

全国高等教育自学考试

数据库原理自学辅导

(修订版)

组编/全国高等教育自学考试指导委员会
主编/丁宝康



辽宁大学出版社

出版前言

为了完善高等教育自学考试教育形式，促进高等教育自学考试的发展，我们组织编写了全国高等教育自学考试自学辅导书。

自学辅导书以全国考委公布的课程自学考试大纲为依据，以全国统编自考教材为蓝本，旨在帮助自学者达到学习目标，顺利通过国家考试。

自学辅导书是高等教育自学考试教育媒体的重要组成部分，我们将根据专业的开考情况和考生的实际需要，陆续组织编写出版文字、音像等多种自学媒体，由此构成与大纲、教材相配套的、完整的自学媒体系统。

全国高等教育自学考试指导委员会

2002年7月

编 者 的 话

《数据库原理》课程是全国高等教育自学考试为计算机及应用、计算机信息管理和计算机通信工程等专业（独立本科段）设置的一门专业基础课。

本书是依据 2000 年出版的《数据库原理》考试大纲和教材编写的，供自考生复习辅导和应考使用的。这次再版时，作了较大的修改。

本书分成两个部分，安排如下：

第一部分为分章辅导。阐述每章的要点、难点和自学重点，对教材中习题作了详细的解答，并给出新的练习题，以熟悉考试题型。

第二部分为四套模拟试题。其目的是为了帮助考生巩固所学知识，检验复习效果，提高应试能力，顺利通过自学考试。

《数据库原理》课程是理论性、实用性和可操作性都很强的一门学科。学好这门课程，对于毕业论文和今后的工作都会带来很大的益处。

学好这门课程的关键有三个：深刻了解教材中的基本概念和术语；深刻理解教材中的例题；操练和掌握教材中的习题。

参与本书改写的有丁宝康、李大学、陈长洪和徐美娟等老师。

限于水平，书中欠妥之处，敬请广大读者和专家指正。

编 者

2002 年 7 月

目 录

数据库原理自学辅导

第一部分 自学重点和习题解答

第1章 数据库概论	3
1.1 要点·难点解析	3
1.1.1 本章要点	3
1.1.2 考核重点(最高能力层次)	4
1.1.3 本章难点	4
1.2 教材中习题1的解答	4
1.3 练习题	8
1.4 练习题参考答案	10
第2章 关系模型	12
2.1 要点·难点解析	12
2.1.1 本章要点	12
2.1.2 考核重点(最高能力层次)	13
2.1.3 本章难点	13
2.2 教材中习题2的解答	14
2.3 练习题	20
2.4 练习题参考答案	23
第3章 关系数据库 SQL 语言	26
3.1 要点·难点解析	26
3.1.1 本章要点	26
3.1.2 考核重点(最高能力层次)	27
3.1.3 本章难点	27
3.2 教材中习题3的解答	27
3.3 练习题	38

3.4 练习题参考答案	40
第4章 关系数据库的模式设计	45
4.1 要点·难点解析	45
4.1.1 本章要点	45
4.1.2 考核重点(最高能力层次)	47
4.1.3 本章难点	48
4.2 教材中习题4的解答	49
4.3 练习题	59
4.4 练习题参考答案	62
第5章 数据库设计	64
5.1 要点·难点解析	64
5.1.1 本章要点	64
5.1.2 考核重点(最高能力层次)	66
5.1.3 本章难点	66
5.2 教材中习题5的解答	67
5.3 练习题	72
5.4 练习题参考答案	76
第6章 数据库保护	80
6.1 要点·难点解析	80
6.1.1 本章要点	80
6.1.2 考核重点(最高能力层次)	81
6.1.3 本章难点	82
6.2 教材中习题6的解答	82
6.3 练习题	90
6.4 练习题参考答案	93
第7章 分布式数据库系统	96
7.1 要点·难点解析	96
7.1.1 本章要点	96
7.1.2 考核重点(最高能力层次)	96
7.1.3 本章难点	97
7.2 教材中习题7的解答	97
7.3 练习题	101
7.4 练习题参考答案	103
第8章 具有面向对象特征的数据库系统	104
8.1 要点·难点解析	104
8.1.1 本章要点	104
8.1.2 考核重点(最高能力层次)	105
8.1.3 本章难点	105

