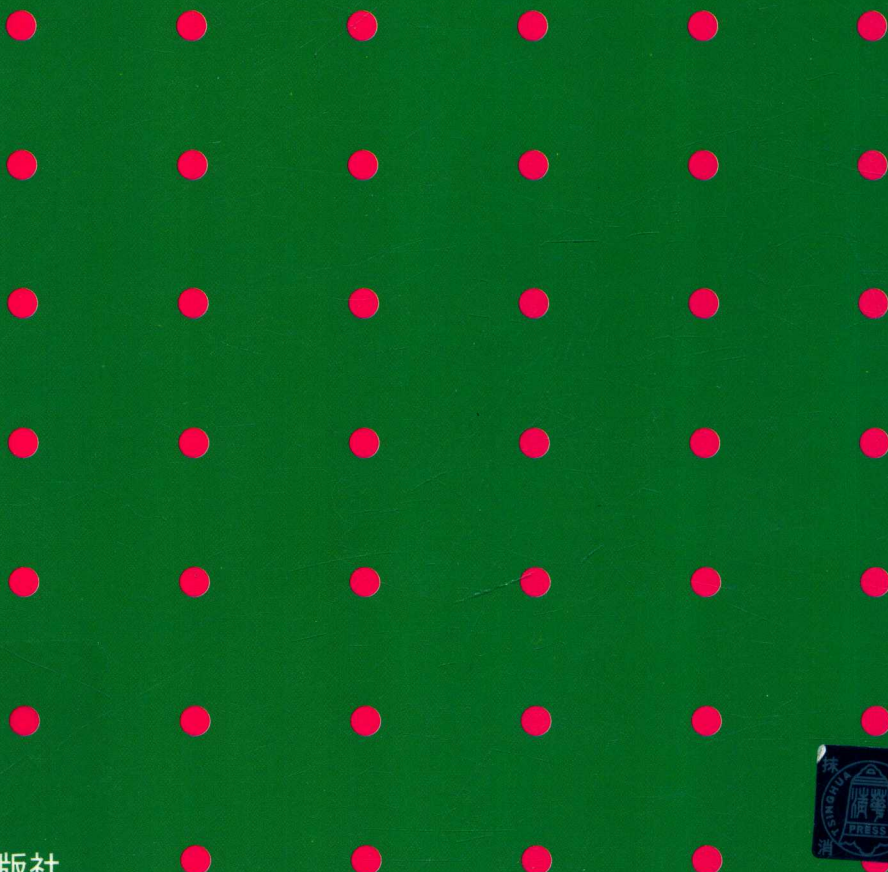


普通高校本科计算机专业特色教材精选 · 算法与程序设计

数据结构 ——从概念到Java实现

王红梅 党源源 刘冰 编著



清华大学出版社

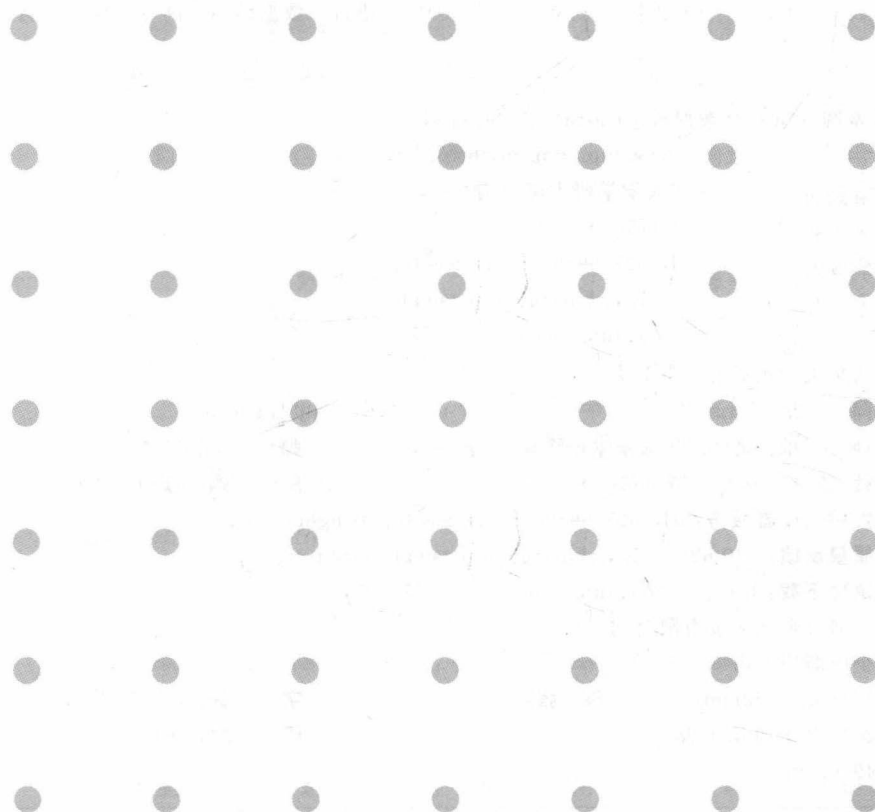


普通高校本科计算机专业特色教材精选·算法与程序设计

数据结构

——从概念到Java实现

王红梅 党源源 刘冰 编著



清华大学出版社
北京

内 容 简 介

数据结构是计算机及相关专业的核心课程,也是计算机及相关专业硕士研究生入学考试的必考科目,而且是理工专业的热门公选课程。本书介绍了数据结构、算法以及抽象数据类型的概念,讲述了线性表、栈和队列、字符串和多维数组、树和二叉树、图等常用数据结构,讨论了查找和排序技术。本书通过合理规划教学内容,梳理了知识单元及其拓扑结构,并兼顾概念层和实现层,既强调数据结构的基本概念和原理及方法,又注重数据结构的程序实现和实际运用,在提炼基础知识的同时,进行了适当的扩展和提高。

本书内容丰富,层次清晰,深入浅出且结合了实例,可作为计算机及相关专业数据结构课程的教材,也可供从事计算机软件开发和应用的工程技术人员参考和阅读。

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:从概念到Java实现/王红梅,党源源,刘冰编著. —北京:清华大学出版社,2019
(普通高校本科计算机专业特色教材精选·算法与程序设计)
ISBN 978-7-302-51340-7

