

数据结构

路莹 主编
刘娜 姚春龙 副主编



清华大学出版社

013057337

21世纪高等学校规划教材 | 计算机科学与技术

TP311.12-43
173

数据结构

路莹 主编

刘娜

姚春龙

副主编

★ 藏书 ★
图书馆



北航

C1662099

清华大学出版社

北京

TP311.12-43
173

内 容 简 介

“数据结构”课程是计算机程序设计的重要理论基础,它不仅是计算机专业重要的专业基础课程与核心课程,同时也是信息管理专业以及数学统计专业的主要课程。

全书共分为9章。第1章绪论,以非数值计算的程序设计实际问题为例,说明什么是数据结构,数据结构的研究内容以及相关概念,最后讨论了算法分析;第2~7章分别论述了线性表、栈和队列、串及模式匹配、数组和广义表、树与二叉树、图等基本数据类型以及相关的数据操作。对于典型操作,书中给出了详尽的算法分析过程和算法描述;第8章和第9章讨论了程序设计中常见的查找和排序问题,并就典型方法进行详尽的算法分析和描述。

本书理论内容阐述详尽,算法分析循序渐进富有逻辑性,算法描述清晰准确,理论知识剖析清楚,且注重解题思路及实用技巧的培养,具有很强的实用性和针对性。书中的算法均采用C语言实现,可直接在任何C环境下调试运行。

本书既可用于高等院校计算机及相关专业的学生学习数据结构课程,也可用于报考计算机专业硕士研究生考试的本科生进行专业课复习,还可以作为相关领域的技术人员和读者自学的参考教材。

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/路莹主编. —北京:清华大学出版社,2013.7

21世纪高等学校规划教材·计算机科学与技术

ISBN 978-7-302-32005-0

I. ①数… II. ①路… III. ①数据结构 IV. ①TP311.12

中国版本图书馆CIP数据核字(2013)第078318号

责任编辑:刘向威 薛 阳

封面设计:傅瑞学

责任校对:梁 毅

责任印制:刘海龙

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:清华大学印刷厂

经 销:全国新华书店

开 本:185mm×260mm

印 张:18

字 数:447千字

版 次:2013年7月第1版

印 次:2013年7月第1次印刷

印 数:1~2000

定 价:33.00元

产品编号:044814-01

编审委员会成员

(按地区排序)

清华大学

周立柱 教授
覃 征 教授
王建民 教授
冯建华 教授
刘 强 副教授
杨冬青 教授
陈 钟 教授
陈立军 副教授
马殿富 教授
吴超英 副教授
姚淑珍 教授
王 珊 教授
孟小峰 教授
陈 红 教授
周明全 教授
阮秋琦 教授
赵 宏 副教授
孟庆昌 教授
杨炳儒 教授
陈 明 教授
艾德才 教授
吴立德 教授
吴百锋 教授
杨卫东 副教授
苗夺谦 教授
徐 安 教授
邵志清 教授
杨宗源 教授
应吉康 教授
乐嘉锦 教授
孙 莉 副教授

北京大学

北京航空航天大学

中国人民大学

北京师范大学

北京交通大学

北京信息工程学院

北京科技大学

石油大学

天津大学

复旦大学

同济大学

华东理工大学

华东师范大学

东华大学

浙江大学	吴朝晖	教授
	李善平	教授
扬州大学	李 云	教授
南京大学	骆 斌	教授
	黄 强	副教授
南京航空航天大学	黄志球	教授
	秦小麟	教授
南京理工大学	张功萱	教授
南京邮电学院	朱秀昌	教授
苏州大学	王宜怀	教授
	陈建明	副教授
江苏大学	鲍可进	教授
中国矿业大学	张 艳	教授
武汉大学	何炎祥	教授
华中科技大学	刘乐善	教授
中南财经政法大学	刘腾红	教授
华中师范大学	叶俊民	教授
	郑世珏	教授
	陈 利	教授
江汉大学	颜 彬	教授
国防科技大学	赵克佳	教授
	邹北骥	教授
中南大学	刘卫国	教授
湖南大学	林亚平	教授
西安交通大学	沈钧毅	教授
	齐 勇	教授
长安大学	巨永锋	教授
哈尔滨工业大学	郭茂祖	教授
吉林大学	徐一平	教授
	毕 强	教授
山东大学	孟祥旭	教授
	郝兴伟	教授
厦门大学	冯少荣	教授
厦门大学嘉庚学院	张思民	教授
云南大学	刘惟一	教授
电子科技大学	刘乃琦	教授
	罗 蕾	教授
成都理工大学	蔡 淮	教授
	于 春	副教授
西南交通大学	曾华燊	教授