8.2 教材中习题 8 的解答	105
8.3 练习题	110
8.4 练习题参考答案	111

第二部分 模拟试题

模拟试题 (一)	115
参考答案	119
模拟试题 (二)	122
参考答案	126
模拟试题 (三)	129
参考答案	132
模拟试题 (四)	137
参考答案	140
2001 年下半年全国高等教育自学考试数据库原理试卷	145
参考答案	148

第一部分

自学重点和习题解答

第 1 章 数据库概论

1.1 要点·难点解析

本章介绍数据库系统的基本概念、结构、功能及组成,使自学者对数据库概貌有个了解。在学习中,自学者应准确使用这些概念。

1.1.1 【本章要点】

1. 数据管理技术的发展

- (1) 人工管理、文件系统、数据库阶段的特点。
- (2) DB、DBMS、DBS 的定义。

2. 数据描述

- (1) 数据描述的三个领域及其术语。
- (2) 物理存储介质层次:六类存储介质。存储器中的数据描述。
- (3) 实体间联系:1:1、1:N、M:N。

3. 数据模型

- (1) 概念数据模型与结构数据模型。
- (2) 结构数据模型的三个组成部分:数据结构、数据操作、数据完整性约束。
- (3) 实体联系模型的主要成分及优点。
- (4) 比较层次模型、网状模型、关系模型、面向对象模型的优缺点。

4. 数据库的体系结构

- (1) 三级结构:外模式、概念模式、内模式。
- (2) 两级映象:外模式/模式映象、模式/内模式映象。
- (3) 两级独立性:逻辑独立性、物理独立性。
- (4) 用户、用户界面。

5. 数据库管理系统 (DBMS)

- (1) DBMS 的六个功能。
- (2) 从模块结构观察, DBMS 由查询处理器和存储管理器两大部分组成。每个部分的主要成分。
- (3) 从用户访问数据的过程,考察 DBMS 起的核心作用。

6. 数据库系统 (DBS)

(1) DBS 的全局结构: 四类数据库用户、DBMS 查询处理器的组成、DBMS 存储管理器的组成、磁盘存储器中的数据结构。

(2) DBS 的效益。

1.1.2 【考核重点 (最高能力层次)】

本章名词、术语甚多, 需要自学者记住。

1. 数据管理发展各阶段的特点。
2. 分清 DB、DBMS、DBS 三个术语, 了解其定义。
3. 数据描述的术语。实体间三种联系。
4. 数据库的体系结构中三级结构、两级映象、两级独立性概念。
5. DBMS 的功能、组成。
6. DBS 的组成、全局结构及效益。

1.1.3 【本章难点】

本章概念较多, 初学时可能不易理解。待全书学完后, 再来温习, 才能有深入的理解。

1.2 教材中习题 1 的解答

一、名词解释:

DB: DB 是统一管理的相关数据的集合。DB 能为各种用户共享, 具有最小冗余度, 数据间联系密切, 而又有较高的数据独立性。

DBMS: DBMS 是位于用户与 OS 之间的一层数据管理软件, 为用户或应用程序提供访问数据库的方法, 包括 DB 的建立、查询、更新及各种数据控制。

DBS: DBS 是实现有组织地、动态地存储大量关联数据, 方便多用户访问的计算机软件、硬件和数据资源组成的系统, 即采用了数据库技术的计算机系统。

1:1 联系: 如果实体集 E1 中每个实体至多和实体集 E2 中一个实体有联系, 反之亦然, 那么 E1 和 E2 的联系称为一对一联系, 记为 “1:1”。

1:N 联系: 如果实体集 E1 中每个实体与实体集 E2 中任意个 (零个或多个) 实体有联系, 而 E2 中每个实体至多和 E1 中一个实体有联系, 那么称 E1 对 E2 的联系是一对多联系, 记为 “1:N”。

M:N 联系: 如果实体集 E1 中每个实体与实体集 E2 中任意个 (零个或多个) 实体有联系, 反之亦然, 那么称 E1 和 E2 的联系是多对多联系, 记为 “M:N”。

数据模型: 对现实世界进行抽象, 表示实体类型及实体间联系的模型称为数据模型。

概念数据模型: 独立于计算机系统的数据模型, 完全不涉及信息在系统中的表示, 强调其语义表达功能。它是现实世界的第一层抽象, 是用户和数据库设计人员交流的工具。著名的概念模型是实体联系模型。

