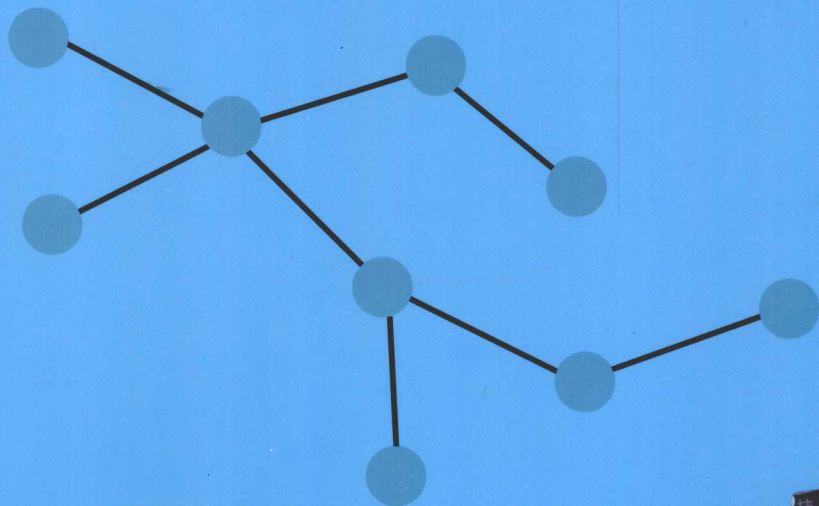


高等院校计算机任务驱动教改教材

数据结构分析与应用 实用教程

陈承欢 编 著



清华大学出版社

高等院校计算机任务驱动教改教材

数据结构分析与应用 实用教程

陈承欢 编 著

清华大学出版社

内 容 简 介

“数据结构”课程是计算机类专业一门非常重要的专业基础课程,它为后续课程提供知识基础、算法设计与实现的方法。本书从“数据结构的分析与应用”角度重构数据结构的教学内容,系统化设计教学模式和教学流程,以“引例剖析”→“知识梳理”→“算法探究”→“典型应用”→“小试牛刀”为主线探究式分析数据结构和设计算法。

本书以有趣的实例导出数据结构,突出数据结构学习的趣味性。以真实的任务应用数据结构,突出数据结构知识的实用性。以形象思维方式理解数据结构,突出数据结构描述的可读性。以规范程序代码实现数据结构,突出数据结构实现的规范性。本着“理论够用、注重实用”的原则,强化知识的条理性和系统性,运用“任务驱动”教学法,着力提升算法分析和程序设计能力,将数据结构灵活应用于各相关领域,引导学习者更好地学以致用。

本书可以作为计算机类各专业和非计算机专业的“数据结构”课程的教材,也可以作为数据结构的培训教材以及自学教材。

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

数据结构分析与应用实用教程/陈承欢编著.--北京:清华大学出版社,2015

高等院校计算机任务驱动教改教材

ISBN 978-7-302-39322-1

I. ①数… II. ①陈… III. ①数据结构—高等学校—教材 IV. ①TP311.12

中国版本图书馆 CIP 数据核字(2015)第 024939 号

责任编辑:张龙卿

封面设计:徐日强

责任校对:刘 静

责任印制:李红英

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

课 件 下 载: <http://www.tup.com.cn>, 010-62795764

印 刷 者:北京富博印刷有限公司

装 订 者:北京市密云县京文制本装订厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:22.75

字 数:548 千字

版 次:2015 年 6 月第 1 版

印 次:2015 年 6 月第 1 次印刷

印 数:1~3000

定 价:39.80 元

产品编号:059761-01

前言

“数据结构”课程是计算机类专业一门非常重要的专业基础课程,它为后续课程提供知识基础、算法设计与实现的方法,让学习者学会把现实世界的问题转化为计算机内部的表示和处理,提高逻辑思维能力。该课程使学习者较为全面地理解数据结构和算法的基本概念,掌握数据的各种逻辑结构、存储结构及其实现方法,能够分析和比较不同数据结构与算法,应用相关知识解决常见的算法设计问题和实际应用问题。

