

三导丛书

数 据 结 构

(C 语言版)

导 教 · 导 学 · 导 考

姚 群 夏清国 编
李 曼 杨 鸽

西北工业大学出版社

【内 容 简 介】 本书是为配合严蔚敏、吴伟民编著的《数据结构》(C 语言版)而编写的辅助教材。书中首先给出重点内容提要 and 知识结构图;其次,从历次本科课程考试及研究生考试中例举出常见题型,并进行分析解答;最后,在每章后面给出了学习效果测试及参考答案,供读者自我测试学习效果。本书对学习数据结构课程的读者来说,是一本针对性很强的辅助教材。

图书在版编目 (CIP) 数据

数据结构 (C 语言版) 导教 · 导学 · 导考/姚群,夏清国编. —
西安:西北工业大学出版社,2004.6

(三导丛书)

ISBN 7 - 5612 - 1754 - 4

.数... . 姚... 夏... .数据结构-高等学校-教学参
考资料 . TP311 . 12

中国版本图书馆 CIP 数据核字 (2004) 第 016540 号

出版发行:西北工业大学出版社

通信地址:西安市友谊西路 127 号 邮编:710072 电话:(029) 88493844

网 址:www.nwpup.com

印 刷 者:陕西向阳印务有限公司

开 本:850 mm × 1 168 mm 1/32

印 张:12.562 5

字 数:412 千字

版 次:2004 年 6 月第 1 版 2005 年 1 月第 2 次印刷

印 数:6 001 ~ 10 000 册

定 价:15.00 元

前 言

数据结构课程是理工科院校计算机专业必修的一门专业基础课，对初学者来说是比较困难、比较抽象的一门课程。为了满足广大读者学习的需求，我们根据多年的教学经验编写了本书。

本书是为配合严蔚敏、吴伟民编著的《数据结构》（C 语言版）而编写的辅助教材，按其章节顺序，分为 12 章，每章按 4 个模块编写：

一、重点内容提要 列出了每章的基本概念、基本术语、数据结构的存储描述、算法及算法分析。

二、重点知识结构图 用框图形式简洁集中地列出各章的知识点。

三、常见题型及典型题精解 根据本科课程考试和考研要求，总结出每章的考点，精选出常见及典型题目，进行详细分析解答。

四、学习效果测试及参考答案 这部分是为读者检查学习效果和应试能力而设计的，通过自测，读者可以进一步加深对所学内容的理解，增强解题能力。

本书从指导课程教学和考试的角度，通过对大量涉及内容广、常见及经典的题型提供算法的思想，并

对算法进行分析，提供了“数据结构”的解题方法、解题规律和解题技巧。这对于提高读者分析问题的能力，理解基本要领和理论，开拓解题思路，会起到良好的效果。对于学习效果测试题，希望读者在学习过程中先独立思考，自己动手解题，然后再对照参考答案检查，不要依赖于解答。

由于水平有限，书中疏漏与不妥之处，恳请读者批评指正。

编 者

2004 年 2 月

目 录

第 1 章	绪论	1
1.1	重点内容提要	1
1.1.1	基本概念和术语	1
1.1.2	算法和算法分析	3
1.2	重点知识结构图	4
1.3	常见题型及典型题精解	4
1.4	学习效果测试及参考答案	9
1.4.1	单项选择题	9
1.4.2	填空题	11
1.4.3	简答题	12
	参考答案	14
第 2 章	线性表	18
2.1	重点内容提要	18
2.1.1	线性表	18
2.1.2	线性表的顺序表示与实现	19
2.1.3	线性表的链式表示与实现	22
2.2	重点知识结构图	28
2.3	常见题型及典型题精解	28

2.4	学习效果测试及参考答案.....	38
2.4.1	单项选择题.....	38
2.4.2	填空题.....	39
2.4.3	简答题.....	41
2.4.4	算法设计题.....	41
	参考答案	44
第3章	栈和队列	64
3.1	重点内容提要.....	64
3.1.1	栈.....	64
3.1.2	队列.....	67
3.2	重点知识结构图.....	72
3.3	常见题型及典型题精解.....	72
3.4	学习效果测试及参考答案.....	80
3.4.1	单项选择题.....	80
3.4.2	填空题.....	82
3.4.3	简答题.....	83
3.4.4	算法设计题.....	84
	参考答案	84
第4章	串	97
4.1	重点内容提要.....	97
4.1.1	串的基本概念.....	97
4.1.2	串的实现与表示.....	98
4.1.3	串的模式匹配算法	108
4.2	重点知识结构图	111
4.3	常见题型及典型题精解	111
4.4	学习效果测试及参考答案	120

