



浙江省普通高校“十三五”新形态教材

数据结构

——C语言描述

融媒体版

刘小晶 / 主 编

朱 蓉 杜卫锋 滕 姿 / 副主编



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

数据结构——C 语言描述(融媒体版)

主 编 刘小晶

副主编 朱 蓉 杜卫锋 滕 姿



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目(CIP)数据

数据结构:C语言描述:融媒体版 / 刘小晶主编.
—杭州:浙江大学出版社,2018.10
ISBN 978-7-308-18507-3

I. ①数… II. ①刘… 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. ①TP311.12②TP312.8

中国版本图书馆 CIP 数据核字(2018)第 182287 号

数据结构——C语言描述(融媒体版)

刘小晶 主编

责任编辑 何 瑜
责任校对 刘 郡
封面设计 春天书装
出版发行 浙江大学出版社
(杭州市天目山路 148 号 邮政编码 310007)
(网址:<http://www.zjupress.com>)
排 版 浙江时代出版服务有限公司
印 刷 浙江新华数码印务有限公司
开 本 787mm×1092mm 1/16
印 张 25.5
字 数 636 千
版 次 2018 年 10 月第 1 版 2018 年 10 月第 1 次印刷
书 号 ISBN 978-7-308-18507-3
定 价 66.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社市场运营中心联系方式:(0571) 88925591;<http://zjdxcs.tmall.com>

内容简介

本书是浙江省普通高校“十三五”新形态教材,内容涵盖了教育部高等学校计算机科学与技术教学指导委员会关于“高等学校计算机科学与技术本科专业规范”中制定的课程体系中的核心知识,并在紧扣考研大纲的前提下剔除了一些难度较大的内容。

本书共 9 章,采用 C 语言作为描述算法的语言,涵盖基本数据结构、排序和查找等主要内容。前 6 章是以逻辑结构为主线系统介绍线性表、栈与队列、串与数组、树和图等各种基本数据结构的特点、存储方式、运算原理、实现方法以及它们在现实中的典型应用;后 3 章是从实现方法的角度系统介绍各类排序、查找的算法,注重对不同数据结构和算法的比较分析及其应用。

本书条理清楚、内容精炼、重点突出,做到循序渐进、深入浅出。强调算法实现方法的分析,通过丰富、典型的实例来强化知识的实际应用,并且融合了互联网技术,将课程的重点、难点知识的讲解全部录制成微视频,并通过二维码的形式嵌入其中,让读者通过观看视频亦可轻松学习,从而使学习更加高效,同时也较好地保证了教材内容的可更新性和可扩展性。

本书可作为普通高等院校“数据结构”课程教材,也可作为工程技术和自学数据结构人员的参考读物。

前言

当今以 MOOC(“慕课”)为代表的在线教育在高等教育领域的迅速兴起,不仅是教育技术的革新,更是引发教育观念、教育体制、人才培养、教学方式的深度变革。传统的课堂教学模式及学习方式正发生重大变化,仅以纸质教材为媒介的课堂教学载体已不能适应当前的教学发展需要。因在“互联网+教育”时代,传统的纸质教材存在着明显的不足:成本高、携带不便、知识信息更新慢、知识容量有限等,而数字化的电子教材则能弥补这些不足,但完全数字化的电子教材在当前国情下,受教师的教学习惯、学生的学习和阅读习惯、网络条件、设备条件等制约,难以在所有高等学校或所有课程中全面使用。为此,在“互联网+”时代,传统纸质教材与数字化教学资源融合形成的新形态教材,已成为教材建设的一种新趋势,也是现实教育的迫切需求。

本教材为浙江省普通高校“十三五”首批新形态教材,它将纸质内容与数字媒体内容进行有机融合,与传统教材相比,无论是课程内容还是内容的呈现形式都进行了一定程度上的重构,使其充分发挥纸质教材和数字媒体的优势,实现知识由静态、抽象化向动态、形象化的转变。具体特点阐述如下:

1. 对教材内容的精心设计

对于知识体系相对固定的《数据结构》课程而言,作者在“互联网+”的思维下重新审视了课程的教学内容,对传统纸质教材的内容进行了重构。而且根据使用者的特点进行了一体化的设计,以确定哪些教学内容适合印在纸质教材上,哪些内容适合以数字媒体形式放在云端。具体设计原则是纸质教材内容必须涵盖教育部计算机科学与技术教指委关于“高等院校计算机科学与技术本科专业规范”中制定的课程体系的核心知识,并在紧扣当时计算机类考研大纲的前提下,把重点、难点、拓展资源做成微视频或其他数字媒体资源,以二维码的形式嵌入纸质教材中,并在所有内容设计上做到以下几点:

(1) 精简内容、强化基础、突出知识的应用性。针对普通高校学生的实际情况,把握“适用”与“够用”的尺度。做到把重点放在基础知识的介绍上,缩减了一些难度较大的内容。并强调其在实际问题中的应用性,充分体现了理论与应用背景的紧密结合。

(2) 理论叙述简洁明了、重点突出、应用实例丰富。各章节都以基本概念为切入点,逐步介绍其特点和基本操作的实现,然后通过应用实例来讲述如何运用所学的原理和方法来解决实际问题,最后附有小结、习题及解答,便于学习总结和提高。这些内容做到“环环相扣,层层推进”,充分体现解析法的精髓,达到通俗易懂、由浅入深的效果,突显了教材对培养读者的算法分析与设计能力,以及知识迁移的能力。

2. 对教材内容呈现形式的精心设计

本教材对传统纸质教材内容的呈现形式也进行了重构,它除了呈现传统纸质的文本内容之外,还对课程中的重点、难点和拓展知识进行了分析与设计,并采用二维码的方式嵌入了对应的数字资源,读者通过使用移动终端扫描书中的二维码,就可获得以视频等媒体形式呈现的教学内容或链接到对应的互动平台,不但可突显其重点、难点知识的讲解,增强学习

者的趣味性和互动性,从而提高课程教学效果;更可适应不同学习方式喜好的人群使用,实现随时随地学习、交流与互动;而且这些内容的更新不再受制于纸质教材的改版更新,可以根据需要随时更新和拓展资源内容。

本教材的内容由9章组成:①绪论,内容主要涉及数据结构课程讨论的内容;数据结构的常用术语及基本概念;数据结构算法的描述和算法分析方法以及相关约定等。②线性表,内容主要涉及线性表的抽象数据类型定义;线性表类型在顺序存储和链式存储两种存储结构下的实现方法以及线性表的应用等。③栈与队列,内容主要涉及栈与队列的抽象数据类型定义;栈与队列在顺序存储和链式存储结构下其基本操作的实现方法以及栈与队列的应用等。④串与数组,内容主要涉及串的基本概念;串的存储结构;串的基本操作实现;数组的定义、基本操作、存储结构以及矩阵的压缩存储和数组的应用等。⑤树与二叉树,内容主要涉及树与二叉树的基本概念和存储结构;树、二叉树和森林的遍历;树、二叉树与森林之间的转换方法以及哈夫曼树与哈夫曼编码等。⑥图,内容主要涉及图的基本概念;图的邻接矩阵和邻接表两种最基本的存储结构;图的广度优先搜索和深度优先搜索两种最基本的遍历方法;有关最小生成树的克鲁斯卡尔(Kruskal)和普里姆(Prim)两种实现算法;拓扑排序和求关键路径等。⑦内排序,内容主要涉及排序的基本概念;常用内部排序(插入排序、交换排序、选择排序、归并排序和基数排序)方法的实现及性能分析以及各种内部排序方法的比较等。⑧外排序,内容主要涉及外部排序的方法;磁盘排序的信息存取、多路归并排序、选择排序和最优归并树等。⑨查找,内容主要涉及查找的基本概念;静态表查找(顺序查找、二分查找和分块查找)、动态表查找(二叉排序树、平衡二叉树、B树和红黑树)、哈希表查找的实现方法及性能分析等。