为了简化算法实现程序的复杂性,化解数据结构分析的抽象性,降低数据结构学习的难度,本书充分应用启发式教学方法和形象思维方法,增加数据结构的实用性,从而提升学习者的学习兴趣。主要有以下几方面的特色与创新。

(1) 从“数据结构的应用与分析”角度重构数据结构的教学内容,系统设计教学模式和教学流程

以“引例剖析”→“知识梳理”→“算法探究”→“典型应用”→“小试牛刀”为主线探究式分析数据结构和设计算法。“引例剖析”环节通过实例分析引出相关数据结构的基本概念和基本应用;“知识梳理”环节对数据结构的基础知识和基本方法进行系统化梳理,为后面的算法设计提供理论指导和方法支持;“算法探究”环节对相关数据结构基本操作的算法设计和程序实现进行探讨和分析;“典型应用”环节综合应用数据结构的基本操作及算法思想分析和解决实际问题,加深对数据结构知识的理解,提升应用所学知识解决实际问题的能力;“小试牛刀”环节是一个自主训练环节,让学习者运用所学知识和方法完成规定任务,验证其知识掌握情况和训练其动手能力。

每个教学环节都设置了合理的教学步骤,“引例剖析”环节分为“任务描述”→“程序编码”→“程序剖析”→“运行结果”4个教学步骤;“算法探究”环节分为“任务描述”→“算法剖析”→“任务实现”→“测试结果”4个教学步骤;“典型应用”环节分为“任务描述”→“任务实现”→“程序剖析”→“测试结果”4个教学步骤,“小试牛刀”环节分为“任务描述”和“编程提示”2个教学步骤。

(2) 突出趣味性、实用性、可读性和规范性

以有趣的实例导出数据结构,突出数据结构学习的趣味性。

以真实的任务应用数据结构,突出数据结构知识的实用性。

以形象思维方式理解数据结构,突出数据结构描述的可读性。

以规范程序代码实现数据结构,突出数据结构实现的规范性。

(3) 本着“理论够用、注重实用”的原则,强化知识的条理性和系统性

本书系统地介绍了各种数据结构,主要包括线性表、栈、队列、树、图、排序、查找等,对不常用的数据结构和算法进行了合理的删减,对于一些较复杂的算法进行简化处理。将基本概念、基本操作、基本方法与算法分析、实现代码相分离,分别在“知识梳理”和“算法探究”两个环节进行分析讲解,以保证理论知识的条理性和系统性、程序代码的可读性和连续性,便于理解代码和测试程序。

(4) 运用“任务驱动”教学法,提升算法分析和程序设计能力

本书总共设置了近百项任务,通过一系列任务,引导学习者循序渐进的学习理论知识,理解算法设计思想,熟悉算法实现方法,在完成各项任务的过程中,巩固和内化理论知识。训练算法设计能力,从而有效提高学习者的实践能力。

(5) 将数据结构灵活应用于各相关领域,引导学习者更好地学以致用

本书精心选用了以下实际问题:学习者数据的查找与排序、图书数据的查找与排序、超长正整数求和运算、两个任意长度整数的加法运算与减法运算、家谱结构的表示与输出、算术表达式的求值、进制转换、约瑟夫环求解、“七乐彩”福利彩票的中奖问题、城市之间通信网络的最小造价问题、城市之间建设高速公路网的最小造价问题、最佳游览线路规划问题、时间最短的乘坐地铁线路求解、消息的加密和解密、文本文件的加密和解密、文件夹与文件结构的表示与输出等,运用恰当的数据结构设计算法、编写代码实现和解决这些实际问题,提高学习者熟练运用理论知识和算法设计技能解决实际应用问题的能力。

(6) 强化良好编程习惯的培养

本书所有的程序都注重规范性、可读性,有良好的示范作用,引导学习者在程序编写过程中养成良好的编程习惯,因为良好的编程习惯、严谨的设计思路、认真的工作态度,能使学习者终身受益。

