

全国高等教育自学考试
计算机信息管理专业指定教材辅导用书

数据组织与管理

主 编 赵东明

副主编 刘 刚 张 岳

主 审 李连友



电子科技大学出版社

全国高等教育自学考试计算机信息管理专业指定教材辅导用书

数据组织与管理

主 编 赵东明

副主编 刘 刚 张 岳

主 审 李连友

电子科技大学出版社

声 明

本书无四川省版权防盗标识,不得销售;版权所有,违者必究,举报有奖。举报电话:(028)6636481
6241146 3201496

全国高等教育自学考试计算机信息管理专业指定教材辅导用书

数据组织与管理

主 编 赵东明

副主编 刘 刚 张 岳

主 审 李连友

出 版: 电子科技大学出版社(成都建设北路二段四号 邮编: 610054)

责任编辑: 朱 丹

发 行: 新华书店

印 刷: 西南冶金地质印刷厂

开 本: 787×1092 1/16 印张 13.5 字数 329 千字

版 次: 1999 年 8 月第一版

印 次: 1999 年 8 月第一次印刷

书 号: ISBN 7—81065—190—0/TP•107

印 数: 1—4000 册

定 价: 16.00 元

内 容 提 要

本书遵照国家高等教育自学考试指导委员会审定的、国家教育委员会批准并颁发试行的《数据组织与管理自学考试大纲》；参照王珊主编的《数据组织与管理》教材；在多年进行本课程社会助学经验的基础上、结合参加每年高等教育自学考试评卷的实践，编写而成。

本书分为三篇，共十四章，每章有三个部分：知识点部分，主要针对本章的考核要求进行重点分析；测试题部分，根据考核要求，结合历年试卷在本章各类题型上的分布以及所涉及的具体内容，给出较有针对性的练习题；参考答案部分，给出相应的参考答案。

编写本书的目的在于，指导学生快速地掌握本课程考核要求的基本知识，能够顺利地通过本课程的考试。本书可作为广大自学者的考前复习指导书，也可作为社会助学者的参考书。

前 言

当今国际间的竞争是综合国力的竞争，关键又是科学技术的竞争，说到底人才的竞争。培养人才要坚持多种形式、多种途径。高等教育自学考试制度，正是为了适应我国社会主义现代化建设培养人才的需要，为了加速我国电子信息技术的广泛应用，为了造就和选拔人才应运而生的。它是个人自学、社会助学、国家考试相结合的一种新的教育方式，是我国高等教育体系的一个重要组成部分。实行高等教育自学考试制度，是宪法规定的“鼓励自学成才”的重要措施，也是造就和选拔人才的一条途径。应考者通过规定的考试课程并经思想品德鉴定达到毕业要求的，可以获得毕业证书，国家承认学历；按照规定享有与普通高等学校毕业生生同等的有关待遇。

为了满足计算机信息管理专业广大自学者的需要，使其有针对性地掌握本专业的有关知识，以便在有限的时间内、用有限的精力自学相应课程并顺利通过考试；同时，也有助于提高社会助学的教学质量，我们在国家高等教育自学考试指导委员会审定的、国家教育委员会批准并颁发试行的该专业自学考试大纲及指定教材的基础上，通过我们的教学实践及深入探讨，编写了这套全国高等教育自学考试计算机信息管理专业指定教材辅导用书。

《数据组织与管理》课程是高等教育自学考试计算机信息管理专业独立本科的一门专业性很强的基础课。第一篇《数据结构》部分，是学习计算机软件课程的综合性的基础课；第二篇《数据库系统基础》是数据管理技术的最新发展，是计算机科学的重要分支。

本书分为三篇，共十四章。第一篇有六章，主要介绍数据结构部分；第二篇有四章，主要介绍数据库系统基础部分；第三篇有四章，主要给出一个关系数据库系统的实例。每章有三个部分：知识点部分，主要针对本章的考核要求进行重点分析；测试题部分，主要是根据考核要求，结合历年试卷在本章各类题型上的分布情况以及所涉及的具体内容，给出较有针对性的练习题；参考答案部分，给出相应的参考答案。本书的目的，在于指导自学者快速地掌握本课程考核要求的基本知识，能够顺利地通过本课程的考试。

本书由郑州大学计算机系赵东明、张岳以及河南省高等教育招生办公室信息科技处刘刚等人编写，其中，赵东明编写了第一章到第九章；张岳编写了第十章；刘刚编写了第十一章到第十四章。