I. ①数… II. ①王… ②党… ③刘… III. ①Java语言—数据结构—高等学校—教材
IV. ①TP312.8

中国版本图书馆CIP数据核字(2018)第229134号

责任编辑:袁勤勇

封面设计:常雪影

责任校对:焦丽丽

责任印制:宋 林

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>, 010-62795954

印 装 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:20 字 数:457千字

版 次:2019年3月第1版 印 次:2019年3月第1次印刷

定 价:49.00元

产品编号:078469-01

前言

PREFACE

数据结构是计算机及相关专业的核心课程,也是计算机及相关专业硕士研究生入学考试的必考科目,而且是理工专业的热门公选课程。作为程序设计的重要补充和延伸,数据结构所讨论的知识内容、蕴含的技术方法、体现的思维方式,无论是进一步学习计算机专业的其他课程,还是从事计算机领域的各项工作,都有着不可替代的作用。

数据结构课程知识丰富,内容抽象,隐藏在各知识单元的概念和方法较多,贯穿于各知识单元的链表和递归更是加重了学习难度。本书的编写者长期从事数据结构的研究和教学,深切理解学生在学习数据结构过程中遇到的问题和困惑,深入探究掌握数据结构的有效途径和方法,深刻思考数据结构对培养程序设计和计算思维的地位和作用,深度把握课程的教学目标和重点难点,在教学内容和教学设计等方面进行了如下处理。

