

计算机核心课程辅导系列教材

数据库原理

题解·综合练习

盛定宇 彭 澎 编著

. 131-44



机械工业出版社
CHINA MACHINE PRESS



计算机核心课程指导系列教材

数据库原理题解·综合练习

盛定宇 彭 澎 编著



机械工业出版社

数据库原理作为研究数据库技术的基本理论已经成为学习信息技术的重要专业课程。在高等教育的有关专业：计算机科学与技术、电子信息科学与技术、信息管理与信息系统等，以及自学考试的计算机科学与应用、计算机信息管理、计算机网络等都开设了相关的课程。全书分为 7 个专题：数据库概述；关系模型；关系数据 SQL 语言；关系数据理论；数据库设计；数据库保护；数据库技术的发展。通过这 7 个专题，向读者介绍了数据库的基本概念和方法。书中还配有大量习题和测试题供读者练习。本书可供大学本科和专科学生学习本课程时参考，也可以作为报考硕士研究生的考生复习本课程的参考书。

图书在版编目 (CIP) 数据

数据库原理题解·综合练习/盛定宇, 彭澎编著. —北京: 机械工业出版社, 2004.7

(计算机核心课程辅导系列教材)

ISBN 7-111-14951-3

I. 数... II. ①盛... ②彭... III. 数据库系统—习题 IV. TP311.13-44

中国版本图书馆 CIP 数据核字 (2004) 第 072069 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策 划: 胡毓坚 赵 慧

责任编辑: 丁 诚

责任印制: 石 冉

三河市宏达印刷有限公司印刷·新华书店北京发行所发行

2004 年 8 月第 1 版·第 1 次印刷

787mm×1092mm 1/16·15 75 印张·384 千字

0001—5000 册

定价: 24.00 元

凡购本图书, 如有缺页、倒页、脱页, 由本社发行部调换

本社购书热线电话 (010) 68993821、88379646

封面无防伪标均为盗版

出版说明

高等教育是培养人才过程中非常重要的一个环节。近年来随着高等教育的不断发展，作为计算机、信息管理等相关专业，每年接受各种教育和培训的人数不下百万，包括在读大专院校的学生、参加高等教育自学考试的学员、参加全国计算机等级考试的人员以及参加研究生入学考试的考生。国家也有许多这方面的考试，如软件人员水平测试、企业信息管理员考试等。

作为计算机、信息管理等专业学生必修的一些专业基础课和专业课程，在学习过程中会遇到很多问题和困难。为了帮助学习者学好和掌握课程内容，正确和深入地理解和把握课程内容，并在参加相关课程的各种考试时能够考出好成绩，我们组织了一批具有相关课程教学经验的一线教师，编写了这套“计算机核心课程辅导系列教材”。

本系列教材中所包括的课程都属于学科重点课程、核心课程和考研课程。每本书都是针对相关课程编写的，不仅仅是与某一具体教材配套的辅导书。本系列教材采用全新的体例结构。内容包括：重点难点分析、练习题及综合练习，并附有参考答案，有些课程还配有实验。编写体例也突破了一般指导书的模式，在对概念、原理内容的描述上尽量采用要点分解方式，使学习者一目了然，直接深入到概念和原理的本质和核心中，便于记忆、掌握和灵活应用，对提高应试成绩具有很大的帮助。由于本系列教材已对课程内容进行了充分的整理和提炼，所以本系列教材还是学生学习过程中最好的“笔记”。考虑到学生在学习中有记笔记的需要，在主要内容、重点内容部分的版面设计中，为读者留出了记笔记的地方，进一步方便了读者的使用。

本系列教材注重概念、原理的准确性，精练性，以及内容的全面性。高校在校生可根据所使用的教材，将其作为辅导材料。对考研的读者，这套辅导教材也是非常好的参考书。另外，本系列教材还适合自考学生和参加各类信息技术考试的人员作参考。

前 言

随着信息技术（包括计算机技术、通信技术和网络技术等）的飞速发展，信息系统渗透到社会的各个领域，作为其核心和基础的数据库技术也得到了越来越广泛的应用。数据库的建设规模、数据库信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