河南省电子信息应用教育中心对本套丛书的编写给予了极大的关怀和支持，并委托全国电子信息应用教育中心教学指导委员会专家组成员、河南省电子信息应用教育中心特邀顾问、郑州大学计算机系李连友教授主审，他对本书提出了许多宝贵意见和建议，谨此一并表示感谢。

本书是在近年来从事《数据组织与管理》自学考试社会助学讲义的基础上整理而成的，由于时间仓促，难免出现疏忽和谬误之处，敬请指正。

编者

1999年5月于郑州

目 录

第一篇 数据结构

第一章 数据组织概述	2
1.1 知 识 点	2
1.1.1 数据的逻辑结构	2
1.1.2 数据的物理结构	3
1.1.3 数据的运算	4
1.2 测 试 题	4
1.3 参 考 答 案	7
第二章 线 性 结 构	9
2.1 知 识 点	9
2.1.1 线性表	9
2.1.2 栈	12
2.1.3 队	14
2.2 测 试 题	16
2.3 参 考 答 案	18
第三章 树 形 结 构	29
3.1 知 识 点	29
3.1.1 树的定义和基本术语	29
3.1.2 树的存储结构	30
3.1.3 树的遍历	30
3.1.4 二叉树	31
3.1.5 二叉排序树	34
3.2 测 试 题	35
3.3 参 考 答 案	37
第四章 排 序	42
4.1 知 识 点	42
4.1.1 概述	42

4.1.2	内排序	42
4.1.3	插入排序	43
4.1.4	交换排序	45
4.1.5	选择排序	46
4.1.6	合并排序	49
4.1.7	外排序	49
4.2	测 试 题	50
4.3	参 考 答 案	53
第五章	找	57
5.1	知 识 点	57
5.1.1	概述	57
5.1.2	线性表的查找	57
5.1.3	哈希表的查找	60
5.2	测 试 题	63
5.3	参 考 答 案	65
第六章	文 件 组 织	72
6.1	知 识 点	72
6.1.1	顺序文件	72
6.1.2	索引文件	73
6.1.3	哈希文件	77
6.1.4	倒排文件	78
6.2	测 试 题	79
6.3	参 考 答 案	80

第二篇 数据库系统基础

第七章	数据库系统介绍	84
7.1	知 识 点	84
7.1.1	数据库系统概述	84
7.1.2	数据模型	86
7.2	测 试 题	88
7.3	参 考 答 案	90
第八章	关系数据库	93
8.1	知 识 点	93
8.1.1	关系模型	93
8.1.2	关系数据语言	94

8.1.3 数据库系统的三级模式与数据独立性	106
8.2 测 试 题	106
8.3 参 考 答 案	111
第九章 数据库保护	117
9.1 知 识 点	117
9.1.1 安全性	117
9.1.2 完整性	118
9.1.3 并发控制	120
9.1.4 恢复	121
9.2 测 试 题	123
9.3 参 考 答 案	126
第十章 数据库设计	130
10.1 知 识 点	130
10.1.1 数据库设计概述	130
10.1.2 需求分析	130
10.1.3 概念结构设计	131
10.1.4 逻辑结构设计	132
10.1.5 数据库的物理设计	133
10.1.6 数据库的实施和维护	134
10.2 测 试 题	134
10.3 参 考 答 案	136

第三篇 关系数据库管理系统实例

第十一章 INFORMIX 概述	141
第十二章 应用开发工具——INFORMIX-SQL	143
第十三章 数据库应用开发举例	148
第十四章 INFORMIX-Online 的基本功能	149
14.1 知 识 点	149
14.2 测 试 题	151
14.3 参 考 答 案	155
附录 A 《数据组织与管理》自学考试大纲	159
A.1 课程性质与设置的目的要求	159
A.2 课程内容与考核目标	160

A.3 有关说明和实施要求	170
附录 B 1997 年上半年全国高等教育自学考试《数据组织与管理》试卷	173
附录 C 1997 年下半年全国高等教育自学考试《数据组织与管理》试卷	184
附录 D 1999 年上半年全国高等教育自学考试《数据组织与管理》试卷	196

