



现代远程教育
精品教材

数据结构

(C++ 语言版)

李合龙 曹晓叶 韦 婷
张 芩 肖南峰 编

武科大图书馆



A1072149

 高等教育出版社
HIGHER EDUCATION PRESS

数据结构

(C++ 语言版)

Shuju Jiegou (C++ Yuyan ban)

李合龙 曹晓叶 韦 婷 编
张 芩 肖南峰



武科大图书馆



A1072149



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内 容 提 要

本书是高等学校计算机基础及应用网络教育教材。全书共分 10 章, 主要内容包括: 绪论、线性表、栈与队列、串、多维数组与广义表、树与二叉树、图、查找表、排序、文件。在每一章中, 都对相应的数据结构的描述、存储结构、基本操作、综合算法做了全面、深入的阐述, 在每一章也都列举了应用相应的数据结构进行算法设计的典型例子, 并且都对该章的基本内容、学习要点、具体要求、重点和难点进行了简要地归纳和总结。另外, 作者精心地为每一章选编了一些习题。本书各章内容翔实, 算法和例题典型, 并给出了对应的 VC++6.0 源程序代码。

本书不仅可作为高等学校继续教育学院、网络教育学院、公开学院计算机专业相关课程本科学生的教材, 也适合作为广大工程技术人员和自学考试人员的参考书。

图书在版编目(CIP)数据

数据结构: C++语言版/李合龙等编. —北京:
高等教育出版社, 2010.9

ISBN 978-7-04-030667-5

I. ①数… II. ①李… III. ①数据结构②C 语言—程
序设计 IV. ①TP311.12②TP312

中国版本图书馆 CIP 数据核字(2010)第 168583 号

策划编辑	吴 勇 罗娅妮	责任编辑	张海波	封面设计	张志奇
版式设计	韩璐儿	责任校对	王效珍	责任印制	毛斯璐

出版发行	高等教育出版社	购书热线	010-58581118
社 址	北京市西城区德外大街 4 号	免费咨询	800-810-0598
邮政编码	100120	网 址	http://www.hep.edu.cn
			http://www.hep.com.cn
		网上订购	http://www.landaco.com
经 销	蓝色畅想图书发行有限公司		http://www.landaco.com.cn
印 刷	北京北苑印刷有限责任公司	畅想教育	http://www.widedu.com
开 本	787×960 1/16	版 次	2010 年 9 月第 1 版
印 张	23	印 次	2010 年 9 月第 1 次印刷
字 数	350 000	定 价	44.00 元(含光盘)

本书如有缺页、倒页、脱页等质量问题, 请到所购图书销售部门联系调换。

版权所有 侵权必究

物料号 30667-00

前 言

“数据结构”是高等学校继续教育学院、网络教育学院、公开学院计算机专业本科学生的一门重要的专业基础课程。本课程主要讲授数据的逻辑结构、存储结构、基本运算、运算实现、算法设计等方面的内容；使学生对线性表、栈、队列、串、数组、树及二叉树、无向图及有向图、静态及动态查找表、文件等各种数据结构有深刻的理解，对各种常见的排序方法有深入的了解。在此基础上，还要求学生系统地掌握在不同的存储结构上利用上述数据结构进行算法设计的方法和技巧。因此，它是一门理论性和实践性都很强的课程。

为了帮助继续教育学院、网络教育学院、公开学院计算机专业的学生更好地掌握该课程的知识，提高算法设计和编程上机的能力，急需为“数据结构”课程编写一本基础扎实、知识面广、适应性强的教材。为此，在华南理工大学继续教育学院、网络教育学院、公开学院精品课程建设基金的资助下，我们编写了这本《数据结构（C++语言版）》教材，主要目的就是加强基础、拓宽知识面、增强适应性，以便使学生能够更深入地理解教材内容，培养和掌握良好的算法设计与程序实现的技能，以及解决实际问题的能力。

本书共分10章，主要内容包括：绪论、线性表、栈与队列、串、多维数组与广义表、树与二叉树、图、查找表、排序、文件。在每一章中，都对相应的数据结构的描述、存储结构、基本操作、综合算法做了全面、深入的阐述，每一章都对基本内容、学习要点、具体要求、重点难点进行了简要地归纳和总结。另外，作者还选编了大量精心设计的习题。本书各章的内容翔实，算法和例题典型，并给出了对应的VC++ 6.0源程序代码。