本教材由陈承欢教授编著,湖南铁道职业技术学院的冯向科、颜谦和、林东升、刘荣胜,南通理工学院的卫丽华,四川航天职业技术学院的郭华,云南林业职业技术学院的远俊红,长沙职业技术学院的殷正坤,湖南工业职业技术学院的刘曼春,安徽水利水电职业技术学院的方跃胜,四川工业科技学院的吕莉,日照职业技术学院的费琳琳,济源职业技术学院的陈平等老师参与了部分章节的编写工作。

由于编者水平有限,书中的疏漏之处敬请专家与读者批评指正,编者的QQ号为1574819688,本教材免费提供电子教案、源代码等相关教学资源。

编者

2015年1月

目 录

单元1 数据结构概述与基本算法分析 1

教学导航 1

引例剖析 1

任务 1-1 分析学生数据的结构特点和存储方式 1

任务 1-2 分析自行车的零部件结构和数据特点 3

任务 1-3 探析城市之间通信网络的最小造价问题 3

知识梳理 5

1.1 数据结构的基本概念 5

1.2 数据类型与抽象数据类型 7

1.3 算法和算法分析 8

算法探究 11

任务 1-4 设计算法实现 3 个整数由小到大排序 11

任务 1-5 设计算法计算正整数的阶乘 12

任务 1-6 设计算法采用递归法计算阶乘 13

任务 1-7 设计算法判断素数 14

典型应用 15

任务 1-8 设计算法计算矩阵的乘法 15

任务 1-9 设计算法实现超长正整数求和运算 17

小试牛刀 18

任务 1-10 分析图书数据的结构特点和存储方式 18

任务 1-11 分析家族的家谱结构和数据特点 19

任务 1-12 分析建设城市之间高速公路网的最小造价问题 19

任务 1-13 探析时间最短的乘坐地铁线路 20

任务 1-14 分析最佳游览线路规划问题 22

任务 1-15 设计算法在数组中查找给定值 22

任务 1-16 设计算法计算阶乘的累加和 23

单元习题 24

单元2 线性表的分析与应用 27

教学导航 27

引例剖析 27

任务 2-1 以顺序表方式在数据表中插入与删除记录数据	27
任务 2-2 以单链表方式在数据表中插入与删除记录数据	30
知识梳理	36
2.1 线性表的基本概念	37
2.2 线性表的基本操作	37
2.3 顺序表	38
2.4 单链表	39
2.5 循环链表	40
2.6 双向链表	40
算法探究	41
任务 2-3 设计算法实现顺序表的基本操作	41
任务 2-4 设计算法实现单链表的基本操作	46
任务 2-5 设计算法实现双向链表的基本操作	53
任务 2-6 设计算法创建循环链表	59
典型应用	61
任务 2-7 应用顺序表实现“七乐彩”福利彩票的生成和中奖查询	61
任务 2-8 应用静态循环链表求解约瑟夫环	64
任务 2-9 应用动态循环链表求解约瑟夫环	66
小试牛刀	69
任务 2-10 设计算法实现顺序表的就地逆置	69
任务 2-11 设计算法实现单链表的逆置操作	70
任务 2-12 设计算法实现双向链表的逆序输出	71
单元习题	72
单元 3 栈和队列的分析与应用	75
教学导航	75
引例剖析	75
任务 3-1 编写程序模拟子弹进出弹夹的过程	75
任务 3-2 编写程序模拟银行排队存取款的过程	78
知识梳理	80
3.1 栈的定义	81
3.2 栈的基本操作	81
3.3 栈的存储结构	82
3.4 队列	82
3.5 队列的基本操作	82
3.6 队列的存储结构	83
算法探究	83
任务 3-3 设计算法实现顺序栈的基本操作	83
任务 3-4 设计算法实现链栈的基本操作	87

