

21世纪

高等院校计算机系列教材

数据结构 (C/C++描述)

阮宏一 主 编

唐宏亮 杨 鹤 张绪辉 贺 睿 副主编



中国水利水电出版社
www.waterpub.com.cn

21 世纪高等院校计算机系列教材

数据结构 (C/C++描述)

阮宏一 主编

唐宏亮 杨 鹤 张绪辉 贺 睿 副主编

中国水利水电出版社

内 容 提 要

本书是为高等学校计算机专业“数据结构”课程编写的教材。本书主要采用 C 语言作为数据结构和算法的描述语言,考虑到算法描述的简洁性和知识的延续性,在本书的算法描述中适当引进了部分 C++ 的基本概念,使算法描述更为简明、清晰。

全书共分 10 章及一个附录。分别介绍数据结构的基本概念;线性结构的相关概念及算法;多维数组、矩阵和广义表的基本概念及算法;非线性结构树、图的基本概念及算法以及查找、文件和内外排序的基本概念及算法,并在附录中给出了有关 C 和 C++ 的相关对照等。

本书给出的绝大多数算法都特别突出了算法设计思想、完整的算法描述及算法分析三个部分。书中各章最后都给出了难易适中的不同类型的习题,供学生课后练习使用。

本书适合作为计算机类各专业的本科或专科教材,也可作为信息类相关专业的选修教材,亦可作为高校相关专业师生、工程技术人员和其他读者的学习参考书。

本书为授课教师免费提供电子教案,此教案用 PowerPoint 制作,可以任意修改。需要者可从中国水利水电出版社网站免费下载,网址为: <http://www.waterpub.com.cn/softdown/>。

图书在版编目(CIP)数据

数据结构: C/C++描述/阮宏一主编. —北京: 中国水利水电出版社, 2007

(21 世纪高等院校计算机系列教材)

ISBN 978-7-5084-4122-1

I. 数... II. 阮... 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. ①TP311.12②TP312

中国版本图书馆 CIP 数据核字 (2006) 第 120446 号

书 名	数据结构 (C/C++描述)
作 者	阮宏一 主 编 唐宏亮 杨 鹤 张绪辉 贺 睿 副主编
出版 发行	中国水利水电出版社 (北京市三里河路 6 号 100044) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 63202266 (总机)、68331835 (营销中心)、82562819 (万水)
经 售	全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	北京市天竺颖华印刷厂
规 格	787mm×1092mm 16 开本 19.25 印张 477 千字
版 次	2007 年 3 月第 1 版 2007 年 3 月第 1 次印刷
印 数	0001—4000 册
定 价	28.00 元