本教材是由课程组成员协作完成:第1~3章及第7~8章由刘小晶教授执笔,第5、9章由朱蓉教授执笔,第4、6章由杜卫锋副教授和滕姿讲师执笔。全书由刘小晶策划并统稿。

本教材在编写过程中,参阅了大量的参考资料,列于书目的参考文献中,在此谨向作者表达衷心的感谢。

由于作者学识所限,书中定有不足之处,敬请读者批评指正,提出宝贵意见,联系方式:lxjing@mail.zjxu.edu.cn。

目录

第 1 章 绪 论	1
1.1 数据结构课程讨论的内容	1
1.1.1 求解问题举例	2
1.1.2 本课程讨论的内容	4
1.2 基本概念与术语	5
1.2.1 数据与数据结构	5
1.2.2 数据类型	9
1.2.3 抽象数据类型	10
1.3 算 法	11
1.3.1 算法的基本概念	11
1.3.2 算法的描述	12
1.3.3 相关约定	14
1.4 算法分析	14
1.4.1 时间复杂度分析	15
1.4.2 空间复杂度分析	20
1.4.3 算法设计举例	21
小 结	25
习题一	26
第 2 章 线性表	28
2.1 线性表的类型定义	28
2.1.1 线性表的基本概念	28
2.1.2 线性表的抽象数据类型描述	29
2.2 线性表的顺序存储及其实现	32
2.2.1 顺序表的顺序存储	32
2.2.2 顺序表上基本操作的实现	35
2.2.3 顺序表应用举例	39

2.3 线性表的链式存储及实现	42
2.3.1 单链表的表示	42
2.3.2 单链表上基本操作的实现	43
2.3.3 单链表应用举例	51
2.3.4 其他链表	55
2.4 顺序表与链表的比较	61
2.5 线性表的应用举例	65
小 结	69
习题二	69
第3章 栈与队列	73
3.1 栈	74
3.1.1 栈的概念	74
3.1.2 栈的抽象数据类型描述	74
3.1.3 顺序栈及其基本操作的实现	75
3.1.4 链栈及其基本操作的实现	79
3.1.5 栈的应用	81
3.2 队 列	94
3.2.1 队列的概念	94
3.2.2 队列的抽象数据类型描述	94
3.2.3 顺序队列及其基本操作的实现	95
3.2.4 链队列及其基本操作的实现	100
3.2.5 其他队列	102
3.2.6 队列的应用	104
3.3 栈与队列的比较	107
3.4 栈与队列的综合应用举例	109
小 结	112
习题三	114
第4章 串与数组	118
4.1 串的类型定义	118
4.1.1 串的基本概念	118
4.1.2 串的抽象数据类型描述	119
4.2 串的存储结构	120
4.2.1 定长顺序存储表示	120
4.2.2 堆分配存储表示	124
4.2.3 块链存储表示	125
4.3 串的模式匹配操作	126
4.3.1 Brute-Force 模式匹配算法	126

4.3.2 KMP 模式匹配算法	128
4.4 串的应用举例	133
4.5 数组的概念及顺序存储结构	137
4.5.1 数组的基本概念	137
4.5.2 数组的抽象数据类型描述	138
4.5.3 数组的顺序存储结构	139
4.6 特殊矩阵的压缩存储	140
4.6.1 对称矩阵的压缩存储	140
4.6.2 三角矩阵的压缩存储	141
4.6.3 对角矩阵的压缩存储	141
4.7 稀疏矩阵的压缩存储	142
4.7.1 三元组顺序表	142
4.7.2 行逻辑链接的顺序表	147
4.7.3 稀疏矩阵的十字链表存储	151
4.8 数组的应用举例	152
小 结	154
习题四	155
第 5 章 树与二叉树	157
5.1 树的基本概念	158
5.2 二叉树	161
5.2.1 二叉树的基本概念	161
5.2.2 二叉树的性质	164
5.2.3 二叉树的存储结构	165
5.3 二叉树的遍历	168
5.3.1 二叉树的遍历方法及其实现	169
5.3.2 二叉树遍历算法的应用举例	177
5.3.3 建立二叉树	181
5.4 哈夫曼树及哈夫曼编码	189
5.4.1 哈夫曼树的基本概念	189
5.4.2 哈夫曼树和哈夫曼编码的构造方法	190
5.4.3 构造哈夫曼树和哈夫曼编码的算法	192
5.5 树与森林	195
5.5.1 树、森林与二叉树之间的转换	195
5.5.2 树的存储结构	198
5.5.3 树和森林的遍历	201
小 结	203
习题五	204