本书不仅可作为高等学校继续教育学院、网络教育学院、公开学院计算机专业相关课程本科学生的教材，也适合作为广大工程技术人员和自学考试

人员的参考书。

曹晓叶编写了第1章和第2章,李合龙编写了第3章、第8章、第9章、第10章,韦婷编写了第4章、第5章、第6章,张芩编写了第7章,肖南峰修订了全部章节;朱志飞和吕建明校对了全部章节和习题。在本书的编写过程中,华南理工大学“智能计算机科研团队”的多位教师也提出了许多的宝贵意见,我们在此向他们和华南理工大学继续教育学院、网络教育学院、公开学院精品课程建设基金表示衷心的感谢。由于作者水平有限,书中难免会存在错误,因此热忱地欢迎广大读者批评指正。

编者

2010年4月

于华南理工大学

目录 CONTENTS

第 1 章 绪论	1
1.1 什么是数据结构	2
1.2 基本概念	5
1.2.1 数据的逻辑结构	6
1.2.2 数据的存储结构	7
1.2.3 数据的运算	8
1.3 数据类型和抽象数据类型	9
1.4 算法和算法分析	11
1.4.1 算法的描述	11
1.4.2 算法设计的要求	12
1.4.3 算法分析	12
本章总结	15
习题 1	15
第 2 章 线性表	19
2.1 线性表的类型定义	20
2.1.1 线性表的定义	20

2.1.2 线性表的抽象数据类型	20
2.2 线性表的顺序存储	23
2.2.1 顺序表	23
2.2.2 基本操作的实现	25
2.3 线性表的链式存储	31
2.3.1 线性链表	31
2.3.2 单链表基本操作的实现	34
2.3.3 静态链表	39
2.3.4 循环链表	39
2.3.5 双向链表	40
2.4 线性表的应用	44
本章总结	50
习题2	50
第3章 栈与队列	54
3.1 栈	55
3.1.1 栈的基本概念	55
3.1.2 栈的抽象数据类型	57
3.1.3 栈的存储实现和运算实现	58
3.2 栈的应用举例	64
3.3 栈与递归的实现	71
3.3.1 递归的概念	71
3.3.2 递归过程和运行时栈	76
3.3.3 递归算法的效率分析	78
3.3.4 设计举例	79
3.4 队列	83
3.4.1 队列的定义及基本运算	83
3.4.2 队列的存储实现及运算实现	84
3.4.3 队列的应用	91
3.4.4 优先级队列	93

本章总结	93
习题 3	94
第 4 章 串	96
4.1 串的逻辑结构	97
4.1.1 串的基本概念	97
4.1.2 串的抽象数据类型	97
4.1.3 串的比较	99
4.1.4 常用的 C++ 串函数	100
4.2 串的存储结构	102
4.2.1 顺序存储结构	102
4.2.2 链式存储结构	104
4.3 串的模式匹配	106
4.3.1 简单串模式匹配算法	107
4.3.2 无回溯的模式匹配算法	109
4.4 串的应用——文本编辑	110
本章总结	111
习题 4	112
第 5 章 多维数组与广义表	114
5.1 数组	115
5.1.1 数组的基本概念	115
5.1.2 数组的抽象数据类型	116
5.1.3 数组的顺序存储结构和寻址公式	117
5.2 矩阵的压缩存储	120
5.2.1 特殊矩阵	121
5.2.2 稀疏矩阵	125
5.3 十字链表	136
5.3.1 十字链表存储结构	136
5.3.2 十字链表的实现	137
5.4 广义表	143