1. 合理规划教学内容。紧扣《高等学校计算机专业核心课程教学实施方案》和《计算机学科硕士研究生入学考试大纲》,涵盖教学方案及考研大纲要求的全部知识点。

2. 遵循认知规律,理清教学主线。根据学生的认知规律和课程的知识结构,按照从已知到未知的思维进程逐步推进教学内容,梳理和规划了各知识单元及其拓扑结构,设计了清晰的教学主线。知识单元及其拓扑结构如图1所示。

3. 提炼基础知识,适当扩展提高。考虑不同学校教学要求的差异以及不同学生学习需求的差别,一方面本着“够用、实用”的原则,抓牢核心概念,提炼基础性知识,贯彻数据结构课程的基本教学要求;另一方面对某些知识点进行了适当的扩充和提高(图1中带星号部分),这部分内容可用于选讲,也可用于学生自学或课外阅读(教学建议:目录中带一个星号可用于选讲,带两个星号可用于课外阅读,其他是必讲的基础知识)。

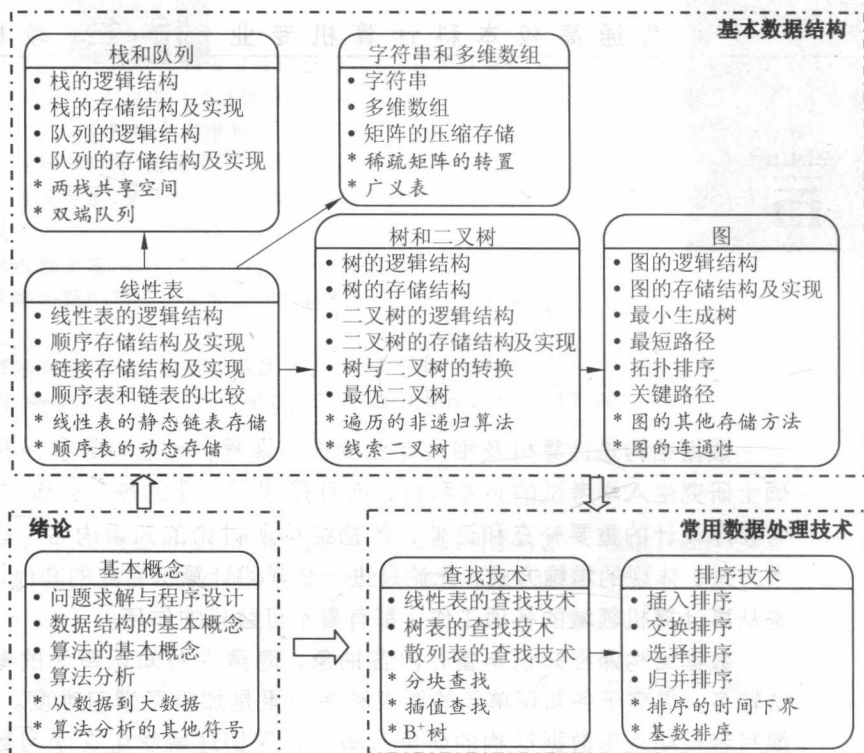


图1 数据结构课程的知识单元及拓扑结构

4. 兼顾概念层和实现层。从抽象数据类型的角度将数据结构的实现过程分为抽象层、设计层和实现层，其中抽象层定义数据结构和基本操作集合，设计层是数据结构的存储表示和算法设计，实现层用 Java 语言实现数据结构，既强调数据结构的基本概念、原理及方法，又注重数据结构的程序实现和实际运用。

5. 展现求解过程，培养计算思维。通过讲思路讲过程讲方法，按照“问题→想法→算法→程序”的模式进行问题求解，采用“阐述基本思想→伪代码描述算法→Java 语言实现算法”的模式进行算法设计与实现，这个过程正是计算思维的应用过程。每章通过两个应用实例展示了问题求解过程以及算法设计过程。