第 6 章 图	207
6.1 图的类型定义	207
6.1.1 图的基本概念	207
6.1.2 图的抽象数据类型描述	211
6.2 图的存储结构	212
6.2.1 邻接矩阵	212
6.2.2 邻接表	217
6.3 图的遍历	220
6.3.1 广度优先搜索	221
6.3.2 深度优先搜索	223
6.3.3 图的遍历方法的应用举例	225
6.4 最小生成树	230
6.4.1 最小生成树的基本概念	230
6.4.2 克鲁斯卡尔算法	232
6.4.3 普里姆算法	232
6.5 最短路径	238
6.5.1 求某个顶点到其余各顶点的最短路径	238
6.5.2 求每一对顶点之间的最短路径	241
6.6 拓扑排序	245
6.6.1 拓扑排序的基本概念	246
6.6.2 拓扑排序的实现	247
6.7 关键路径	249
小 结	256
习题六	257
第 7 章 内排序	260
7.1 排序的基本概念	260
7.2 插入排序	262
7.2.1 直接插入排序	262
7.2.2 希尔排序	265
7.3 交换排序	266
7.3.1 冒泡排序	267
7.3.2 快速排序	268
7.4 选择排序	271
7.4.1 直接选择排序	272
7.4.2 树形选择排序	273
7.4.3 堆排序	274
7.5 归并排序	279

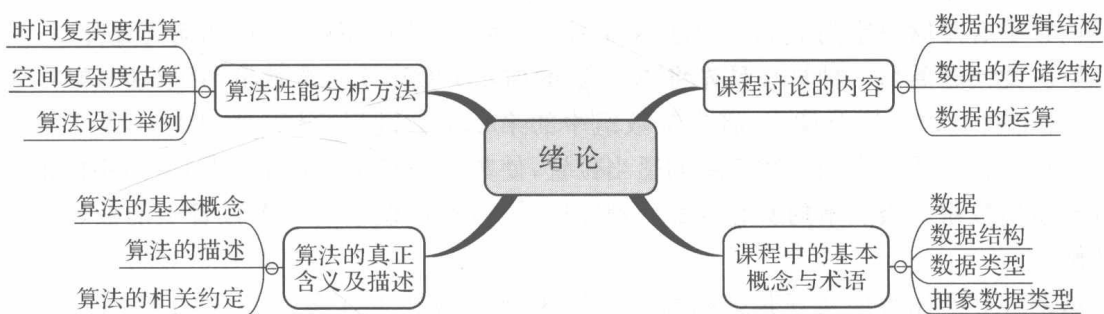
7.6 基数排序	282
7.6.1 多关键字排序	282
7.6.2 链式基数排序	283
小 结	284
习题七	286
 第8章 外排序	 289
8.1 外排序方法	289
8.2 磁盘排序	291
8.2.1 磁盘信息的存取	291
8.2.2 多路平衡归并	292
8.2.3 置换—选择排序	293
8.2.4 最优归并树	296
小 结	300
习题八	301
 第9章 查 找	 304
9.1 查找的基本概念	304
9.2 静态表查找	305
9.2.1 顺序查找	306
9.2.2 二分查找	307
9.2.3 分块查找	310
9.3 动态表查找	312
9.3.1 二叉排序树	312
9.3.2 平衡二叉树	322
9.3.3 B ₋ 树和 B ⁺ 树	325
9.3.4 红黑树简介	333
9.4 哈希表查找	336
9.4.1 哈希表的定义	336
9.4.2 常用的哈希函数	337
9.4.3 处理冲突的方法	339
9.4.4 哈希表的查找和性能分析	348
小 结	349
习题九	350
 附录 参考答案	 353
习题一	353
习题二	355
习题三	360

习题四.....	368
习题五.....	370
习题六.....	374
习题七.....	382
习题八.....	386
习题九.....	389
参考文献.....	395

第 1 章 绪 论

随着计算机和信息技术的飞速发展,计算机应用远远超出了单纯进行数值计算的范畴,计算机技术已渗透到了国民经济的各行各业和人们日常生活的方方面面。由于现实世界中产生的大量数据只有经过计算机存储才能进行后续处理与利用,因此,数据的表示是计算机处理问题的基础。从表面上看,程序设计和编写的主要目的是为了实现对某类数据的处理;而实质上,更为重要的是要考虑如何在计算机中组织这些数据,以支持后续高效的处理过程。因此,深入研究各种数据的逻辑结构及其在计算机中的表示和实现,科学、合理地指导程序设计,这正是我们学习“数据结构”这门课程的目的。

