦门大学	冯少荣	教授
厦门大学嘉庚学院	张思民	教授
云南大学	刘惟一	教授
电子科技大学	刘乃琦	教授
	罗 蕾	教授
成都理工大学	蔡 淮	教授
	于 春	副教授
西南交通大学	曾华桑	教授

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《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》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail: weijj@tup.tsinghua.edu.cn

前言

数据结构是计算机专业和信息管理专业一门重要的专业基础课,只要涉及程序的地方,均要用到数据结构的知识,许多课程(如操作系统、编译原理、数据库和人工智能等)也要用到数据结构的知识。其主要内容是研究和解决非数值计算的问题,讨论如何合理地组织、存储和处理数据的方法与技术;其主要任务是使读者掌握数据组织、存储和处理的常用方法和常用的算法思想及在实际中的应用技巧,为今后学习后续专业课和进行软件开发打下良好的基础。

本书的编写者长期从事数据结构的教學,对课程的难点和重点有比较深切的体会。在总结讲授数据结构课程的基础上,对多年传授的数据结构教学内容进行合理的重组,既强调数据结构的原理和方法,又注重其实践性与实用性。本书希望能够引起读者对数据结构课程的兴趣,提高对数据结构课程的重要性、必要性的认识,特别希望能够提高数据结构课程的教学实效,让读者满意并有收获。

本书详细介绍了各种基本的数据结构,包括线性结构、数组和广义表、树(层次)结构和图结构;基本的存储结构,包括顺序(数组)结构、链接结构、索引结构和散列结构,以及其中的一种或多种结构的组合及应用。本书具有以下特色:

(1) 采用 C++ 版,注重面向对象的编程,强调算法的逻辑性、严谨性、技巧性以及时间复杂度等,弱化 C++ 语言中的继承、派生等机制。

(2) 本着“加强基础,注重算法,因材施教,案例驱动”的教育理念,在算法的设计上做了大胆改革,为了更好地理解理论知识,每章之后都配有应用实例,提倡从实用性和实践性的角度学习数据结构。

(3) 完全包含 2011 年计算机统考大纲中数据结构指定的相关知识点。

(4) 通俗易懂、图文并茂,详略得当,便于从事计算机工作的人员自学。

本书第 1,5,10 章及所有的课后习题由阿雅娜编写,第 2,3,4 章由张巨萍编写,第 6,8 章由陈宝平编写,第 7,9 章由孙宝军编写。陈宝平对全书的文字风格、内容、算法实现等方面进行了细致的修改和统稿。作者在此对赵俊岚教授、乔晓华教授和王彪教授表示衷心的感谢,他们对此书提出的建议和意见都弥足可贵。同时感谢王景龙、魏娜、李泽申等同学,仔细阅读本书的预印版,验证了书稿的部分算法并调试了代码。

本书力求以通俗的语言和实例描述数据结构课程的内容。实例丰富,图表清晰,深入浅出,易于读者融会贯通。根据 C++ 的特点设计与实现了各种数据结构的存储结构和算法,有助于读者进一步掌握 C++ 的编程思想、方法和技术内涵。

由于编者水平有限,书中错误和不妥之处在所难免,恳请读者与同行批评指正。

陈宝平

目 录

第 1 章 绪论	1
1.1 为什么要学习数据结构	1
1.2 什么是数据结构	2
1.2.1 数据的逻辑结构	4
1.2.2 数据的存储结构	6
1.2.3 抽象数据类型	7
1.3 算法与算法分析	8
1.3.1 算法	8
1.3.2 算法的设计要求	10
1.3.3 算法效率的量度	10
1.3.4 算法的设计方式	15
习题	16
第 2 章 线性表	20
2.1 线性表的逻辑结构	20
2.1.1 线性表的定义	20
2.1.2 线性表的抽象数据类型定义	21
2.2 线性表的顺序表示和实现	23
2.2.1 顺序存储结构的定义	23
2.2.2 基本操作在顺序表中的实现	24
2.2.3 顺序存储结构的特点	27
2.3 线性表的链式表示和实现	28
2.3.1 单链表	28
2.3.2 双向链表	32
2.3.3 循环链表	33
2.3.4 链式存储结构的特点	34
2.4 一元多项式求和	35
2.4.1 一元多项式的表示	35
2.4.2 一元多项式的求和	36
习题	38
第 3 章 栈和队列	40
3.1 栈	40



