

普通高等院校
应用型本科计算机专业系列教材

SHUJU JIEGOU YU SUANFA
YINGYONG JIAOCHENG (C YUYAN BAN)

数据结构与算法…… 应用教程（C语言版）

主 编 / 彭 娟 杨 勇
副主编 / 袁开友 周龙福 郭 静
何春梅 郑孝宗



重庆大学出版社

<http://www.cqup.com.cn>

普通高等院校
应用型本科计算机专业系列教材

SHUJU JIEGOU YU SUANFA
YINGYONG JIAOCHENG (C YUYAN BAN)

数据结构与算法·… 应用教程（C语言版）

主 编 / 彭 娟 杨 勇
副主编 / 袁开友 周龙福 郭 静
何春梅 郑孝宗

重庆大学出版社

内容简介

本书通过具体的编程应用案例,系统地介绍了各种类型数据结构的逻辑结构、存储结构及相关的算法。全书共分10章,内容包括绪论、线性表、栈和队列、串、数组与广义表、树与二叉树、图、查找和排序、常用算法及其应用等,此外本书还附录了“应用实践”拓展训练内容及参考代码,供学生参考和练习。全书采用C语言应用案例驱动教学,讲解数据结构、算法及应用,内容翔实,层次清晰,实例丰富,讲解深入浅出。

本书作为计算机及相关专业本科数据结构课程的教材,也适合各类成人教育相关课程使用,还可以供从事计算机软件开发和应用的工程技术人员阅读、参考。

图书在版编目(CIP)数据

数据结构与算法应用教程: C语言版 / 彭娟, 杨勇主编. —重庆: 重庆大学出版社, 2017.2

普通高等院校应用型本科计算机专业系列教材

ISBN 978-7-5689-0397-4

I. ①数… II. ①彭… III. ①数据结构—高等学校—教材 ②算法分析—高等学校—教材 ③C语言—程序设计—高等学校—教材
IV. ①TP311.12 ②TP312.8

中国版本图书馆CIP数据核字(2017)第013209号

普通高等院校应用型本科计算机专业系列教材 数据结构与算法应用教程(C语言版)

主 编 彭 娟 杨 勇

责任编辑: 章 可 版式设计: 章 可

责任校对: 邬小梅 责任印制: 张 策

*

重庆大学出版社出版发行

出版人: 易树平

社址: 重庆市沙坪坝区大学城西路21号

邮编: 401331

电话: (023) 88617190 88617185 (中小学)

传真: (023) 88617186 88617166

网址: <http://www.cqup.com.cn>

邮箱: fxk@cqup.com.cn (营销中心)

全国新华书店经销

重庆升光电力印务有限公司印刷

*

开本: 787mm × 1092mm 1/16 印张: 17.75 字数: 379千

2017年2月第1版 2017年2月第1次印刷

ISBN 978-7-5689-0397-4 定价: 38.00元

本书如有印刷、装订等质量问题,本社负责调换

版权所有,请勿擅自翻印和用本书
制作各类出版物及配套用书,违者必究

前言

数据结构和算法是计算机科学与工程的基础性学科，是开发高效计算机程序以解决各领域应用问题的核心。寻求和实现数学模型的过程使计算机算法与数据结构密切相关，算法依赖于具体的数据结构，数据结构也直接关系到算法的选择和效率。数据结构和算法课程的学习和应用，不仅可以使学生掌握计算机基础课程的基本方法，更是训练学生计算思维的有效途径。

数据结构和算法是介于数学、计算机硬件和计算机软件三者之间的一门核心课程。这门课程的内容不仅是一般程序设计（特别是非数值程序设计）的基础，而且是设计和实现编译程序、操作系统、数据库系统及其他系统程序的重要基础。学习数据结构的目的是使学生全面理解数据结构和算法的概念、掌握基本的数据结构和算法的特点以及数据结构与算法的主要原理和方法。通过理论学习和应用实践，提高学生的计算思维和使用计算机解决问题的能力。