出版说明

随着我国改革开放的进一步深化,高等教育也得到了快速发展,各地高校紧密结合地方经济建设发展需要,科学运用市场调节机制,加大了使用信息科学等现代科学技术提升、改造传统学科专业的投入力度,通过教育改革合理调整和配置了教育资源,优化了传统学科专业,积极为地方经济建设输送人才,为我国经济社会的快速、健康和可持续发展以及高等教育自身的改革发展做出了巨大贡献。但是,高等教育质量还需要进一步提高以适应经济社会发展的需要,不少高校的专业设置和结构不尽合理,教师队伍整体素质亟待提高,人才培养模式、教学内容和方法需要进一步转变,学生的实践能力和创新精神亟待加强。

教育部一直十分重视高等教育质量工作。2007年1月,教育部下发了《关于实施高等学校本科教学质量与教学改革工程的意见》,计划实施“高等学校本科教学质量与教学改革工程”(简称“质量工程”),通过专业结构调整、课程教材建设、实践教学改革、教学团队建设等多项内容,进一步深化高等学校教学改革,提高人才培养的能力和水平,更好地满足经济社会发展对高素质人才的需要。在贯彻和落实教育部“质量工程”的过程中,各地高校发挥师资力量强、办学经验丰富、教学资源充裕等优势,对其特色专业及特色课程(群)加以规划、整理和总结,更新教学内容、改革课程体系,建设了一大批内容新、体系新、方法新、手段新的特色课程。在此基础上,经教育部相关教学指导委员会专家的指导和建议,清华大学出版社在多个领域精选各高校的特色课程,分别规划出版系列教材,以配合“质量工程”的实施,满足各高校教学质量和教学改革的需要。

为了深入贯彻落实教育部《关于加强高等学校本科教学工作,提高教学质量的若干意见》精神,紧密配合教育部已经启动的“高等学校教学质量与教学改革工程精品课程建设工作”,在有关专家、教授的倡议和有关部门的大力支持下,我们组织并成立了“清华大学出版社教材编审委员会”(以下简称“编委会”),旨在配合教育部制定精品课程教材的出版规划,讨论并实施精品课程教材的编写与出版工作。“编委会”成员皆来自全国各类高等学校教学与科研第一线的骨干教师,其中许多教师为各校相关院、系主管教学的院长或系主任。

按照教育部的要求,“编委会”一致认为,精品课程的建设工作从开始就要坚持高标准、严要求,处于一个比较高的起点上。精品课程教材应该能够反映各高校教学改革与课程建设的需要,要有特色风格、有创新性(新体系、新内容、新手段、新思路,教材的内容体系有较高的科学创新、技术创新和理念创新的含量)、先进性(对原有的学科体系有实质性的改革和发展,顺应并符合21世纪教学发展的规律,代表并引领课程发展的趋势和方向)、示范性(教材所体现的课程体系具有较广泛的辐射性和示范性)和一定的前瞻性。教材由个人申报或各校推荐(通过所在高校的“编委会”成员推荐),经“编委会”认真评审,最后由清华大学出版

社审定出版。

目前,针对计算机类和电子信息类相关专业成立了两个“编委会”,即“清华大学出版社计算机教材编审委员会”和“清华大学出版社电子信息教材编审委员会”。推出的特色精品教材包括:

(1) 21 世纪高等学校规划教材·计算机应用——高等学校各类专业,特别是非计算机专业的计算机应用类教材。

(2) 21 世纪高等学校规划教材·计算机科学与技术——高等学校计算机相关专业的教材。

(3) 21 世纪高等学校规划教材·电子信息——高等学校电子信息相关专业的教材。

(4) 21 世纪高等学校规划教材·软件工程——高等学校软件工程相关专业的教材。

(5) 21 世纪高等学校规划教材·信息管理与信息系统。

(6) 21 世纪高等学校规划教材·财经管理与应用。

(7) 21 世纪高等学校规划教材·电子商务。

(8) 21 世纪高等学校规划教材·物联网。

清华大学出版社经过三十多年的努力,在教材尤其是计算机和电子信息类专业教材出版方面树立了权威品牌,为我国的高等教育事业做出了重要贡献。清华版教材形成了技术准确、内容严谨的独特风格,这种风格将延续并反映在特色精品教材的建设中。

清华大学出版社教材编审委员会

联系人:魏江江

E-mail: weijj@tup.tsinghua.edu.cn

前言

随着计算机的普及,计算机软件的应用渗透各行各业,日益普遍。在进行程序设计时,对程序处理的复杂数据进行系统的分析和结构设计,是系统化程序设计、正确解决问题的关键,是开发出高质量、高效率的程序的重要环节。因此,“数据结构”课程已经成为计算机及其相关专业的核心课程,是从事计算机软件开发、应用人员必备的专业基础。同时,自2009年起,“数据结构”成为全国硕士研究生入学统一考试计算机科学与技术学科联考的重点课程,成为衡量计算机学科学生学习成效的一个重要标准。为了满足数据结构课程的日常学习和晋级考试的需求,我们总结多年的教学经验、细化学习技巧,精心编写了本书,旨在系统、全面地讲解数据结构课程的研究内容。本书系统全面,重点突出,理论知识剖析清楚,且注重解题思路及实用技巧的培养,具有以下几个特色:

(1) 深入浅出,通俗易懂。如何系统而全面地掌握数据结构的解题思路和算法设计思想是学习数据结构课程的难点,而有效理解数据表示和数据处理、正确分析算法设计的要点、建立算法设计思路成为学好本课程的关键。本书详细介绍了线性表、栈和队列、串、数组和广义表、树和图等数据结构,以及在程序设计过程中经常遇到的查找和排序问题。全书共分9章,每章从典型的实例分析入手,从应用出发,系统地进行理论阐述并配以精确的算法分析与描述,帮助读者快速理解数据结构中的各个知识点、掌握重点内容、突破学习瓶颈,从而使读者更好地应对各种应用需求。

(2) 本书中的算法例程均由作者采用C语言编写,可在C环境下直接调试运行。目前,同类图书中的算法描述大多比较粗略,且采用伪代码描述,没有使用真正的计算机语言实现,不便于学生理解和参考。本书中的主要算法在逻辑分析之后,均采用C语言编写实现,学生在学习完理论知识后可直接调用算法程序调试运行,从而获得直观印象,帮助学生更好地理解算法细节,激发学生的学习热情。

(3) 理论联系实际。书中每章节后都附有典型习题,帮助学生更好地检测学习效果。同时,本书还配套有待出版的《数据结构习题精解与考试指导》一书,该书通过典型的习题精解、大量习题测试和综合上机题目推荐,为学生学习该课程以及通过更高一级的学习考试提供支持。

本书由大连工业大学路莹主编,大连工业大学姚春龙、刘娜副主编。其中第1章、第7章、第8章、第9章由路莹编写,第2章、第3章由姚春龙编写,第4章、第5章、第6章由刘娜编写,全书由路莹统编。

适合阅读本书的读者非常广泛,包括在读的本专科、中专职高技校等计算机专业学生、想转行做软件开发的非专业人员、欲考取计算机专业研究生的应届或在职人员,以及工作后需要补学或温习数据结构和算法的程序员等各类读者。