凡购买我社图书,如有缺页、倒页、脱页的,本社营销中心负责调换

版权所有·侵权必究

出版精品教材

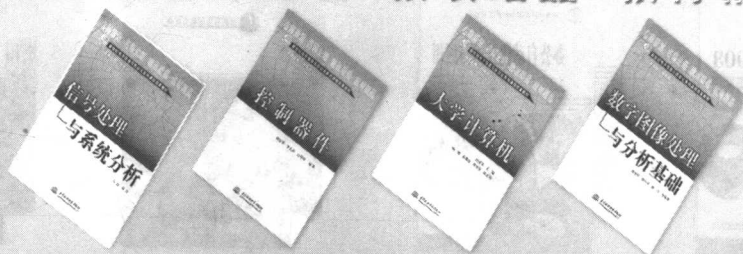
服务高校师生

面向21世纪 免费电子教案 教材立体化配套 免费提供样书

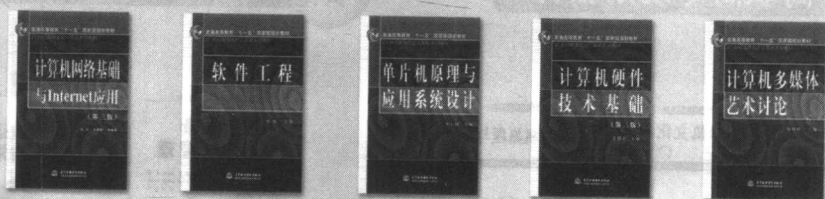


新世纪电子信息与自动化系列课程改革教材

名师策划 名师主理 教改结晶 教材精品



普通高等院校“十一·五”国家规划教材



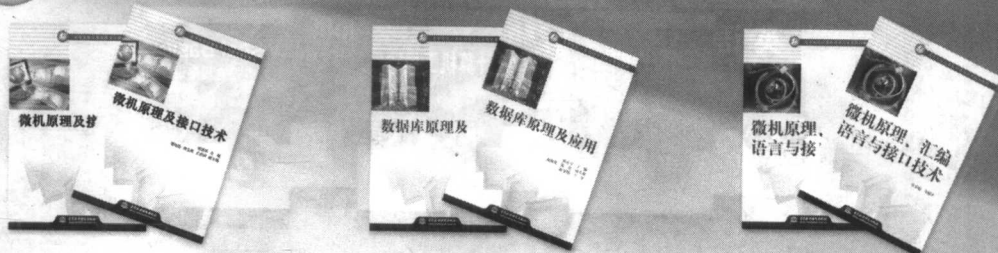
高等院校规划教材

适应高等教育的跨越式发展 符合应用型人才的培养要求

本套丛书是由一批具备较高的学术水平、丰富的教学经验、较强的工程实践能力的学术带头人和主要从事该课程教学的骨干教师在对应用型人才与研究人才在培养目标、课程体系 and 内容编排上的区别，精心策划出来的。丛书共分3个层面，百余种。



高等院校计算机科学与技术规划教材



21
世纪

高职高专新概念教材

免费提供电子教案 免费源代码下载

本套教材已出版百余种，发行量均达万册以上，深受广大师生和读者好评，近期根据作者自身教学体会以及各学校的使用建议，大部分教材推出第二版对全书内容进行了重新审核与更新，使其更能跟上计算机科学的发展、跟上高职高专教学改革的要求。



21
世纪

高职高专规划教材

21
世纪

高职高专计算机科学规划教材



十五

软件职业技术学院“十一五”规划教材

本套丛书特点：

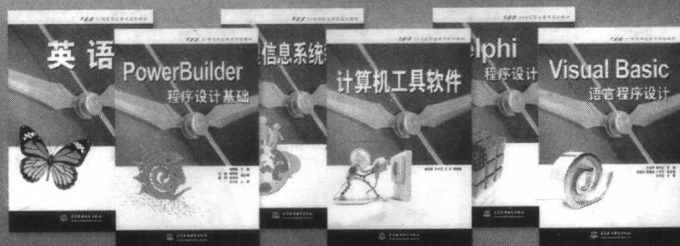
- 以实际工程项目为引导来说明各知识点，使学生学为所用。
- 突出实习实训，重在培养学生的专业能力和实践能力。
- 内容衔接合理，采用项目驱动的编写方式，完全按项目运作所需的知识体系设置结构。
- 配套齐全，不仅包括教学用书，还包括实习实训材料、教学课件等，使用方便。

强调实践 面向就业 产学结合



21
世纪

职业教育规划教材



前 言

“数据结构”是学习计算机软件的一门重要基础课程。计算机科学各领域及相关的应用软件都要用到各种数据结构。“数据结构”课程的目的是使学生学会分析各种数据结构及其内在的逻辑关系，掌握它们在计算机中的存储表示及其相应的算法，并初步掌握算法执行的时间和空间分析技术。由于它在计算机学科的各门课程中具有承前启后的作用，因此学好数据结构这门课程至关重要。