任务 3-5 设计算法实现顺序队列的基本操作	91
任务 3-6 设计算法实现链队的基本操作	96
典型应用	100
任务 3-7 应用顺序栈实现十进制转换为其他进制	100
任务 3-8 应用顺序栈实现超长整数的加法运算	103
任务 3-9 应用顺序栈求算术表达式的值	108
任务 3-10 应用顺序队列实现消息的加密和解密	116
任务 3-11 应用链队输出符合规定要求的符号三角形	118
小试牛刀	122
任务 3-12 应用顺序栈将字符串逆序输出	122
任务 3-13 应用顺序栈检测括号是否匹配	123
任务 3-14 应用链栈判断字符串是否为回文字符串	124
任务 3-15 应用顺序队列模拟医院排队看病	126
单元习题	127
单元 4 树结构的分析与应用	130
教学导航	130
引例剖析	130
任务 4-1 编写程序模拟家谱结构建立与遍历二叉树	131
知识梳理	135
4.1 二叉树的基本概念	135
4.2 树的基本概念	137
4.3 树的表示	138
4.4 二叉树的主要性质	139
4.5 二叉树的存储结构	140
4.6 树的存储结构	142
4.7 二叉树的基本操作及实现	144
4.8 二叉树的遍历	145
4.9 树的基本操作	148
4.10 树的遍历	149
4.11 线索二叉树的定义及结构	149
4.12 哈夫曼树及其应用	150
算法探究	151
任务 4-2 设计算法建立二叉树及实现其基本操作	151
任务 4-3 设计算法实现二叉树的多种遍历方式	157
任务 4-4 设计算法建立二叉线索树及实现其基本操作	166
任务 4-5 设计算法建立树及实现其基本操作	172
任务 4-6 设计算法建立哈夫曼树	181
任务 4-7 设计算法应用哈夫曼树构造哈夫曼编码方案	183

10	典型应用	188
80	任务 4-8 应用二叉树和栈求表达式的值	188
001	任务 4-9 应用哈夫曼编码实现文本文件的加密和解密	197
001	小试牛刀	203
801	任务 4-10 设计算法判断一棵二叉树是否为完全二叉树	203
801	任务 4-11 应用哈夫曼编码进行解码	204
811	单元习题	205
811	单元 5 图结构的分析与应用	210
881	教学导航	210
881	引例剖析	210
181	任务 5-1 建立地铁站点的邻接矩阵	210
881	知识梳理	214
181	5.1 图的基本概念	214
061	5.2 图的基本操作	217
061	5.3 图的存储表示方法	217
061	5.4 图的遍历	219
061	5.5 图的最小生成树	220
181	5.6 图的最短路径	220
881	5.7 AOV 网与拓扑排序	222
881	算法探究	223
181	任务 5-2 设计算法建立有向图的邻接矩阵表示	223
881	任务 5-3 设计算法建立无向图的邻接表表示	225
081	任务 5-4 设计算法实现邻接矩阵表示的图的深度优先遍历	228
081	任务 5-5 设计算法实现邻接表表示的图的深度优先遍历	230
981	任务 5-6 设计算法实现邻接矩阵表示的图的广度优先遍历	233
181	任务 5-7 设计算法实现邻接表表示的图的广度优先遍历	235
881	典型应用	238
881	任务 5-8 应用普里姆算法求解最小生成树	238
081	任务 5-9 应用克鲁斯卡尔算法求解最小生成树	243
081	任务 5-10 应用迪杰斯特拉算法求解单源图的最短路径	248
081	任务 5-11 应用弗洛伊德算法求解无向图的最短路径	253
181	任务 5-12 编写程序实现拓扑排序算法并输出拓扑序列	257
181	小试牛刀	261
181	任务 5-13 建立高速公路线路图的邻接矩阵表示	261
981	任务 5-14 建立有向图的邻接表表示	261
881	任务 5-15 求解高速公路线路图的最小生成树	262
181	任务 5-16 求解花费时间最短的游览线路	262
881	单元习题	262

