



21

世纪计算机辅导系列丛书

数据结构

课程辅导与 习题解析

● 胡元义 邓亚玲 徐睿琳 编著

12

人民邮电出版社

POSTS & TELECOMMUNICATIONS PRESS

图书在版编目(CIP)数据

数据结构课程辅导与习题解析 / 胡元义编著. —北京: 人民邮电出版社, 2003.3

(21 世纪计算机辅导系列丛书)

ISBN 7-115-10924-9

I. 数... II. 胡... III. 数据结构—高等学校—教学参考资料 IV. TP311.12

中国版本图书馆 CIP 数据核字(2003)第 013161 号

内容提要

数据结构课程具有概念性强、灵活和不易掌握等特点, 学习起来难度较大。本书配合教学内容, 从学生“学”的角度提供了全面的辅导。全书共分 10 章, 基本涵盖了数据结构的全部内容, 每章包括“重点内容讲解”、“典型例题解析”、“习题与答案”三大部分, 带领读者经历从“学习理论”到“结合实际理解理论”再到“自己亲自动手解决问题”的学习过程, 意在帮助读者深刻理解本课程涉及的原理和概念, 掌握数据结构知识, 从而透彻地领悟数据结构的精髓。本书的算法全部用 C 语言描述。

书中的例题与习题大多选自研究生和本科生的考试习题, 也包括作者结合多年教学经验设计出来的典型范例, 内容丰富且全面, 具有一定的知识水平和代表性。

本书可作为计算机专业学生学习辅导书, 也可作为研究生入学考试的复习参考书, 还可供计算机专业开发人员参考阅读。

21 世纪计算机辅导系列丛书

数据结构课程辅导与习题解析

- ◆ 编 著 胡元义 邓亚玲 徐睿琳
责任编辑 王文娟
- ◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街 14 号
邮编 100061 电子函件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
读者热线 010-67132692
北京汉魂图文设计有限公司制作
北京隆昌伟业印刷有限公司印刷
新华书店总店北京发行所经销
- ◆ 开本: 787×1092 1/16
印张: 22
字数: 532 千字
印数: 1-5 000 册
- 2003 年 3 月第 1 版
2003 年 3 月北京第 1 次印刷

ISBN7-115-10924-9/TP · 3243

定价: 30.00 元

本书如有印装质量问题, 请与本社联系 电话: (010) 67129223

前 言

如果把程序设计比作一棵大树，数据结构无疑就是大树的躯干，可以说程序设计的精髓就是数据结构。计算机各个领域都要用到各种数据结构，只有较好地掌握了数据结构知识，才能在程序设计的编程实现中游刃有余，从而在计算机应用领域的研制开发中做到胸有成竹。

数据结构课程是计算机专业的一门核心课程，它是在长期的程序设计实践中提炼、升华而成的，反过来又运用于程序设计；它同时又是操作系统、编译原理等计算机核心课程的基础，在计算机专业课程的学习中起着承上启下的作用。数据结构是一门应用广泛并且最具有实用价值的课程。不掌握数据结构知识，就难以成为一个合格的软件工程师或计算机工作者。

由于数据结构的原理和算法都比较抽象，且数据结构又具有概念性强、内容灵活、不易掌握和难度大等特点，因此学习起来就更加困难。本书通过课程辅导与习题解析的方式，引导学生由基本概念出发寻找数据结构问题普遍的求解思路和方法，并由此深化对概念的理解，达到举一反三、提高分析问题与解决问题的能力目的。

本书共分 10 章，算法均用 C 语言描述。第 1 章简要介绍了数据结构的基本概念与算法基础；第 2 章的线性表重点介绍了两种存储结构——顺序表和单链表的有关概念及题解；第 3 章栈和队列，对栈和队列的概念、特点及应用做了介绍；第 4 章涉及到串的概念，特别是串的模式匹配的有关内容；第 5 章介绍了数组和广义表的有关概念和题解；第 6 章树和二叉树则重点介绍了二叉树的有关概念和典型题解；第 7 章重点介绍了图的有关知识和应用；第 8 章给出了各种静态和动态查找表的方法；第 9 章详细讨论了各种排序的方法和特点；第 10 章综述了基本文件的组织结构。

本书所选习题概念性强、覆盖面广，所选习题均具有典型性和代表性，内容全面。为了帮助读者汲取数据结构的精髓，提高解题技能，大部分章节的典型例题划分为三类：第一类为概念题，题目都是涉及本章内容的典型概念，这对于加深理解并掌握本章知识是必不可少的；第二类为基本题，这些题目涵盖了本章的基本内容和重点内容；第三类为算法设计题，这些题目对提高学生动手能力和提高程序设计素质密切相关。