数据库原理是研究数据库技术的基本理论已经成为学习信息技术的重要专业课程。在高等教育的有关专业：计算机科学与技术、电子信息科学与技术、信息管理与信息系统等，以及自学考试的计算机科学与技术、计算机信息管理、计算机网络等都开设了相关的课程。

数据库原理所涉及的内容，不仅是这些相关专业必备的知识基础，也是从事信息产业工作人员的必备知识与技能，同时也是进一步深入研究数据管理与应用的出发点。

本书分为 7 个专题：数据库概述；关系模型；关系数据库 SQL 语言；关系数据理论；数据库设计；数据库保护；数据库技术的发展，每个专题都对所涉及的概念作概要的叙述，然后通过实际例题加深对概念的理解，在一些更深入研究的问题上，提出主要思路。

在书中配置了大量的习题，读者通过解题可进一步理解概念与方法，建议读者完成习题后再核对答案。本书针对大学专科、本科学生和报考硕士研究生的需要，内容按深浅程度分为几个层次，供读者选择。

本书编写中，作者参考了一些相关书籍与教材中的内容与习题，在此向有关参考资料的作者表示深切的感谢。

数据库技术涉及的领域很广，新的技术与方法也在不断出现，由于本人水平有限，其中难免出现错误和不完善之处。真诚希望广大专家和读者给予批评指教，笔者将非常感谢。

编 者

目 录

前言	问题	107
第1章 数据库概述	练习题	111
第1节 数据管理技术的发展	参考答案	116
第2节 数据描述	第5章 数据库设计	119
第3节 数据模型	第1节 数据库设计介绍	120
第4节 数据库体系结构	第2节 规划	122
第5节 数据库管理系统	第3节 需求分析	122
第6节 数据库系统	第4节 概念设计	126
练习题	第5节 逻辑设计	132
参考答案	第6节 物理设计	142
第2章 关系模型	第7节 实现与维护	143
第1节 关系模型的基本概念	练习题	143
第2节 关系代数	参考答案	149
第3节 关系演算	第6章 数据库保护	156
第4节 查询优化	第1节 数据库的恢复	157
练习题	第2节 数据库的并发控制	160
参考答案	第3节 数据库的完整性	163
第3章 关系数据库 SQL 语言	第4节 数据库的安全性	168
第1节 SQL 介绍	练习题	173
第2节 SQL 数据定义	参考答案	174
第3节 SQL 数据查询	第7章 数据库技术的发展	176
第4节 SQL 数据更新	第1节 面向对象数据库系统	177
第5节 嵌入式 SQL	第2节 分布式数据库系统	186
练习题	第3节 网络环境下的数据库体系	191
参考答案	第4节 数据仓库与数据挖掘	198
第4章 关系数据理论	练习题	211
第1节 关系数据模型的构造	参考答案	214
第2节 关系模型的范式	综合练习	217
第3节 多值依赖与 4NF	综合练习一	218
第4节 关系模式规范化的步骤	综合练习二	220
第5节 函数依赖集合的规范化	综合练习三	223
第6节 关系模型候选码的确定	综合练习一答案	229
第7节 关系模式分解的规则	综合练习二答案	233
第8节 连接依赖与 5NF	综合练习三答案	237
第9节 数据规范化理论证明	参考文献	243

第1章

数据库概述

- 数据管理技术的发展
- 数据描述
- 数据模型
- 数据库体系结构
- 数据库管理系统
- 数据库系统

数据库原理所涉及的内容,不仅是这些相关专业必备的知识基础,也是从事信息产业工作人员的必备知识与技能,同时也是进一步深入研究数据管理与应用的出发点。本章从数据管理技术的发展过程,分析数据管理技术的特点,重点讲述了数据库系统的特点和数据模型的基本概念。

第1节 数据管理技术的发展

笔记

1. 数据管理技术的发展经历了三个阶段

(1) 人工管理阶段(20世纪50年代中期以前),其特点是:

- 1) 数据不保存在计算机中。
- 2) 没有专用的软件对数据进行管理。
- 3) 只有程序的概念,没有文件的概念。

(2) 文件系统阶段(20世纪50年代后期至20世纪60年代中期),其特点是:

- 1) 数据可以脱机长期保存在外存储器中。
- 2) 数据具有了与物理结构不同的逻辑结构。
- 3) 出现了多样化的文件组织方式。
- 4) 数据可以为多个不同的程序重复使用。

其缺点是:

- 1) 数据冗余较大。
- 2) 存在数据不一致性的问题。
- 3) 数据之间缺乏联系。

(3) 数据库阶段(20世纪60年代末开始),其特点是:

- 1) 数据结构化。
- 2) 数据共享性高、冗余度小、易扩充。
- 3) 数据独立性高,包括物理独立性和逻辑独立性。
- 4) 统一的数据管理和控制,包括数据安全性、数据完整性、并发控制和数据库恢复。

5) 数据的最小存取单位是数据项。

2. 需要明确并区分的几个概念

(1) 数据库(DB):数据库是长期存储在计算机存储设备内、有组织的、共享的数据集合。这些数据按一定的数据模型组织、描述、存储,具有较小冗余度、较高的数据独立性和易扩充性,并可为各种用户共享。

(2) 数据库管理系统(DBMS):位于用户和操作系统之间的,实现各项数据管理功能的软件。

(3) 数据库系统(DBS):实现有组织地、动态地存储大量关联数据,方便用户访问的计算机硬件、软件、数据资源,以及相关工作人员组成的系统。

(4) 数据库技术:研究数据库的结构、存储、管理和使用的软件学科。

第2节 数据描述

笔 记

1. 数据描述的三个领域

- (1) 现实世界：存在于人们头脑之外的客观世界。
- (2) 信息世界：现实世界在人们头脑里的反映，被人们用文字和符号记载下来。其主要描述工具是实体、属性和联系。
- (3) 机器世界：在计算机中以数据存储的形式存储信息世界的信息。主要存储方式是数据项、记录和文件。

2. 数据库中类型与值的区别

类型是概念的内涵（结构、框架），值是概念的外延（具体内容、实例）。

3. 数据描述的两形式

- (1) 物理描述：物理数据是实际存放在设备上的数据，对物理数据存放方式的描述称为数据的物理描述。
- (2) 逻辑描述：抽象的概念化数据，是用户（程序员或操作用户）观点的数据描述。

4. 数据存储介质与存储描述

(1) 数据存储介质中，按存取速度从高到低排列的一般顺序是：高速缓冲存储器、主（内）存储器、快擦写存储器、（硬）磁盘存储器、光存储器、磁带。随着数据存储技术的改进，这个次序也会有一些变化。例如软盘存储器的速度一般比较低，但现在已经开始广泛使用的可移动磁盘存储器，速度也很快，而且容量也较大。

(2) 数据物理存储的描述从小到大的顺序是：位、字节、字、块、桶、卷。

5. 实体（Entity）

客观存在并可相互区别的事物称为实体。实体可以是具体的人、事、物，也可以是抽象的概念或联系，例如，一个供应商、一个学生、一个职工、一门课、一个部门、学生的一次选课、部门的一次订货、老师与系的工作关系（即某位老师在某系工作）等都是实体。

6. 属性（Attribute）

实体所具有的某一特性称为属性。一个实体可以由若干个属性来刻画。例如，学生实体可以由学号、姓名、性别、出生年份、系、入学时间等属性组成（19960011，周志平，男，1971，计算机系，1989）。这些属性组合起来表征了一个学生。

7. 联系（Relationship）

在现实世界中，事务内部以及事务之间是有联系的，这些联系在信

息世界中反映为实体内部的联系和实体之间的联系。实体内部的联系通常是指组成实体的各属性之间的联系。两个实体型之间的联系可以分为三类。

(1) 一对一联系 (1:1)