单元 6 排序的分析与应用	266
教学导航	266
引例剖析	266
任务 6-1 应用直接插入排序法对磁盘数据进行排序	266
任务 6-2 应用选择排序法对商品数据进行排序	269
知识梳理	273
6.1 排序的基本概念	273
6.2 插入排序	274
6.3 交换排序	276
6.4 选择排序	278
6.5 二路归并排序	279
算法探究	280
任务 6-3 设计算法实现插入排序	280
任务 6-4 设计算法实现希尔排序	282
任务 6-5 设计算法实现冒泡排序	284
任务 6-6 设计算法实现快速排序	286
任务 6-7 设计算法实现简单选择排序	290
任务 6-8 设计算法实现堆排序	292
任务 6-9 设计算法实现二路汇并排序	298
典型应用	300
任务 6-10 应用冒泡排序法和选择排序法对图书数据进行排序	300
任务 6-11 应用希尔排序法和堆排序法对学生数据进行排序	304
小试牛刀	307
任务 6-12 编写程序应用插入排序法实现图书销量的降序排列	307
任务 6-13 编写程序应用选择排序法实现成绩的降序排列	308
任务 6-14 编写程序应用冒泡排序法实现学生姓名的降序排列	309
单元习题	309
单元 7 查找的分析与应用	315
教学导航	315
引例剖析	315
任务 7-1 查找指定名称的手机数据	315
知识梳理	317
7.1 查找的基本概念	317
7.2 静态查找	319
7.3 二叉排序树及查找	322
7.4 哈希表查找	323
算法探究	327

任务 7-2 设计算法应用顺序查找法查找指定数据	327
任务 7-3 设计算法应用折半查找法查找指定数据	329
任务 7-4 设计算法建立二叉排序树并查找指定数据	332
任务 7-5 设计算法应用哈希表实现数据的查找	335
典型应用	338
任务 7-6 应用顺序查找法查找图书数据	338
任务 7-7 应用二叉排序树查找学生数据	339
小试牛刀	341
任务 7-8 应用顺序查找法查找指定学号的学生数据	341
任务 7-9 应用折半查找法查找指定姓名的学生数据	342
单元习题	343
附录 A 数据结构综合训练	345
附录 B 数据结构常见术语中英文对照表	349
参考文献	351

单元 1 数据结构概述与基本算法分析

数据结构与数学、计算机硬件和软件有十分密切的关系,数据结构技术也广泛应用于信息科学、系统工程、应用数学、工程技术等领域。计算机系统软件和应用软件都要用到各种类型的数据结构。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。当我们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,最后编写、调试与测试程序进行,直至得到最终的解答。由于当时所涉及的运算对象是简单的整型、实型或逻辑型数据,所以程序设计者的主要精力是集中在程序设计的技巧上,而无须重视数据结构。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,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,才能有效地解决问题。

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、图之类的数据结构。学习数据结构的目的是了解计算机处理对象的特性,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。

教学导航

教学目标	(1) 熟悉数据、数据元素、数据结构、逻辑结构、存储结构等基本概念
	(2) 掌握逻辑结构和存储结构的分类以及基本特点
	(3) 掌握算法的概念及其特性
	(4) 理解和掌握时间复杂度的分析方法,了解空间复杂度的分析方法
教学方法	任务驱动法、分组讨论法、探究学习法、讲授课
课时建议	4 课时

引例剖析

任务 1-1 分析学生数据的结构特点和存储方式

【任务描述】

部分学生数据如表 1-1 所示,分析这些学生数据的结构特点、C 程序中的数据类型及其存储方式。

表 1-1 部分学生数据

序号	学 号	姓名	性别	出生日期	成绩
1	20141001	Tom	男	1996-07-15	92.5
2	20141002	Lily	女	1996-04-06	75
3	20141003	Sami	男	1995-03-21	83
4	20141004	Nina	女	1998-01-26	98
5	20141005	Ailsa	男	1996-12-18	68.5
6	20141006	Marcy	男	1996-09-27	87

【任务实现】

表 1-1 中每一行学生数据称为一条记录,也称为一个数据元素,由序号、学号、姓名、性别、出生日期、成绩等数据项组成,每一列是一条记录中的数据项。C 程序中“序号”可以定义为字符型或整型、“成绩”可以定义为实型、“学号”、“姓名”、“性别”和“出生日期”都由一个以上的字符组成的字符串,可以定义为字符型数组。由于每条记录的数据项的数据类型不完全相同,而每条记录的数据项具有相关性,C 程序中一般定义为结构。

表 1-1 所示的学生数据可以采用索引存储结构,即在存储数据元素的同时,还建立附加的索引表,索引表由若干索引项组成,如图 1-1 所示。这种存储方式,逻辑上相邻的数据元素其物理位置不一定相邻,索引项可以采用顺序存储结构。

序号	存放地址	学号	姓名	其他字段
1	10101100	→ 20141001	Tom	...
2	10101110	→ 20141002	Lily	...
3	10110000	→ 20141003	Sami	...
4	10110010	→ 20141004	Nina	...
5	10110100	→ 20141005	Ailsa	...
6	10110110	→ 20141006	Marcy	...