数据结构的原理及算法较为抽象，而该课程一般在本科低年级开设，我们在多年的教学过程中深深感到，对于只具有计算机程序设计基础知识的学生，要很好地理解和掌握其中的原理比较困难，特别是理解数据的存储结构、相关的算法，以及将算法转化成对应的应用程序，这些都是学习的难点。这本《数据结构》教材是作者在多年教授“数据结构”课程并指导学生上机实践所积累的知识与经验的基础之上编写而成的，希望能够帮助学生在尽可能短的时间内对数据结构的知识与应用有一个比较全面、深入和系统的认识。本书给出的绝大多数算法都特别突出了算法设计思想、完整的算法描述及算法分析三个部分，既便于学生将算法转换成 C 或 C++ 程序，也能为培养学生在实际应用中分析问题和解决问题的能力打下良好的基础。

本书主要采用 C 语言作为数据结构和算法的描述语言，考虑到算法描述的简洁性和知识的延续性，我们在算法描述中适当地引进了部分 C++ 的基本概念。例如，C++ 语言的引用调用、参数传递方式等，使得算法描述更为简明清晰。读者只需通过本书的附录，对 C++ 的引用调用的参数传递方式作一个了解，就能读懂书中的所有算法。

全书共分 10 章及一个附录。第 1 章介绍数据结构的基本概念、抽象数据类型及算法分析方法；第 2 章至第 4 章介绍线性结构的相关概念及算法，包括线性表、栈、队列和串；第 5 章介绍较为复杂的数据结构，包括多维数组、稀疏矩阵和广义表；第 6 章、第 7 章分别介绍了非线性结构树及图；第 8 章至第 10 章介绍常用的查找、排序算法及文件结构。附录中给出了本书所涉及的有关 C 和 C++ 相关知识的对照，有助于前期学过 C 或 C++ 语言的学生顺利阅读本书。

本书编写力求概念准确、文字流畅、算法精湛、突出重点、注重应用。其中各章都给出了难易适中的不同类型的习题，供学生课后练习使用。本书适合作为计算机类专业的本科或专科教材，也可作为信息类相关专业的选修教材。本书授课的总学时可定为 108 学时，其中讲授 72 学时，上机实践 36 学时。

本书由阮宏一主编，并负责全书的总体策划与统稿、定稿工作，唐宏亮、杨鹤、张绪辉任副主编。各章主要编写人员分工如下：第 1、2、5、7 章由阮宏一、张绪辉、贺睿编写，第 3、4、6 章由唐宏亮、王芳、朱勤童编写，第 8 章至第 10 章由杨鹤、李红玲、杜超编写。参加本书编写的还有文忠林、强士端、赵晶、王大雷、李玲、黄朝援、江涛、刘平、

张伟、刘舒、陈灿等。

本书在编写过程中还得到了雷建军老师、罗忠老师的大力支持和帮助，并提出了很好的意见和建议，在此谨向他们表示衷心感谢。

本书为授课教师免费提供电子教案，此教案采用 PowerPoint 制作，可以任意修改。需要者可从中国水利水电出版社网站免费下载，网址为：<http://www.waterpub.com.cn/softdown/>。

由于时间仓促及作者水平有限，书中难免存在欠妥和疏漏之处，敬请广大读者批评指正。作者的 E-mail 地址为：hyruan@hubce.edu.cn、hyyh001@126.com。

编 者

2006 年 11 月

目 录

前言

第 1 章 绪论	1
1.1 数据结构的概念	1
1.1.1 什么是数据结构	1
1.1.2 为什么要学习数据结构	2
1.1.3 基本概念和术语	5
1.2 抽象数据类型	9
1.2.1 数据类型	10
1.2.2 抽象数据类型	10
1.2.3 抽象数据类型的表示和实现	11
1.3 算法和算法分析	12
1.3.1 算法的特性	13
1.3.2 算法描述	13
1.3.3 算法性能分析与度量	14
习题 1	16
第 2 章 线性表	19
2.1 线性表的类型定义	19
2.1.1 抽象数据类型线性表的定义	19
2.1.2 基于 ADT 线性表的算法设计	20
2.2 线性表的顺序存储及实现	22
2.2.1 顺序表的定义及存储	22
2.2.2 顺序表基本运算的实现	24
2.2.3 顺序表应用举例	28
2.3 线性表的链式存储及实现	29
2.3.1 单链表的定义及存储	30
2.3.2 单链表基本运算的实现	31
2.3.3 循环链表	36
2.3.4 双向链表	38
2.3.5 静态链表	40
2.4 线性表应用举例	42
2.4.1 一元多项式的表示	42
2.4.2 一元多项式相加	43