如果对于实体集 A 中的每一个实体, 实体集 B 中至多有 n 个实体与之联系, 反之亦然, 则称实体集 A 与实体集 B 具有一对一联系, 记为 1:1。

例如, 确定部门实体与经理实体之间存在一对一联系, 意味着一个部门只能有一个经理管理, 而一个经理只管理一个部门。

(2) 一对多联系 (1:n)

如果对于实体集 A 中的每一个实体, 实体集 B 中有 n 个实体 ($n \geq 0$) 与之联系, 反之, 对于实体集 B 中的每一个实体, 实体集 A 中至多只有一个实体与之联系, 则称实体集 A 与实体集 B 有一对多联系, 记为 1:n。

例如, 一个部门中有若干名职工, 而每个职工只在一个部门中工作, 则部门与职工之间具有一对多联系。

(3) 多对多联系 (m:n)

如果对于实体集 A 中的每一个实体, 实体集 B 中有 n 个实体 ($n \geq 0$) 与之联系, 反之, 对于实体集 B 中的每一个实体, 实体集 A 中也有 m 个实体 ($m \geq 0$) 与之联系, 则称实体集 A 与实体集 B 具有多对多联系, 记为 m:n。

例如, 一门课程同时有若干个学生选修, 而一个学生可以同时选修多门课程, 则课程与学生之间具有多对多联系。

实际上, 一对一联系是一对多联系的特例, 而一对多联系又是多对多联系的特例。实体型之间的这种一对一、一对多、多对多联系不仅存在于两个实体型之间, 也存在于两个以上的实体型之间。

若实体集 $E_1, E_2, \dots, E_n$ 存在联系, 对于实体集 E_j ($j=1, 2, \dots, i-1, i+1, \dots, n$) 中的给定实体, 最多只和 E_i 中的一个实体相联系, 则我们说 E_i 与 $E_1, E_2, \dots, E_{i-1}, E_{i+1}, \dots, E_n$ 之间的联系是一对多的。例如, 对于课程、教师与参考书三个实体型, 如果一门课程可以有若干个教师讲授, 使用若干本参考书, 而每一个教师只讲授一门课程, 每一本参考书只供一门课程使用, 则课程与教师、参考书之间的联系是一对多的。

多实体型之间一对一、多对多联系的定义及其例子请读者自行给出。

同一个实体集内的各实体之间也可以存在一对一、一对多、多对多的联系。例如, 学生实体集内部具有领导与被领导的联系, 即某一学生 (班干部) 领导若干名学生, 而一个学生仅被另外一个学生直接领导, 因此这是一对多的联系。

第3节 数据模型

笔记

数据模型是现实世界数据特征的抽象, 构造一些相关的数据组织的集合。任何一种数据模型都是严格定义的概念的集合, 这些概念必须能够精确地描述系统的静态特性、动态特性和完整性约束条件。因此数据模型通常都是由数据结构、数据操作和完整性约束三个要素组成。

数据库是某个企业、组织或部门所涉及的数据的一个综合, 它不仅要反映数据本身的内容, 而且要反映数据之间的联系。由于计算机不可能直接处理现实世界中的具体事物, 所以人们必须事先把具体事物转换成计算机能够处理的数据。在数据库中用数据模型这个工具来抽象、表示和处理现实世界中的数据和信息。通俗地讲数据模型就是现实世界中的模拟。

1. 数据模型的基本概念

现有的数据库系统都是基于某种数据模型的。数据模型是数据库系统的数学形式框架, 用来描述数据的一组概念和定义。它包括以下方面的内容。

- 数据的静态特征。它包括对数据结构和数据间联系的描述。
- 数据的动态特征。一组定义在数据上的操作, 包括操作的含义、操作符、运算规则及其语言等。
- 数据的完整性约束。这是一组规则, 数据库中的数据必须满足这组规则。

数据模型应满足三方面要求, 一是能比较真实地模拟现实世界; 二是容易为人所理解; 三是便于在计算机上实现。一种数据模型要很好地满足这三方面的要求, 在目前尚很困难。在数据库系统中应针对不同的使用对象和应用目的, 采用不同的数据模型。