6. 明确重点，化解难点。每一章开篇即给出了本章的重点难点，以及各知识点的教学要求，在具体阐述时还有针对性的处理方法。针对数据结构内容抽象的特点，全书设计了大量插图，将抽象内容进行了具体化处理，降低了理解问题的复杂性。

总之，本书在概念的描述、实例的选择、知识的前后衔接、内容的组织结构，以及教学内容的理解、教学目标的实现、教学意图的融入、教学方法的运用等方面进行了系统思考和统筹设计，力图通过本书为读者构建多层次的知识体系。在问题求解层面，以数据表示和数据处理为主线，给出“问题→想法→算法→程序”的思维模式；在算法设计层面，通过伪代码描述算法，强调计算思维的培养；在算法分析层面，理解什么是“好”算法，给出算法分析的基本方法；在存储结构层面，通过存储示意图理解数据表示，再给出存储结构定义；在程序实现层面，给出所有数据结构的 Java 程序实现以及

使用范例；在数据结构和算法的运用层面，通过应用实例理解如何为求解问题设计适当的数据结构，如何基于数据结构设计算法，从而将数据结构、算法和程序设计有机地融合在一起。书中所有代码均在 Java 8 环境下调试通过。

参加本书编写的还有王涛、肖巍等。

由于作者的知识与写作水平有限，书稿虽再三斟酌几经修改，仍难免有缺点和错误，欢迎专家和读者批评指正。作者的电子邮箱是 wanghm@ccut.edu.cn。

作 者

2018 年 10 月

目 录

CONTENTS

第 1 章 绪论	1
1.1 问题求解与程序设计	2
1.1.1 程序设计的一般过程	2
1.1.2 数据结构在程序设计中的作用	5
1.1.3 算法在程序设计中的作用	6
1.1.4 本书讨论的主要内容	7
1.2 数据结构的基本概念	8
1.2.1 数据结构	8
1.2.2 抽象数据类型	11
1.3 算法的基本概念	13
1.3.1 算法及算法的特性	13
1.3.2 算法的描述方法	14
1.4 算法分析	16
1.4.1 算法的时间复杂度	16
1.4.2 算法的空间复杂度	18
1.4.3 算法分析举例	18
1.5 扩展与提高	21
1.5.1 从数据到大数据	21
1.5.2 算法分析的其他渐进符号	22
思想火花——概率算法	23
习题	24
第 2 章 线性表	27
2.1 引言	28
2.2 线性表的逻辑结构	29
2.2.1 线性表的定义	29

2.2.2	线性表的抽象数据类型	29
2.2.3	线性表的 Java 接口定义	31
2.3	线性表的顺序存储结构及实现	31
2.3.1	顺序表的存储结构	31
2.3.2	顺序表的实现	32
2.3.3	顺序表的使用	37
2.4	线性表的链接存储结构及实现	38
2.4.1	单链表的存储结构	39
2.4.2	单链表的实现	41
2.4.3	单链表的使用	47
2.4.4	双链表	48
2.4.5	循环链表	50
2.5	顺序表和链表的比较	51
2.6	扩展与提高	52
2.6.1	线性表的静态链表存储	52
2.6.2	顺序表的动态分配方式	54
2.7	应用实例	55
2.7.1	约瑟夫环问题	55
2.7.2	一元多项式求和	57
	思想火花——好算法是反复努力和重新修正的结果	61
	习题	62
	实验题	65
第3章	栈和队列	67
3.1	引言	68
3.2	栈	69
3.2.1	栈的逻辑结构	69
3.2.2	栈的顺序存储结构及实现	70
3.2.3	栈的链接存储结构及实现	73
3.2.4	顺序栈和链栈的比较	76
3.3	队列	76
3.3.1	队列的逻辑结构	76
3.3.2	队列的顺序存储结构及实现	77
3.3.3	队列的链接存储结构及实现	82
3.3.4	循环队列与链队列的比较	85
3.4	扩展与提高	85
3.4.1	两栈共享空间	85
3.4.2	双端队列	87