图 1-1 索引存储结构

图 1-1 所示的学生数据各个数据元素之间存在“一对一”的线性关系,除了第一个数据元素外,其他各个元素有唯一的前趋;除了最后一个数据元素外,其他各个元素有唯一的后继。

当在学生管理信息系统中需要查找某个学生的有关信息的时候,或者想查询某个专业或年级所有学生有关信息的时候,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,可以在学生管理信息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。由这些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的是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。

任务 1-2 分析自行车的零部件结构和数据特点

【任务描述】

自行车外观图如图 1-2 所示,一般可以将自行车整车分为“车架部件”、“车把部件”、“车轮部件”、“脚蹬部件”和“制动部件”等组成部分,每个部件可分为多个子部件或零件,绘制自行车的分层结构图,并分析其数据特点和存储方式。

【任务实现】

自行车的分层结构图如图 1-3 所示,这种分层结构类似一棵根在上、枝和叶在下的倒立树,该结构清晰地描述了自行车各部件、子部件和零件之间的关系,即“一对多”关系,称为树状数据结构。在该分层结构中,除了根元素外,其他元素都有唯一的前趋,所有数据元素都可以有多个后继,即数据元素之间存在一对多的关系。



图 1-2 自行车外观图

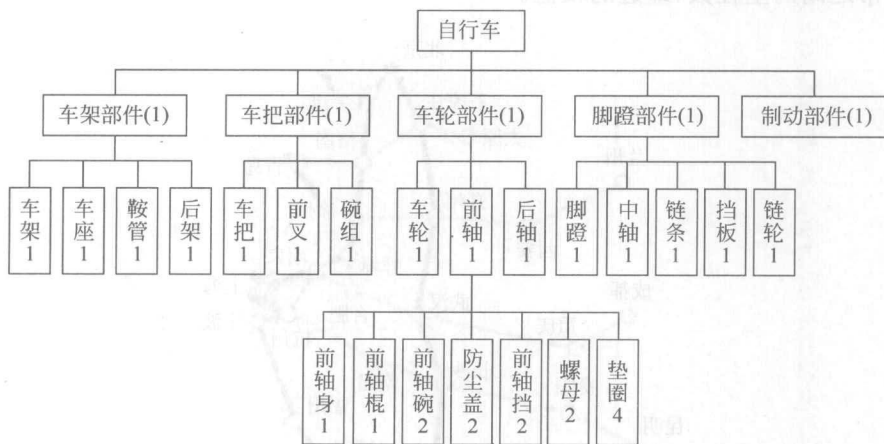


图 1-3 自行车分层结构图

任务 1-3 探析城市之间通信网络的最小造价问题

【任务描述】

全国 16 座城市之间的里程数如表 1-2 所示,各城市的地理位置和主要的高速铁路示意图如图 1-4 所示,如果需要在这些城市之间建设专用高铁通信网络,要求其中任意两个城市之间都有直接的或间接的通信线路,并且在已知城市之间直接通信线路预算造价的前提下,绘制的高铁专用通信网络初步规划示意图,尝试规划这些城市之间的高铁专用通信网络使

其造价最低。

表 1-2 全国 16 座城市之间的里程

单位: 公里

城市—城市	距离	城市—城市	距离	城市—城市	距离
北京 1—郑州 2	695	上海 8—杭州 9	202	杭州 9—长沙 4	989
郑州 2—武汉 3	514	杭州 9—宁波 10	149	长沙 4—昆明 14	1625
武汉 3—长沙 4	413	宁波 10—福州 11	564	南京 7—武汉 3	516
长沙 4—深圳 5	854	福州 11—深圳 5	1060	武汉 3—重庆 15	1159
北京 1—徐州 6	814	徐州 6—郑州 2	349	重庆 15—成都 16	504
徐州 6—南京 7	348	郑州 2—西安 12	511		
南京 7—上海 8	301	西安 12—兰州 13	676		