为了便于读者正确理解概念，掌握解题技巧，各章的典型例题大多给出了详尽的解题过程，对有代表性的习题和疑难习题，也给出了详细的分析和说明。此外，针对某些难题，书中还给出了一些新的解题思路及方法。

最后，每章还给出了供读者演练解题能力的习题，并附有习题答案。

希望读者通过对本书的学习，能够更全面、更透彻地理解和掌握数据结构这门课程。

由于编者水平有限，书中难免存在差错，敬请广大读者批评指正。此外，吴孟斌同志也参与了本书的编写工作。本书责任编辑的电子邮件地址为 wangwenjuan@ptpress.com.cn，衷心希望与各位读者进行交流。

编者
2002 年 11 月

目 录

第 1 章 基本概念	1
1.1 重点内容讲解	1
1.1.1 数据结构概念	1
1.1.2 逻辑结构的 4 种基本形态	1
1.1.3 数据存储结构的基本组织方式	2
1.1.4 算法的定义和描述	2
1.1.5 算法分析和复杂度计算	3
1.2 典型例题解析	4
1.2.1 概念题	4
1.2.2 基本题	8
1.3 习题与答案	12
1.3.1 习题	12
1.3.2 习题答案	15
第 2 章 线性表	17
2.1 重点内容讲解	17
2.1.1 线性表的存储结构	17
2.1.2 线性表的插入和删除运算	19
2.2 典型例题分析	20
2.2.1 概念题	20
2.2.2 基本题	25
2.2.3 算法设计	28
2.3 习题与答案	34
2.3.1 习题	34
2.3.2 习题答案	36
第 3 章 栈与队列	43
3.1 重点内容讲解	43
3.1.1 栈	43
3.1.2 队列	44
3.1.3 栈、递归与非递归	45
3.2 典型例题分析	46
3.2.1 概念题	46
3.2.2 基本题	51
3.2.3 算法设计	54

3.2.4	递归转化为非递归示例.....	58
3.3	习题与答案.....	66
3.3.1	习题.....	66
3.3.2	习题答案.....	68
第4章	串.....	75
4.1	重点内容讲解.....	75
4.1.1	串的概念与存储结构.....	75
4.1.2	简单模式匹配.....	76
4.1.3	无回溯的 KMP 匹配.....	77
4.2	典型例题解析.....	78
4.2.1	概念题.....	78
4.2.2	基本题.....	81
4.2.3	算法设计.....	86
4.3	习题与答案.....	89
4.3.1	习题.....	89
4.3.2	习题答案.....	91
第5章	数组与广义表.....	95
5.1	重点内容讲解.....	95
5.1.1	数组的概念与存储.....	95
5.1.2	矩阵的压缩存储.....	96
5.1.3	广义表的概念与存储.....	99
5.2	典型例题分析.....	100
5.2.1	概念题.....	100
5.2.2	基本题.....	105
5.2.3	算法设计.....	109
5.3	习题与答案.....	114
5.3.1	习题.....	114
5.3.2	习题答案.....	117
第6章	树和二叉树.....	121
6.1	重点内容讲解.....	121
6.1.1	树的定义与概念.....	121
6.1.2	二叉树的概念、性质及存储.....	122
6.1.3	二叉树的遍历.....	124
6.1.4	线索二叉树.....	125
6.1.5	最优二叉树(哈夫曼树).....	127
6.1.6	树和森林.....	129

6.2	典型例题解析.....	131
6.2.1	概念题.....	131
6.2.2	基本题.....	141
6.2.3	算法设计.....	155
6.3	习题与答案.....	164
6.3.1	习题.....	164
6.3.2	习题答案.....	170
第7章	图.....	179
7.1	重点内容讲解.....	179
7.1.1	图的基本概念.....	179
7.1.2	图的存储结构.....	180
7.1.3	图的遍历.....	182
7.1.4	最小生成树.....	183
7.1.5	拓扑排序和关键路径.....	184
7.1.6	最短路径.....	186
7.2	典型例题解析.....	188
7.2.1	概念题.....	188
7.2.2	基本题.....	196
7.2.3	算法设计.....	212
7.3	习题与答案.....	218
7.3.1	习题.....	218
7.3.2	习题答案.....	223
第8章	查找.....	229
8.1	重点内容讲解.....	229
8.1.1	静态查找表.....	229
8.1.2	动态查找表.....	231
8.2	典型例题解析.....	240
8.2.1	概念题.....	240
8.2.2	基本题.....	248
8.2.3	算法设计.....	260
8.3	习题与答案.....	264
8.3.1	习题.....	264
8.3.2	习题答案.....	268
第9章	排序.....	275
9.1	重点内容讲解.....	275
9.1.1	插入排序.....	275