4.4.1	填空题	120
4.4.2	简答题	121
4.4.3	算法设计题	121
	参考答案.....	122
第 5 章	数组和广义表.....	129
5.1	重点内容提要	129
5.1.1	数组的定义	129
5.1.2	数组的存储结构	129
5.1.3	广义表的定义	138
5.1.4	广义表的存储结构	139
5.1.5	广义表的递归算法	141
5.2	重点知识结构图	144
5.3	常见题型及典型题精解	144
5.4	学习效果测试及参考答案	152
5.4.1	单项选择题	152
5.4.2	填空题	153
5.4.3	简答题	154
5.4.4	算法设计题	155
	参考答案.....	156
第 6 章	树和二叉树.....	170
6.1	重点内容提要	170
6.1.1	树	170
6.1.2	二叉树	171
6.1.3	遍历二叉树和线索二叉树	178
6.1.4	树和森林	180
6.1.5	赫夫曼树及应用	183

6.2	重点知识结构图	185
6.3	常见题型及典型题精解	185
6.4	学习效果测试及参考答案	197
6.4.1	单项选择题	197
6.4.2	填空题	199
6.4.3	简答题	200
6.4.4	算法设计题	202
	参考答案.....	204
第7章	图.....	218
7.1	重点内容提要	218
7.1.1	图的基本概念	218
7.1.2	图的存储结构	219
7.1.3	图的遍历	221
7.1.4	图的连通性及最小生成树	223
7.1.5	有向无环图及其应用	225
7.1.6	最短路径	227
7.2	重点知识结构图	228
7.3	常见题型及典型题精解	228
7.4	学习效果测试及参考答案	242
7.4.1	单项选择题	242
7.4.2	填空题	245
7.4.3	简答题	247
7.4.4	算法设计题	251
	参考答案.....	251
第8章	动态存储管理.....	264
8.1	重点内容提要	264

8.1.1	基本概念	264
8.1.2	可利用空闲表及分配方法	265
8.1.3	边界标识法	266
8.1.4	伙伴系统	268
8.1.5	无用单元收集	269
8.1.6	存储紧缩	270
8.2	重点知识结构图	271
8.3	常见题型及典型题精解	271
8.4	学习效果测试及参考答案	277
	测试题.....	277
	参考答案.....	277
第 9 章	查找.....	279
9.1	重点内容提要	279
9.1.1	基本概念	279
9.1.2	静态查找表	279
9.1.3	动态查找表	282
9.1.4	哈希表	287
9.2	重点知识结构图	290
9.3	常见题型及典型题精解	290
9.4	学习效果测试及参考答案	299
9.4.1	单项选择题	299
9.4.2	填空题	301
9.4.3	简答题	302
9.4.4	算法设计题	303
	参考答案.....	304

第 10 章	内部排序	318
10.1	重点内容提要	318
10.1.1	排序的基本概念	318
10.1.2	插入排序	318
10.1.3	交换排序	321
10.1.4	选择排序	323
10.1.5	归并排序	325
10.1.6	基数排序	327
10.2	重点知识结构图	329
10.3	常见题型及典型题精解	329
10.4	学习效果测试及参考答案	345
10.4.1	单项选择题	345
10.4.2	填空题	347
10.4.3	简答题	348
10.4.4	算法设计题	349
	参考答案	349
第 11 章	外部排序	364
11.1	重点内容提要	364
11.1.1	外部排序方法	364
11.1.2	多路平衡归并的实现	364
11.1.3	置换 - 选择排序	365
11.1.4	最佳归并树	366
11.2	重点知识结构图	366
11.3	常见题型及典型题精解	367
11.4	学习效果测试及参考答案	368
	测试题	368

参考答案.....	369
第 12 章 文件	374
12.1 重点内容提要.....	374
12.1.1 文件的基本概念.....	374
12.1.2 顺序文件.....	375
12.1.3 索引文件.....	375
12.1.4 索引顺序文件.....	375
12.1.5 直接存取文件（散列文件）	377
12.1.6 多关键字文件.....	377
12.2 重点知识结构图.....	378
12.3 常见题型及典型题精解.....	378
12.4 学习效果测试及参考答案.....	383
12.4.1 单项选择题.....	383
12.4.2 填空题.....	385
12.4.3 简答题.....	385
参考答案.....	387