习题 2	46
第 3 章 栈和队列	48
3.1 栈	48
3.1.1 栈的定义及基本运算	48
3.1.2 栈的顺序表示与实现	49
3.1.3 栈的链式表示与实现	51
3.2 栈的应用举例	53
3.2.1 括号匹配的检验	53
3.2.2 行编辑程序	55
3.2.3 表达式求值	56
3.3 栈与递归	58
3.3.1 递归的实现	59
3.3.2 递归设计	60
3.4 队列	62
3.4.1 队列的定义及基本运算	62
3.4.2 队列的顺序表示和实现	63
3.4.3 队列的链式表示和实现	66
3.4.4 队列的应用举例	67
习题 3	69
第 4 章 串	72
4.1 串的定义	72
4.2 串的存储及基本运算	74
4.2.1 串的定长顺序存储表示	74
4.2.2 串的堆分配存储表示	78
4.2.3 串的块链存储表示	81
4.3 串的模式匹配算法	82
4.3.1 简单的模式匹配算法	82
4.3.2 无回溯的模式匹配算法	84
4.4 串的应用	87
习题 4	89
第 5 章 多维数组、矩阵和广义表	90
5.1 多维数组	90
5.1.1 数组的定义	90
5.1.2 数组的存储表示	91
5.1.3 数组基本运算的实现	93
5.2 特殊矩阵	96
5.2.1 对称矩阵	96

5.2.2	三角矩阵	97
5.2.3	对角矩阵	98
5.3	稀疏矩阵	98
5.3.1	稀疏矩阵的定义	99
5.3.2	三元组顺序表	100
5.3.3	十字链表	103
5.4	广义表	108
5.4.1	广义表的定义和基本运算	109
5.4.2	广义表的存储结构	110
5.4.3	广义表基本运算的实现	113
习题 5		117
第 6 章	树和二叉树	120
6.1	树的定义及其存储结构	120
6.1.1	树的定义及基本术语	120
6.1.2	树的存储结构	123
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	128
6.2.3	二叉树的存储结构	129
6.3	遍历二叉树和线索化二叉树	131
6.3.1	遍历二叉树	131
6.3.2	线索化二叉树	135
6.4	树、森林和二叉树的关系	139
6.4.1	树、森林与二叉树的转换	139
6.4.2	树和森林的遍历	141
6.5	哈夫曼树及其应用	142
6.5.1	哈夫曼树	142
6.5.2	哈夫曼树的应用	144
6.5.3	哈夫曼算法的实现	146
习题 6		149
第 7 章	图	151
7.1	图的基本概念	151
7.1.1	图的定义	151
7.1.2	图的基本术语	153
7.2	图的存储结构	157
7.2.1	邻接矩阵	157
7.2.2	邻接表	159

7.2.3	十字链表	161
7.3	图的遍历及图的连通分量	163
7.3.1	深度优先搜索	163
7.3.2	广度优先搜索	165
7.3.3	图的连通分量	167
7.4	生成树和最小生成树	168
7.4.1	生成树和生成森林	168
7.4.2	最小生成树	170
7.4.3	关节点和重连通分量	175
7.5	最短路径	178
7.5.1	从某个源点到其余各顶点的最短路径	178
7.5.2	每一对顶点之间的最短路径	181
7.6	拓扑排序与关键路径	183
7.6.1	拓扑排序	184
7.6.2	关键路径	186
习题 7		191
第 8 章	查找	194
8.1	基本概念	194
8.2	静态查找	195
8.2.1	顺序查找	195
8.2.2	折半查找	197
8.2.3	分块查找	200
8.3	动态查找	201
8.3.1	二叉排序树	202
8.3.2	平衡二叉树	207
8.3.3	B_树和 B ⁺ 树	216
8.4	哈希查找	223
8.4.1	哈希表的基本概念	223
8.4.2	哈希函数的构造方法	224
8.4.3	处理冲突的方法	226
8.4.4	哈希表的查找与分析	228
8.4.5	哈希算法举例	229
习题 8		232
第 9 章	内排序	234
9.1	基本概念	234
9.2	插入排序	235
9.2.1	直接插入排序	235