由于作者水平有限,书中难免有不妥与疏漏之处,敬请广大读者批评指正。

编 者

2013年1月

目 录

第 1 章 绪论	1
1.1 什么是数据结构	1
1.2 基本术语	4
1.3 算法和算法分析	6
1.3.1 算法	7
1.3.2 算法设计的要求	7
1.3.3 算法效率的度量	7
小结	9
习题	9
第 2 章 线性表	11
2.1 线性表的定义和基本操作	11
2.1.1 线性表的定义	11
2.1.2 线性表的抽象数据类型	12
2.2 线性表的顺序存储结构	12
2.2.1 顺序表的结构	13
2.2.2 顺序表基本操作的实现	14
2.2.3 顺序表的一个简单应用	19
2.3 线性表的链式存储结构	22
2.3.1 线性链表	23
2.3.2 线性链表的一个应用实例	29
2.3.3 循环链表和双向链表	34
2.3.4 循环链表的一个应用实例	37
小结	40
习题	41
第 3 章 栈和队列	45
3.1 栈	45
3.1.1 栈的定义	45
3.1.2 栈的抽象数据类型	45
3.1.3 栈的表示和实现	46
3.2 栈的应用举例	51

3.2.1	数制转换	51
3.2.2	括号匹配检测	53
3.2.3	表达式求值	54
3.2.4	迷宫求解	60
3.3	栈与递归	66
3.3.1	函数的嵌套调用	66
3.3.2	递归调用	66
3.4	队列	71
3.4.1	队列的定义	71
3.4.2	队列的抽象数据类型	71
3.4.3	链队列——队列的链式表示和实现	72
3.4.4	循环队列——队列的顺序表示和实现	75
3.4.5	一个队列的应用实例	78
小结		83
习题		83
第4章	串及模式匹配	86
4.1	串类型的定义	86
4.2	串的存储结构及其运算	87
4.2.1	串的定长顺序存储	87
4.2.2	堆的分配存储结构	89
4.2.3	串的块链存储结构	91
4.3	串的模式匹配	91
4.3.1	简单的模式匹配算法	92
4.3.2	改进后的模式匹配算法	94
4.4	串操作应用举例	100
小结		107
习题		107
第5章	数组和广义表	108
5.1	数组的定义和运算	108
5.2	数组的顺序存储结构	109
5.3	矩阵的压缩存储	110
5.3.1	特殊矩阵	110
5.3.2	稀疏矩阵	114
5.4	广义表	118
5.4.1	广义表的定义	118
5.4.2	广义表的运算	119
5.5	广义表的存储结构	119

小结	121
习题	121
第6章 树与二叉树	122
6.1 树的基本概念	122
6.2 二叉树	124
6.2.1 二叉树的定义	124
6.2.2 二叉树的性质	125
6.2.3 二叉树的存储结构	126
6.3 遍历二叉树和线索二叉树	129
6.3.1 二叉树的遍历方法与算法实现	129
6.3.2 二叉树的非递归算法实现	131
6.3.3 由遍历序列恢复二叉树	135
6.3.4 线索二叉树	137
6.4 树和森林	138
6.4.1 树的存储表示	138
6.4.2 森林与二叉树的转换	140
6.4.3 树和森林的遍历	142
6.5 哈夫曼树及其应用	144
6.5.1 最优二叉树	144
6.5.2 哈夫曼编码	146
6.6 回溯法与树的遍历	149
小结	150
习题	150
第7章 图	152
7.1 图的基本概念	152
7.2 图的存储结构及基本操作	155
7.2.1 邻接矩阵	155
7.2.2 邻接表	158
7.2.3 有向图的十字链表	162
7.2.4 邻接多重表	166
7.3 图的遍历	167
7.3.1 深度优先搜索(Depth_First Search)	168
7.3.2 广度优先搜索(Breadth_First Search)	170
7.4 图的连通性问题	174
7.4.1 无向图的连通性	174
7.4.2 有向图的连通性	174
7.4.3 生成树和生成森林	175