本教材主要面向计算机及相关专业的本科学生。教材共 10 章：第 1 章绪论；第 2 章线性表；第 3 章栈和队列；第 4 章串；第 5 章数组与广义表；第 6 章树和二叉树；第 7 章图；第 8 章查找；第 9 章排序；第 10 章常用算法及其应用。本书以案例驱动教学，除了注重理论体系的完整性外，特别强调应用能力的培养，通过每章应用案例及实践的学习掌握理论知识、建立程序逻辑、提升编程能力。本书主要具有以下特点：

1. 应用教学驱动案例引领理论讲授，帮助学生循序渐进地学习。

本书每一章都通过分析来自学生身边的应用案例，引出体系相对完整的理论知识，通过精心组织设计的主题和应用，由浅入深、循序渐进地讲解和扩展知识，帮助学生逐步学习及理解数据结构相关知识。

2. 注重数据结构算法思想的培养，帮助学生建立程序逻辑。本书在介绍每种数据结构及其算法时，注重编程思想和程序逻辑的培养，对每种数据结构、算法思想和对应案例都配有大量图表进行详细分析和说明，大部分章节的主要算法配有流程图，帮助学生理解算法思想、建立程序逻辑。

3. 强调应用能力的培养, 帮助学生提升编程能力。

本书所有案例, 在讲清知识、建立思想、理顺逻辑的基础上, 都用完整的 C 代码加以实现, 不但有助于学生深入理解数据结构知识, 而且有助于提升学生实际的应用编程能力, 增强了学生的编程兴趣和成就感。

4. 强化“应用实践”训练, 帮助学生拓展应用编程能力。

本书贯彻分层教学的思想, 在附录中增设了“应用实践”拓展训练内容, 要求能力较强的学生应用前面重点章节的相关知识实现编程, 从而拓展训练学生的应用编程能力和知识迁移能力, 同时附有应用实践参考代码, 供学生参考和练习。本书的教学学时以 60~80 学时为宜, 最好采用一体化教学, 让学生以“学中做, 做中学”的方式开展学习。教师可根据实际学时和学生具体情况等自行调整教学进度和内容。

本书由彭娟老师负责设计全书体系及统稿工作, 并编写了第 6 章、第 9 章和第 10 章, 杨勇编写了第 2 章、第 3 章、第 7 章、附录 A, 袁开友编写了第 4 章、第 8 章和附录 B, 周龙福编写了第 1 章, 郑孝宗编写了第 5 章, 郭静和何春梅编写了附录 C。在教材的编写过程中, 重庆工程学院软件学院的领导和老师对该书的编写提出了很多宝贵的意见和建议, 还得到重庆工程学院教务处的大力支持, 以及广大软件学院学生的积极支持, 在此表示衷心的感谢。

由于编者水平和时间方面的限制, 书中难免存在疏漏与不足之处, 敬请广大读者和同行专家批评指正。

编者

2017 年 1 月

目录 CONTENTS

■ 第1章 绪论	1
1.1 引言	1
1.2 基本概念和术语	4
1.3 算法的概念及其特性	6
1.3.1 算法的定义	6
1.3.2 算法的组成要素	7
1.3.3 算法的基本性质	8
1.4 算法设计的要求	8
1.5 算法的描述方法	9
1.6 算法分析和评价	11
1.6.1 算法的时间复杂度	12
1.6.2 算法的空间复杂度	14
本章小结	15
习题	15
■ 第2章 线性表	19
2.1 线性表的概念与基本操作	19
2.2 顺序表	20
2.2.1 顺序表的概念	21
2.2.2 顺序表基本操作及实现	21
2.3 单链表	28
2.3.1 单链表的概念	28
2.3.2 单链表基本操作及实现	29
2.4 循环链表	34
2.4.1 循环链表的概念	34
2.4.2 循环链表基本操作及实现	34
2.5 双链表	37
2.5.1 双链表的概念	37
2.5.2 双链表基本操作及实现	37