3.5 应用举例	88
3.5.1 括号匹配问题	88
3.5.2 表达式求值	89
思想火花——好程序要能识别和处理各种输入	92
习题	93
实验题	95
第4章 字符串和多维数组	97
4.1 引言	98
4.2 字符串	99
4.2.1 字符串的逻辑结构	99
4.2.2 字符串的存储结构	101
4.2.3 模式匹配	101
4.3 多维数组	105
4.3.1 数组的逻辑结构	105
4.3.2 数组的存储结构与寻址	106
4.4 矩阵的压缩存储	107
4.4.1 特殊矩阵的压缩存储	107
4.4.2 稀疏矩阵的压缩存储	110
4.5 扩展与提高	112
4.5.1 稀疏矩阵的转置运算	112
4.5.2 广义表	114
4.6 应用实例	118
4.6.1 发纸牌	118
4.6.2 八皇后问题	119
思想火花——用常识性的思维去思考问题	121
习题	121
实验题	123
第5章 树和二叉树	125
5.1 引言	126
5.2 树的逻辑结构	127
5.2.1 树的定义和基本术语	127
5.2.2 树的抽象数据类型	128
5.2.3 树的遍历操作定义	129
5.3 树的存储结构	130
5.3.1 双亲表示法	130
5.3.2 孩子表示法	131

5.3.3 孩子兄弟表示法	131
5.4 二叉树的逻辑结构	133
5.4.1 二叉树的定义	133
5.4.2 二叉树的基本性质	134
5.4.3 二叉树的抽象数据类型定义	136
5.4.4 二叉树的遍历操作定义	137
5.4.5 二叉树的 Java 接口定义	139
5.5 二叉树的存储结构及实现	139
5.5.1 顺序存储结构	139
5.5.2 二叉链表	140
5.5.3 三叉链表	145
5.6 森林	146
5.6.1 森林的逻辑结构	146
5.6.2 树、森林与二叉树的转换	147
5.7 最优二叉树	149
5.7.1 哈夫曼算法	149
5.7.2 哈夫曼编码	152
5.8 扩展与提高	154
5.8.1 二叉树遍历的非递归算法	154
5.8.2 线索二叉树	157
5.9 应用实例	161
5.9.1 堆与优先队列	161
5.9.2 并查集	165
思想火花——调试程序与魔术表演	167
习题	168
实验题	169
第 6 章 图	171
6.1 引言	172
6.2 图的逻辑结构	173
6.2.1 图的定义和基本术语	173
6.2.2 图的抽象数据类型定义	176
6.2.3 图的遍历操作	176
6.2.4 图的 Java 接口定义	178
6.3 图的存储结构及实现	179
6.3.1 邻接矩阵	179
6.3.2 邻接表	183
6.3.3 邻接矩阵和邻接表的比较	188

6.4 最小生成树	188
6.4.1 Prim 算法	189
6.4.2 Kruskal 算法	192
6.5 最短路径	196
6.5.1 Dijkstra 算法	196
6.5.2 Floyd 算法	199
6.6 有向无环图及其应用	201
6.6.1 AOV 网与拓扑排序	201
6.6.2 AOE 网与关键路径	204
6.7 扩展与提高	206
6.7.1 图的其他存储方法	206
6.7.2 图的连通性	208
6.8 应用实例	210
6.8.1 七巧板涂色问题	210
6.8.2 医院选址问题	211
思想火花——直觉可能是错误的	213
习题	214
实验题	217
第7章 查找技术	219
7.1 引言	220
7.1.1 查找的基本概念	220
7.1.2 查找算法的性能	221
7.2 线性表的查找技术	221
7.2.1 线性表查找结构的定义	221
7.2.2 顺序查找	222
7.2.3 折半查找	222
7.3 树表的查找技术	225
7.3.1 二叉排序树	225
7.3.2 平衡二叉树	231
7.3.3 B 树	235
7.4 散列表的查找技术	239
7.4.1 散列查找的基本思想	239
7.4.2 散列函数的设计	240
7.4.3 处理冲突的方法	242
7.4.4 散列查找的性能分析	245
7.4.5 开散列表与闭散列表的比较	246
7.5 各种查找方法的比较	247