不同的数据模型实际上是提供给我们模型化数据和信息的不同工具。根据模型应用的不同目的, 可以将这些模型划分为两类, 它们分属于两个不同的层次。第一类模型是概念模型, 也称信息模型, 它是按用户的观点对数据和信息建模。另一类模型是结构模型, 主要包括网状模型、层次模型、关系模型和面向对象模型等, 它是按计算机系统的观点对数据建模。数据模型是数据库系统的核心和基础。

2. 数据模型三要素

(1) 数据结构

数据结构用于描述系统的静态特性。数据结构是所研究的对象类型 (Object Type) 的集合。这些对象是数据库的组成成分, 它们包括两类, 一类是与数据类型、内容、性质有关的对象, 例如网状模型中的数据项、记录, 关系模型中的域、属性、关系等; 一类是与数据之间联系有关的

笔 记

对象,例如网状模型中的系型(Set Type)。

数据结构是刻画一个数据模型性质最重要的方面。因此在数据库系统中,通常按照其数据结构的类型来命名数据模型。例如,层次结构、网状结构和关系结构的数据模型分别命名为层次模型、网状模型和关系模型。

(2) 数据操作

数据操作用于描述系统的动态特性。数据操作是指对数据库中各种对象(型)的实例(值)允许执行的操作的集合,包括操作及有关的操作规则。数据库主要有检索和更新(包括插入、删除、修改)两大类操作。数据模型必须定义这些操作的确切含义、操作符号、操作规则(如优先级)以及实现操作的语言。

(3) 数据的约束条件

数据的约束条件是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的制约和存储规则,用以限定符合数据模型的数据库状态以及状态的变化,以保证数据的正确、有效和相容。数据模型应该反映和规定本数据模型必须遵守的基本的通用的完整性约束条件。例如,在关系模型中,任何关系必须满足实体完整性和参照完整性两个条件。

此外,数据模型还应该提供定义完整性约束条件的机制,以反映具体应用所涉及的数据必须遵守的特定的语义约束条件。

3. 实体联系模型(E-R模型)

(1) E-R图的组成

- 实体型:用矩形表示,矩形框内写明实体名。
- 属性:用椭圆形表示,并用无向线段将其与相应的实体连接起来。
- 联系:用菱形表示,菱形框内写明联系名,并用无向线段分别与有关实体连接起来,同时在旁边标上联系的类型(1:1, 1:n或m:n)。

需要注意的是,联系本身也是一种实体型,也可以有属性。如果一个联系具有属性,则这些属性也要用无向线段与该联系连接起来。

图 1-1 用 E-R 图描述了上述有关两个实体型之间的三类联系、三个实体型之间的一对多联系和一个实体型内部的一对多联系例子。

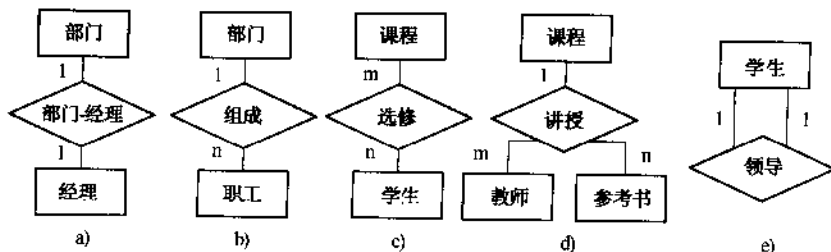


图 1-1 实体间联系

a) 1:1 联系 b) 1:n 联系 c) m:n 联系

d) 三个实体间的 m:n 联系 e) 同一实体间的 1:1 联系

(2) 实体关联图

实体关联图是使用信息结构描述现实世界数据视图的有效工具。E-R 图提供了表示实体型、属性和联系的方法，其主要特点是表达直观，易于理解，修改方便，并且便于向数据库逻辑模型转换。因此实体关联图是数据库数据模型的雏形和基础，在数据库设计中具有重要的作用。