本章小结	42
习题	42
■ 第3章 栈和队列	44
3.1 栈	44
3.1.1 栈的概念及基本操作	44
3.1.2 顺序栈	45
3.1.3 链栈	49
3.1.4 递归和栈	53
3.2 队列	55
3.2.1 队列概念及基本操作	55
3.2.2 顺序队列	56
3.2.3 循环队列	59
3.2.4 链式队列	61
本章小结	65
习题	66
■ 第4章 串	69
4.1 串的概念和基本操作	69
4.2 串的实现	70
4.2.1 顺序定长存储及实现	70
4.2.2 串的堆存储及实现	72
4.2.3 串的链式存储及实现	74
4.3 串的模式匹配算法	76
4.3.1 Brute-Force 模式匹配算法	77
4.3.2 KMP 模式匹配算法	78
本章小结	81
习题	81
■ 第5章 数组与广义表	83
5.1 数组	83
5.1.1 数组的概念	83
5.1.2 数组的存储	83
5.1.3 矩阵的压缩	86

5.2 广义表	97
5.2.1 广义表的定义	97
5.2.2 广义表的特性	98
5.2.3 广义表的存储结构	99
本章小结	100
习题	101
■ 第6章 树和二叉树	102
6.1 树	103
6.1.1 树的概念及基本术语	103
6.1.2 树的逻辑表示方法	104
6.1.3 树的存储结构	104
6.2 二叉树	106
6.2.1 二叉树的概念	106
6.2.2 二叉树的性质	107
6.2.3 二叉树存储结构	109
6.2.4 二叉树的遍历	110
6.3 线索二叉树	113
6.3.1 线索二叉树的概念	113
6.3.2 线索化二叉树	113
6.3.3 遍历线索化二叉树	116
6.4 哈夫曼树	117
6.4.1 哈夫曼树概述	117
6.4.2 哈夫曼树的构造算法	117
6.4.3 哈夫曼编码	119
6.5 树、森林和二叉树	120
6.5.1 二叉树与树、森林之间的转换	120
6.5.2 树、森林的遍历	122
本章小结	124
习题	125
■ 第7章 图	128
7.1 图的概念及基本术语	129
7.2 图的存储结构	130

7.2.1	邻接矩阵表示法	131
7.2.2	邻接表表示法	132
7.3	图的遍历	133
7.3.1	深度优先搜索法	134
7.3.2	广度优先搜索法	136
7.4	图的应用	139
7.4.1	最小生成树	139
7.4.2	最短路径	144
7.4.3	AOV 网与拓扑排序	148
7.4.4	AOE 网与关键路径	152
	本章小结	156
	习题	157
■ 第8章 查找		159
8.1	查找的概念及基本术语	159
8.2	线性表查找	160
8.2.1	顺序查找	160
8.2.2	二分查找	162
8.2.3	分块查找	164
8.2.4	顺序表三种查找方法的比较	169
8.3	树表的查找	169
8.4	哈希表的查找	176
	本章小结	181
	习题	182
■ 第9章 排序		184
9.1	排序的概念及基本术语	184
9.2	插入排序	185
9.2.1	直接插入排序	185
9.2.2	二分插入排序	188
9.2.3	希尔排序	191
9.3	交换排序	193
9.3.1	冒泡排序	193
9.3.2	快速排序	195

9.4	选择排序	197
9.4.1	直接选择排序	197
9.4.2	堆排序	199
9.5	归并排序	201
9.6	基数排序	203
9.7	各种内排序方法的比较和选择	207
	本章小结	208
	习题	209
■	第 10 章 常用算法及其应用	212
10.1	分治算法	212
10.1.1	分治算法概述	212
10.1.2	分治算法的基本步骤	213
10.1.3	分治算法应用实例	213
10.2	动态规划算法	214
10.2.1	动态规划算法概述	214
10.2.2	动态规划算法的基本步骤	215
10.2.3	动态规划算法应用实例	216
10.3	贪心算法	219
10.3.1	贪心算法概述	219
10.3.2	贪心算法的基本步骤	220
10.3.3	贪心算法应用实例	221
10.4	回溯算法	222
10.4.1	回溯算法概述	222
10.4.2	回溯算法的基本步骤及程序模式	223
10.4.3	回溯算法应用实例	224
10.5	分支界限算法	226
10.5.1	分支界限算法概述	226
10.5.2	分支界限算法的基本步骤	227
10.5.3	分支界限算法应用实例	227
	本章小结	230
	习题	231
■	附录 A 应用实践	232
	实践 1 顺序表的建立与基本操作	232