9.2.2	折半插入排序	237
9.2.3	希尔排序	238
9.3	选择排序	240
9.3.1	直接选择排序	240
9.3.2	堆排序	242
9.4	交换排序	247
9.4.1	冒泡排序	247
9.4.2	快速排序	249
9.5	归并排序	252
9.6	基数排序	255
9.7	各种内排序方法比较	260
习题 9		261
第 10 章	文件与外排序	263
10.1	文件	263
10.1.1	外存信息的存取	263
10.1.2	文件的基本概念	264
10.1.3	顺序文件	266
10.1.4	索引文件	267
10.1.5	索引顺序文件	268
10.1.6	直接存取文件（散列文件）	272
10.1.7	多关键字文件	273
10.2	外排序	276
10.2.1	外排序的方法	276
10.2.2	多路平衡归并及实现	277
10.2.3	置换—选择排序	282
10.2.4	最佳归并树	287
习题 10		289
附录		291
参考文献		296

第 1 章 绪论

计算机科学是一门研究数据表示和数据处理的科学。数据是计算机化的信息，它是计算机可以直接处理的最基本和最重要的对象。无论是科学计算还是数据处理、是过程控制还是对文件的存储和检索及数据库技术等计算机应用领域中，都是计算机对数据进行加工处理的过程。因此，要设计出一个结构好、效率高的程序，必须研究分析待处理的数据的特性及数据间的相互关系及其对应的存储表示，并利用这些特性和关系设计出相应的算法和程序。本章将介绍有关数据和数据结构的相关概念，抽象数据类型及其表示，算法的概念及评价标准。

1.1 数据结构的概念

1.1.1 什么是数据结构

对于什么是**数据结构** (Data Structure)，目前还没有一个公认的标准定义，一般地，数据结构就是指数据间的相互关系。随着计算机应用技术的发展，不但处理的数据量越来越大，而且要处理的数据结构也越来越复杂。

对于任何一种数据结构，都自然地联系着对这种类型的数据所需要的运算，以及为了在计算机上执行这些运算而如何在计算机中存储这些数据这一问题。所以，数据结构的**概念**一般包括数据之间的相互关系、数据在计算机中的存储表示及数据的运算三个方面，简称为数据的**逻辑结构**、数据的**存储结构**及数据的**运算**，它们是数据结构的三要素。因此，人们称按某种逻辑关系组织起来的一批数据，按一定的存储方式存储于计算机中，并在这些数据之上定义了一个运算的集合，称为一个**数据结构**。

数据结构与数学、计算机硬件和软件有着十分密切的关系。它是介于数学、计算机硬件和计算机软件之间的一门计算机科学与技术专业的核心课程，是编译原理、操作系统、数据库、人工智能等后续课程的基础。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技术领域。

数据结构作为一门独立的课程在国外是从 1968 年才开始的，但在此之前其有关内容已散见于编译原理及操作系统之中。20 世纪 60 年代中期，美国的一些大学开始设立有关课程，但当时的课程名称并不叫数据结构。1968 年美国唐·欧·克努特教授开创了数据结构的最初体系，他所著的《计算机程序设计技巧》第一卷《基本算法》是第一本较系统地阐述数据的逻辑结构和存储结构及其操作的著作。从 20 世纪 60 年代末到 70 年代初，出现了大型程序，软件也相对独立，结构程序设计成为程序设计方法学的主要内容，人们越来越重视数据结构。从 20 世纪 70 年代中期到 80 年代，各种版本的数据结构著作相继出现。