结构数据模型：涉及到计算机系统和 DBMS 的数据模型，它是现实世界的第二层抽象，有严格的形式化定义，以便于在计算机系统中实现。

层次模型：用树型结构表示实体类型及实体间联系的数据模型称为层次模型。

网状模型：用有向图结构表示实体类型及实体间联系的数据模型称为网状模型。

关系模型：用二维表格表示实体集，外键表示实体间联系的数据模型。

外模式（子模式）：用户与 DBS 的接口，是用户用到的那部分数据的描述。

概念模式（模式）：是数据库中全部数据的整体逻辑结构的描述。

内模式（存储模式）：数据库在物理存储方面的描述。

模式/内模式映象：用于定义概念模式和内模式间的对应性。

外模式/模式映象：用于定义外模式和概念模式间的对应性。

数据独立性：应用程序和数据库的数据结构之间相互独立，不受影响。

物理数据独立性：如果数据库的内模式（存储结构）要作修改，那么模式/内模式映象要作相应的修改，使概念模式尽可能保持不变，对外模式 and 应用程序的影响就更小。

逻辑数据独立性：如果数据库的概念模式要作修改，那么外模式/模式映象要作相应的修改，使外模式尽可能保持不变，对应用程序的影响就更小。

宿主语言：使用数据库的应用程序用 COBOL、FORTRAN、C 一类高级语言编写，这些语言在 DBS 中称为宿主语言。

DDL：定义数据库三级结构的语言。

DML：实现对数据库检索和更新的语言。

交互型 DML：可自成系统在终端上直接对数据库进行操作的 DML。

嵌入型 DML：嵌入在宿主语言中使用的 DML。

过程性 DML：编程时，需要指出“做什么”、“怎么做”的 DML。由用户进行数据导航。

非过程性 DML：编程时，只需指出“做什么”，不需指出“怎么做”的 DML。由系统进行数据导航。

DD（数据字典）：DBS 中存放三级结构的定义的 DB 称为 DD。

DD 系统：管理 DD 的实用程序称为 DD 系统。

二、文件系统阶段的数据管理有什么缺陷？试举例说明。

答案：

有三个缺陷：数据冗余性、数据不一致性、数据联系弱。

例如某单位添置了一台计算机，各部门纷纷在计算机中建立了文件。譬如有职工档案文件、职工工资文件和职工保健文件。但是每一职工的电话号码在三个文件中重复出现。这就是“数据冗余”；如果某职工的电话号码要修改，就要修改三个文件中的数据，否则会引起同一数据在三个文件中不一样；产生上述问题的原因是三个文件中数据没有联系。

三、数据库阶段的数据管理有什么特点？

答案：

有五点：复杂的数据模型、较高的数据独立性、方便的用户接口、数据控制功能（恢复、并发控制、完整性和安全性）、系统的灵活性。

四、你是怎样理解实体、属性、记录、字段这些概念的类型和值的差别？试举例说明。

答案：

类型是概念的内涵、定义，值是概念的外延、内容。

譬如“学生”是记录类型，而“张三”、“李四”为具体的学生，是记录值。在传统的数据库技术中，类型和值是严格加以区分的。

五、逻辑记录与物理记录，逻辑文件与物理文件有些什么联系和区别？

答案：

逻辑数据是用户用以操作的数据形式，是抽象的概念化数据。物理数据是实际存放在存储设备上的数据。逻辑数据与物理数据之间可以差别很大。可以用这样的思想解释题目（略）。

六、为某百货公司设计一个 ER 模型。

百货公司管辖若干连锁商店，每家商店经营若干商品，每家商店有若干职工，但每个职工只能服务于一家商店。

实体类型“商店”的属性有：商店编号、店名、店址、店经理。实体类型“商品”的属性有：商品编号、商品名、单价、产地。实体类型“职工”的属性有：职工编号、职工名、性别、工资。在联系中应反映出职工参加某商店工作的开始时间、商店销售商品的月销售量。