【本章主要知识导图】



1.1 数据结构课程讨论的内容

Pascal 语言之父、结构化程序设计的先驱、著名的瑞士科学家沃思(Niklaus Wirth)教授曾出版过一本著名的书籍《算法+数据结构=程序》,他认为:程序是计算机指令的组合,用来控制计算机的工作流程,完成一定的逻辑功能以实现某种任务;算法是程序的逻辑抽象,是解决某类客观问题的策略;数据结构是现实世界中的数据及其之间关系的反映,它可以从逻辑结构和存储(物理)结构两个层面刻画,其中客观事物自身所具有的结构特点,被称为逻辑结构,而具有这种逻辑结构的数据在计算机存储器中的组织形式则被称为存储结构。可见,这个等式反映了算法、数据结构对于程序设计的重要性。通过学习本课程,读者将体

会到在解决一个现实问题时应该如何采用一种合适的方法(算法),又应该如何编写一个高效率程序,并通过计算机实现问题的求解。

1.1.1 求解问题举例

人们利用计算机是为了能快速解决实际的应用问题,而程序是人们编写的,它是计算机按照人的旨意进行问题处理的指令的集合。因而要利用计算机实现问题的求解,就需要完成一个从问题到程序的实现过程。此过程的主要步骤归纳如下:



C01-1-1

(1) **确定问题求解的数学模型(或逻辑结构)**:对问题进行深入分析,确定处理的数据对象是什么,再根据数据对象的逻辑关系给出其数学模型。

(2) **确定存储结构**:根据数据对象的逻辑结构及其所需完成的功能,选择一种合适的组织形式将数据对象映射到计算机的存储器中。

(3) **设计算法**:讨论要解决问题的策略,即确定求解问题的具体步骤。

(4) **编程并测试结果**:根据算法编写程序并上机测试,直至得到问题的最终解。

由此可见,程序设计的本质在于解决两个主要问题:一是根据实际问题选择一种“好”的数据结构;二是设计一个“好”的算法。后者的好坏在很大程度上取决于前者。例如:

问题一: N 个数的选择问题。假设有 N 个数,要求找出 N 个数中第 k 大的那个数。

要解决这个问题,首先要考虑的就是这 N 个数的取值范围是什么。也就是说,只有明确了这 N 个数在计算机中是如何组织的,才能决定采用何种求解问题的策略。

解决此问题的一种算法是将 N 个数读入一个数组中,再通过冒泡排序方法对 N 个数以递减顺序排序,最后在已排序数组中的第 k 个元素就是所要求的解。这种策略实施的前提是内存中有足够的空间能够容纳这 N 个数。然而,若空间不够,又该如何处理呢?

解决此问题的另一种算法是先将 N 个数的前 k 个数读入一个数组,并以递减顺序进行排序,再将剩下的数逐个读入,将它与数组中的第 k 个数进行比较。如果它不大于第 k 个数,则忽略;否则就将其插入数组中的适当位置,使数组仍然保持有序,同时数组中的最后一个数被挤出数组。当剩下的所有数都处理完毕时,数组中位于第 k 个位置上的数就是所要求的解。

这两种算法用程序实现其编码都很简单,但哪种算法更好呢?当 N 和 k 都充分大(大于 10^6)时,这两种算法的程序在合理时间内都不能结束,它们需要计算机处理若干天后才能计算完毕,这是不切实际的,所以它们都不能被认为是好的算法。在第7章将给出解决这个问题的更好算法。

问题二:生产订单的自动查询问题。在一个订单管理系统中,有一个实体“生产订单统计表”,实体中包括若干个订单记录,它们按照每个生产订单的订单号递增顺序排列,如表1-1所示。现要求查找“张三”制作的所有订单信息。

表 1-1 生产订单统计表

生产订单号	行号	物料编码	物料名称	开工时间	完工时间	计量单位	生产数量	制单人
00000010	1	12100	长针	2017-09-13	2017-09-14	根	60.00	张三
00000010	2	12100	长针	2017-09-14	2017-09-15	根	390.00	张三

续表

生产订单号	行号	物料编码	物料名称	开工时间	完工时间	计量单位	生产数量	制单人
00000011	1	10000	电子挂钟	2017-09-18	2017-09-20	个	60.00	李四
00000011	2	10000	电子挂钟	2017-09-19	2017-09-22	个	390.00	李四
00000012	1	12400	盘面	2017-09-13	2017-09-14	个	60.00	李四
00000012	2	12400	盘面	2017-09-14	2017-09-15	个	390.00	李四
00000013	1	12000	钟盘	2017-09-14	2017-09-18	个	60.00	张三
00000013	2	12000	钟盘	2017-09-15	2017-09-19	个	390.00	张三