第 1 章 绪 论

1.1 重点内容提要

1.1.1 基本概念和术语

1 数据

数据是对客观事物的符号表示,在计算机科学中是指所有能输入到计算机中并被计算机程序处理的符号的总称。

2 数据元素

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。

3 数据项

数据项是数据的不可分割的最小单位。

4 数据对象

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。

5 数据结构

数据结构是相互之间存在一种或多种特定关系的数据元素的集合。数据结构包括三方面的内容:数据的逻辑结构、数据的存储结构和数据的运算。

6 数据的逻辑结构

数据的逻辑结构是指数据元素之间的逻辑关系,即从逻辑关系上描述数据。它与数据的存储无关,是独立于计算机的。

通常有下列四种基本结构。

(1) 集合

结构中的数据元素之间除了“同属于一个集合”的关系外,别无其它关系。

(2) 线性结构

结构中的数据元素之间存在一对一的关系。若结构是非空集,则有且仅有一个开始结点和一个终端结点,并且除开始结点无直接前趋和终端结点无直接后继外,其它所有结点都只有一个直接前趋和一个直接后继。

(3) 树形结构

结构中的数据元素之间存在一对多的关系。若结构是非空集,则除第一个结点外,其它所有结点都只有一个直接前趋,除叶子结点外,其它所有结点可能有多个直接后继。

(4) 图状结构或网状结构

结构中的数据元素之间存在多对多关系。若结构是非空集,所有结点都可能多个直接前趋和多个直接后继。

7 数据的存储结构

数据的存储结构是指数据元素及其关系在计算机存储器内的表示(也称为映像)。数据的存储结构是逻辑结构用计算机语言的实现,它依赖于计算机语言。

通常有下列四种存储映像方法。

(1) 顺序存储方法

该方法是把逻辑上相邻的结点存储在物理位置上相邻的存储单元里,结点间的逻辑关系由存储单元的邻接关系来体现,由此得到的存储结构称为顺序存储结构,通常顺序存储结构是借助于程序语言的数组来描述的。

(2) 链接存储方法

该方法不要求逻辑上相邻的结点在物理位置上也相邻,结点间的逻辑关系是由附加的指针字段表示的,由此得到的存储表示称为链式存储结构,通常要借助于程序语言的指针类型来描述它。

(3) 索引存储方法

该方法通常是在存储结点信息的同时,还建立附加的索引表。索引表中的每一项称为索引项,索引项的一般形式是:(关键字,地址)。其中关键字惟一标识结点,地址作为指向结点的指针。

(4) 散列存储方法

该方法的基本思想是根据结点的关键字直接计算出该结点的存储地址。

8 数据的运算

数据的运算是根据数据的逻辑结构上定义的操作算法,如检索、插入、删除、更新和排序等。

9 数据类型

(1) 原子类型

原子类型是其值不可再分的数据类型。

(2) 结构类型

结构类型是其值可以再分解为若干成分(分量)的数据类型。

(3) 抽象数据类型

抽象数据类型是抽象数据组织和与之相关的操作。

1.1.2 算法和算法分析

1 算法

算法是对特定问题求解步骤的一种描述,它是指令的有限序列,其中每条指令表示一个或多个操作。

算法有以下 5 个主要特征。

有穷性:一个算法必须总是(对任何合法的输入)在执行有穷步之后结束,且每一步都可在有穷时间内完成。

确定性:算法中每一条指令必须有确切的含义,确保不会产生二义性。并且,在任何条件下,算法只有惟一的一条执行路径,即对于相同的输入只能得出相同的输出。

可行性:一个算法是能行的,即算法中描述的操作都是可以通过已实现的基本运算执行有限次来实现的。

输入性:一个算法有零个或多个的输入。

输出性:一个算法有一个或多个的输出。

2 算法效率的度量

(1) 时间复杂度

一个语句的频度,是指该语句在算法中被重复执行的次数。算法中所有语句的频度之和记做 $T(n)$,它是该算法所求解问题规模 n 的函数。当问题的规模趋向无穷大时, $T(n)$ 的数量级称为渐近时间复杂度,简称为时间复杂度,记做 $T(n) = O(f(n))$ 。