9.1.2	交换排序	277
9.1.3	选择排序	278
9.1.4	归并排序	280
9.1.5	基数排序	280
9.1.6	各种内部排序方法的比较	281
9.1.7	外部排序	282
9.2	典型例题解析	284
9.2.1	概念题	284
9.2.2	基本题	293
9.2.3	算法设计	303
9.3	习题与答案	312
9.3.1	习题	312
9.3.2	习题答案	316
第 10 章	文件	323
10.1	重点内容讲解	323
10.1.1	文件的基本概念	323
10.1.2	顺序文件	323
10.1.3	索引文件	324
10.1.4	散列文件	325
10.1.5	多关键字文件	325
10.2	典型例题分析	326
10.2.1	概念题	326
10.2.2	基本题	330
10.2.3	算法设计	334
10.3	习题与答案	336
10.3.1	习题	336
10.3.2	习题答案	338
参考文献		341

第 1 章

基本概念

1.1 重点内容讲解

1.1.1 数据结构概念

1. 数据与数据元素

数据是计算机程序所加工处理的描述客观事物的符号表示。

数据元素是数据的基本单位，是数据集合中的一个个体。有些情况下也把数据元素称为结点、记录、表目等。一个数据元素可由一个或多个数据项组成，数据项是有独立含义的数据最小单位，有时把它称作域、字段等。

2. 数据结构

数据结构是存在一种或多种特定关系的数据元素的集合，即数据的组织形式。数据结构包括三个方面的内容：数据元素的逻辑结构、存储结构和相适应的运算（操作）。

数据的逻辑结构是从逻辑关系上描述数据来抽象地反映数据之间的逻辑关系，它与数据的存储无关，是独立于计算机的。因此，数据的逻辑结构可以看作是从具体问题中抽象出来的数学模型。

数据的存储结构是指数据的逻辑结构在计算机存储器中的映像表示，即在能够反映数据逻辑关系的原则下，数据在存储器中的存储方式。

数据的运算是在数据上所施加的一系列操作，称为抽象运算，它只考虑这些操作的功能，而暂不考虑如何完成；只有在确定了存储结构后，才会具体实现这些操作。也即，抽象运算是定义在逻辑结构上而实现在存储结构上。

1.1.2 逻辑结构的 4 种基本形态

现实世界中非数值数据的逻辑关系可以划分为以下 4 种基本结构，如图 1-1 所示。

- (1) 集合结构：数据元素之间除了“属于同一集合”的联系之外，没有其他关系。
- (2) 线性结构：数据元素之间存在一对一的关系。
- (3) 树形结构：数据元素之间存在一个对多个的关系。

(4) 图结构 (或称网状结构): 数据元素之间存在多对多的关系。

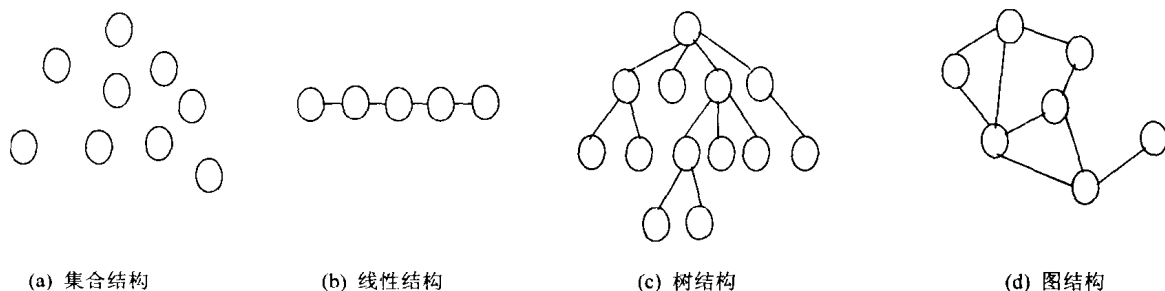


图 1-1

由于集合结构的简单性和松散性, 所以通常只讨论其他三种逻辑结构。数据的逻辑结构可分为线性结构和非线性结构两类, 若数据元素之间的逻辑关系可以用一个线性序列简单地表示出来, 则称为线性结构, 否则称为非线性结构。树形结构和图结构就属于非线性结构。