数据结构的教程 第一章

点 册 一 1.1

第一篇

数据结构教程 1.1

数据结构教程 1.1

数 据 结 构

第一章 数据组织概述

1.1 知 识 点

1.1.1 数据的逻辑结构

1. 基本概念和术语

所谓数据的逻辑结构，即数据之间的逻辑关系的描述。

数据是指所有能输入到计算机中并能被计算机程序处理的符号总称。简言之，数据是对客观事物的符号表示。

数据是有一定层次的。数据由数据项(或组项)组成。

数据项是数据中不可分割的具有独立含义的最小单位。数据项简称为“项”或“字段”。项有名和值之分，如工资表中的“姓名”是数据项名，“李民”是数据项值。

组合项由一个或多个数据项或组项组成。如工资表中的“扣除”组合项可以由“房租”、“水电费”、“煤气费”三个基本数据项组成。

记录是由若干相关的项或组项组成的数据单位，它常常是我们在研究某一问题时作为整体考虑和处理的基本单位。记录有型和值之分，记录型是组成该记录的项名的有序集合，记录值则是与项名对应的一组值的集合。

文件是具有相同性质的记录的集合。根据记录的长度固定与不固定，文件分为定长文件和变长文件。

关键码是指能唯一确定一个记录的数据项或数据项的组合。关键码又称为主码。

结点是在研究数据的逻辑结构时，为了方便讨论，作为整体进行考虑和处理的数据单位。通常把一个记录抽象为一个结点，但并非在任何情况下结点都代表一个记录。

2. 数据的逻辑结构

数据的逻辑结构就是结点的有限集合和结点间关系的有限集合。或可形式地定义为一个二元组： $B = (A, R)$ ，其中 A 是结点的非空的有限集合， R 是定义在 A 上的关系的有限集合。

理解数据结构中前驱、后继、起始结点和终止结点。给出数据结构的二元组形式，会画出它的图形表示，反之亦然。

如果 $B=(A, R)$ 是一数据结构。设 $r \in R$, A 中二结点 a, b 有 $\langle a, b \rangle \in r$, 则称结点 a 为结点 b 的前驱, 反之, 结点 b 为结点 a 的后继。

若 A 中不存在结点 d , 使 $\langle d, a \rangle \in r$, 则结点 a 是关系 r 中的一个起始结点。

若 A 中不存在结点 c , 使 $\langle b, c \rangle \in r$, 则结点 b 是关系 r 中的一个终止结点。

按结点间的逻辑关系不同, 数据结构可分为线性结构和非线性结构两大类。其中非线性结构又可分为树形结构和网状结构。

线性结构所表示的结点间的关系是一对一的, 只有一个起始结点和一个终止结点, 其他结点只有一个前驱和一个后继。

树形结构所表示的结点间的关系是一对多的, 只有一个起始结点但可以有多个终止结点, 其他结点只有一个前驱和可以有多个后继。

网状结构所表示的结点间的关系是多对多的, 可以有多个起始结点也可以有多个终止结点, 其他结点可以有多个前驱, 也可以有多个后继。

1.1.2 数据的物理结构

数据的物理结构是指如何把结点的值和结点间的关系存储在计算机中。数据的物理结构又称为数据的存储结构。

数据的逻辑结构是独立于计算机的, 而数据的存储结构是依赖于计算机的。

1. 数据结构在内存中的实现

数据结构在内存中的存储方式主要有顺序存储和链接存储。

(1) 顺序存储

顺序存储只能表示结点间唯一的线性关系。顺序存储只需存储结点的值, 无需存储结点间的关系, 结点在内存中的相对位置就表示了结点间的逻辑关系。

在顺序存储中, 若第一个结点的起始地址为 da_1 , 每个结点值需占 M 个存储单元, 则第 I 个结点的起始地址为 $da_1 + (i-1) * m$ 。

(2) 链接存储

链接存储是将结点所占的存储单元分成两部分, 一部分叫数据区, 用来存储结点的值; 另一部分叫指针区, 用来存储表示结点间关系的指针(指针通常是指后继结点的地址, 有时也包括前驱结点的地址或其他特殊结点的地址)。

2. 数据结构在外存中的实现

数据结构在外存中的存储都是以文件作为组织方式。

文件组织要解决的问题就是如何确定一个寻找数据记录地址的方法, 以便能对存储在大量数据方便、快速地进行存取。

(1) 外存储器的特性

磁带是顺序存取的存储设备; 磁盘、光盘是直接存取的存储设备。

存放在顺序存取的外存上的文件只能组织成顺序文件, 存放在直接存取的外存上的文件可组织成任何结构方式。

评价一个文件结构的效率主要看时间效率和空间效率。

(2)如何选择合适的文件结构

文件组织的选择主要取决于对文件执行何种操作和所用的外存种类。或者说,文件组织的选择取决于外存的种类、询问的类型和操作方式。

外存主要有顺序存取的存储器和直接存取的存储器;询问的类型主要有简单询问、范围询问、函数询问、布尔询问、单码询问和多码询问六类;操作方式主要有检索和更新(插入、删除、修改)两大类操作。