【任务实现】

根据表 1-2 所示 16 座城市之间的里程数和图 1-4 所示的城市地理分布图和主要的高速铁路示意图, 绘制如图 1-5 所示的高速铁路专用网络初步规划示意图。图中圆圈表示一个城市, 即一个结点; 两个圆圈之间的连线表示对应城市之间的高速公路, 即边; 连线上的数字表示城市之间的里程数, 即边的权值。

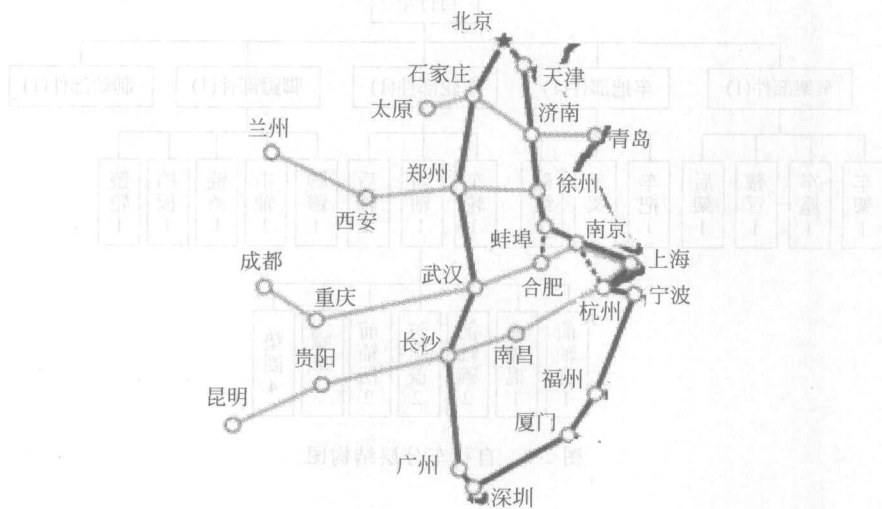


图 1-4 16 座城市的地理分布图

规划高速铁路专用网络, 我们最关心的问题如下。

(1) 两座城市之间是否有直接或间接的通信线路。

(2) 两座城市之间有多条通路的情况下, 哪条通信线路最短或所花费的建设费用最低。

假设各城市之间建设高铁专用通信线路每公里造价相同, 建设高铁专用通信网的总造价取决于通信网络的总里程, 如果能求解通信网络的最短里程, 就可以保证总造价最小。这就是典型的最小生成树问题, 此时路径的长度不是路径上的边的数目, 而是路径上的边的权值的总和, 求解图的最小生成树是一个很有实际应用价值的问题。