1.1.3 数据存储结构的基本组织方式

数据元素之间的关系在计算机中有以下 4 种基本存储结构表示方式。

(1) 顺序存储结构: 借助数据元素在存储器中的相对位置来表示数据元素之间的逻辑关系; 通常顺序存储结构是利用程序语言中的数组来描述。

(2) 链式存储结构: 在数据元素上附加指针域, 并借助指针来指示数据元素之间的逻辑关系; 通常利用程序语言的指针类型来描述。

(3) 索引存储结构: 在存储所有数据元素信息的同时建立附加索引表。索引表表项的一般形式是: 关键字、地址; 关键字是数据元素的某个数据项的值, 通过关键字可以找到相关的数据元素的存储地址。

(4) 散列存储结构 (也称为哈希存储结构): 此方法的基本思想是根据数据元素的关键字直接计算出相应的存储地址。

实质上索引存储结构和散列存储结构可看作前两种结构的衍生。上述 4 种基本存储结构既可单独使用也可组合使用, 可根据具体要求而定。

1.1.4 算法的定义和描述

算法是建立在数据结构基础上对特定问题求解步骤的一种描述, 是若干条指令组成的有限序列。算法必须满足以下性质:

- (1) 确定性: 算法中的每一条指令必须有确切的含义, 无二义性。
- (2) 有穷性: 算法指令是一有限序列, 且算法必须在一定时间内完成。
- (3) 可行性: 算法所描述的操作都可以通过可用的基本运算实现。
- (4) 输入: 一个算法可以有零个或多个输入。
- (5) 输出: 一个算法有一个或多个输出, 它们是与输入有特定关系的量。

一个算法可采用自然语言如英语、汉语描述, 也可以采用图形方式如流程图、拓扑图描述; 如果算法用各种计算机语言描述, 那么就表现为一个程序。

对某个特定问题的求解，究竟采用何种数据结构及选择什么算法，需要看问题的具体要求 and 现实环境的各种条件，然后把数据结构与算法有机地结合起来才能设计出高质量的程序来，算法必须采用与之相适应的数据结构才能有效地求解问题。

注意，算法与程序的主要区别在于一个程序通常是用某种程序语言书写的一个计算过程，而一个算法则并不一定表现为一个计算机程序，它可以用不同方式和不同语言来描述。

1.1.5 算法分析和复杂度计算

算法设计主要考虑可解算法的设计，而算法分析则是研究和比较各种算法的性能与优劣。算法的时间复杂度（或称时间复杂性）和空间复杂度（或称空间复杂性）是算法分析的两个主要方面，其目的主要是考察算法的时间和空间效率，以求改进算法或对不同的算法进行比较。

在算法分析和设计中，沿用实用性的复杂度概念，把求解问题的关键操作，如加法、减法和比较的运算指定为基本操作，然后把算法执行基本操作的次数定义为算法的时间复杂度，算法执行期间占用的存储单元则定义为算法的空间复杂度。

某个指定语句的频度是指它在算法中被重复执行的次数。算法中所有语句的频度之和记做 $T(n)$ ，它是该算法所求解问题规模 n 的函数。当问题的规模趋向无穷大时， $T(n)$ 的数量级（阶）称为时间复杂度，记作 $T(n)=O(f(n))$ 。

上述表达式中“O”的文字含义是 $T(n)$ 的量级，其严格的数学定义是：若 $T(n)$ 和 $f(n)$ 是定义在正整数集合上的两个函数，则存在正的常数 C 和 n_0 ，使得当 $n \geq n_0$ 时，都满足：

$$0 \leq T(n) \leq C \cdot f(n)。$$

程序分析法则为：

- （1）执行一条读写或赋值语句用 $O(1)$ 时间；
- （2）依次执行一系列语句所用时间采用求和准则；
- （3）判断语句的耗时主要是执行语句所用的时间，检验条件还需 $O(1)$ ；
- （4）循环语句的运行时间为多次叠代中执行循环体以及检验循环条件的耗时，常用乘法准则。