现在，从分析一个实例开始，讨论进行设计实体联系图的方法。

假设五个实体型即学生、班级、课程、教师、参考书分别具有下列属性。

学生：学号、姓名、性别、年龄

班级：班级编号、所属专业系

课程：课程号、课程名、学分

教师：职工号、姓名、性别、年龄、职称

参考书：书号、书名、内容提要、价格

这五个实体的属性用 E-R 图表示，如图 1-2 所示。这五个实体之间的联系可以用 E-R 图表示，如图 1-3 所示。注意，选修和组成两个联系又都分别具有各自的属性。

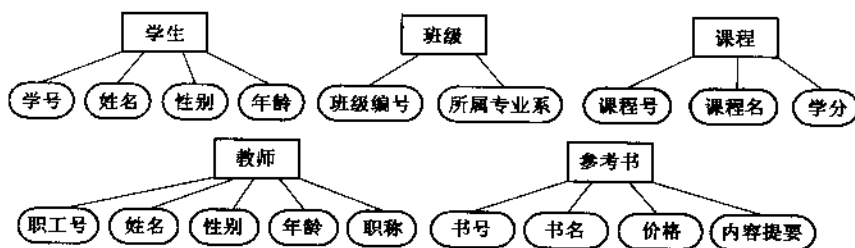


图 1-2 实体属性图

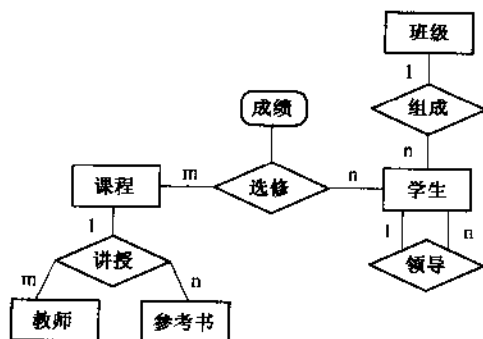


图 1-3 实体联系图

将图 1-2 与图 1-3 合并在一起（图 1-4）就是一个完整的关于学校课程管理的概念模型了。但在实际当中，在一个概念模型中涉及的实体和实体的属性较多时，为了清晰起见，往往采用将实体及其属性与实体及其联系分别用两张 E-R 图表示的方法。

实体-联系方法（E-R 方法）是抽象和描述现实世界的有力工具。用 E-R 图表示的概念模型独立于具体的 DBMS 所支持的数据模型，是各种数据模型的共同基础，因而比数据模型更一般、更抽象、接近现实世界。

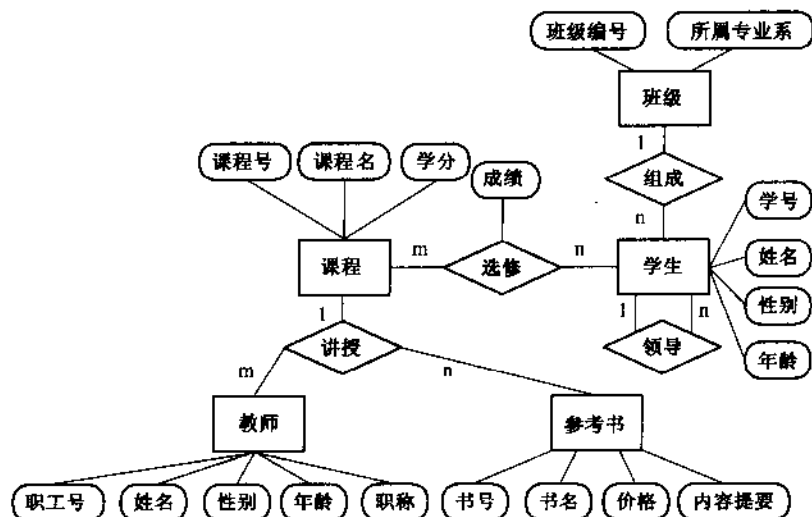


图 1-4 完整的实体联系图

4. 主要结构数据模型

不同的数据模型具有不同的数据结构形式。目前最常用的数据模型有层次模型（Hierarchical Model）、网状模型（Network Model）、关系模型（Relational Model）和面向对象数据模型（Object Oriented Model）。其中层次模型和网状模型统称为非关系模型。非关系模型的数据库系统在 20 世纪 70 年代与 80 年代初非常流行，在数据库系统产品中占据了主导地位，现在已逐渐被关系模型的数据库系统取代。