7.6 扩展与提高	247
7.6.1 顺序查找的改进——分块查找	247
7.6.2 折半查找的改进——插值查找	248
7.6.3 B 树的改进——B ⁺ 树	249
思想火花——把注意力集中于主要因素,不要纠缠于噪声	250
习题	251
实验题	253
第 8 章 排序技术	255
8.1 引言	256
8.1.1 排序的基本概念	256
8.1.2 排序算法的性能	257
8.1.3 排序类的定义	257
8.2 插入排序	258
8.2.1 直接插入排序	258
8.2.2 希尔排序	260
8.3 交换排序	261
8.3.1 起泡排序	261
8.3.2 快速排序	263
8.4 选择排序	266
8.4.1 简单选择排序	266
8.4.2 堆排序	268
8.5 归并排序	273
8.5.1 二路归并排序的递归实现	273
8.5.2 二路归并排序的非递归实现	275
8.6 各种排序技术的使用	277
8.7 各种排序方法的比较	277
8.8 扩展与提高	279
8.8.1 排序问题的时间下界	279
8.8.2 基数排序	280
思想火花——学会“盒子以外的思考”	283
习题	284
实验题	286
附录 A 预备知识	289
A.1 数学术语	289
A.2 级数求和	289
A.3 集合	290

A.4 关系	291
附录 B Java 语言基本语法	293
B.1 程序文件结构	293
B.2 数据类型	294
B.3 Java 编程规范	295
B.4 控制语句	295
B.5 函数	296
B.6 类与对象	297
B.7 接口	298
B.8 异常处理	299
附录 C 中英文词汇对照表	301
参考文献	305

第1章

绪论

CHAPTER 1

本章概述	用计算机求解任何问题都离不开程序设计,程序设计的关键是数据表示和数据处理。数据若要被计算机处理,首先必须能够存储在计算机的内存中,这项任务称为数据表示,其核心是数据结构;一个实际问题的求解必须满足多项处理要求,这项任务称为数据处理,其核心是算法。数据结构课程讨论数据表示和数据处理的基本思想和方法。 本章通过一个问题实例的求解过程给出程序设计的一般过程,说明数据结构和算法在程序设计中的作用,介绍数据结构和算法的基本概念,说明用大 O 记号进行算法分析的基本方法。				
教学重点	数据结构的基本概念;数据的逻辑结构、存储结构以及二者之间的关系;算法及算法的特性;大 O 记号				
教学难点	抽象数据类型;算法的时间复杂度分析				
教学内容和 教学目标	知 识 点	教 学 要 求			
		了解	理解	掌握	熟练掌握
	程序设计的一般过程	√			
	数据结构在程序设计中的作用		√		
	算法在程序设计中的作用		√		
	数据结构的基本概念				√
	抽象数据类型		√		
	算法及算法的特性				√
	算法的描述方法			√	
	算法的时间复杂度			√	
算法的空间复杂度		√			

1.1 问题求解与程序设计

计算机科学致力于研究用计算机求解人类生产生活中的各种实际问题,只有最终在计算机上能够运行良好的程序才能解决特定的实际问题,因此程序设计的过程就是利用计算机求解问题的过程。

1.1.1 程序设计的一般过程

用计算机求解任何问题都离不开程序设计,但是计算机不能分析问题并产生问题的解决方案,必须由人(即程序设计者)分析问题,确定问题的解决方案,采用计算机能够理解的指令描述这个问题的求解步骤(即编写程序),然后让计算机执行程序最终获得问题的解。用计算机求解问题(即程序设计)的一般过程^①如图 1-1 所示。