若算法的两个部分的时间复杂度为 $T_1(n)=O(f(n))$ 和 $T_2(n)=O(g(n))$ ，则大“O”下的求和准则为：

$$T_1(n)+T_2(n)=O(\max f(n), g(n))$$

大“O”下的乘法准则为：

$$T_1(n) \times T_2(n)=O(f(n) \times g(n))$$



1.2 典型例题解析

1.2.1 概念题

单项选择题

1. 研究数据结构就是研究_____。(全国 2001 年自考题)
 - a. 数据的逻辑结构
 - b. 数据的存储结构
 - c. 数据的逻辑结构和存储结构
 - d. 数据的逻辑结构、存储结构及其数据在运算上的实现
2. 下面关于算法的说法, 错误的是_____。(南京理工大 2000 年研究生试题)
 - a. 算法最终必须由计算机程序实现
 - b. 为解决某问题的算法与为该问题编写的程序含义是相同的
 - c. 算法的可行性是指指令不能有二义性
 - d. 以上几个都是错误的
3. 计算机中的算法指的是解决某一个问题的有限运算序列, 它必须具备输入、输出、_____等 5 个特性。(武汉大学 2000 年研究生试题)
 - a. 可执行性、可移植性和可扩充性
 - b. 可执行性、有穷性和确定性
 - c. 确定性、有穷性和稳定性
 - d. 易读性、稳定性和确定性
4. 以下属于逻辑结构的是_____。(西安电子科大 2001 年研究生试题)
 - a. 顺序表
 - b. 哈希表
 - c. 有序表
 - d. 单链表
5. 具有线性结构的数据结构是_____。(电子科大 1996 年研究生试题)
 - a. 图
 - b. 树
 - c. 广义表
 - d. 栈
6. 数据的_____包括集合、栈、树和图结构 4 种基本类型。
 - a. 存储结构
 - b. 逻辑结构
 - c. 基本运算
 - d. 算法描述
7. 数据的存储结构包括顺序、链接、散列和_____4 种基本类型。
 - a. 向量
 - b. 数组
 - c. 集合
 - d. 索引
8. 通常从正确性、易读性、健壮性、高效性等 4 个方面评价算法(包括程序)的质量。以下解释错误的是_____。
 - a. 正确性算法应能正确地实现预定的功能(即处理要求)
 - b. 易读性算法应易于阅读和理解, 以便于调试、修改和扩充
 - c. 健壮性当环境发生变化时, 算法能适当地做出反应或进行处理, 不会产生不需要的运行结果
 - d. 高效性即达到所需要的时间性能
9. 关于逻辑结构, 以下说法错误的是_____。



- a. 逻辑结构与数据元素本身的形式、内容无关
- b. 逻辑结构与数据元素的相对位置有关
- c. 逻辑结构与所含结点个数无关
- d. 一些表面上很不相同的数据可以有相同的逻辑结构

10. 根据数据元素之间关系的不同特性, 以下4类基本逻辑结构反映了4类基本数据组织形式。下列解释错误的是_____。

- a. 集合中任何两个结点之间都有逻辑关系, 但组织形式松散
 - b. 线性结构中结点按逻辑关系依次排列成一条“锁链”
 - c. 树型结构具有分支、层次特性, 其形态有点像自然界中的树
 - d. 图状结构中各个结点按逻辑关系互相缠绕, 任何两个结点都可以邻接
11. 下面程序的时间复杂度为_____。

```
for i=0; i<m; i++)
    for (j=0; j<n; j++)
        A[i][j]=i*j;
```

- a. $O(m^2)$
- b. $O(n^2)$
- c. $O(m \times n)$
- d. $O(m+n)$

【解答】

1. 数据结构包含逻辑结构、存储结构及数据的抽象运算这三部分内容, 故选 d。
2. 算法可以有各种描述方法以满足不同的要求, 即可以用伪码语言或框图等其他形式描述; 有时也用伪码语言描述一些只含抽象操作的抽象算法。此外, 算法可行性是指每一条指令都应在有限时间内完成。因此, a、b、c 项都是错误的, 故选 d。
3. 选 b。
4. 顺序表、单链表均为线性表的存储结构, 而哈希表是查找表的一种, 只有有序表属于逻辑结构, 故选 c。
5. 栈是运算受限的线性表, 其数据结构为线性结构, 故选 d。
6. 选 b。
7. 选 d。
8. 高效性即达到所需要的时空性能; 故 d 错, 即选 d。
9. 数据的逻辑结构是对数据之间关系的描述, 它与数据元素之间的相对位置无关; 故 b 错, 即选 b。
10. 集合结构的数据元素之间除了“属于同一集合”的联系之外没有其他关系; 故 a 错, 即选 a。
11. 第一个 for 循环执行 m 次, 第二个 for 循环执行 n 次, 两个 for 循环嵌套起来共执行 $m \times n$ 次, 故选 c。