实践 2	单链表的建立与基本操作	233
实践 3	顺序栈的建立与基本操作	233
实践 4	顺序队列的建立与基本操作	233
实践 5	线性表查找	234
实践 6	内排序	234
附录 B	各章习题参考答案	236
附录 C	应用实践参考代码	257
实践 1	顺序表的建立与基本操作	257
实践 2	单链表的建立与基本操作	260
实践 3	顺序栈的建立与基本操作	263
实践 4	顺序队列的建立与基本操作	265
实践 5	线性表查找	268
实践 6	内排序	270
参考文献	274

第1章 | 绪论

1.1 引言

在计算机科学中，数据结构与算法是计算机科学与工程的基础性学科，是开发高效计算机程序以解决各领域应用问题的核心。数据结构与算法是计算机学科中最基础的课程，是介于数学、计算机硬件和计算机软件三者之间的一门核心课程。这门课程的内容不仅是一般程序设计（特别是非数值程序设计）的基础，而且是设计和实现编译程序、操作系统、数据库系统及其他系统程序的重要基础。

计算机解决一个具体问题时，大致需要经过下列几个步骤：首先要从具体问题中抽象出一个适当的数学模型，然后设计一个解此数学模型的算法（Algorithm），最后编出程序、进行测试、调整直至得到最终解答。寻求数学模型的实质是分析问题，从中提取操作的对象，并找出这些操作对象之间的关系，然后用数学的语言加以描述。计算机算法与数据的结构密切相关，算法无不依附于具体的数据结构，数据结构直接关系到算法的选择和效率。

由于早期所涉及的运算对象是简单的整型、实型或布尔类型数据，所以程序设计者的主要精力是集中于程序设计的技巧上，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，非数值计算显得越来越重要。据统计，当今处理非数值计算性问题占用了90%以上的机器时间。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。下面所列举的就是属于这一类的具体问题。

【例1.1】学生信息检索系统。

当我们需要查询某个学生的有关情况时；或者想查询某个专业或年级的学生的有关情况时，只要建立了相关的数据结构，按照某种算法编写相关程序，就可以实现计算机自动检索。由此，可以在学生信息检索系统中建立一张按学号顺序排列的学生信息表（见图1.1（a））以及分别按姓名（见图1.1（b）），专业（见图1.1（c）），年级（见图1.1（d））顺序排列的索引表。由这4张表构成的文件便是学生信息检索的数学模型，计算机的主要操作就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。

学号	姓名	性别	专业	年级
980001	吴承志	男	计算机科学与技术	98级
980002	李淑芳	女	信息与计算科学	98级
990301	刘丽	女	数学与应用数学	99级
990302	张会友	男	信息与计算科学	99级
990303	石宝国	男	计算机科学与技术	99级
000801	何文颖	女	计算机科学与技术	2000级
000802	赵胜利	男	数学与应用数学	2000级
000803	崔文靖	男	信息与计算科学	2000级
010601	刘丽	女	计算机科学与技术	2001级
010602	魏永鸣	男	数学与应用数学	2001级