3.1.1	栈的抽象数据类型定义	40
3.1.2	栈的实现	41
3.2	栈的应用举例	44
3.3	栈与递归	53
3.4	队列	57
3.4.1	队列的抽象数据类型定义	57
3.4.2	队列的实现	57
3.4.3	队列的应用	62
习题		65
第4章	串	68
4.1	串类型的定义	68
4.2	串的存储结构	70
4.2.1	串的顺序存储结构	70
4.2.2	堆分配存储表示	72
4.2.3	串的块链存储表示	74
4.3	串的模式匹配算法	75
4.3.1	求子串的定位函数	75
4.3.2	模式匹配的一种改进算法	77
4.4	串的应用	80
习题		81
第5章	数组和广义表	82
5.1	数组	82
5.1.1	数组的定义	82
5.1.2	数组的存储	82
5.1.3	特殊矩阵	84
5.1.4	稀疏矩阵	86
5.2	广义表	93
5.2.1	广义表的定义	93
5.2.2	广义表的存储结构	94
5.2.3	广义表的递归算法	96
5.2.4	广义表的应用	100
习题		100
第6章	树与二叉树	102
6.1	树的定义与基本术语	102
6.2	二叉树	104
6.2.1	二叉树的定义	104

6.2.2	二叉树的性质	105
6.2.3	二叉树的存储结构	106
6.3	二叉树的遍历	108
6.3.1	递归遍历二叉树	108
6.3.2	应用二叉树遍历的实例	110
6.4	线索二叉树	113
6.5	树与森林	115
6.5.1	树的存储表示	115
6.5.2	森林与二叉树的转换	117
6.5.3	树的遍历	118
6.5.4	森林的遍历	119
6.6	树的应用	119
6.6.1	堆	119
6.6.2	哈夫曼树与编码	124
	习题	128
第7章	集合与搜索	133
7.1	集合及其表示	133
7.1.1	集合的定义	133
7.1.2	集合的抽象数据类型	134
7.1.3	用位向量实现集合	134
7.2	静态搜索结构	136
7.2.1	搜索的定义	136
7.2.2	静态搜索结构	137
7.2.3	顺序搜索	138
7.2.4	基于有序顺序表的折半搜索	139
7.2.5	分块搜索	142
7.3	二叉搜索树	144
7.3.1	二叉搜索树的定义	144
7.3.2	二叉搜索树的搜索	146
7.3.3	二叉搜索树的插入	147
7.3.4	二叉搜索树的建立	148
7.3.5	二叉搜索树的删除	149
7.4	AVL 树	151
7.4.1	AVL 树的定义	151
7.4.2	最小不平衡二叉树	152
7.4.3	不平衡二叉树的调整方法	153
7.4.4	建立平衡二叉树举例	158
7.5	应用举例计算机登录验证	160



习题.....	163
第 8 章 图.....	164
8.1 图的定义	164
8.1.1 图的定义与相关术语.....	164
8.1.2 图的抽象数据类型.....	166
8.2 图的存储结构	166
8.2.1 数组表示法.....	166
8.2.2 邻接表表示法.....	168
8.2.3 邻接多重表表示法.....	171
8.2.4 十字链表法.....	172
8.3 图的遍历	172
8.3.1 深度优先遍历.....	173
8.3.2 广度优先遍历.....	174
8.4 图的最小生成树	176
8.4.1 Prim 算法	176
8.4.2 Kruskal 算法	179
8.5 最短路径	180
8.5.1 单源最短路径.....	181
8.5.2 每对顶点的最短路径.....	183
8.6 拓扑排序	185
8.7 关键路径	187
8.8 应用实例	191
习题.....	193
第 9 章 排序.....	196
9.1 概述	196
9.2 插入排序	198
9.2.1 直接插入排序.....	198
9.2.2 折半插入排序.....	199
9.2.3 希尔排序.....	200
9.3 交换排序	202
9.3.1 冒泡排序.....	202
9.3.2 快速排序.....	204
9.4 选择排序	206
9.4.1 直接选择排序.....	206
9.4.2 堆排序.....	207
9.5 归并排序	211
9.5.1 归并排序概述.....	211