算法的时间复杂度不仅仅依赖于问题的规模,也取决于输入实例的初始状态。一个问题的输入实例是满足问题陈述中所给出的限制和为计算该问题的解所需要的所有输入构成的。

最坏时间复杂度是指在最坏情况下算法的时间复杂度。

平均时间复杂度是指所有可能的输入实例均以等概率出现的情况下,算法的期望运行时间。

上述表达式中“ O ”的含义是 $T(n)$ 的数量级,其严格的数学定义是:若 $T(n)$ 和 $f(n)$ 是定义在正整数集合上的两个函数,则存在正的常数 C 和 n_0 ,使得当 $n \geq n_0$ 时,都满足 $0 \leq T(n) \leq C \times f(n)$ 。

一般总是考虑在最坏的情况下的时间复杂度,以保证算法的运行时间不会比它更长。

(2) 空间复杂度

算法的空间复杂度 $S(n)$, 定义为该算法所耗费的存储空间,它是问题规模 n 的函数。渐进空间复杂度也常常简称为空间复杂度,记作 $S(n) = O(f(n))$ 。

1 2 重点知识结构图

绪论 { 基本概念(数据、数据元素、数据项、数据对象、数据结构)
数据的逻辑结构(集合、线性结构、树形结构、图状结构)
数据的存储结构(顺序、链式、索引、散列存储结构)
数据类型(原子类型、结构类型、抽象数据类型)
算法(什么是算法、算法具有的特性)
算法效率的度量(时间复杂度、空间复杂度)

1 3 常见题型及典型题精解

例 1.1 逻辑结构和存储结构之间的关系?

【例题解答】 对于已经建立的逻辑结构是设计人员根据解题需要选定的数据组织形式,因此建立的机内表示应遵循选定的逻辑结构,所建立数据的机内表示称为数据存储结构。

例 1.2 常用的存储表示方法有哪几种?

【例题解答】 常用的存储表示方法有 4 种:

顺序存储方法:它是把逻辑上相邻的结点存储在物理位置相邻的存储单元里,结点的逻辑关系由存储单元的邻接关系来体现,由此得到的存储结构称为顺序存储结构。

链式存储方法:它不要求逻辑上相邻的结点在物理位置上亦相邻,结点之间的逻辑关系是由附加的指针字段表示的。由此得到的存储结构称为链式存储结构。

索引存储方法:除建立存储结点信息外,还建立附加的索引表来标识结点的地址。

散列存储方法:根据结点的关键字直接计算出该结点的存储地址。

例 1.3 设有数据逻辑结构为: $line = (D, R)$, 其中

$D = \{01, 02, 03, 04, 05, 06, 07, 08, 09, 10\}$

$R = \{r\}$

$r = \{ \langle 05, 01 \rangle, \langle 01, 03 \rangle, \langle 03, 08 \rangle, \langle 08, 02 \rangle, \langle 02, 07 \rangle, \langle 07, 04 \rangle, \langle 04, 06 \rangle, \langle 06, 09 \rangle, \langle 09, 10 \rangle \}$

试分析该数据结构属于哪种逻辑结构。

【例题解答】 对应的图形如图 1.1 所示。

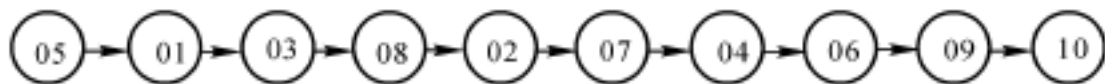


图 1.1 数据的线性结构示意图

在 $line$ 中,每个数据元素有且仅有一个直接前驱元素(除结构中第一个元素 05 外),有且仅有一个直接后继元素(除结构中最后一个元素 10 外)。这种数据结构的特点是数据元素之间的 1 对 1(1-1)关系,即线性关系,因此本题所给定的数据结构为线性结构。

例 1.4 设有数据逻辑结构为: $tree = (D, R)$, 其中

$D = \{01, 02, 03, 04, 05, 06, 07, 08, 09, 10\}$

$R = \{r\}$

$r = \{ \langle 01, 02 \rangle, \langle 01, 03 \rangle, \langle 01, 04 \rangle, \langle 02, 05 \rangle, \langle 02, 06 \rangle, \langle 03, 07 \rangle, \langle 03, 08 \rangle, \langle 03, 09 \rangle, \langle 04, 10 \rangle \}$

