Access 2016
数据库应用教程



主　编◎李　军

北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

内 容 提 要

本书以 Access 2016 为背景，从数据库的基础理论开始，由浅入深、循序渐进，系统地介绍了 Access 2016 的主要功能和使用方法。本书通过大量的例题，讲解了 Access 数据库技术的相关知识以及使用 Access 2016 开发数据库应用系统的过程。本书主要内容包括数据库系统基础知识，Access 2016 的数据库、表、查询、窗体、报表、宏、VBA 等。为了使读者能够更好地掌握本书内容，本书配有丰富的例题；为了及时检验学习效果，每章后面均配有大量的习题并给出了参考答案。本书可与《Access 2016 数据库实践教程》配套使用。

本书采用“理论与实践相结合”的模式，将课堂教学内容与实验教学内容有机结合。本书例题从始至终用的都是“学生管理”数据库，每个例题都是编者 20 多年数据库教学中积累的材料。本书内容充实，突出操作性和实践性，实例丰富，可以使学生尽快掌握 Access 的基本功能和操作，掌握 Access 的编程功能和技巧。

本书既可以作为高等学校非计算机专业计算机公共基础课程的教材，也可以作为全国计算机等级考试——二级 Access 数据库程序设计的辅导与自学教材，还可以作为数据库开发人员的参考书。

版权专有　侵权必究

图书在版编目（CIP）数据

Access 2016 数据库应用教程 / 李军主编. --北京 : 北京理工大学出版社, 2023.2

ISBN 978-7-5763-2140-1

Ⅰ. ①A… Ⅱ. ①李… Ⅲ. ①关系数据库系统-高等学校-教材 Ⅳ. ①TP311.132.3

中国国家版本馆 CIP 数据核字（2023）第 034882 号

出版发行 / 北京理工大学出版社有限责任公司
社　　址 / 北京市海淀区中关村南大街 5 号
邮　　编 / 100081
电　　话 / （010）68914775（总编室）
（010）82562903（教材售后服务热线）
（010）68944723（其他图书服务热线）
网　　址 / http：//www.bitpress.com.cn
经　　销 / 全国各地新华书店
印　　刷 / 唐山富达印务有限公司
开　　本 / 787 毫米×1092 毫米　1/16
印　　张 / 19.25　　责任编辑 / 李　薇
字　　数 / 523 千字　　文案编辑 / 李　硕
版　　次 / 2023 年 2 月第 1 版　2023 年 2 月第 1 次印刷　　责任校对 / 刘亚男
定　　价 / 89.00 元　　责任印制 / 李志强

图书出现印装质量问题，请拨打售后服务热线，本社负责调换

前言

数据库应用技术是计算机应用的重要组成部分，尤其是随着大数据时代的到来，该课程不仅是计算机专业的必修课程，同时也是高等教育非计算机专业的必修课程之一，并成为高等学校非计算机专业继“大学计算机”课程之后的重要课程。Access 2016 是一个关系型数据库管理系统，是 Microsoft Office 2016 的组件之一，它可以有效地组织、管理和共享数据库中的数据，并将数据库与 Web 结合在一起。

本书在写作模式上以应用为目的，系统详细地介绍了数据库的基础理论知识，以及 Access 2016 数据库中的各种对象和 VBA 编程等内容。全书共分为 8 章，包括数据库系统概述、创建与管理数据库、创建数据表和关系、查询、窗体、报表、宏、VBA 编程基础。本书致力于将知识传授、能力培养和素质教育融为一体，实现理论教学与实践教学相结合，激发学生的创新意识，在部分章节加入了一些思政元素内容。

本书内容介绍通俗易懂、简明扼要，有利于教师的教学和读者的自学。为了让读者能够在较短的时间内掌握教材的内容，及时检验自己的学习效果，巩固和加深对所学知识的理解，每章后面均附有习题，可扫二维码查看相应的习题参考答案。

本书由经验丰富的一线教师编写完成，其中第 1、2、3、4、5、7、8 章及部分附录由李军编写，第 6 章由梁静毅编写，苏晓勤参与了部分附录的编写。本书由李军统稿，在编写过程中，王梦倩、李艳琴、张志鑫、杜利农、郭淑珍、王桂荣、王钢老师做了大量的工作并提供了宝贵的经验，在此一并表示感谢。另外还要感谢北京理工大学出版社编辑的悉心策划和指导。