7.5 有向无环图及其应用	185
7.5.1 拓扑排序	186
7.5.2 关键路径	192
7.6 最短路径	195
7.6.1 从一个源点到其余各顶点的最短路径	195
7.6.2 每一对顶点间的最短路径	197
小结	199
习题	199
第8章 查找	200
8.1 基本概念	200
8.2 静态查找表	201
8.2.1 顺序查找	201
8.2.2 有序表的查找	203
8.2.3 索引顺序表的查找	207
8.3 动态查找表	210
8.3.1 二叉排序树	210
8.3.2 平衡二叉树	221
8.3.3 B-树和 B+树	230
8.4 哈希表	236
8.4.1 哈希表和哈希查找	237
8.4.2 常用的哈希函数	238
8.4.3 处理冲突的方法	241
8.4.4 哈希表的查找及其分析	244
小结	245
习题	245
第9章 内部排序	247
9.1 排序的基本概念	247
9.2 插入排序	248
9.2.1 直接插入排序	248
9.2.2 希尔排序	251
9.3 交换排序	253
9.3.1 冒泡排序	253
9.3.2 快速排序	255
9.4 选择排序	259
9.4.1 简单选择排序	259
9.4.2 堆排序	261
9.5 归并排序	265

9.6 基数排序	267
9.6.1 多关键字排序的算法思想	267
9.6.2 链式基数排序	268
9.7 各种内部排序算法的比较	270
9.7.1 选择排序算法的依据	270
9.7.2 选择排序算法的结论	270
小结	271
习题	271

第1章

绪论

数据结构不仅是计算机专业的核心基础课程,也是其他非计算机专业的重要选修课程,是程序设计的基础。它介于数学、计算机硬件和计算机软件三者之间,是计算机系统软件和应用软件研制者的必修课程,是学习软件知识和提高软件设计水平的重要基础。

1.1 什么是数据结构

问题 1: 计算圆的面积。

问题 2: 简单的学生成绩管理系统。

对于问题 1,我们只需知道求解圆的面积公式等于 $\pi \times R \times R$,即可通过定义圆的半径变量、输入半径值、套用面积公式非常轻松地计算出给定圆的面积。

对于问题 2,我们需要分析学生都有哪些属性信息,需要管理学生的哪些信息,如学生学号、学生姓名、学生籍贯、学生奖惩信息、学生专业、学生高数成绩、学生英语成绩、学生程序设计成绩、学生操作系统成绩……。通过对这些学生属性信息进行分析,我们才能根据程序功能要求,选取合适的学生属性信息,并将学校里几千名甚至几万名的学生信息很好地组织起来,然后按照一定的处理方法编写出功能正确的学生成绩管理系统。

通过上面两个问题的分析,我们发现:与早期计算机主要用于处理复杂的科学计算不同(如导弹飞行轨迹、桥梁应力等),目前,计算机主要用于处理控制、管理等非数值计算问题,如图书管理系统、多媒体播放系统、网络监控系统。计算机加工处理的对象也从纯粹的数值发展到字符、图像、语音等具有复杂结构的数据,这类问题的数据量大、数据间的相互关系多样,且不能依靠明确的数学公式加以描述和处理。为了正确处理这类数据,必须预先对程序处理的复杂数据进行系统的分析和结构设计,研究待处理对象的特性以及各处理对象之间存在的关系,并根据其关系制订合理、有效的处理方法。这是系统化程序设计、正确解决问题的关键,是开发出高质量、高效率的程序的重要环节。

下面,我们继续来看三个案例。

案例 1: 图书管理系统。要通过计算机实现快速、准确的图书信息管理,需要建立相应的数据表来保存图书信息,然后按照某种算法编写计算机程序,最终实现图书信息更新、图书信息查找、图书信息删除等管理功能。为此,建立图书信息表、图书类别表、作者信息表、出版单位信息表,如表 1.1~表 1.4 所示。

表 1.1 图书信息表

图书编号	ISBN	书名	作者	图书类别	出版单位编号	...
100000011	978-7-32-14751-0	数据结构	严蔚敏	J0003	QH001	
420000020	978-7-5366-7987-0	藏地密码	何马	WX011	CQ001	
010000001	978-7-04-010820-0	高等数学	同济大学应用数学系	S0001	GJ001	
020000001	978-7-81-080159-1	大学英语综合教程	李荫华	F0001	SW001	