试分析该数据结构属于哪种逻辑结构。

【例题解答】 对应的图形如图 1.2 所示。

图 1.2 像倒着画的一棵树,在这棵树中,最上面的一个没有前驱只有后继,称做树根结点,最下面一层的只有前驱没有后继,称做树叶结点。在一棵树中,每个结点有且只有一个前驱结点(除树根结点外),但可以有任意多个后继结点

(树叶结点可看做具有 0 个后继结点)。这种数据结构的特点是数据元素之间的 1 对 N ($1 \leq N$) 关系 ($N \geq 0$), 即层次关系, 因此本题所给定的数据结构为树形结构。

例 1.5 设有数据逻辑结构为: $\text{graph} = (D, R)$, 其中

$$D = \{01, 02, 03, 04, 05, 06, 07\}$$

$$R = \{r\}$$

$$r = \{ \langle 01, 02 \rangle, \langle 02, 01 \rangle, \langle 01, 04 \rangle, \langle 04, 01 \rangle, \langle 02, 03 \rangle, \langle 03, 02 \rangle, \langle 02, 06 \rangle, \langle 06, 02 \rangle, \langle 02, 07 \rangle, \langle 07, 02 \rangle, \langle 03, 07 \rangle, \langle 07, 03 \rangle, \langle 04, 06 \rangle, \langle 06, 04 \rangle, \langle 05, 07 \rangle, \langle 07, 05 \rangle \}$$

试分析该数据结构属于哪种逻辑结构。

【例题解答】 对应的图形如图 1.3 所示。

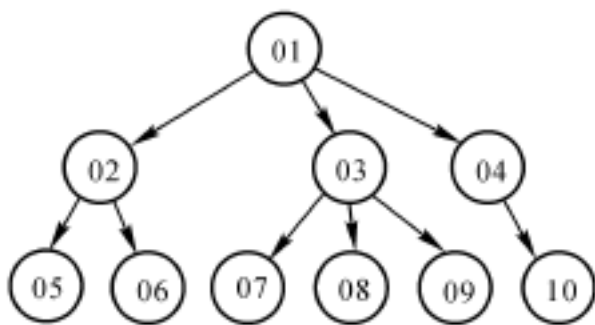


图 1.2 数据的树形结构示意图

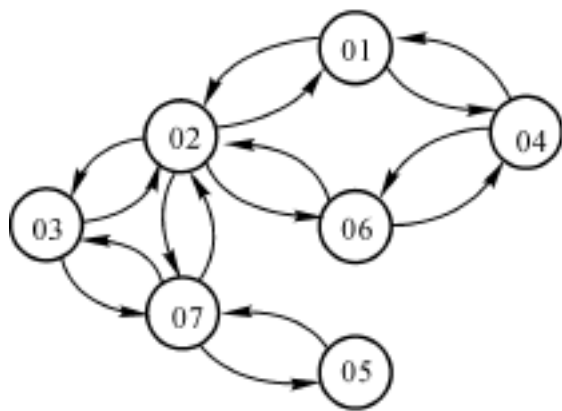


图 1.3 数据的图形结构示意图

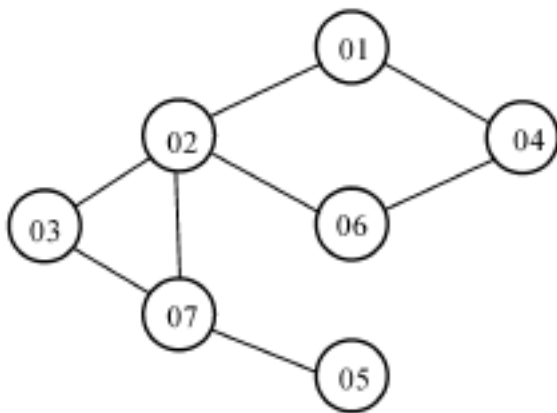


图 1.4 图 1.3 的等价表示

从图 1.3 可以看出, r 是 D 上的对称关系, 为了简化起见, 我们把 $\langle x, y \rangle$ 和 $\langle y, x \rangle$ 这两个对称序偶用一个无序对 (x, y) 或 (y, x) 来代替; 在图形表示中, 我们把 x 结点和 y 结点之间两条相反的有向边用一条无向边来代替。这样 r 关系可改写为:

$$r = \{(01, 02), (01, 04), (02, 03), (02, 06), (02, 07), (03, 07), (04, 06), (05, 07)\}$$