目前，数据结构的发展并未终结，一方面，面向各专门领域中特殊问题的数据结构还有待研究和发展，如多维图形数据结构等；另一方面，从抽象数据类型和面向对象的观点来讨论数据结构已成为一种新的趋势，越来越被人们所重视。

数据结构课程集中讨论软件开发过程中的算法设计及实现，同时为了构造好的算法，还需考虑数据结构及其实现的评价与选择。因此，数据结构课程的内容主要包括三个层次的 5 个“要素”，如表 1-1 所示。

表 1-1 数据结构课程内容体系

层次 \ 方面	数据表示	数据处理
	逻辑结构	基本运算
抽象	逻辑结构	基本运算
实现	存储结构	算法
评价	不同数据结构的比较及算法分析	

数据结构课程的核心是分解与抽象。通过分解可以将问题划分出数据结构的三个层次；再通过抽象，舍弃数据元素的具体内容，就得到数据的逻辑结构。类似地，通过分解将处理要求划分成各种功能，再通过抽象舍弃实现细节，就得到基本运算的定义。上述两个方面的结合可使问题变换为数据结构。这是一个从具体（即具体问题）到抽象（即数据结构）的过程。然后，通过增加对实现细节的考虑进一步得到存储结构和算法，从而完成设计任务。这是一个从抽象（即数据结构）到具体（即具体实现）的过程。熟练地掌握这两个过程是数据结构课程在专业技能培养方面的基本目标。

1.1.2 为什么要学习数据结构

在计算机发展的初期，人们使用计算机的目的主要是处理数值计算问题。当使用计算机来解决一个具体问题时，一般需要经过下列几个步骤：首先要从该具体问题中抽象出一个适当的数学模型，然后设计或选择一个解此数学模型的算法，最后编出程序进行调试、测试，直至得到最终的解答。由于早期所涉及的运算对象是简单的整型、实型或布尔类型数据，所以程序设计者的主要精力是集中于程序设计的技巧上，而无需重视数据结构。随着计算机应用领域的扩大和软、硬件的发展，非数值计算问题越来越显得重要。据统计，当今处理非数值计算性问题占用了 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加以描述。因此，解决这类问题的关键不再是数学分析和计算方法，而是要设计出合适的数据结构及相应的算法，才能有效地解决问题。下面所列举的几个例子就是属于这一类的具体问题。

例 1.1 学生信息检索系统。当需要查找某个学生的基本情况的时候；或者想查找某个专业或年级的学生的相关情况的时候，只要建立了相关的数据结构，按照某种算法编写相关程序，就可以实现计算机自动检索。由此，可以在学生信息检索系统中建立一张按学号顺序排列的学生信息表和分别按姓名、专业、年级顺序排列的索引表，如表 1-2 所示。由这四张表构成的文件便是学生信息检索的数学模型，计算机的主要操作便是按照某个特定要求（如给定姓名）对学生信息文件进行查询。

诸如此类的还有电话自动查号系统、考试查分系统、仓库库存管理系统等。在这类文档管理的数学模型中，计算机处理的对象之间通常存在着一种简单的线性关系，这类数学模型可称为线性的数据结构。

表 1-2 学生信息查询系统中的数据结构

(a) 学生信息表

学号	姓名	性别	专 业	年 级	...
000001	石 磊	男	计算机科学与技术	2000 级	...
000002	李 婷	女	经济管理	2000 级	...
010301	刘 芳	女	数学与应用数学	2001 级	...
010302	张大友	男	经济管理	2001 级	...
010303	石 心	男	计算机科学与技术	2001 级	...
020801	何 颖	女	计算机科学与技术	2002 级	...
020802	赵胜利	男	数学与应用数学	2002 级	...
020803	崔冬冬	男	经济管理	2002 级	...
030601	刘 芳	女	计算机科学与技术	2003 级	...
030602	魏一鸣	男	数学与应用数学	2003 级	...