20 世纪 80 年代以来，面向对象的方法和技术在计算机各个领域，包括程序设计语言、软件工程、信息系统设计、计算机硬件设计等各方面都产生了深远影响，也促进了数据库中面向对象数据模型的研究和发展。

(1) 层次数据模型

层次模型是数据库系统中最早出现的数据模型，它用树形结构表示各类实体以及实体间的联系。现实世界中许多实体之间的联系本来就呈现出一种很自然的层次关系，如行政机构、家族关系等。层次模型数据库系统的典型代表是 IBM 公司的 IMS（Information Management Systems）数据库管理系统，这是一个曾经广泛使用的数据库管理系统。

(2) 网状数据模型

在现实世界中实体间的联系更多的是非层次关系，用层次模型表示非树形结构是很不直接的，采用网状模型作为数据的组织方式，可以克服这一弊病。网状数据模型的典型代表是 DBTG 系统，也称 CODASYL

系统。

层次数据模型和网状数据模型都是早期的数据库数据模型，数据库系统与文件系统的主要区别就是不仅定义数据的存储而且还定义存储数据之间的联系，所谓“层次”和“网状”就是指这种联系的方式。

(3) 关系数据模型

关系模型与以往的模型不同，它是建立在严格的数学理论的基础上的。在用户看来，一个关系模型的逻辑结构是一张二维表，它由行和列组成。

关系模型是目前最重要的一种模型。关系数据库系统采用关系模型作为数据的组织方式。美国 IBM 公司的研究员 E.F.Codd 于 1970 年发表题为“大型共享系统的关系数据库的关系模型”的论文，文中首次提出了数据库系统的关系模型。20 世纪 80 年代以来，计算机厂商新推出的数据库管理系统（DBMS）几乎都支持关系模型，非关系系统的产品也大都加上了关系接口。数据库领域当前的研究工作都是以关系方法为基础。

在关系数据模型中不仅仅有一种数据结构——二维表，而且在数据操作中也具有“非过程化”的特点，因此易学易用，成为当前数据库应用系统的主流。

(4) 面向对象数据模型

面向对象数据模型是近些年出现的一种新的数据模型，它比前三种数据模型具有对现实世界更广泛、更深入、更丰富的表达能力。所谓对象是对现实世界事物的高度抽象，每个对象是状态和行为的封装。对象的状态是属性的集合，行为是在该对象上操作方法的集合。因此面向对象的模型不仅可以处理各种复杂多样的数据结构，而且具有数据与行为相结合的特点。目前面向对象的方法已经逐渐成为系统开发、设计的全新思路。但这种方法模型相对复杂，实现有一定难度。在当前信息处理技术中，关系数据模型仍然是数据库数据模型的主流，即使使用面向对象的模型也往往采用关系数据模型的方法和工具。