(a) 学生信息表

崔文靖	8
何文颖	6
李淑芳	2
刘丽	3, 9
石宝国	5
魏永鸣	10
吴承志	1
赵胜利	7
张会有	4

(b) 姓名索引表

计算机科学与技术	1, 5, 6, 9
信息与计算科学	2, 4, 8
数学与应用数学	3, 7, 10

(c) 专业索引表

2000级	6, 7, 8
2001级	9, 10
98级	1, 2, 3
99级	4, 5

(d) 年级索引表

图 1.1 学生信息查询系统中的数据

【例 1.2】人机对弈问题。

人机对弈是一个古老的人工智能问题，其解题思路是将对弈的策略事先存入计算机，包括对弈过程中所有可能出现的情况和响应的对策。在决定对策时，根据当前局势发展的趋势作出最有利的选择。因此计算机操作的对象（数据元素）是对弈过程中的每一步棋盘状态（格局），元素之间的关系由比赛规则决定。通常这个关系不是线性的，从一个格局可能派生出多个格局，因此常用树形结构来表示，如图 1.2 所示。

【例 1.3】教学计划编排问题。

一个教学计划包含许多课程，在教学计划包含的许多课程之间，有些必须按规定的先后次序进行，有些则没有次序要求。即有些课程之间有先修和后续的关系，有些课程可以任意安排次序。这种各个课程之间的次序关系可用一个称为图的数据结构来表示，如图 1.3

所示。有向图中的每个顶点表示一门课程，如果从顶点 v_i 到 v_j 之间存在有向边 $\langle v_i, v_j \rangle$ ，则表示课程 i 必须先于课程 j 进行。

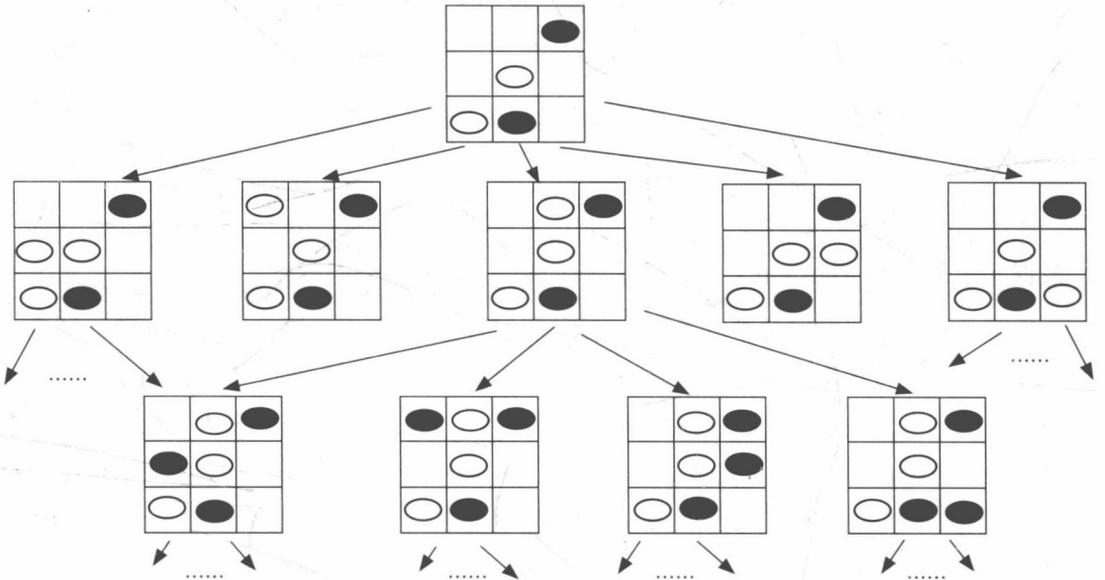
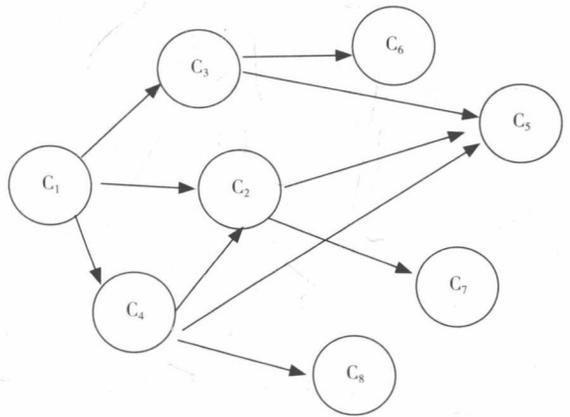