在研究如何查找满足指定条件的订单信息时,首先必须考虑生产订单统计表在计算机中是如何组织的、每一订单包括哪些信息项、各订单之间又是按什么顺序存放的,是按订单号递增的顺序存放?还是按物料名称的顺序存放?或是按制单人的姓氏顺序存放?然后才能根据特定的查找要求去确定某种查找方法。

上述问题中要求查找“张三”制作的所有订单信息,也就是说若在生产订单统计表中有该人制作的订单,则输出相关的所有订单信息;否则就输出该人制作的订单不存在。要解决这个问题,首先应将生产订单统计表中的数据映射到计算机的存储器中,从而形成生产订单统计表的存储结构。如何设计查找算法则取决于统计表的存储结构。一种算法可以将表中数据顺序地存储到计算机中,查找时从表的第一行记录开始到最后一行记录为止,依次核对制单人的姓名是否存在与“张三”相同的记录。若存在,则可获得该人制作的一个订单行信息;若找遍整个表均无此制单人,则说明没有该人制作的订单。这种算法在订单数不多的情况下是可行的,但是当订单数很多时是不实用的。另一种算法可以先将生产订单统计表顺序地存放,再按制单人的姓氏建立一张索引表,存储结构如图 1-1 所示。查找时首先在索引表中核对姓氏,然后根据索引表中的地址找到对应的指针桶,再根据指针桶中的地址直接到生产订单统计表的指定地址中查对制单人的姓名。注意,这时已经不需要查找其他不同姓氏的行信息了。相比之下,第二种查找算法比第一种算法更为高效。在第 6 章将给出更多的实现查找操作的算法。

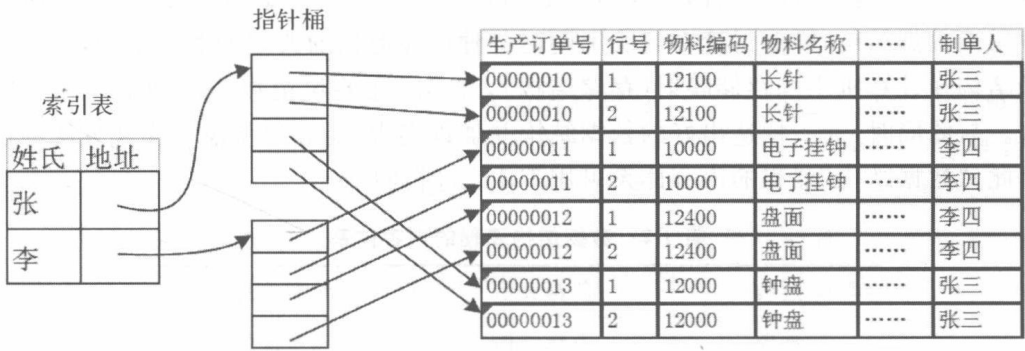


图 1-1 生产订单统计表的索引存储结构

问题三:城市网络的铺设问题。假设需为城市的 n 个小区之间铺设网络线路(任意两个小区之间都可以铺设), n 个小区只需铺设 $n-1$ 条线路,就能使 n 个小区网络相通。然而,

由于地理环境不同等因素导致各条线路所需投资不同,问题是采用怎样的设计方案才能使总投资最低?

根据各小区之间的地理位置关系及各线路所需的投资情况,确定这个问题的数学模型可用一个图来描述。假设该图如图 1-2 所示,其中顶点表示城市,顶点之间的连线及其上面的数值表示可以铺设的线路及所需经费。求解该问题的算法可以简单地描述为:在可以铺设的所有线路中选取 $n-1$ 条,使得这 $n-1$ 条线路既能连通 n 个小区,又使总投资最低。实际上,这是一个求“图的最小生成树”的问题,图 1-2 对应的最小生成树可用图 1-3 来描述,这个图给出的就是一种求解问题的设计方案。有关求图的最小生成树的具体算法将在第 6 章中进行详细讨论。