由于编者水平有限，书中难免存在疏漏和不足之处，恳请读者批评指正，以便于本书的修改和完善。如有问题，可以通过 E-mail:tjlijun@ tjcu. edu. cn 与编者联系。

编　者

2022 年 10 月

目录

第 1 章　数据库系统概述 …… 001

1.1　数据库系统基础知识 …… 001
1.2　数据模型 …… 008
1.3　关系型数据库 …… 013
1.4　数据库设计基础 …… 017

第 2 章　创建与管理数据库 …… 024

2.1　Access 2016 系统概述 …… 024
2.2　Access 2016 的数据库 …… 027
2.3　Access 2016 的数据库对象 …… 030
2.4　在导航窗格中操作数据库对象 …… 031
2.5　管理数据库 …… 037

第 3 章　创建数据表和关系 …… 042

3.1　数据表 …… 042
3.2　创建表 …… 044
3.3　维护表 …… 067
3.4　表的基本操作 …… 076
3.5　创建表间关系 …… 083
3.6　数据表的导入和导出 …… 092

第 4 章　查询 …… 102

4.1　认识查询 …… 102
4.2　使用查询向导创建查询 …… 104
4.3　使用设计视图创建查询 …… 114
4.4　基本查询 …… 122
4.5　操作查询 …… 133
4.6　SQL 查询 …… 138

4.7 其他 SQL 语句 …… 145

第 5 章 窗体 …… 154

5.1 窗体概述 …… 154
5.2 创建窗体 …… 157
5.3 使用设计视图创建窗体 …… 163
5.4 主/子窗体设计 …… 191
5.5 窗体条件格式的使用 …… 196

第 6 章 报表 …… 201

6.1 报表概述 …… 201
6.2 创建报表 …… 208
6.3 报表设计 …… 218

第 7 章 宏 …… 228

7.1 宏概述 …… 228
7.2 创建宏 …… 231
7.3 宏的应用 …… 255
7.4 宏的运行与调试 …… 257

第 8 章 VBA 编程基础 …… 263

8.1 VBA 编程环境 …… 263
8.2 VBA 模块简介 …… 267
8.3 VBA 程序设计基础 …… 268
8.4 VBA 程序流程设计 …… 277
8.5 过程声明、调用与参数传递 …… 290
8.6 VBA 事件驱动机制 …… 295

参考文献 …… 301

第 1 章　数据库系统概述



本章学习目标：

- 掌握信息、数据、数据管理的基本概念
- 掌握数据模型及其特点
- 了解数据库系统的特点和发展过程
- 了解并掌握关系的定义及关系的性质

随着计算机科学的飞速发展，计算机已被广泛应用于社会的各个领域，计算机的广泛应用被认为是人类进入信息时代的标志。在信息时代，人们利用计算机对大量的数据进行加工处理。在处理过程中，用于复杂科学计算的工作较少，而大量的工作用于在相关的数据中提取信息。为了有效地使用存放在计算机系统中的大量数据，必须采用一整套科学的方法，对数据进行组织、存储、维护和使用，即数据处理。在数据处理过程中应用了数据库技术。

数据库系统产生于 20 世纪 70 年代初，它的出现，既促进了计算机技术的高速发展，又形成了专门的信息处理理论和数据库管理系统。因此，数据库系统是计算机技术和信息时代相结合的产物，是数据处理的核心，是研究数据共享的一门科学，是现代计算机系统软件的重要组成部分之一。

1.1　数据库系统基础知识

1.1.1　信息、数据和数据处理

要了解数据处理就要了解什么是信息、数据和数据处理。

1. 信息

信息（Information）是对客观事物属性的反映。它所反映的是客观事物的某一属性或某一时刻的表现形式，如成绩的好坏、温度的高低、质量的优劣等。因此，信息是经过加工处理并对人类客观行为和决策产生影响的数据表现形式。

信息有以下 4 个特征。

（1）信息是可以感知的。人类对客观事物的感知，可以通过感觉器官，也可以借助于各种仪器设备。不同的信息源有不同的感知形式，如书上的信息可以通过视觉器官感知，广播中的信息可以通过听觉器官感知。