多项选择题

1. 数据元素是_____。

(陕西省 1998 年自考题)

- a. 数据集合中的一个个体
- b. 数据的基本单位
- c. 数据的最小单位
- d. 一个结点
- e. 一个记录

2. 数据结构被形式地定义为 (K, R) ，其中 K 是 ① 的有限集， R 是 K 上的 ② 有限集。

- a. 算法 b. 数据元素 c. 数据操作 d. 逻辑结构
e. 操作 f. 映像 g. 存储 h. 关系

3. 线性结构的顺序存储结构是一种 ① 的存储结构，线性表的链式存储结构是一种 ② 的存储结构。

- a. 随机存取 b. 顺序存取 c. 索引存取 d. 散列存取

4. 算法分析的目的是 ①，算法分析的两个主要方面是 ②。

- ① a. 找出数据结构的合理性 b. 研究算法中的输入和输出关系
c. 分析算法的效率以求改进 d. 分析算法的易懂性和文档性
② e. 空间复杂性和时间复杂性 f. 正确性和简单性
g. 可读性和文档性 h. 数据复杂性和程序复杂性

5. 计算机算法指的是 ①，它必须具备输入、输出和 ② 等 5 个特性。

- ① a. 计算方法 b. 排序方法
c. 解决问题的有限运算序列 d. 调度方法
② e. 可执行性、可移植性和可扩充性 f. 可行性、确定性和有穷性
g. 确定性、有穷性和稳定性 h. 易读性、稳定性和安全性

【解答】

1. 选 b、d、e。

2. 选 b、h

3. 顺序存储结构是一种随机存取结构，链式存储结构是一种顺序（一个结点一个结点的查找）存取结构，故选：①a；②b。

4. 算法分析的目的是考察算法的时间和空间效率，以求改进算法或对不同的算法进行比较。因此选：①c；②e。

5. 选：①c；②f。

例题 1.3

填空题

1. 一个数据结构在计算机中的_____称为存储结构。（华中理工大 2000 年研究生试题）

2. 对于给定的 n 个元素，可以构造出的逻辑结构有_____、_____、_____和_____4 种。
（中科院计算所 1999 年研究生试题）

3. 数据是描述客观事物的数、字符以及所有_____计算机中并被计算机程序所_____符号的集合。
（陕西省 1997 年自考题）

4. 从一维数组 $a[n]$ 中顺序查找出一个最大值元素的平均时间复杂性为_____，读取一个二维数组 $b[m,n]$ 中任一元素的时间复杂性为_____。

5. 线性结构中元素之间存在_____关系，树形结构中元素之间存在_____关系，图形结构中元素存在_____关系，而集合结构中元素之间不存在_____关系。

6. 数据结构是研究数据的_____和_____以及它们之间的相互关系，并对这种结构定义相应的_____并设计出相应的_____。

7. 由一种____结构和一组____构成的整体是实际问题的一种数学模型, 这种数学模型的建立、选择和实现是数据结构的核心问题。

8. 通常从____、____、____、____等4个方面评价算法(包括程序)的质量。

9. 数据的____结构与数据元素本身的内容和形式无关。

10. 一个算法的时空性能是指该算法的____和____; 前者是算法包含的____, 后者是算法需要的____。

11. 程序段“ $i=1; \text{while}(i \leq n) \quad i=i*2;$ ”的时间复杂性为_____。

【解答】

1. 表示(或映像)

2. 集合 线性结构 树形结构 图结构

3. 能够输入到 处理的

4. 最大值元素可能的查找次数为 $1, 2, \dots, n$, 取平均值所 $\frac{n}{2}(n+1)$, 故平均时间复杂性为 $O(n)$; 由于数组元素可以随机读取, 故读取二维数组 $b[m, n]$ 中任一元素的时间复杂性为 $O(1)$, 即填 $O(n) \quad O(1)$ 。

5. 一对一 一对多 多对多 逻辑

6. 物理结构 逻辑结构 运算 算法

7. 逻辑 基本运算

8. 正确性 易读性 健壮性 高效性

9. 逻辑

10. 时间性能(或时间效率) 空间性能(或空间效率) 计算量 存储量

11. 如果循环体为 $i=i+1$ 或 $i=i+2$ 这类形式, 则应填 $O(n)$; 但此处为 $i=i*2$, 即循环次数 k 满足 $2^k=n$, 因此: $k=\log_2 n$ 。故填 $O(\log_2 n)$ 。