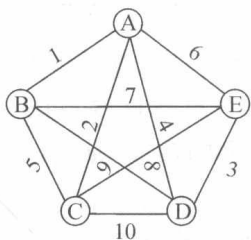


图 1-2 5 个小区的地理位置关系

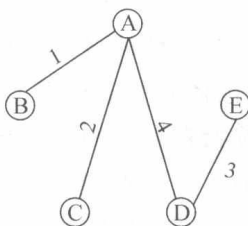


图 1-3 最小生成树



C01-1-2

1.1.2 本课程讨论的内容

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、计算机硬件和软件之间的一门计算机类和电子信息类相关专业的核心课程之一,是高级程序设计语言、编译原理、操作系统、数据库、人工智能等课程的基础。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。

从上节讨论的几个求解问题可知,用计算机来解决一个具体问题,总是围绕以下三个主要步骤进行:

- (1)抽象出所求解问题中需要处理的数据对象的逻辑结构(数学模型)。
- (2)根据问题求解的功能特性,实现对抽象数据的存储结构表述。
- (3)确定并实现为求解问题而需进行的操作或运算。

事实上,只有这三者的结合,才能清晰地刻画出数据结构的本质特性。因此,通常在本课程中讨论数据结构,既要讨论在解决问题时各种可能遇到的典型的逻辑结构,又要讨论这些逻辑结构在计算机中的存储映射(存储结构)。此外,还要讨论数据结构上的相关操作及其实现。与此同时,为了构造出好的数据结构并加以实现,还必须考虑其实现方法的性能优劣。因此,“数据结构”课程的内容体系可用表 1-2 进行归纳。

表 1-2 数据结构课程的内容体系

过程	数据表示	数据处理
抽象	逻辑结构	基本操作
实现	存储结构	算法
评价	不同数据结构的比较和算法性能分析	

1.2 基本概念与术语

在本节中将对数据结构课程中一些常用的基本概念与术语给出详细的解释,这些基本概念与术语将贯穿数据结构课程学习的整个过程。

1.2.1 数据与数据结构

数据(Data) 数据是信息的载体,是对客观事物的符号表示,能够被计算机程序识别、存储、加工和处理。因此,数据就是所有能够有效地输入计算机中并且能够被计算机处理的符号的总称;也可谓计算机程序处理对象的集合,是计算机程序加工的“原料”。例如,一个利用数值分析方法求解代数方程的程序,其处理对象是整数和实数等数值数据;一个编译程序或文字处理程序的处理对象是字符串。数据还包括图像、声音、视频等非数值数据。数据结构课程中讨论的对象主要是非数值数据。



C01-2-1

数据元素(Data Element) 数据元素是数据中的一个“个体”,是数据的基本组织单位。在计算机程序中通常将它作为一个整体进行考虑和处理。在不同条件下,数据元素又可称为结点、顶点和记录。如表 1-1 中的一行数据可称为一个数据元素或一条记录。在树或图中,一个数据元素用一个圆圈表示。如图 1-4 中,每一个圆圈所表示的就是一个数据元素,或被称为一个顶点。

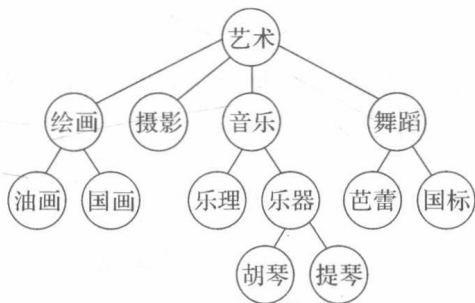


图 1-4 艺术分类结构

数据项(Data Item) 数据项是数据元素的组成部分,是具有独立含义的标识单位。一个数据元素可以由若干个数据项组成。如表 1-1 中的每一列,“生产订单号”“物料名称”等都是一个数据项。数据项又可分为两种:一种是简单数据项;另一种是组合数据项。如图 1-5 是一个教师的数据元素描述。其中“姓名”“单位名称”“职务”“职称”“工作业绩”是简单数据项,它们在数据处理时不能再被分割;而“出生日期”则是一个组合数据项,它可以进一步被划分为“年”“月”和“日”等更小的数据项。

姓名	出生日期	单位名称	职务	职称	工作业绩
----	------	------	----	----	------

图 1-5 一个教师的数据元素描述

数据对象(Data Object) 数据对象是性质相同的数据元素的集合。例如,在对生产订