5.4.1 广义表的基本概念	143
5.4.2 广义表的存储结构	146
本章总结	151
习题 5	151
第 6 章 树与二叉树	154
6.1 树	155
6.1.1 树的基本概念	155
6.1.2 树的抽象数据类型	156
6.2 树的存储结构与遍历	158
6.2.1 树的存储结构	158
6.2.2 树和森林的遍历	163
6.3 二叉树	164
6.3.1 二叉树的基本概念	164
6.3.2 二叉树的性质	167
6.4 二叉树的存储结构	169
6.4.1 顺序存储结构	169
6.4.2 链式存储结构	170
6.5 二叉树的实现	172
6.5.1 二叉树的类定义	172
6.5.2 基本操作的实现	174
6.6 二叉树的遍历	179
6.6.1 遍历二叉树的定义	179
6.6.2 遍历二叉树的递归算法	180
6.6.3 遍历二叉树的非递归算法	182
6.7 二叉树的线索化	186
6.7.1 线索二叉树的定义	186
6.7.2 线索二叉树的类定义	187
6.7.3 基本操作的实现	189
6.8 树和森林与二叉树的转换	192

6.8.1 树与二叉树的转换	192
6.8.2 森林与二叉树的转换	194
6.9 哈夫曼树及其应用	197
6.9.1 哈夫曼树	197
6.9.2 哈夫曼编码	199
本章总结	205
习题 6	206
第 7 章 图	209
7.1 图的基本概念与操作	210
7.1.1 图的基本概念	210
7.1.2 图的抽象数据类型	213
7.2 图的存储结构	215
7.2.1 邻接矩阵法	215
7.2.2 邻接表	218
7.2.3 十字链表 (有向图)	220
7.2.4 邻接多重表 (无向图)	222
7.3 图的遍历	224
7.3.1 概述	224
7.3.2 深度优先搜索	225
7.3.3 广度优先搜索	227
7.3.4 图的 (强) 连通分量	228
7.4 最小生成树	229
7.4.1 Prim 算法	230
7.4.2 Kruskal 算法	232
7.5 有向无环图及其应用	234
7.5.1 拓扑排序与 AOV 网	234
7.5.2 关键路径与 AOE 网	237
7.6 最短路径	242
7.6.1 单源点到其余各顶点的最短路径	243

7.6.2 所有顶点对之间的最短路径	246
本章总结	249
习题7	249
第8章 查找表	252
8.1 基本概念	253
8.2 静态查找表	254
8.2.1 顺序表上的查找	256
8.2.2 有序表上的二分查找	259
8.2.3 索引顺序表上的查找	262
8.3 动态查找表	264
8.3.1 二叉排序树	264
8.3.2 平衡二叉树	273
8.3.3 B 树	276
8.3.4 B + 树	281
8.4 哈希表	284
8.4.1 哈希表的基本概念	284
8.4.2 哈希函数构造方法	286
8.4.3 哈希冲突解决方法	290
8.4.4 哈希表类设计	292
本章总结	297
习题8	297
第9章 排序	300
9.1 基本概念	301
9.2 插入排序	303
9.2.1 直接插入排序	304
9.2.2 希尔排序	307
9.3 交换排序	309
9.3.1 冒泡排序	309
9.3.2 快速排序	311

9.4 选择排序	314
9.4.1 直接选择排序	315
9.4.2 堆排序	317
9.5 归并排序	322
9.6 基数排序	324
9.7 各种内部排序算法的比较	328
9.8 外部排序简介	329
本章总结	331
习题 9	331
第 10 章 文件	334
10.1 基本概念	335
10.1.1 文件及其类别	335
10.1.2 文件的逻辑结构和物理结构	336
10.1.3 外存储器简介	338
10.2 顺序文件	338
10.3 索引文件	340
10.4 索引顺序文件	343
10.4.1 ISAM 文件	343
10.4.2 VSAM 文件	346
10.5 散列文件	348
10.6 多关键字文件	350
10.6.1 多重表文件	350
10.6.2 倒排文件	351
本章总结	352
习题 10	353
参考文献	354