(b) 姓名索引表

崔冬冬	8
何 颖	6
李 婷	2
刘 芳	3, 9
石 磊	1
石 心	5
魏一鸣	10
赵胜利	7
张大友	4

(c) 专业索引表

计算机科学与技术	1, 5, 6, 9
经济管理	2, 4, 8
数学与应用数学	3, 7, 10

(d) 年级索引表

2000 级	1, 2
2001 级	3, 4, 5
2002 级	6, 7, 8
2003 级	9, 10

例 1.2 计算机和人对弈问题。计算机之所以能够和人对弈是因为人们将对弈的策略已事先存入计算机。对弈的过程是在一定的规则下随机进行的，所以在对弈过程中所有可能发生的情况以及相应的对策都必须存入计算机。因此，在人机对弈问题中计算机处理的对象是对弈过程中可能出现的棋盘状态，称之为格局。

例如，图 1-1 (a) 所示的为井字棋的一个格局。井字棋由两人对弈，棋盘为 3×3 的方格，当一方的三个棋子占据同一行或同一列或同一对角线时为胜。图 1-1 (b) 为图 1-1 (a) 所示的格局派生出的 5 个格局，而从每一个新的格局又可派生出 4 个可能出现的格局，依此类推。因此，若将对弈开始到结束的过程中所有可能出现的格局都组织到一张图上，则可得到一棵倒长的“树”。“树根”是对弈开始之前的棋盘格局，而所有的“叶子”就是可能出现的结局，对弈的过程走的就是从树根到树叶的一条路径。“树”也是一种数据结构。

例 1.3 教学计划编排问题。一个教学计划包含许多课程的设置及安排等。在教学计划包含的许多课程之间，有些有先后次序的要求，有些则没有，如图 1-2 (a) 所示。这种各个课程之间的次序关系可以用一个称作“图”的数据结构来表示，如图 1-2 (b) 所示。图中的每

个顶点表示一门课程，有向边表示课程进行的顺序。如果从顶点 C_i 到 C_j 之间存在有向边 $\langle C_i, C_j \rangle$ ，则表示课程 i 必须先于课程 j 进行。

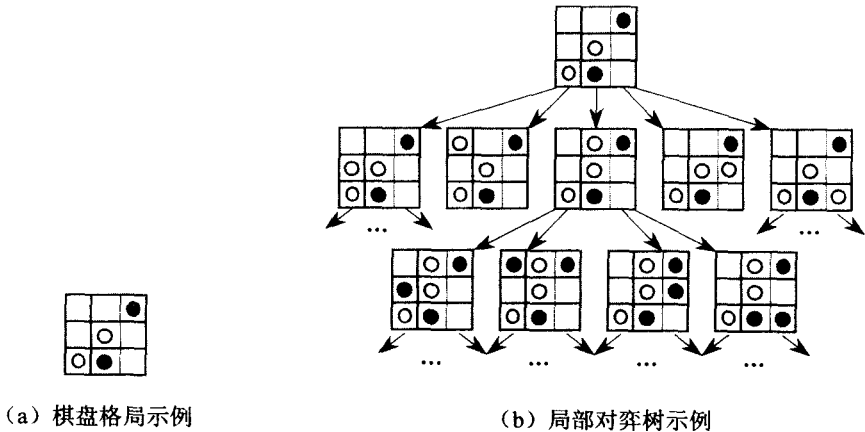
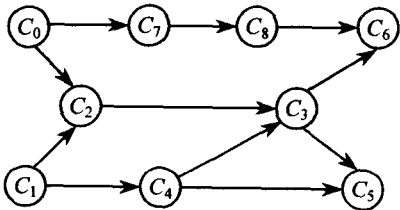