试画出反映商店、商品、职工实体类型及其联系类型的 ER 图，并将其转换成关系模式集。

答案：

ER 图如图 1-1 所示。

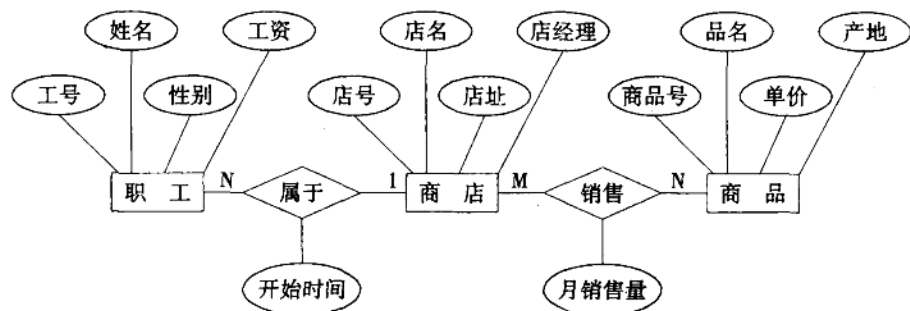


图 1-1

关系模式集：

职工（工号，姓名，性别，工资，店号，开始时间）

商店（店号，店名，店址，店经理）

商品（商品号，品名，单价，产地）

销售（店号，商品号，月销售量）

七、试述 ER 模型、层次模型、网状模型、关系模型和面向对象模型的主要特点。

答案：

ER 模型直接表示实体类型及实体间联系,与计算机无关,用户容易接受。

层次模型的数据结构为树结构,记录之间联系通过指针实现,查询效率较高,但 DML 属于过程化 DML,操作复杂。

网状模型的数据结构为有向图,记录之间联系也是通过指针实现,查询效率较高,并且容易实现 M:N 联系,但 DML 属于过程化语言,编程较复杂。

关系模型的数据结构为二维表格,容易为初学者理解。记录之间联系通过外键实现。DML 属于非过程化语言,编程较简单。

面向对象模型能完整描述现实世界的数据结构,具有丰富的表达能力,能表达嵌套、递归的数据结构。但涉及知识面较广,用户较难理解。

八、试述概念模式在数据库结构中的重要地位。

答案:

数据按外模式的描述提供给用户,按内模式的描述存储在磁盘中,而概念模式提供了连接这两级的相对稳定的中间观点,并使得两级的任何一级的改变都不受另一级的牵制。

九、数据独立性与数据联系这两个概念有什么区别?

答案:

数据独立性是指应用程序和数据库的数据结构之间相互独立,不受影响。而数据联系是指记录之间的联系。

十、试述 DBMS 在用户访问数据库过程中起的作用。

答案:

实现数据的三级结构的转换。把逻辑数据转换成物理数据,把物理数据转换成逻辑数据。

十一、试述 DBMS 的主要功能。

答案:

数据库的定义、操纵、保护、存储管理、维护、数据字典等六个功能。(详细解释略)

十二、试叙述 DBMS 对数据库的保护功能。

答案:

包括恢复、并发控制、完整性控制、安全性控制四个方面。(详细解释略)

十三、试叙述 DBMS 对数据库的维护功能。

答案:

DBMS 有许多实用程序提供给 DBA 运行 DBS 时使用,主要的实用程序有数据装载程序、备份程序、文件重组程序、性能监控程序等四个。

十四、从模块结构看, DBMS 有哪些部分组成?

答案:

DBMS 由两大部分组成:

①查询处理器,有四个主要成分:DDL 编译器、DML 编译器、嵌入型 DML 的预编译器、查询运行核心程序。

②存储管理器,有四个主要成分:授权和完整性管理器、事务管理器、文件管理器、缓冲区管理器。

十五、DBS 由哪几个部分组成？

答案：

DBS 由数据库、硬件、软件、DBA 四个部分组成。（详细解释略）

十六、什么样的人 DBA？DBA 应具有什么素质？DBA 的职责是什么？

答案：

DBA 是控制数据整体结构的人，负责 DBS 的正常运行。

DBA 应具有下列素质：熟悉企业全部数据的性质和用途；对用户的要求有充分的了解；对系统的性能非常熟悉。

DBA 承担创建、监控和维护整个数据库结构的责任。具体有五点：定义概念模式、定义内模式、根据需求修改概念模式和内模式、对数据库访问的授权、完整性约束的说明。

十七、试对 DBS 的全局结构作详细解释。

答案：

①数据库用户有四类：DBA、专业用户、应用程序员、最终用户。

②DBMS 的查询处理器有四个模块：DML 编译器、嵌入型 DML 预编译器、DDL 编译器、查询运行核心程序。

③DBMS 的存储管理器有四个模块：授权和完整性管理器、事务管理器、文件管理器、缓冲区管理器。

④磁盘存储器中有四种数据结构：数据文件、数据字典、索引文件、统计数据组织。

十八、使用 DBS 的用户有哪几类？

答案：使用 DBS 的用户有 DBA、专业用户、应用程序员和最终用户等四类。

十九、DBMS 的查询处理器有哪些功能？

答案：

有四个：DDL 编译、DML 编译、嵌入型 DML 的预处理、执行由 DML 编译器产生的低层指令。

二十、DBMS 的存储管理器有哪些功能？

答案：

有四个：授权和完整性管理、事务管理、文件管理、缓冲区管理。

二十一、磁盘存储器中有哪四类主要的数据结构？

答案：磁盘存储器中有数据文件、数据字典、索引文件和统计数据组织等四类数据结构

1.3 练习题

一、单项选择题

1. 文件系统中的数据独立性是指 ()
A. 文件独立性 B. 硬件独立性 C. 设备独立性 D. 软件独立性
2. 在 DBS 中，“数据独立性”和“数据联系”这两个概念之间的联系是 ()

- A. 没有必然的联系 B. 同时成立或不成立
C. 前者蕴涵后者 D. 后者蕴涵前者
3. DB 中数据导航是指 ()
A. 数据之间的联系 B. 数据之间指针的联系
C. 从已知数据找未知数据的过程 D. 数据之间的组合方式
4. 在数据库方式下, 信息处理中占据中心位置的是 ()
A. 磁盘 B. 程序 C. 数据 D. 内存
5. DB 的三级体系结构是对_____抽象的三个级别。 ()
A. 存储器 B. 数据 C. 程序 D. 外存
6. DB 的三级模式结构中最接近外部存储器的是 ()
A. 子模式 B. 外模式 C. 概念模式 D. 内模式
7. DBS 具有“数据独立性”特点的原因是因为在 DBS 中 ()
A. 采用磁盘作为外存 B. 采用三级模式结构
C. 使用 OS 来访问数据 D. 用宿主语言编写应用程序
8. DBMS 要在 OS 的支持下才能工作, 两者的关系是 ()
A. 相互调用 B. 前者调用后者 C. 后者调用前者 D. 两者并发运行

二、填空题

1. 在 DBS 中, 第三级存储器是指_____和_____。
2. “块”是内存和外存交换信息的最小单位, 内外存信息交换是由_____实现的。
3. 数据库设计人员与用户进行交流的工具是_____, 它是对现实世界的第一层抽象。
4. 结构数据模型应包含_____、_____和_____等三个部分。
5. 层次模型、网状模型中记录之间的联系是通过_____实现的。
6. 面向对象模型中两个最基本的概念是_____和_____, 这两个概念相当于 ER 模型中的实体和实体集。
7. 数据库的三级模式结构是对_____的三个抽象级别。
8. 在 DBS 中, 数据按_____的描述提供给用户, 按_____的描述存储在磁盘中, 而_____提供了连接这两级的相对稳定的中间观点, 并使得两级的任何一级的改变都不受另一级的牵制。
9. DBS 中存放三级结构定义的 DB 称为_____。
10. 从模块结构考察, DBMS 由两大部分组成: _____和_____。
11. 从用户访问数据的过程可以看出, DBMS 的主要职责是_____。
12. DBS 是_____, _____、_____和_____的集合体。
13. 在 DBS 中, DB 在磁盘上的基本组织形式是_____, 这样可以充分利用 OS _____的功能。
14. DBMS 是位于_____和_____之间的一层数据管理软件。

三、问答题

1. 与文件的结构相比, 数据库的结构有些什么不同?

2. 根据对现实世界抽象层次的不同将数据模型分为哪两种? 各有什么特点。
3. 在各种结构数据模型中, 数据之间的联系是怎么实现的?
4. 应用程序、DB 的三级模式结构和磁盘上的数据之间有什么联系?
5. 什么是“DB 的系统缓冲区”?