例题 1.4

判断题

1. 顺序存储方式只能用于存储线性结构。 (中科院软件所 1999 年研究生试题) ()

2. 数据元素是数据的最小单位。 (北邮 1998 年研究生试题) ()

3. 算法可以用不同的语言描述, 如果用 C 语言或 PASCAL 语言等高级语言来描述, 则算法实际上就是程序了。 (西交大 1996 年研究生试题) ()

4. 数据结构是带有结构的数据元素的集合。 ()

5. 数据的逻辑结构是指各数据元素之间的逻辑关系, 是用户根据需要而建立的。 ()

6. 数据结构、数据元素、数据项在计算机中的映像(或表示)分别称为存储结构、结点、数据域。 ()

7. 数据的物理结构是指数据在计算机内实际的存储形式。 ()

【解答】

1. 错。顺序存储方式也可用来存储树形结构, 如完全二叉树的数组存储和堆排序中堆的存储。

2. 错。数据元素是数据的基本单位, 数据元素可以由数据项组成, 且数据项是数据的最

小单位。

3. 错。算法用各种计算机语言描述则表现为一个程序但并不等于程序，因为程序逻辑不一定能满足有穷性。

4. 正确。若把数据结构进行形式化描述，则可从逻辑上认为数据结构 DS 是数据元素的集合 D 和 D 上关系的集合 R 所构成的二元组： $DS = (D, R)$ ，这里关系集合 R 被用作结构的定义手段。

5. 正确。

6. 正确。数据元素的映像——结点可能除了数据域之外还包含指针域来表示数据元素间的结构关系。

7. 正确。

1.2.2 基本题

例题 1.5

简述数据结构的 4 种基本关系并画出它们的关系图。

【解答】

数据结构的 4 种基本关系为：

- (1) 集合 集合中的数据元素之间除了“同属于一个集合”的关系外再无其他关系；
- (2) 线性结构 结构中的数据元素之间存在一个对一个的关系；
- (3) 树形结构 结构中的数据元素之间存在一个对多个的关系；
- (4) 图结构 结构中的数据元素之间存在多个对多个的关系。

上述 4 种基本结构的关系图如图 1-2 所示。

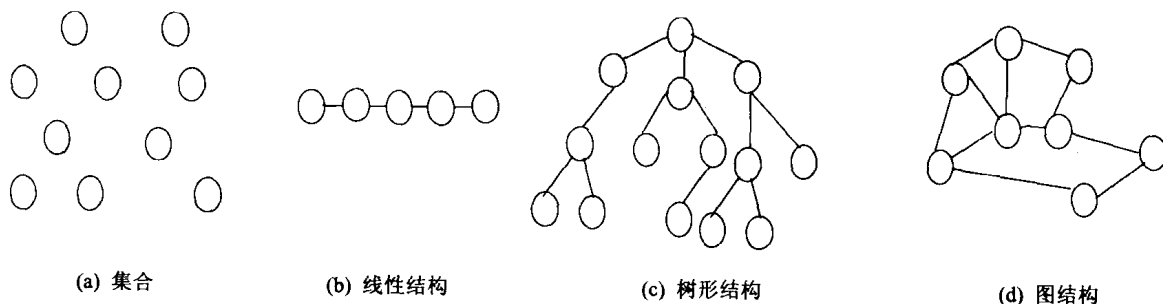


图 1-2 4 种基本逻辑结构示意图

当你为解决某一问题而选择数据结构时，应从哪些方面考虑？

【解答】

通常从两方面考虑：第一是算法所需的存储空间量；第二是算法所需的时间。对算法所需的时间又涉及以下几点：

- (1) 程序运行时所需输入的数据总量；
- (2) 对源程序进行编译所需的时间；



- (3) 计算机执行每条指令所需的时间;
- (4) 程序中指令重复执行的次数。

试叙述算法的定义及其重要特性。

【解答】

算法是特定问题求解步骤的一种描述,它是指令的有限序列,其中每一条指令表示一个或多个操作。算法的重要特性如下所述。

- (1) 有穷性: 一个算法必须(对任何合法的输入值)在执行有穷步之后结束,且每一步都可以在有穷时间内完成。
- (2) 确定性: 算法中每一条指令必须有确切的含义而不会产生二义性,且在任何条件下算法只有唯一的一条执行路径,即对相同的输入只能得到相同的输出。
- (3) 可行性: 指算法中描述的操作都可以通过已经实现的基本运算执行有限次来实现。
- (4) 输入: 一个算法有零个或多个输入,这些输入取自于特定的对象的集合。
- (5) 输出: 一个算法有一个或多个输出,这些输出是同输入有某种特定关系的量。