图 1.2 人机对弈图

课程编号	课程名称	先修课程
C_1	计算机导论	无
C_2	数据结构	C_1, C_4
C_3	汇编语言	C_1
C_4	C 程序设计语言	C_1
C_5	计算机图形学	C_2, C_3, C_4
C_6	接口技术	C_3
C_7	数据库原理	C_2
C_8	编译原理	C_4

(a) 计算机专业的课程设置



(b) 表示课程之间优先关系的有向图

图 1.3 教学计划编排问题的数据结构

由以上 3 个例子可见，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。因此，可以说数据结构课程主要是研究非数值计算的程序设计问题中所出现的计算机操作对象以及它们之间的关系和操作。

学习数据结构的目的是为了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。与此同时，通过算法训练来提高学生的思维能力，通过程序设计的技能训练来促进学生的综合应用能力和专业素质的提高。

1.2 基本概念和术语

在系统地学习数据结构知识之前，先对一些基本概念和术语赋予确切的含义。

1. 数据 (Data)

数据是人们利用文字符号、数字符号以及其他规定的符号对现实世界的事物及其活动所做的描述。在计算机科学中，数据的含义非常广泛，将一切能够输入计算机，并被计算机程序处理的信息，包括文字、表格、图像等，都称为数据。例如，一个学生信息的表格。

2. 数据元素 (Data Element)

数据元素是组成数据的基本单位。在不同的条件下，数据元素又可称为元素、结点、顶点、记录等。在程序中通常把结点作为一个整体进行考虑和处理。例如，在如图 1.4 所示的学生信息中，为了便于处理，把其中的每一行（代表一个学生）作为一个基本单位来考虑，故该数据由 30 个结点构成。一般情况下，一个结点中含有若干个字段（也称为数据项）。例如，在图 1.4 中，每个结点都由学号、姓名、分组、年龄和住址 5 个字段构成。字段是构成数据的最小单位。

3. 数据对象 (Data Object) 或数据元素类 (Data Element Class)

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。例如：整数数据对象是集合 $N=\{0, \pm 1, \pm 2, \dots\}$ ，字母字符数据对象是集合 $C=\{'A', 'B', \dots, 'Z'\}$ ，如图 1.4 所示的表也可看成一个数据对象。由此可以看出，不论数据元素集合是无限集（如整数集）、有限集（如字符集），还是由多个数据项组成的复合数据元素（如学籍表），只要性质相同，都是同一个数据对象。

学号 (ID)	姓名 (Name)	分组 (Group)	年龄 (Age)	住址 (Addr)
120010101	李华	100	16	四川成都
120010102	王丽	010	15	重庆万州
120010103	张阳	011	19	陕西西安
120010104	赵斌	012	16	重庆云阳
120010105	孙琪	020	18	四川广安
120010106	马丹	021	19	陕西宝鸡
120010107	刘畅	030	20	重庆黔江
⋮	⋮	⋮	⋮	⋮
120010130	黄凯	032	17	江苏南京

图 1.4 学生信息表

4. 数据结构 (Data Structure)

数据结构是研究数据元素 (Data Element) 之间抽象化的相互关系和这种关系在计算

机中的存储表示（即所谓数据的逻辑结构和物理结构），并对这种结构定义相适应的运算，设计出相应的算法，而且确保经过这些运算后所得到的新结构仍然是原来的结构类型。根据数据元素间关系的不同特性，通常有下列4类基本的结构：

- 集合结构：在集合结构中，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。
- 线性结构：该结构的数据元素之间存在着一对一的关系。
- 树形结构：该结构的数据元素之间存在着一对多的关系。
- 图形结构：该结构的数据元素之间存在着多对多的关系，图形结构也称为网状结构。