图 1-1 井字棋对弈示例

课程编号	课程名称	先修课程
C_0	高等数学	无
C_1	计算机导论	无
C_2	离散数学	C_0, C_1
C_3	数据结构	C_2, C_4
C_4	程序设计语言	C_1
C_5	编译原理	C_3, C_4
C_6	操作系统	C_3, C_8
C_7	电子线路基础	C_0
C_8	计算机组成原理	C_7

(a) 计算机专业的部分课程设置



(b) 表示课程之间优先关系的有向图

图 1-2 教学计划编排问题的数据结构

由以上 3 个例子可见，描述这类非数值计算问题的数学模型不再是数学方程，而是诸如表、树和图之类的数据结构。因此，可以说数据结构是一门研究非数值计算的程序设计问题中所出现的计算机操作对象以及它们之间的关系和操作等问题的学科。

学习数据结构的目的是为了了解计算机处理对象的特性，学会将实际问题中所涉及的处理对象在计算机中表示出来并对它们进行处理。

1.1.3 基本概念和术语

在系统地学习数据结构知识之前,先对一些基本概念和术语赋予确切的含义。

数据(Data)是对客观事物的符号表示,是信息的载体。在计算机科学中,所谓数据就是指能够被计算机识别、存储和加工处理的对象,是所有能输入到计算机并被计算机程序处理的符号的总称。数据可以是数值数据,也可以是非数值数据。数值数据是一些整数、实数或复数,主要用于工程计算、科学计算和商务处理等;非数值数据包括字符、文字、图形、图像、语音等。

数据元素(Data Element)是数据的基本单位,在计算机程序中通常被作为一个整体进行考虑和处理。在不同的情况下,数据元素又可称为元素、结点、顶点、记录等。例如,学生信息检索系统中学生信息表中的一个记录、人机对弈问题中状态树的一个棋盘格局、教学计划编排图中的一个顶点等,都被称为一个数据元素。

数据项(Data Item)有时一个数据元素可由若干个数据项组成。例如,学生信息检索系统中学生信息表的每一个数据元素就是一个学生记录,它包括学生的学号、姓名、性别、专业等数据项。数据项是数据不可分割的最小单位。通常,在解决实际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。

数据对象(Data Object)或**数据元素类(Data Element Class)**是具有相同性质的数据元素的集合。在某个具体问题中,数据元素都具有相同的性质(元素值不一定相等),属于同一数据对象(数据元素类)。数据元素是数据对象或数据元素类的一个实例。例如,在教学计划编排图中,所有的顶点同属于一个数据对象。顶点 C_1 和顶点 C_2 各自代表一门课程,是该数据对象或数据元素类中的两个实例,其数据元素的值分别为“计算机导论”和“离散数学”。

数据的逻辑结构(Data Logical Structure)是指元素和元素之间的逻辑关系。在任何问题中,相关数据元素集合中的元素之间都不会是孤立的,在它们之间都存在着这样或那样的特定关系,这种数据元素之间的特定关系称为数据的逻辑结构,简称数据结构。根据数据元素间关系的不同特性,通常有下列4类基本的数据结构。

(1) **集合结构**。结构中的数据元素间的特定关系是同属于一个集合。集合是元素关系极为松散的一种结构。

(2) **线性结构**。结构中的数据元素之间存在着一个对一个的关系。二维表就是典型的线性结构。

(3) **树型结构**。结构中的数据元素之间存在着一个对多个的关系。

(4) **图型结构**。结构中的数据元素之间存在着多个对多个的关系,图形结构也称作网状结构。图1-3为表示上述4类基本结构的示意。

在形式上,数据的逻辑结构通常可以采用一个二元组

$$S = (D, R)$$

来表示,其中, D 是数据元素的有限集, R 是 D 上关系的有限集。

例1.4 若有5个人,分别记为 a 、 b 、 c 、 d 、 e 。其中 a 是 b 的父亲, b 是 c 的父亲, c 是 d 的父亲, d 是 e 的父亲。如果只讨论他们之间所存在的父子关系,则可以用下面的二元组形式化地表达为:

$$S = (P, R)$$