图 1-5 中的各个结点表示的数据元素之间存在“多对多”的任意关系,称为图状数据结构。在该图结构中各个结点(数据元素)可以有多个前趋或多个后继。

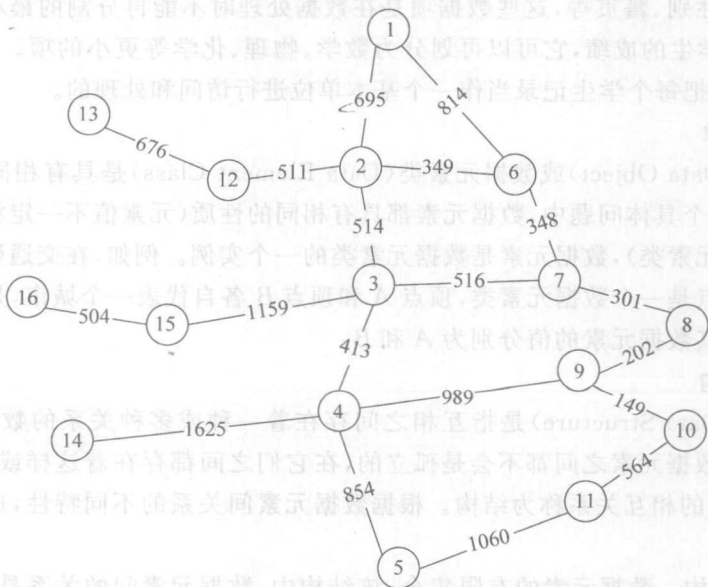


图 1-5 高铁专用通信网络初步规划示意图

由以上分析可知,各座城市之间的关系可用一个称作图的数据结构来表示。图中的每个顶点表示一座城市,各城市之间的通信线路只需连通即可,没有先后顺序关系。

知识梳理

1.1 数据结构的基本概念

在系统地学习数据结构知识之前,先明确一些基本术语的确切含义。

1. 数据

数据(Data)是信息的载体,它能够被计算机识别、存储和加工处理,它是计算机程序加工的“原料”,计算机应用程序处理各种各样的数据。计算机科学中,所谓数据就是计算机加工处理的对象,它可以是数值数据,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、音频、视频、动画等。

2. 数据元素

数据元素(Data Element)是数据的基本单位。在不同的条件下,数据元素又可分为元素(Element)、结点(Node)、顶点(Vertex)、记录(Record)等。例如,学生管理信息系统中学生信息表中的一个记录、八皇后问题中状态树的一个状态、教学计划编排问题中的一个顶点等,都被称为一个数据元素。有时,一个数据元素可由若干个数据项(Data Item)组成,例

如,学生管理信息系统中学生信息表的每一个数据元素就是一条学生记录。它包括学生的学号、姓名、性别、籍贯、出生年月、成绩等数据项。这些数据项可以分为两种:一种叫作基本项,如学生的性别、籍贯等,这些数据项是在数据处理时不能再分割的最小单位;另一种叫作组合项,如学生的成绩,它可以再划分为数学、物理、化学等更小的项。通常,在解决实际应用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。

3. 数据对象

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),数据元素是数据元素类的一个实例。例如,在交通咨询系统的交通网中,所有的顶点是一个数据元素类,顶点A和顶点B各自代表一个城市,是该数据元素类中的两个实例,其数据元素的值分别为A和B。

4. 数据结构

数据结构(Data Structure)是指互相之间存在着一种或多种关系的数据元素的集合。在任何问题中,数据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关系,这种数据元素之间的相互关系称为结构。根据数据元素间关系的不同特性,通常有下列4种基本的结构。

(1) 集合结构。数据元素的有限集合,该结构中,数据元素间的关系是“属于同一个集合”,集合是元素关系极为松散的一种结构。

集合是一种常用的数据表示方法,对集合可以作多种操作,假设集合S由若干个元素组成,可以按照某一规则把集合S划分成若干个互不相交的子集合,例如,集合 $S=\{1,2,3,4,5,6,7,8,9,10\}$,可以被分成如下三个互不相交的子集合:

$$S_1=\{1,2,4,7\}, \quad S_2=\{3,5,8\}, \quad S_3=\{6,9,10\}$$

集合 $\{S_1, S_2, S_3\}$ 就被称为集合S的一个划分。

此外,在集合上还有最常用的一些运算,例如集合的交、并、补、差以及判定一个元素是否是集合中的元素等。

由于集合是数据元素之间关系极为松散的一种结构,因此也可用其他结构来表示它,例如可以用树结构表示集合。

(2) 线性结构。数据元素的有序集合,该结构的数据元素之间存在着一对一的关系,线性表、栈、队列、字符串都属于线性结构。

(3) 树形结构。数据元素的层次结构,该结构的数据元素之间存在着一对多的关系。

(4) 图形结构。该结构的数据元素之间存在着多对多的关系,图形结构也称作网状结构。

树形结构和图形结构都属于非线性结构。

5. 数据的逻辑结构与存储结构

数据结构包括数据的逻辑结构和数据的物理结构。数据的逻辑结构是指数据元素之间的关系,是从逻辑关系角度描述数据,可以看作从具体问题抽象出来的数学模型,它与数据的存储无关。我们研究数据结构的目的是为了在计算机中实现对它的操作,为此还需要研究如何在计算机中表示一个数据结构。数据元素及数据元素之间的逻辑关系在计算机存储器内的表示(又称映像)称为数据的物理结构,或称存储结构。它所研究的是数据结构在计