第 1 章 绪 论

1946 年世界上第一台计算机诞生,至今已有 60 多年的历史。在这期间,计算机的发展和应用已经渗透到了人类社会的各个领域,计算机加工和处理的对象也从纯粹的数值发展到了字符、图像、声音等各种具有一定结构的数据。为了更好地进行程序设计,提高计算机在解决复杂问题时的处理效率,研究数据的特性和数据之间存在的关系至关重要。“数据结构”作为计算机科学与技术领域中的一项专业基础课,主要研究数据的特性和数据之间的关系,以及如何在计算机中有效地存取数据和处理数据。因此,“数据结构”是设计和实现编译程序、操作系统、数据库系统和大型应用程序等的重要基础,它也是介于数学、计算机硬件和计算机软件之间的一门核心课程,并将随着人们对各个领域中的计算问题的不断深入研究而继续发展。

1.1 什么是数据结构

使用计算机解决具体问题，一般需要经过以下几个步骤：首先从具体问题抽象出一个适当的数学模型，然后设计或寻找一个能求解此数学模型的算法，最后编写程序并进行调试、测试直至得到最终解答。在计算机发展的初期，人们使用计算机主要是进行科学计算，所涉及的数据对象比较单纯，程序设计以算法为中心，而无须重视数据结构。随着计算机应用领域的扩大和硬件的发展，非数值计算问题显得越来越重要。这类问题涉及的数据结构更为复杂，无法用简单的数学方程加以描述。为了编写出一个好的程序，必须分析待处理的对象的特性以及各处理对象之间存在的关系，设计出合理的数据结构。

例 1-1 学生信息检索自动化问题。

如果我们需要查找某个学生的有关情况，或者想查询某个专业或年级的学生的有关情况，只要建立了相关的数据结构，按照某种算法编写了相关程序，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排序的学生信息表和分别按姓名、专业、年级排序的索引表，如图 1.1 所示。由这 4 张表构成的文件便是学生信息检索的数学模型，计算机的

记录号	学号	姓名	性别	专业	年级
001	08101	张雨	女	计算机科学与技术	2008 级
002	08105	李鹏	男	计算机软件	2008 级
003	09103	何文	男	信息安全	2009 级
...	...	...	...	...	...

姓名	记录号
何文	003
李鹏	002
张雨	001
...	...

年级	记录号
2008 级	001, 002
2009 级	003
...	...

专业	记录号
计算机科学与技术	001
计算机软件	002
信息安全	003
...	...

图 1.1 学生信息查询系统文件示例

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。诸如此类的还有书目检索系统、电话号码查询系统等。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最简单的线性关系，这类数学模型可称为线性的数据结构。

例 1-2 四皇后问题。

在一个 4×4 的棋盘上放置 4 个皇后，使得任意两个皇后在行、列和斜方向上都不在一条线上。在四皇后问题中，处理过程不是根据某种确定的计算法则，而是利用试探和回溯的探索技术来求解。为了求得布局合理，在计算机中要存储布局的当前状态。从最初的布局状态开始，一步步地进行试探，每试探一步形成一个新的状态，整个试探过程形成了一棵隐含的状态树，如图 1.2 所示。回溯法求解过程实质上就是一个遍历状态树的过程。在