9.5.2 递归的归并排序算法	212
9.6 基数排序	213
9.6.1 多关键码排序	213
9.6.2 链式基数排序	214
9.7 各种排序方法的比较讨论	216
9.8 外部排序的方法	217
习题	220
第 10 章 索引结构和散列	222
10.1 静态索引结构	222
10.1.1 线性索引	222
10.1.2 倒排表	223
10.1.3 m 路静态索引树	225
10.2 动态索引结构	225
10.2.1 动态的 m 路静态索引树	225
10.2.2 B_树	226
10.2.3 B_树的插入	229
10.2.4 B_树的删除	232
10.2.5 B ₊ 树	234
10.3 散列	235
10.3.1 散列函数	236
10.3.2 开散列方法	238
10.3.3 闭散列方法	239
10.3.4 散列表的实现	241
10.3.5 散列表分析	243
习题	244
参考文献	245

第 1 章

绪论

由唐纳德·克努特(Donald Ervin Knuth)所著的洋洋数百万言的多卷本《计算机程序设计的艺术》(The Art of Computer Programming)堪称计算机科学技术的经典巨著。其第一卷《基本算法》于1968年出版,全面、系统地阐述了数据的逻辑结构和存储结构,为数据结构奠定了基础。随后一些大学开始将《数据结构》作为一门课程讲授,并且列入了教学计划。迄今为止,《数据结构》的研究和发展已经经历了数十载。期间,尼克劳斯·沃斯(Niklaus Wirth)编写的《算法+数据结构=程序》(Algorithms+Data Structure=Programs)、约翰·霍普克罗夫特(John Edward Hopcroft)编写的《数据结构和算法》(Data Structure and Algorithms)等都为数据结构的发展做出了卓越的贡献。如今,虽然数据结构理论已经较为成熟,但随着计算机技术的发展,对于它的研究从未停止过。如何将数据结构应用于专业领域的研究还在不断发展中,如多维图形数据结构、动画数据结构、音频数据结构等。

1.1 为什么要学习数据结构

第一台得到应用的计算机 ENIAC(Electrical Numerical Integrator And Calculator,电子数字积分计算机)于1946年在美国诞生,第一家计算机研发和生产的企业埃克特-莫奇莱计算机公司(EMCC)于1948年成立,第一个成功的高级程序设计语言 FORTRAN 在1957年开始商业发行,第一个商业成功的文字处理软件——WordStar 在20世纪80年代风行于全球,第一个专门用于网格计算的数据库——Oracle 数据库在1983年发布了具有可移植性的 Oracle 第3版,1991年开始,Internet 真正开始商业化运行。计算机渐渐渗透到现实世界的各个领域当中,汹涌而来的信息数据使人无所适从,管理它们、使它们井井有条、易于使用变得非常困难。这些信息数据不再仅仅局限于简单的数值,还包括更为复杂的字符、图像、表格等。为了更好地用计算机处理这些具有一定组织结构的数据,就必须分析它们的特性、相互关系及相应的存储方式,在这个基础上设计出相应的算法和程序,这正是数据结构这门课程的研究内容。

从课程体系来说,数据结构是一门综合性很强的专业基础课,不仅是计算机专业教学计划中的核心课程之一,而且还是其他非计算机专业的主要选修课之一。它是介于数学、计算机硬件和计算机软件三者之间的一门核心课程,是计算机系统软件和应用软件研制者的必修课程。在数学方面,数据结构这个概念包含了图论、树、网络、集合论、关系等现在《离散数学》中讲述的内容;在硬件方面,编码理论、存储装置和存储方法等与数据结构有着密切的

关系；在软件方面，计算机应用软件和系统软件如操作系统、编译系统和数据库等都需要很多数据结构相关知识，如编译技术要使用栈，操作系统会用到队列、存储管理，数据库系统将运用线性表、散列表、链表以及索引技术等。

本书中介绍了一些数据结构中最常用的数据结构，阐明数据结构的逻辑关系，分析它们在计算机中的存储表示，并结合各种数据结构讨论可以实现的算法。通过对本书的学习，结合实际编程练习，读者将有能力根据实际问题选择恰当的数据结构以及控制相应算法的效率。