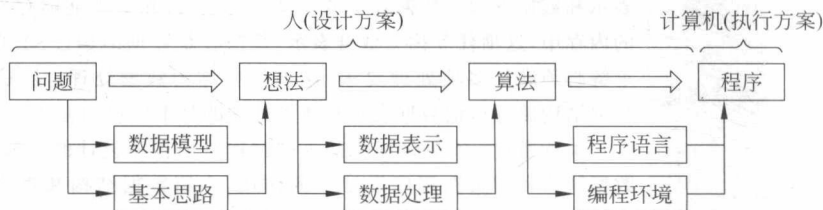


图 1-1 程序设计的一般过程

由问题到想法需要分析待处理的数据以及数据之间的关系,抽象出具体的数据模型并形成问题求解的基本思路。对于数值问题抽象出的数据模型通常是数学方程,对于非数值问题抽象出的数据模型通常是表、树、图等数据结构。

算法用来描述问题的解决方案,是具体的、机械的操作步骤。利用计算机解决问题最重要的步骤是将人的想法描述成算法,也就是从计算机的角度设想计算机是如何一步一步完成这个任务的。由想法到算法需要完成数据表示和数据处理,即描述问题的数据模型,将数据模型从机外表示转换为机内表示;描述问题求解的基本思路,将问题的解决方案形成算法。

由算法到程序需要将算法的操作步骤转换为某种程序设计语言对应的语句,转换所依据的规则就是某种程序设计语言的语法,换言之,就是在某种编程环境下用程序设计语言描述要处理的数据以及数据处理的过程。

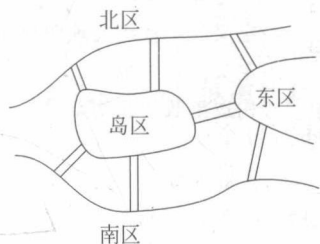
著名的计算机学者沃思^②给出了一个公式:算法+数据结构=程序。从这个公式可以看到,数据结构和算法是构成程序的两个重要的组成部分。一个“好”程序首先是将问

^① 在作者编著的算法与程序设计系列教材《程序设计基础——从问题到程序》《数据结构——从概念到 C 实现》《数据结构——从概念到 C++ 实现》《数据结构——从概念到 Java 实现》《算法设计与分析》中,都是通过这个问题求解过程培养学生的计算思维能力和程序设计能力。

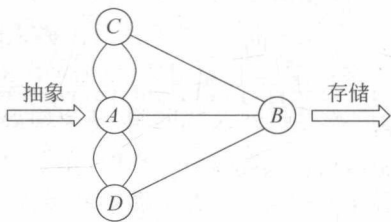
^② 沃思(Niklaus Wirth)1934 年生于瑞士。1968 年设计并实现了 Pascal 语言,1971 年提出了结构化程序设计,1976 年设计并实现了 Modula 2 语言。除了程序设计语言之外,沃思在其他方面也有许多创造,如扩充了著名的巴科斯范式、发明了语法图等。1984 年获图灵奖。

题抽象为一个适当的数据模型并转换为机内表示,然后基于该数据结构设计一个“好”算法,或者说,学习数据结构的意义在于编写高质量、高效率的程序。下面以著名的哥尼斯堡七桥问题为例,说明程序设计的一般过程。

【问题】 哥尼斯堡七桥问题(以下简称七桥问题)。17世纪的东普鲁士有一座哥尼斯堡城(现在叫加里宁格勒,在波罗的海南岸),城中有一座岛,普雷格尔河的两条支流环绕其旁,并将整个城市分成北区、东区、南区和岛区4个区域,全城共有七座桥将4个城区连接起来,如图1-2(a)所示。于是,产生了一个有趣的问题:一个人是否能在一次步行中经过全部的七座桥后回到出发点,且每座桥只经过一次。