表 1.2 图书类别表

类别编号	类别名称	类别编号	类别名称
F0001	大学英语教材类	S0001	高等数学教材类
J0003	计算机教材类	WX011	文学小说类

表 1.3 作者信息表

作者编号	作者姓名	作者工作单位	...
10012	严蔚敏	清华大学	
35012	何马	自由职业	
10533	李荫华	上海外国语学院	
10611	同济大学应用数学系	同济大学	

表 1.4 出版单位信息表

出版单位编号	出版单位名称	出版单位编号	出版单位名称
QH001	清华大学出版社	GJ001	高等教育出版社
C001	重庆出版集团	SW001	上海外语教育出版社

根据这 4 张表构成的图书处理数学模型,系统可以按读者要求实现图书检索功能。比如,浏览全部图书、按图书类别查询图书、按作者信息查询图书、按出版社信息查询图书……与此类似的还有通讯录系统、企业员工管理系统、银行账目管理系统等。在这类非数值计算的数据处理系统中,计算机处理的对象之间存在一种简单的线性管理,这类数学模型称为线性的数据结构。

案例 2: 人-机对弈问题。对弈的过程是一个在一定规则约束下的随机过程,为保证人机对弈正常进行,在和人弈之前,需要事先将对弈策略存入计算机。制订对弈策略时,需要对对弈过程中所有可能发生的情况和相应的对策都考虑周全,并根据棋盘当时的状态,预测棋局的发展态势。如此,才能保证人机对弈直至最后结局。在对弈过程中,计算机操作的对象是对弈过程中可能出现的棋盘状态,即“格局”。

图 1.1 是“五子棋”的对弈树。上面的大图是“五子棋”的一个格局,依据比赛规则,该格局可以派生出多个“子格局”,下面的小图是其中的 5 个子格局。格局之间的关系是由比赛规则决定的,每个格局又可以派生出多个子格局。从图示可以看出:它是一种“非线性”关系,我们称之为“树”。纵观整个对弈过程,从开始到结束,可以将每个格局及其子格局的这

种“多分支”关系描述成一棵“倒长”的树,“树根”就是对弈开始前的棋盘格局,每个“树叶”就是可能出现的结局,对弈的过程就是从树根沿分支到达某个树叶的过程。“树”是用来处理某些非数值计算问题的数学模型,是一种非线性的数据结构。

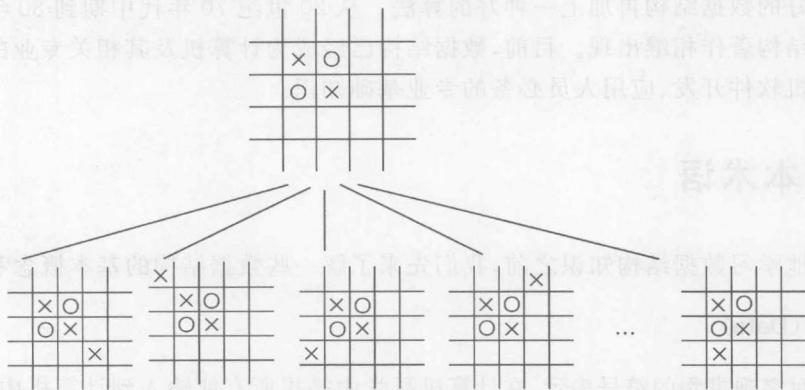


图 1.1 “五子棋”对弈树

案例 3: 城市公共交通咨询系统。城市公交线路繁多、四通八达。要想从某一地点到另一地点,通常有多条线路可选。这些路线的选择,通常受转车次数、路程时间、路程花费等因素制约。为此,很多城市都建设了城市公共交通咨询系统。

图 1.2 描述了某个城市的部分公共交通路线。假设,某名乘客要从地点 A 乘车到地点 G,请求咨询系统给出合适的乘车路线。为了解决这个问题,系统必须建设城市公共交通的数学模型,有效地描述各个地点及其之间的路况、公交线路设置情况,以及每条线路的花费(距离、乘车时间、乘车花费等)。对于这类问题,通常采用“图/网”来描述路况——将地点抽象成一个点,地点之间的线路抽象成一条线,线路的花费用线上的数字描述,这些点线就组成了一个图/网。然后,再根据图/网的算法求解乘客对乘车路线的要求,并给出合理的乘车路线建议。图/网是一种复杂的、用来处理某些非数值计算问题的非线性数据结构。