1.2 什么是数据结构

在日常生活中，人们会遇到各种信息，如在路上进行交通指挥的交警的手势，人们通过电子邮件或聊天工具交流的思想，顾客在商场购买物品后所做的记录等。这些信息必须转换成数据才能在计算机中进行处理。因此，在学习数据结构这门课程之前，需要定义和说明一些术语。

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。我们平时所遇到的数据主要有两类，一类是数值型数据，包括整数、实数、复数，主要在工程和科学计算，以及商业事务处理中使用；另一类是非数值数据，主要包括文本、表格、图形和语音等数据。

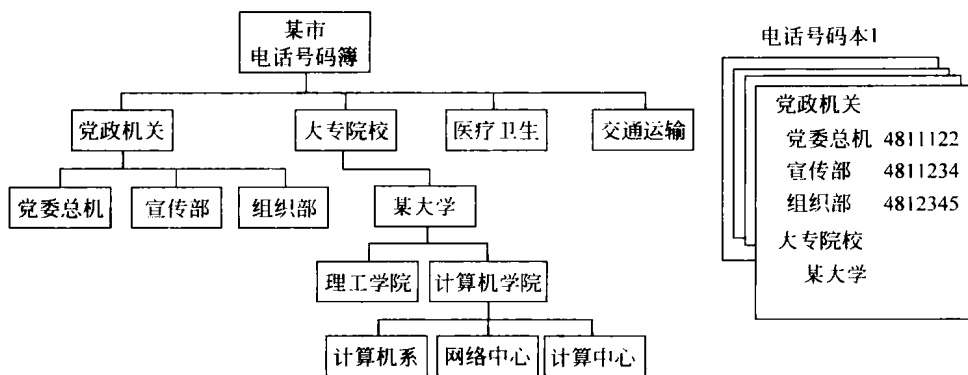
数据元素(data element)简称元素，在用计算机进行处理的时候通常将它作为一个相对独立的单位。如对于一个二维表格数据来说，每行信息就是它的数据元素；对于一个字符串数据来说，每个字符就是它的数据元素；对于一维数组来说，每个下标位置所存储的值就是它的数据元素。一个数据元素通常由一个或多个数据项组成，每个数据项可以是简单数据项，即不可再分，如一个数值、字符等；也可以是组合数据项，如一个字符串、数组、记录、对象等。数据和数据元素是相对而言的，如对于一条记录信息来说，它是所属文件的一个数据元素，而它相对于所含的数据项而言又可看成数据。

数据处理(data processing)是指利用计算机对数据进行处理的过程，这里所指的处理可以是存储、检索、插入、删除、排序、统计、输入、输出等。

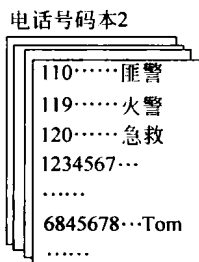
1.1 节提到随着计算机应用领域的扩大，计算机所处理数据的结构越来越复杂，非数值计算问题显得越来越重要。据统计，目前 85% 以上的机器时间用于处理非数值计算性问题。下面列举一些具体问题。

【例 1.1】 通常情况下，在制作一个电话号码簿时，会按照组织机构来编排，如图 1.1(a) 所示。这种编排形式的优点是，在确定要查找的号码隶属于哪个部门的情况下，可以迅速地查找出电话号码。但是，知道某个电话号码，想查找出它属于哪个部门，就不容易了。所以电话号码簿的编排可以有另一种形式，即按照电话号码的大小进行排列，标注出其隶属部门，如图 1.1(b) 所示。数据的安排直接影响到工作效率。

【例 1.2】 通信网络问题。假设要在 n 个城市之间建立通信联络网，如图 1.2 所示，图中用 $\bigcirc$ 表示城市，两个 $\bigcirc$ 之间连线上的数字表示在两个城市之间铺设通信线路所需要的代价。连通 n 个城市最少需要 $n-1$ 条线路，在建立通信网络的时候需要在众多线路中选出最小代价并且能连通 n 个城市的 $n-1$ 条线路。