（2）信息是可以存储、传递、加工和再生的。人类可以利用大脑记忆信息，可以利用语言、文字、图像和符号等记载信息，可以借助纸张、各种存储设备长期保存信息，可以利用电视、广播和网络传播信息，可以对信息进行加工、处理后得到其他信息。

（3）信息源于物质和能量。信息不能脱离物质而存在，信息的传递需要物质载体，信息的获取和传递需要消耗能量。没有物质载体，信息就不能存储和传递。

（4）信息是有用的。它是人们活动所必需的知识，利用信息能够克服工作中的盲目性，增加主动性和科学性，利用有用的信息，人们可以科学地处理事情。

2. 数据

数据（Data）是信息的具体表现形式，是反映客观事物属性的符号化表示和记录，如年龄“20”岁，分数“98”分，出生日期“1999 年 10 月 01 日”等。数据所反映的事物属性是它的内容，而符号是它的表现形式。

数据不仅包括数字、字母、文字和其他特殊符号组成的文本形式数据，而且还包括图形、图像、动画、影像、声音等多媒体数据。从计算机角度看，数据泛指那些可以被计算机接受并处理的符号。

3. 信息和数据的关系

信息和数据既有联系，又有区别。数据是信息的载体，信息是数据的内涵。数据是物理性的，是被加工的对象，而信息是对数据加工的结果，是观念性的，并依赖于数据而存在。数据表示了信息，而信息只有通过数据形式表现出来，才能被人们理解和接受。信息是有用的数据，数据如果不具有知识性和有用性，则不能称为信息。

从某种意义上讲，数据就是信息，信息就是数据，二者在一定的条件下，可以相互转换。

4. 数据处理

数据处理（Data Process）也称为信息处理，是指利用计算机对各种类型的数据进行采集、整理、存储、分类、排序、检索、维护、加工、统计和传输等操作，使之变为有用信息的一系列活动的总称。就是从某些已知的数据出发，推导加工出一些新的数据，这些新的数据又表示了新的信息。因此，数据处理也称为信息处理。信息处理的真正含义是为了产生信息而处理数据。数据处理是数据升华的过程。

1.1.2 数据库技术的产生与发展

随着计算机技术，特别是数据库技术的产生与发展，数据处理过程也发生了巨大的变化，其核心就是数据管理。数据管理指的是对数据进行分类、组织、编码、存储、检索和维护等。数据处理和数据管理是相互联系的，数据管理技术的优劣，将直接影响数据处理的效率。

数据管理技术的发展经历了人工管理、文件系统、数据库 3 个阶段。

1. 人工管理阶段

在人工管理阶段（20 世纪 50 年代中期以前），计算机主要用于科学计算。外部存储器只有磁带、卡片和纸带，软件只有汇编语言，尚无数据管理方面的软件。数据处理的方式基本上是批处理。这个时期数据管理的特点如下。

（1）数据不保存。因为当时计算机主要用于科学计算，对于数据保存的需求尚不迫切。需要时把数据输入内存，运算后将结果输出，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。在应用程序中，不仅要管理数据的逻辑结构，还要设计其物理结构、存取方法、输入/输出方法等。当存储改变时，应用程序中存取数据的子程序就需随之改变。