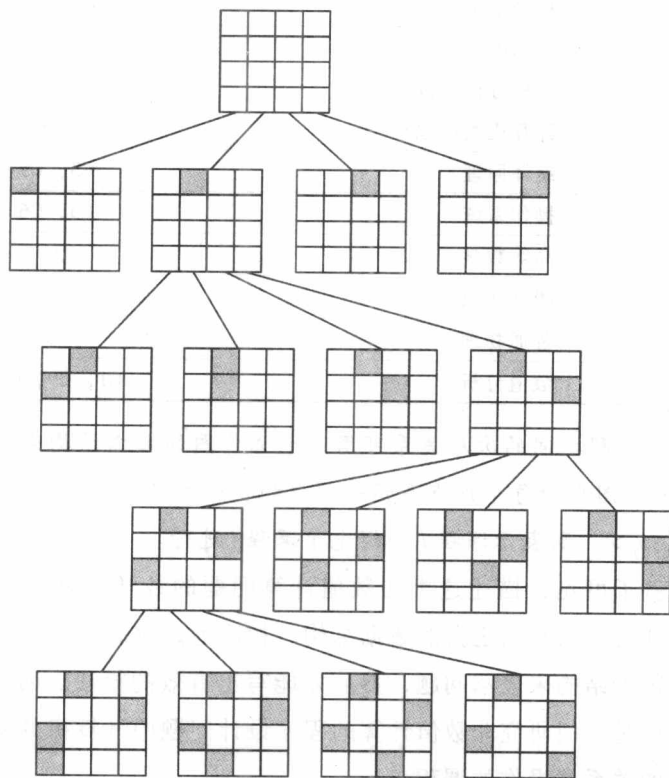


图 1.2 四皇后问题中隐含的状态树

这个问题中所出现的树也是一种数据结构，它可以应用在许多非数值计算的问题中。

例 1-3 制定教学计划问题。

一个教学计划包含许多课程，在制定教学计划时，需要考虑各门课程的开设顺序。有些课程需要先导课程，有些课程则不需要，而有些课程又是其他课程的先导课程。例如，计算机专业课程的开设情况如表 1.1 所示。

表 1.1 计算机专业学生的必修课程

课程编号	课程名称	需要的先导课程编号
C1	程序设计基础	无
C2	离散数学	C1
C3	数据结构	C1, C2
C4	汇编语言	C1
C5	算法分析与设计	C4
C6	计算机组成原理	C11
C7	编译原理	C3, C5
C8	操作系统	C3, C6
C9	高等数学	无
C10	线性代数	C9
C11	普通物理	C9
C12	数值分析	C1, C9, C10

这种各个课程之间的次序关系可用一个称作图的数据结构来表示，如图 1.3 所示。有向图中的每个顶点表示一门课程，如果从顶点 v_i 到 v_j 之间存在有向边 $\langle v_i, v_j \rangle$ ，则表示课程 i 必须先于课程 j 进行。

由上述例子可见，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。相应地，解决问题的关键步骤是设计合适的数据结构来表示问题，然后才能写出有效的算法。因此，简单来说，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课程。

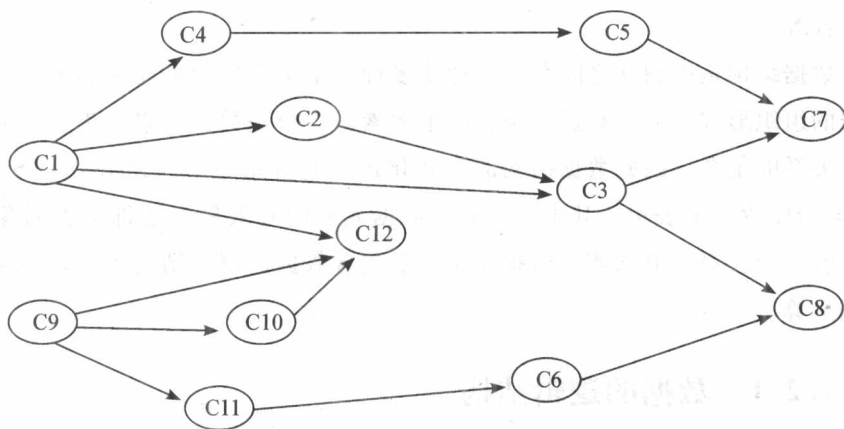


图 1.3 制定教学计划问题的数据结构

1.2 基本概念

在系统地学习数据结构知识之前，我们将对一些概念和术语赋以确定的含义，这些概念和术语将在以后的章节中用到。

数据是信息的载体，是客观事物的符号表示。数据能够被计算机识别、存取和处理，数据也是计算机程序加工和处理的“原料”。例如，实数、字符串、图像和声音等都称为数据。

数据项是具有独立含义的最小标识单位。例如，字段、域、属性等。

数据元素是数据的基本单位。在计算机程序中通常作为一个整体进行考虑和处理。一个数据元素可由若干个数据项组成，按其组成可分为简单型数据元素和复杂型数据元素两种。简单型数据元素由一个数据项组成，复杂型数据元素由多个数据项组成，它通常反映一个事物的多方面信息。例如，例 1-1 学生信息表中的学生信息为一个数据元素，而学生信息中的每一项（如姓名、专业等）为一个数据项。

数据对象是性质相同的数据元素的集合，是数据的一个子集。例如，26 个英文字母构成的字符集合，一个学校全体学生或教师构成的学生集合或教