第 4 节 数据库体系结构

数据库系统体系结构如图 1-5 所示。

1. 三级模式结构

(1) 外部级：每个用户的数据视图的描述，称为“外模式”，也称“子模式”，它是概念模式的逻辑子集。

(2) 概念级：数据库全局逻辑结构的描述，称为“概念模式”，简称“模式”。

(3) 内部级：数据库存储结构的描述，称为“内模式”。

笔 记

笔 记

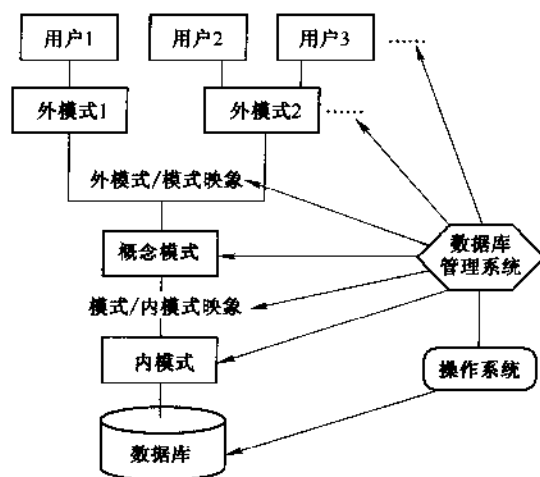


图 1-5 数据库系统体系结构

需要特别说明的是，此处的“概念模式”与数据库设计中的“概念模式”不是一个概念，相当于数据库设计中“逻辑模式”的概念。

2. 三级结构和两级映像

每级模式的定义有相应的数据定义语言（DDL）来实现，即模式 DDL、子模式 DDL 和内模式 DDL。

外模式与模式之间通过外模式 / 模式映像来实现转换，模式与内模式之间通过模式 / 内模式映像来实现转换。

3. 两级数据独立性

(1) 数据的物理独立性：当数据物理结构改变时，通过修改模式 / 内模式映像使得数据的逻辑结构不变，因而应用程序不变由此实现了数据的物理独立性。

(2) 数据的逻辑独立性：当数据全局逻辑结构改变时，通过修改外模式/模式映像使得数据的局部逻辑结构不变，因而应用程序不变，由此实现了逻辑的物理独立性。

4. 用户及用户界面

用户：使用数据库的应用程序或联机终端的人员。

用户界面：用户与数据库系统的分界线，这个分界线定在外部级上，用户对外模式是可知的，界面之下对用户都是不可知的。

用户使用数据操作语言（DML）操作数据库的数据处理。DML 可以自成系统，用户可以直接使用 DML 对数据库操作，这是自含型（交互型）DML；DML 也可以嵌入到某中高级语言（如 COBOL、FORTRAN、C 等语言）中使用，这些高级语言称为主语言，DML 称为嵌入型（宿主型）DML，其特点是主语言发挥其过程化处理特点，DML 执行数据库数据处理功能。

5. 数据库模式与数据库

数据库模式是数据库的结构描述，而不是数据库本身，是相对稳定的。数据库是在一定的数据库模式下数据实际存储的一个实例，其内容是随时间变化的。

第5节 数据库管理系统

数据库管理系统（DBMS）是管理、控制数据库系统和执行数据库操作的核心软件，数据库系统一切有关数据的处理都是通过这个软件实现的，根据数据模型不同，DBMS 也具有不同的类型。

1. DBMS 的功能

(1) 数据库的定义功能：数据定义包括定义构成数据库结构的模式、存储模式和外模式；定义各个外模式与模式之间的映射；定义模式与存储模式之间的映射；定义有关的约束条件；为保证数据库中数据具有正确语义而定义的完整性规则；为保证数据库安全而定义的用户口令和存取权限等。通过使用 DDL 实现。

(2) 数据库的操纵功能：数据操纵包括对数据库数据的检索、插入、修改和删除等基本操作。通过使用 DML 实现。

(3) 数据库的保护功能：有效地保护数据库资源，包括以下内容。

- 1) 数据库的恢复。
- 2) 数据库的并发控制。
- 3) 数据库的安全性控制。
- 4) 数据库的完整性控制。

(4) 数据库的存储管理：数据库中需要存放多种数据，如数据字典、用户数据、存取路径等，DBMS 负责分门别类地组织、存储和管理这些数据，确定以何种文件结构和存取方式物理地组织这些数据，如何实现数据之间的联系，以便提高存储空间利用率以及提高随机查找、顺序查找、增、删、改等操作的时间效率。

(5) 数据库的维护功能：供数据库管理员（DBA）维护数据库系统使用，主要包括以下程序。

- 1) 数据装载程序。
- 2) 数据备份程序。
- 3) 文件重组程序。
- 4) 性能监控程序。

(6) 数据字典（DD）：数据库系统中存放三级结构定义的数据库。数据字典是数据库系统的重要工具，数据字典与 DBMS 的密切配合，才能完成 DBMS 对数据库系统中数据资源的应用和管理。

DD 有两类：只能被用户和 DBA 访问而不能被 DBMS 软件访问的

笔 记