（3）数据不具有独立性。数据的独立性是指逻辑独立性和物理独立性。当数据的类型、格式或输入/输出方式等逻辑结构或物理结构发生变化时，必须对应用程序作出相应的修改。

（4）数据是面向程序的。一组数据只对应于一个应用程序，即使两个应用程序都涉及某些相同数据，也必须各自定义，无法相互利用。因此，在程序之间有大量的冗余数据。

在人工管理阶段，数据与程序的关系如图 1.1 所示。

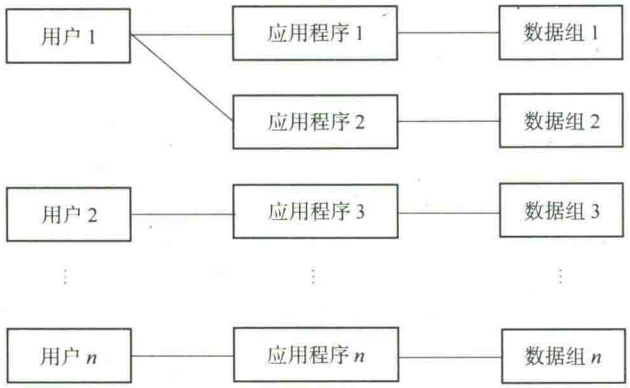


图 1.1 人工管理阶段数据与程序的关系

2. 文件系统阶段

在文件系统阶段（20 世纪 50 年代后期到 60 年代中期），计算机不仅用于科学计算，还用于信息管理。此时，外部存储器已有磁盘、磁鼓等直接存取的存储设备，软件领域出现了高级语言和操作系统。操作系统中的文件系统是专门的数据管理软件。这时可以把相关的数据组成一个文件存放在计算机中，在需要时只要提供文件名，计算机就能从文件系统中找出所要的文件，把文件中存储的数据提供给用户进行处理。

1）特点

（1）数据以文件形式可长期保存在外部存储器的磁盘上。应用程序可对文件进行大量的检索、修改、插入和删除等操作。

（2）文件组织已多样化。计算机中文件有索引文件、顺序存取文件和直接存取文件等，因而对文件中的记录可顺序访问，也可随机访问，便于存储和查找数据。

（3）数据与程序间有一定的独立性。数据由专门的软件即文件系统进行管理，程序和数据间由软件提供的存取方法进行转换，数据存储发生变化不一定影响程序的运行。

（4）对数据的操作以记录为单位。这是由于文件中只存储数据，不存储文件记录的结构描述信息。文件的建立、存取、查询、插入、删除、修改等所有操作，都要用程序来实现。

2）存在的问题

在文件系统阶段，用户虽有了一定的方便，但仍存在以下 3 个主要问题。

（1）数据冗余度大。由于各数据文件之间缺乏有机的联系，造成每个应用程序都有对应的文件，所以有可能同样的数据在多个文件中被大量地重复存储，数据不能共享。

（2）数据独立性低。数据和程序相互依赖，一旦改变数据的逻辑结构，必须修改相应的应用程序。而应用程序发生变化，如改用另一种程序设计语言来编写程序，也必须修改数据结构。

（3）数据一致性差。由于相同数据的重复存储、各自管理，因此在进行更新操作时，容易造成数据的不一致。

这样，文件系统仍然是一个不具有弹性的无结构的数据集合。文件之间是孤立的，不能反映现实世界中事物之间的内在联系。在文件系统阶段，数据与程序的关系如图 1.2 所示。

3. 数据库阶段

数据管理技术进入数据库阶段是在 20 世纪 60 年代末。由于计算机应用于管理的规模更加庞大，数据量急剧增加；硬件方面出现了大容量磁盘，使计算机联机存取海量数据成为可能；硬件价格下降，而软件价格上升，使开发和维护系统软件的成本增加。因此，文件系统的数据管理方法已无法适应开发应用系统的需要，为满足多用户、多个应用程序共享数据的需求，出现了统一管理数据的专门软件系统，即数据库管理系统，这使利用数据库技术管理数据变成了现实。

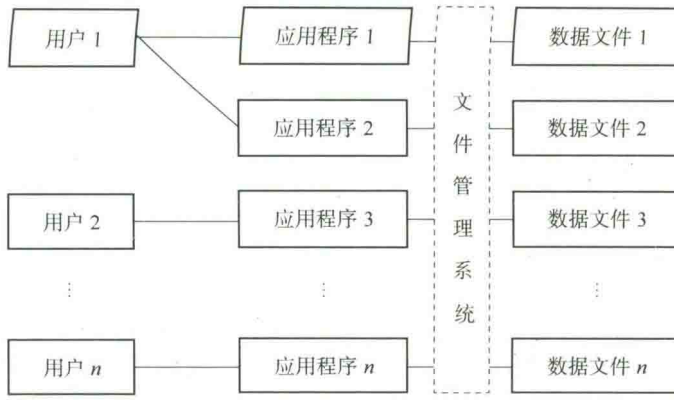


图 1.2 文件系统阶段数据与程序的关系

数据库有以下 4 个特点。

(1) 数据共享性高、冗余度低。这是数据库阶段的最大改进，数据不再面向某个应用程序而是面向整个系统，当前所有用户可同时访问数据库中的数据。这样便减少了不必要的数据冗余，节约了存储空间，同时也避免了数据之间的不相容性与不一致性。