名词解释: (1) 抽象数据类型; (2) 算法的时间复杂度。

【解答】

(1) 抽象数据类型是指一个数学模型以及定义在该模型上的一组操作。抽象数据类型的定义仅取决于它的一组逻辑特性,而与其在计算机内部如何表示和实现无关,即无论其内部结构如何变化,只要它的数学特性不变就不影响其外部使用,抽象数据类型和数据类型实质上是一个概念。此外,抽象数据类型的范畴更广,它不再局限于各机器已定义并实现的数据类型,还包括用户在设计软件系统时自行定义的数据类型。

(2) 一般情况下,算法中基本操作重复执行的次数是问题规模 n 的某个函数 $f(n)$,算法的时间量度记作:

$$T(n)=O(f(n))$$

它表示随问题规模 n 的增大,算法执行时间的增长率和 $f(n)$ 的增长率相同,称作算法的渐进时间复杂度,简称时间复杂度。

斐波那契数列 F_n 定义如下:

$$F_0=0, \quad F_1=1, \quad F_n=F_{n-1}+F_{n-2} \quad n=2,3,\dots$$

请就此斐波那契数列,回答下列问题:

- (1) 在递归计算 F_n 的时候,需要对较小的 F_{n-1} 、 F_{n-2} 、 $\dots$ 、 F_1 、 F_0 精确计算多少次?
- (2) 如果用大 O 表示法,试给出递归计算 F_n 时递归函数的时间复杂度。

【解答】

- (1) 由斐波那契 F_n 的定义可得:



$$\begin{aligned}
 F_n &= F_{n-1} + F_{n-2} \\
 &= 2F_{n-2} + F_{n-3} \\
 &= 3F_{n-3} + 2F_{n-4} \\
 &= 5F_{n-4} + 3F_{n-5} \\
 &= 8F_{n-5} + 5F_{n-6} \\
 &\dots \\
 &= PF_1 + qF_0
 \end{aligned}$$

设 F_m 的执行次数为 B_m ($m=0, 1, 2, \dots, n-1$)，由上面等式的第一项可知， F_{n-1} 被执行一次，即 $B_{n-1}=1$ ； F_{n-2} 被执行两次，即 $B_{n-2}=2$ ； $\dots$ ；直至 F_1 被执行 P 次、 F_0 被执行 q 次，即 $B_1=P$ ， $B_0=q$ 。通过上面等式可以看出， B_m 的执行次数为前两等式第一因式系数之和，即 $B_m=B_{m-1}+B_{m-2}$ ，再结合 $B_{n-1}=1$ 和 $B_{n-2}=2$ ，这也是一个斐波那契数列。可以解得：

$$B_m = \frac{\sqrt{5}}{5} \left[\left(\frac{1+\sqrt{5}}{2} \right)^{n-m+2} - \left(\frac{1-\sqrt{5}}{2} \right)^{n-m+2} \right] \quad (m=0, 1, \dots, n-1)$$

(2) 递归计算 F_n 时递归函数的时间复杂度为 $O(n)$ 。

对下面程序段：

```

for (i=0; i<=n; i++)
    for (j=0; j<=i; j++)
        for (k=0; k<=j; k++)
            s=s+1;
    
```

试分析每一条语句执行的次数及时间复杂度。

【解答】

对第一层 for 语句，判断需执行 $n+1$ 次，而进入循环体（即第二层 for 语句）共 n 次；对于第二层 for 语句，由于循环终值 i 受第一层 for 语句 i 值每次改变的制约，其判断次数和进入次数如下：

$$\text{判断次数 } 2+3+\dots+n+n+1 = \frac{n}{2}(n+3) \quad \text{进入次数 } 1+2+\dots+n-1+n = \frac{n}{2}(n+1)$$

对于第三层 for 语句，其执行次数同时受第一层 for 语句的 i 值和第二层 for 语句的 j 值控制，其进入次数如图 1-3 所示。

$i=$	1	2	3	4		n
$j=1$	1	1	1	1		1
$j=2$		2	2	2		2
$j=3$			3	3		3
$j=4$				4		4
$\vdots$					$\ddots$	$\vdots$
$j=n-1$					$n-1$	$n-1$
$j=n$						n

图 1-3 第三层 for 语句进入循环体的次数

因此，第三层 for 语句进入循环体，也就是赋值语句 $S=S+1$ 的执行次数为：