(a) 七桥问题



(b) 七桥问题的数据模型

	0(A)	1(B)	2(C)	3(D)
0(A)	0	1	2	2
1(B)	1	0	1	1
2(C)	2	1	0	0
3(D)	2	1	0	0

(c) 对应的二维数组mat[4][4]

图 1-2 七桥问题的数据抽象过程

【想法】 将城区抽象为顶点,用A、B、C、D表示4个城区,将桥抽象为边,用7条边表示七座桥,抽象出七桥问题的数据模型如图1-2(b)所示,从而将七桥问题抽象为一个数学问题:求经过图中每条边一次且仅一次的回路,后来人们称为欧拉回路。欧拉回路的判定规则是:

- (1) 如果没有一个城区通奇数桥,则无论从哪里出发都能找到欧拉回路。
- (2) 如果通奇数桥的城区多于两个,则不存在欧拉回路。
- (3) 如果只有两个城区通奇数桥,则不存在欧拉回路,但可以从这两个城区之一出发找到欧拉路径(不要求回到出发点)。

由上述判定规则得到求解七桥问题的基本思路:依次计算与每个顶点相关联的边数,根据边数为奇数的顶点个数判定是否存在欧拉回路。

【算法】 将顶点A、B、C、D编号为0、1、2、3,用二维数组mat[4][4]存储七桥问题的数据模型,如果顶点 $i(0 \leq i \leq 3)$ 和顶点 $j(0 \leq j \leq 3)$ 之间有 k 条边,则元素mat[i][j]的值为 k ,如图1-2(c)所示。求解七桥问题的关键是求与每个顶点相关联的边数,即在二维数组mat[4][4]中求每一行元素之和。算法描述如下:

算法: LoddVertexNum

输入: 二维数组 mat[4][4], 顶点个数 n

输出: 通奇数桥的顶点个数 count

1. count 初始化为 0;
2. 下标 i 从 0~n-1 重复执行下述操作:
 - 2.1 计算第 i 行元素之和 degree;

```

2.2 如果 degree 为奇数,则 count++;
3. 返回 count;

```

【程序】 定义 EulerCircuit 类,成员变量定义实现欧拉回路问题模型存储, getOddVertexNum 函数实现奇数边顶点个数计算,主函数实例化 EulerCircuit 对象,通过 getOddVertexNum 函数计算图模型中通奇数桥顶点个数,再根据欧拉规则判定是否存在欧拉回路。程序如下:

```

public class EulerCircuit {
    int[][] mat; //定义二维数组实现图存储
    int vertexNum; //定义顶点个数
    public EulerCircuit(int[][] mat, int vertexNum) { //构造函数问题初始化
        this.mat=mat;
        this.vertexNum=vertexNum;
    }
    public int getOddVertexNum() {
        int count=0; //累加器记录奇数顶点个数
        for(int i=0; i<vertexNum; i++) { //依次累加每一行元素
            int degree=0; //记录通过顶点 i 的边数
            for(int j=0; j<vertexNum; j++) {
                degree=degree+mat[i][j];
            }
            if(degree % 2 !=0) //顶点度为奇数
                count++;
        }
        return count; //结束函数,返回奇数顶点个数
    }
    public static void main(String[] args) {
        //七桥问题数组
        int[][] mat={{0, 1, 2, 2},{1, 0, 1, 1},{2, 1, 0, 0},{2, 1, 0, 0}};
        //生成 EulerCircuit 对象
        EulerCircuit eulerCircuit=new EulerCircuit(mat, 4);
        //据函数获取奇数桥个数
        int oddVertexNum=eulerCircuit.getOddVertexNum();
        if(oddVertexNum>=2) //两个以上顶点通奇数桥
            System.out.print("有"+oddVertexNum+"个顶点通奇数桥,无欧拉回路。");
        else
            System.out.print("存在欧拉回路。");
    }
}

```