(2) 数据结构化。数据结构化是指按照某种数据模型，将应用的各种数据组织到一个结构化的数据库中。在数据库中数据的结构化，不仅要考虑某个应用的数据结构，还要考虑整个系统的数据结构，并且要能够表示出数据之间的有机关联。

(3) 数据独立性高。数据的独立性是指数据的逻辑独立性和物理独立性。

数据的逻辑独立性是指当数据的总体逻辑结构改变时，数据的局部逻辑结构不变。由于应用程序是依据数据的局部逻辑结构编写的，因此应用程序不必修改，从而保证了数据与程序间的逻辑独立性。

数据的物理独立性是指当数据的存储结构改变时，数据的逻辑结构不变，从而应用程序也不必改变。

表 1.1 列出了数据库系统与一般文件应用系统的主要性能差异，通过该表可看出数据库系统的特点。

表 1.1 数据库系统与一般文件应用系统的主要性能差异

序号	文件应用系统	数据库系统
1	文件中的数据由特定的用户专用	数据库内的数据由多个用户共享
2	每个用户拥有自己的数据，导致数据重复存储	原则上可消除重复。为方便查询，允许少量数据重复存储，但冗余度可以控制
3	数据从属于程序，二者相互依赖	数据独立于程序，强调数据的独立性
4	各数据文件彼此独立，从整体看是“无结构”的	各文件的数据相互联系，从整体看是“有结构”的

(4) 有统一的数据控制功能。数据库为多个用户和应用程序所共享，对数据的存取往往是并发的，即多个用户可以同时存取数据库中的数据，甚至可以同时存取数据库中的同一个数据。为确保数据库数据的正确有效和数据库系统的有效运行，数据库管理系统提供下述 4 个方面的数据控制功能。

①数据的安全性控制。数据的安全性控制是指防止不合法使用数据造成数据的泄露和破坏，保证数据的安全和机密。例如，系统提供口令检查或其他手段来验证用户身份，防止非法用户使

用系统；也可以对数据的存取权限进行限制，只有通过检查后才能执行相应的操作。

②数据的完整性控制。数据的完整性控制是指系统通过设置一些完整性规则以确保数据的正确性、有效性和相容性。正确性是指数据的合法性，如年龄属于数值型数据，只能包含0，1，…，9这些阿拉伯数字，不能包含字母或特殊符号。有效性是指数据是否在其定义的有效范围内，如月份只能用1~12之间的正整数表示。相容性是指表示同一事实的两个数据应相同，否则就不相容，如一个人不能有两个性别。

③并发控制。并发控制是指防止多用户同时存取或修改数据库时，因相互干扰而提供给用户不正确的数据，并使数据库受到破坏。

④数据恢复。当数据库被破坏或数据不可靠时，系统有能力将数据库从错误状态恢复到最近某一时刻的正确状态。

在数据库阶段，程序与数据之间的关系如图1.3所示。

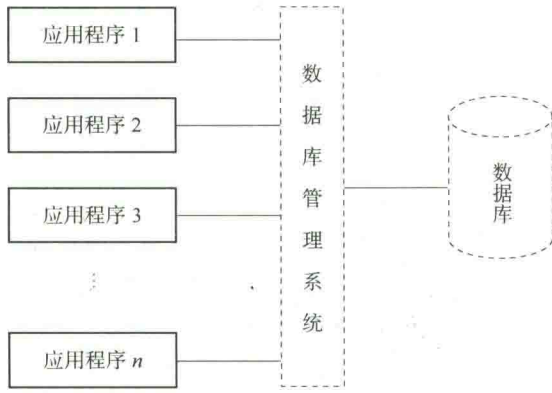


图1.3 数据库阶段程序与数据之间的关系

1.1.3 数据库系统的组成

1. 有关数据库系统的几个基本概念

1）数据库（DataBase，DB）

数据库是存储在计算机存储设备上的、结构化的相关数据的集合，这些数据被数据库管理系统按一定的组织形式存储在各个数据文件中。数据库中的数据具有较小的冗余度、较高的数据独立性和易扩展性，具有完善的自我保护能力和数据恢复能力，并能够提供数据共享。