(a) 按照组织机构来排列电话号码



(b) 按照电话号码大小来排列电话号码

图 1.1 电话号码簿排列图

【例 1.3】 人机对弈问题。假设 A 为玩家, B 为电脑, A 与 B 下棋对弈。首先 A 下棋时对应局面是 1, 由此局面出发 B 有 n 种符合规则的方式下棋, 走棋后对应的局面为 $2, 3, 4, \dots, n+1$; 如果 B 选择了形成局面 2 的方式走棋, 则 A 在下一步有 m 种符合规则的方式下棋, 对应的局面为 $1+1, 1+2, \dots, 1+m$; 如果 A 选择了形成局面 $1+1$ 的方式走棋, B 也有若干种符合规则的走棋方式, 以此类推。这样 A 和 B 轮流下棋, 棋盘局面的发展变化如图 1.3 所示, 这种关系可以由一种叫做树的数据结构来表示。

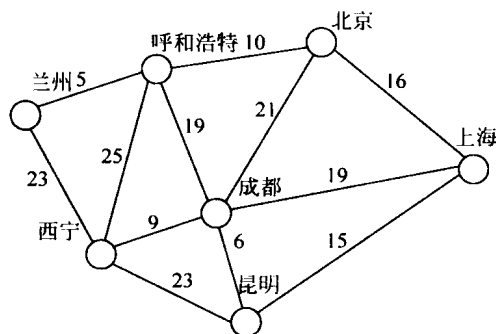


图 1.2 通信网络图

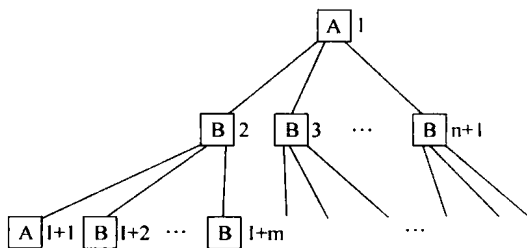


图 1.3 博弈树

通过以上 3 个实例可以清楚地认识到, 用简单的数值和数学方程已经不足以表示和计算这类非数值问题, 而要使用更为复杂的图、表和树之类的数据结构。所以, 简单来讲, 数据结构是一门研究非数值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

作的学科。其中,数据以及它们之间的关系可以用数据的逻辑结构来表示,在计算机上进行存储的方式用数据的存储结构表示。我们时常把数据的逻辑结构简称为数据结构,而在讨论数据的存储结构时则必须指明是数据的存储结构。

1.2.1 数据的逻辑结构

数据的逻辑结构是指从解决问题的需要出发,为实现必要的功能所建立的数据结构,它属于用户的视图,是面向问题的,如在招生系统中建立的按考分排列的考生记录的有序表格。数据的逻辑结构包括集合、线性、树、图等基本结构,由它们的组合和嵌套可以形成更复杂的结构。

为了描述数据的逻辑结构,需要对数据结构进行二元组表示。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:

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

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。

1. 集合

集合结构是指数据中各元素之间没有任何次序。如一个容器中的所有乒乓球,一个俱乐部里的所有成员,可以认为它们之间没有任何次序,它们均为集合结构。

2. 线性结构

线性结构是指数据中各元素之间具有1对1的先后次序关系。如在一个列车时刻表中,各车次记录之间是按照发车时间的先后次序排列的;在一个人事职工表中,各职工记录之间是按照职工编号的先后次序排列的。所以,它们的表结构都是线性结构。

3. 树结构

树结构是指数据中各元素之间具有1对多的先后次序关系,并且只有一个元素称为树根结点,其余均为树枝结点和树叶结点。如在一个企业的组织机构中,总经理只有一个,相当于树根;其下属多个部门,每个部门又各有一个部门经理,相当于树枝;每个部门又有多名员工,归部门经理领导,相当于树叶。所以,企业的组织结构是一个树结构。

4. 图结构

图结构是指数据中各元素之间具有多对多的关系。这是数据结构中最复杂的结构。如在一个城市交通图中,所有道路连接成一个复杂的图结构,交叉路口就是图中的顶点,每条道路就是图中的边,从一个交叉路口可以到达与其连接的其他交叉路口。

各种数据结构类型的图形表示如图1.4所示。

从图形表中可以看出:

(1) 集合结构中的元素是各自独立的,元素之间没有联系。

(2) 线性结构中的元素是一个接一个串联起来的,它有一个头元素A和一个尾元素G,其余为中间元素;每个中间元素既有前驱元素,又有后继元素,如B的前驱元素为A,后继元素为C;C的前驱元素为B,后继元素为D,……,头元素A没有前驱元素,只有后继元素B;尾元素G只有前驱元素F,没有后继元素。