简单询问是指查询关键码(或某一数据项)等于给定值的记录。

范围询问是指查询关键码(或某一数据项)属于某个区域内的记录。

函数询问是指按给定关键码(或某一数据项)的某个函数来查询。

布尔询问是指把多个条件用布尔运算组合起来的查询。

单码询问是指对一个码值的查询,这个码可以是主码,也可以是其他数据项。

多码询问是指同时查询若干个“查找码”。由于多码查询要用布尔条件组合成一个条件,故多码询问也是布尔询问。

常用的文件存储结构有:顺序结构、计算寻址结构、索引结构和表结构。

1.1.3 数据的运算

数据的运算是定义在数据的逻辑结构之上,实现是在数据的物理结构之上的数据操作。

常用的运算有:查找、更新、插入、删除和排序(或分类)五种。

数据的运算操作常采用类 PASCAL 语言作为描述算法的工具。类 PASCAL 是标准 PASCAL 语言的一种扩充,但不能直接上机运行。

算法的效率分为时间效率和空间效率。

通常从算法中选取一种对于所研究的问题来说是基本运算的操作,以该操作的重复执行次数为算法的时间量度。

空间效率主要考虑存储数据结构本身以及实现算法所要求的辅助空间。

$T(n) = O(f(n))$ 是指随着算法的问题规模 n 的增大,算法执行时间 $T(n)$ 的增长率和 $f(n)$ 的增长率相同。

算法的时间效率评价主要用“大O”表示法。

数据结构的实质就是研究数据的逻辑结构、物理结构和定义在数据的逻辑结构上的运算操作。

1.2 测试题

一、填空题

- 1.所谓逻辑结构,是对数据之间_____的描述。
- 2.数据是对客观事物的_____表示。
- 3.数据是有层次的,常用的一种划分是_____、_____、_____、_____。

4. _____是数据的不可分割的最小单位。

5. _____是具有相同性质的记录的集合。

6.能唯一确定一个记录的数据项或数据项的组合称为_____。

7.在研究数据的逻辑结构时, 作为整体进行考虑和处理的数据单位称为_____。

8.一个数据结构 B 可形式地用一个二元组表示: $B=(A, R)$ 。

其中, A 是_____, R 是_____。

9.设 $B=(A, R)$ 是一个数据结构, $r \in R$, 若 A 中的两结点 a, b 有 $\langle a, b \rangle \in r$, 称结点 a 为结点 b 的_____, 结点 b 为结点 a 的_____。

10.若 A 中不存在结点 d, 使 $\langle d, a \rangle \in r$, 则结点 a 是关系 r 中的一个_____。若 A 中不存在结点 c, 使 $\langle b, c \rangle \in r$, 则结点 b 是关系 r 中的一个_____。

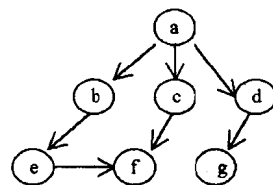


图 1-1

11.设数据结构的图形如图 1-1 所示, 若用二元组 $B=(A, R)$ 表示, 则:

$A = \{ \text{_____} \};$

$R = \{ \text{_____} \}.$

12.按结点间的逻辑关系不同, 数据结构可分为_____和_____两大类。其中, _____又可分为_____和_____。

13.线性结构中结点间关系是_____的; 树形结构中结点间关系是_____的; 网状结构中结点间的关系是_____的。

14.数据结构在计算机中的存储, 包括对_____的存储和对_____的存储。

15.数据的_____结构是独立于计算机的, 数据的_____结构是依赖于计算机的。

16.数据结构在内存的存储方式主要有_____和_____; 在外存的存储主要是以_____为基础。

17.当数据结构在内存中是顺序存储时, 结点在内存中的_____表示了结点间的逻辑关系。

18.当数据结构在内存是链接存储时, 结点所占用的存储单元分为两部分, 一部分存放_____, 另一部分存放_____。

19.磁带的存取方式是_____, 磁盘的存取方式是_____。存放在磁带上的文件只能组织成_____。

20.文件结构(或组织)的选择取决于_____和_____。

21.常用的文件结构有_____, _____, _____和_____。

22.数据结构中常用的运算操作有_____, _____, _____, _____和_____。

23.算法的效率分为_____效率和_____效率。

24.算法的空间效率主要考虑_____和_____。

25.数据结构的实质是包括_____, _____和_____三个方面的内容。

二、选择题

1. 在研究数据的逻辑结构时，常把作为整体进行考虑和处理的数据单位称为()。
A. 文件 B. 记录 C. 结点 D. 数据项
2. 设 $B=(A, R)$ 是一数据结构， $r \in R$ ，若 A 中不存在任何结点 d ，使得 $\langle d, a \rangle \in r$ ，则结点 a 是关系 r 中的()。
A. 起始结点 B. 终止结点 C. 前驱 D. 后继
3. 文件是具有相同性质的()的集合。
A. 数据项 B. 记录 C. 结点 D. 主码
4. 在树形结构中，结点间的关系是()。
A. 一对一的 B. 一对多的 C. 多对多的 D. 没关系
5. 按结点间的逻辑关系不同，数据结构可分为线性结构和非线性结构两大类。其中非线性结构又可分为()。
A. 树形结构和层次结构 B. 线性结构和网状结构
C. 线性结构和树形结构 D. 树形结构和网状结构
6. 一个数据结构定义为二元组 $B=(A, R)$ ，这里 A 和 R 分别为()。
A. 结点的非空有限集合和结点间关系的集合
B. 结点的集合和结点间关系的集合
C. 结点的集合和结点间关系的非空集合
D. 结点的非空有限集合和结点间关系的非空有限集合
7. 数据的存储结构是()于计算机的。
A. 独立 B. 依赖 C. 不同 D. 其他
8. 已知对数据结构采用了顺序存储，且第一个结点的起始地址为 1001，若每个结点的值需占 2 个存储单元，则第三个结点的起始地址为()。
A. 1003 B. 1005 C. 1006 D. 1007
9. 对文件按单个数据项进行检索，即查找某个数据项等于给定值的询问类型叫()。
A. 单码询问 B. 多码询问 C. 范围询问 D. 函数询问
10. 对文件的存储方式是选择一个函数，对记录的关键码进行转换，用所得的函数值作为存放该记录的地址。这种文件的存储结构称为()。
A. 顺序结构 B. 表结构 C. 计算寻址结构 D. 索引结构
11. 对文件查询关键码属于某个区域内的记录，这类询问称为()。
A. 范围询问 B. 函数询问 C. 布尔询问 D. 简单询问

三、简答题

1. 数据结构主要介绍哪三个方面的问题？
2. 数据结构主要分哪三类？其特点是什么？
3. 什么是“算法的时间复杂度”和“大 O”表示法？
4. 文件的存储结构采用的基本方法有哪些？

1.3 参 考 答 案

一、填空题

- 1.逻辑关系
- 2.符号
- 3.数据项 组项 记录 文件
- 4.数据项
- 5.文件
- 6.关键码(或主码)
- 7.结点
- 8.结点的非空有限集合 定义 A 上的关系的非空有限集合
- 9.前驱 后继
- 10.起始结点 制止结点
11. $A = \{a, b, c, d, e, f, g\}$
 $R = \{ \langle a, b \rangle, \langle a, c \rangle, \langle a, d \rangle, \langle b, e \rangle, \langle c, f \rangle, \langle d, g \rangle, \langle e, f \rangle \}$
- 12.线性结构 非线性结构 非线性结构 树形结构 网状结构
- 13.一对一 一对多 多对多
- 14.结点 结点间关系
- 15.逻辑 物理
- 16.顺序存储 链接存储 文件组织
- 17.相对位置
- 18.结点的值 结点间关系的指针
- 19.顺序存取 直接存取 顺序文件
- 20.所用的外存种类 所需执行的操作
- 21.顺序结构 计算寻址结构 索引结构 表结构
- 22.查找 更新 插入 删除 排序
- 23.时间 空间
- 24.存储数据结构本身 实现算法所要求的辅助空间
- 25.逻辑结构 物理结构 定义在逻辑结构上的一组运算操作

二、选择题

- | | | | | |
|------|-----|-----|-----|------|
| 1.C | 2.A | 3.B | 4.B | 5.D |
| 6.D | 7.B | 8.B | 9.A | 10.C |
| 11.A | | | | |

三、简答题

1.答:

数据结构主要研究数据的逻辑结构、物理结构及定义在逻辑结构上的一组运算操作。