图 1.5 为表示上述 4 类基本结构的示意图。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。

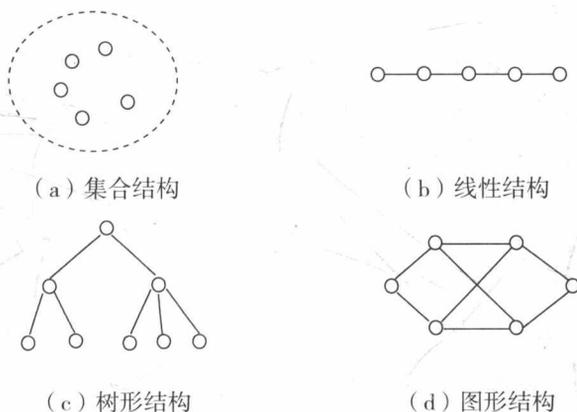


图 1.5 4 类基本结构的示意图

从上述的数据结构的概念中可知，一个数据结构有两个要素：一个是数据元素的集合；另一个是关系的集合。在形式上，数据结构通常可以采用一个二元组来表示。

数据结构的定义形式为：数据结构是一个二元组。

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

其中：D 是数据元素的有限集，R 是 D 上关系的有限集。

5. 数据的逻辑结构

数据的逻辑结构（Logical Structure）是指数据结构中元素之间的逻辑关系，它是从具体问题中抽象出来的数学模型，是独立于计算机存储器的（与具体的计算机无关）。数据的逻辑结构可分为 4 种基本类型：集合结构、线性结构、树形结构和图形结构，表和树是最常用的两种高效数据结构，许多高效的算法可以用这两种数据结构来设计实现。表是线性结构（一对一关系），树（一对多）和图（多对多）是非线性结构。

6. 数据存储结构

数据存储结构（Storage Structure）是数据的逻辑结构在计算机内存中的存储方式，又

称为物理结构。数据存储结构要用计算机编程语言来实现,因而依赖于具体的计算机语言,其可分为顺序存储或链式存储两种方式。

- 顺序存储:是把逻辑上相邻的元素存储在物理位置相邻的存储单元中,由此得到的存储方式称为顺序存储结构。顺序存储结构是一种最基本的存储表示方法,通常借助于程序设计语言中的数组来实现。

- 链式存储:是对逻辑上相邻的元素不要求其物理位置相邻,元素间的逻辑关系通过附设的指针字段来表示,由此得到的存储方式称为链式存储结构。链式存储结构通常借助于程序设计语言中的指针类型来实现。

除了通常采用的顺序存储方式和链式存储方式外,有时为了查找的方便还采用索引存储方式和散列存储方式。

索引存储:在数据文件的基础上增加一个索引文件,通过索引表建立索引,可以将一个顺序表分成几个顺序子表,其目的是在查询时提高查找效率,避免盲目查找。

散列存储:是通过数据元素与存储地址之间建立某种映射关系,使每个数据元素与每个存储地址之间尽量达到一对一的关系。

7. 数据处理

数据处理是指对数据进行查找、插入、删除、合并、排序、统计以及简单计算等的操作过程。在早期,计算机主要用于科学和工程计算,进入20世纪80年代以后,计算机主要用于数据处理。据有关统计资料表明,现在计算机用于数据处理的时间比例达到80%以上,随着时间的推移和计算机应用的进一步普及,计算机用于数据处理的时间比例必将进一步增大。

8. 数据类型

数据类型(Data Type)是和数据结构密切相关的一个概念,在用高级程序设计语言编写的程序中,每个变量、常量或表达式都对应一个确定的数据类型。数据类型可分为两类:一类是非结构的原子类型,如基本数据类型(整型、实型、字符型等);另一类是结构类型,它可以由多个结构类型组成,并可以分解。结构类型的分解量可以是结构的,也可以是非结构的。例如数组的值由若干分量组成,每个分量可以是整型,也可以是数组等结构类型。

1.3 算法的概念及其特性

1.3.1 算法的定义

算法(algorithm)是指在解决问题时,按照某种机械的步骤一定可以得到问题的结果(有解时给出问题的解,无解时给出无解的结论)的处理过程。当面临某个问题时,需要找到用计算机解决这个问题的方法和步骤,算法就是解决这个问题的方法和步骤的描述。