2）数据库系统（DataBase System，DBS）

数据库系统是指引入数据库后的计算机系统。它主要由五部分组成：硬件系统、数据库、数据库管理系统、数据库应用系统、用户，如图1.4所示。

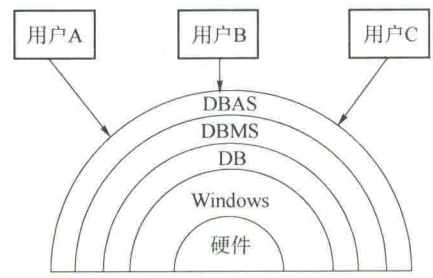


图1.4 数据库系统组成

3）数据库管理系统（DataBase Management System，DBMS）

数据库管理系统是数据库系统中对数据进行管理的软件系统，位于用户与操作系统之间。数据库管理系统可以对数据库的建立、使用和维护进行管理，可以使数据库中的数据具有最小的冗余度，并对数据库中的数据提供安全性和完整性等统一控制机制，方便用户以交互命令方式或程序方式对数据库进行操作。

DBMS是数据库系统的核心组成部分，用户对数据库的定义、查询、更新等各种操作都是通过DBMS进行的。

4）数据库应用系统（DataBase Application System，DBAS）

数据库应用系统是指系统开发人员利用数据库系统资源开发出来的、面向某一类实际应用问题的应用软件系统，如以数据库为基础的教学管理系统、财务管理系统、图书管理系统等。一个数据库应用系统通常由数据库和应用程序组成，是在数据库管理系统支持下设计和开发出来的。

5）用户（Users）

用户是指使用和管理数据库的人，他们可以对数据库进行存储、维护和检索等操作。数据库系统中用户可分为以下 3 类。

（1）终端用户。终端用户主要是指使用数据库的各级管理人员、工程技术人员等，一般来说，他们是非计算机专业人员。

（2）应用程序员。应用程序员负责为终端用户设计和编制应用程序，以便终端用户对数据库进行操作。

（3）数据库管理员。数据库管理员是指对数据库进行设计、维护和管理的专门人员。

2. 数据库系统的发展

经过三十余年的发展，数据库系统已走过了第一代（格式化数据库系统）、第二代（关系数据库系统），现正向第三代（对象-关系数据库系统）迈进。

1）格式化数据库系统

格式化数据库系统是对第一代数据库系统的总称，其中又包括层次型数据库系统与网状型数据库系统两种类型，这一代数据库系统具有以下特征。

（1）采用“记录”为基本的数据结构。在不同的“记录型”（Record Type）之间，允许存在相互联系。

“层次模型”（Hierarchical Model）的总体结构为树形，在不同记录型之间只允许存在单线联系，如图 1.8 所示；“网状模型”（Network Model）的总体结构为网状结构，在两个记录型之间允许存在两种或多于两种的联系，如图 1.9 所示。前者适用于管理具有家族系统结构的数据库，后者则更适用于管理在数据之间具有复杂联系的数据库。

（2）无论是层次模型还是网状模型，一次查询只能访问数据库中的一个记录，存取效率不高。对于具有复杂联系的系统，用户查询时还需详细描述数据的访问路径（存取路径），操作也比较麻烦。因此，自关系数据库系统兴起后，格式化数据库系统已逐渐被关系数据库系统所取代，目前仅在一些大中型计算机系统中使用。

2）关系数据库系统（Relational DataBase System，RDBS）

早在 1970 年，IBM 公司 San Jose 研究实验室的研究员科德（E. F. Codd）就在一篇论文中提出了“关系模型”（Relational Model）的概念，从而开创了关系数据库理论的研究。

20 世纪 70 年代中期，国外已有商品化的 RDBS 问世，数据库系统随之进入了第二代。80 年代后，RDBS 在包括 PC 在内的各类型计算机上实现，目前在 PC 上使用的数据库系统主要是 RDBS。

与第一代数据库系统相比，RDBS 具有以下优点。

（1）采用人们习惯使用的表格作为基本的数据结构，通过公共的关键字段来实现不同二维表之间（或“关系”之间）的数据联系。关系模型呈二维表形式，简单明了，使用与学习都很方便。

（2）一次查询仅用一条命令或语句，即可访问整个“关系”（或二维表），因而查询效率较高，不像第一代数据库系统那样每次仅能访问一个记录。在 RDBS 中，通过多表联合操作，还能对有联系的若干二维表实现“关联”查询。