1.4 练习题参考答案

一、单项选择题

1. C 2. A 3. C 4. C 5. B 6. D 7. B 8. B

二、填空题

1. 光盘 磁带 2. OS 的文件系统 3. 概念模型 (或 ER 模型) 4. 数据结构 数据操作 数据完整性约束 5. 指针 6. 对象 类 7. 数据 8. 外模式 内模式 概念模式 9. 数据字典 (或 DD) 10. 查询处理器 存储管理器 11. 实现数据三级结构的转换 12. 数据库 硬件 软件 DBA 13. 文件 管理外存 14. 用户 OS

三、问答题

1. 答: 与文件结构相比, 数据库结构主要有下面三点不同:

(1) 数据结构化。数据结构用数据模型表示, 整体结构化。数据模型不仅描述数据本身的特点, 还要描述数据之间的联系。

(2) 数据独立性。文件只有设备独立性, 而数据库还具有逻辑独立性和物理独立性。

(3) 访问数据的单位。访问文件中的数据, 以记录为单位。访问数据库中的数据, 以数据项 (字段) 为单位, 增加了系统的灵活性。

2. 答: 分为概念数据模型和结构数据模型两种。前者用于概念设计, 后者用于逻辑设计。

概念模型是一种独立于硬件和软件的模型。这种模型是从用户的观点对数据建模, 必须充分反映用户的需求, 并得到用户的确认才能定下来。它是现实世界的第一层抽象, 是用户和数据库设计人员之间进行交流的工具, 其典型代表是 ER 模型。

结构模型是用于描述 DB 的逻辑结构, 与 DBMS 有关。这种模型是从计算机的观点对数据建模。它是现实世界的第二层抽象, 是数据库设计人员和应用程序员之间进行交流的工具。其典型代表是层次、网状、关系和面向对象模型。

3. 答: (1) 在层次、网状模型中, 数据之间联系用指针实现。指针值是记录的地址。这种方法查询速度较快, 但带来一个缺点, 查询语言是过程性的, 应用程序的编写比较复杂。

(2) 在关系模型中, 数据之间联系通过关键码来体现。在关系中通过外键和主键的联系实现两个记录的联系。这种方法, 查询语言是非过程性的, 编程比较简单。但缺点是查询速度较慢。随着硬件性能的改善, “慢”的现象正在消失。

(3) 在面向对象模型中, 数据联系用“对象标识”实现。对象标识是指针一级的概念, 但它是一个强有力的数据操纵原语。引用对象时, 不是引用对象本身, 而是引用“对象标识”。

4. 答：应用程序、DB 的三级模式结构和磁盘上的数据之间记录的联系，可用图 1-2 表示。

在图中，标出了每个层次上记录类型的称呼。这里应注意两点：

(1) 三级结构上的记录类型的结构不一致，通过两级映象实现格式转换。

(2) 外模式中的外部记录类型应与应用程序中的记录类型一致。内模式中的内部记录类型应与磁盘上的物理文件的记录类型一致。

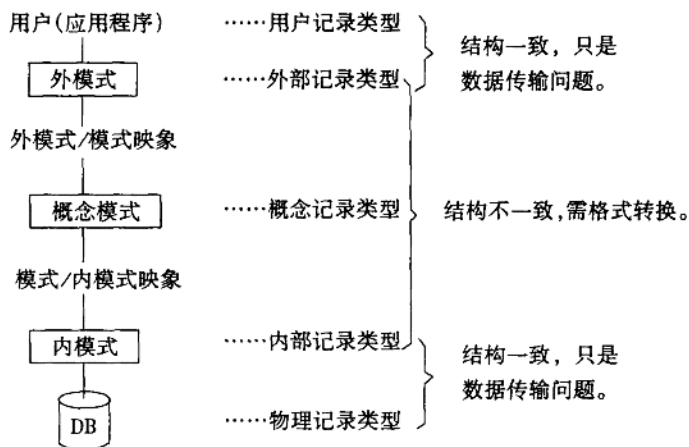


图 1-2

5. 答：在应用程序运行时，DBMS 在内存为其开辟一个 DB 的系统缓冲区，用于“数据的传输和格式的转换”。