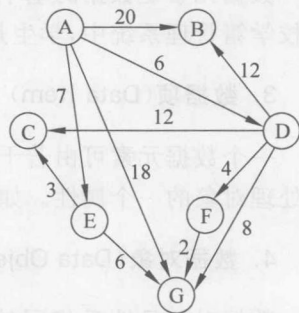


图 1.2 城市公共交通示意图

通过上述 3 个案例可见,求解这类非数值计算问题的关键不再是数学公式,而是诸如表、树和图之类的数据结构。总地来说,数据结构是一门研究非数值计算的程序设计问题中数据对象之间的相互关系及其处理方法的课程。科学计算或数据处理、过程控制、对文件的存储和检索及数据库技术等计算机应用,都是对数据进行加工处理的过程。因此,要设计出一个结构好而且效率高的程序,必须研究数据的特性、数据间的相互关系及其对应的存储表示,并利用这些特性和关系设计出相应的算法和程序。

“数据结构”课程最早是在 1968 年作为独立的课程在国外开设的。1968 年,美国计算机协会 (Association of Computing Machinery, ACM) 颁布了建议性的计算机教学计划,该计划中把数据结构作为一门独立的课程。同年,著名的计算机科学家、美国的 D. E. Knuth 教授论著的《计算机程序设计技巧》第一卷《基本算法》出版,该书第一次全面系统地阐述了数据的逻辑结构和存储结构,以及各种典型操作算法,为数据结构奠定了理论基础。著名的

计算机科学家 N. Wirth 编写的《算法+数据结构=程序》一书中更明确地指出,在程序设计中,数据结构与算法同等重要。从 20 世纪 60 年代末到 70 年代初,随着计算机的广泛应用,结构化程序设计成程序设计方法学的主要内容,人们越来越重视“数据结构”,认为好的程序其实质就是好的数据结构再加上一种好的算法。从 20 世纪 70 年代中期到 80 年代初,各种版本的数据结构著作相继出现。目前,数据结构已经成为计算机及其相关专业的核心课程,是从事计算机软件开发、应用人员必备的专业基础知识。

1.2 基本术语

在系统地学习数据结构知识之前,我们先来了解一些数据结构的基本概念和术语。

1. 数据(Data)

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。它是计算机程序加工的“原料”,如整数、实数、字符串、图像、声音等。

2. 数据元素(Data Element)

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。如学校学籍管理系统中,学生是被处理对象,一个学生信息就是一个数据元素。

3. 数据项(Data Item)

一个数据元素可由若干个数据项组成,是数据的不可分割的最小单位。每个数据项描述处理对象的一个属性。如学生信息中的学号、学生姓名等即为一个数据项。

4. 数据对象(Data Object)

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。如整数数据对象集合 $N=\{0,+/-1,+/-2,\dots\}$,学籍管理系统中的学生信息数据库文件。

5. 数据结构(Data Structure)

一个数据对象中,数据元素不是孤立存在的,它们之间存在某种结构关系,这种数据元素相互之间的结构关系简称为结构。

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。使用一个二元组来描述数据结构,即:

$$\text{Data_Structure}=(D,S)$$

其中, D 是元素有限集, S 是关系有限集。下面利用一个更直观的例子来简要说明:

例 1: 用数据结构表示一周中的 7 天。则:

$$\text{Data_Structure}=(D,S)$$

其中, $D=\{\text{Mon,Tue,Wen,Thu,Fri,Sat,Sun}\}$,

$$S=\{<\text{Mon,Tue}>,<\text{Tue,Wen}>,<\text{Wen,Thu}>,<\text{Thu,Fri}>,<\text{Fri,Sat}>,<\text{Sat,Sun}>,<\text{Sun,Mon}>\}$$