3）对象-关系数据库系统（Object-Relational DataBase System，ORDBS）

RDBS 管理的信息，可包括字符型、数值型、日期型等多种类型，但本质上都属于单一的文

本信息。多媒体应用的扩大，对数据库提出了新的要求，希望数据库系统能存储图形、声音等复杂的对象，并能实现复杂对象的复杂行为。将数据库技术与面向对象技术相结合，便顺理成章地成为研究数据库技术的新方向，构成了第三代数据库系统的基础。

20 世纪 80 年代中期以来，人们对于面向对象的数据库系统（Object-Oriented DataBase System，OODBS）的研究十分活跃。1989 年和 1990 年，《面向对象数据库系统宣言》和《第三代数据库系统宣言》相继发表，后者主要介绍对象-关系数据库系统。一批代表新一代数据库系统的软件产品也陆续推出。由于 ORDBS 是建立在 RDBS 技术之上的，可以直接继承 RDBS 的原有技术和用户基础，因此其发展比 OODBS 更为顺利，正在成为第三代数据库系统的主流。

3. 数据库系统的分类

1987 年，著名的美国数据库专家厄尔曼（J. D. Ullman）教授在一篇题为《数据库理论的过去和未来》的论文中，曾把数据库理论概括为 4 个分支：关系数据库理论、面向对象数据库理论、分布式数据库理论、演绎数据库理论。今天，关系数据库已经得到广泛的应用，并成为当今数据库系统的主流。其余 3 个分支，在过去十余年间也取得了不小的进展，并在理论研究的基础上开发出各种实用的数据库系统，下面对其进行详细介绍。

1）面向对象数据库

数据库的发展时代是根据所采用的数据模型划分的。这里的数据模型，首先是指把数据组织起来所采用的数据结构，同时也包含数据操作和数据完整性约束等要素。与第一代数据库常见的层次模型和网状模型相比，关系模型不仅简单易用，理论也比较成熟，但如果用它来存储和检索包括图形、文本、声音、图像在内的多媒体数据，就显得不太方便了。因此，当面向对象技术兴起后，人们就探索用对象模型来组织多媒体数据库，推动并促进了第三代数据库的诞生。

多媒体数据库是面向对象数据库的重要实例，它管理的数据不仅容量大，而且长短不一，检索方法也从传统数据库的“精确查询”，改变为以“非精确匹配和相似查询”为主的“基于内容”的检索。20 世纪 90 年代，一些著名的第二代数据库如 Oracle、Sybase 等都在原来关系模型的基础上引入了对象机制，扩展了对多媒体数据的管理功能。

2）分布式数据库

如果说多媒体应用促进了面向对象数据库的发展，而网络的应用与普及，则推动了分布式数据库的发展。在早期的数据库中，数据都是集中存放的，即所谓的集中式数据库。分布式数据库则把数据分散地存储在网络的多个节点上，彼此用通信线路连接。例如，一家银行有众多储户，如果他们的数据集中存放在一个数据库中，则所有的储户在存、取款时都要访问这个数据库，网络通信量必然很大；若改用分布式数据库，将储户的数据分散地存储在离各自住所最近的储蓄所，则大多数时候数据可就近存取，仅有少数时候数据需远程调用，从而大大减少了网络上的数据传输量。现在在 Internet/Intranet 上流行的 Web 数据库，就是分布式数据库的实例，它使全城（市）的储户通过同一家银行的任何一个储蓄所，都能够实现通存通兑。

分布式数据库也是多用户数据库，可供多个用户同时在网络上使用。但多用户数据库并非总是分布式存储的。以飞机订票系统为例，它允许乘客在多个售票点进行订票，但同一家航空公司的售票数据通常是集中存放的，而不是分散存放在各个售票点上。

3）演绎数据库

传统数据库存储的数据都代表已知的事实，演绎数据库则除存储事实外，还能存储用于逻辑推理的规则。例如，某演绎数据库存储有“校长领导院长”的规则。如果该数据库中同时存有“甲是校长”“乙是院长”等数据，那么它就能推理得出“甲领导乙”的新事实。

由于这类数据库是由“事实+规则”所构成的，因此有时也称为基于规则的数据库或逻辑数据库。它所采用的数据模型则称为逻辑模型或基于逻辑的数据模型。随着人工智能不断走向实用化，对演绎数据库的研究也日趋活跃。演绎数据库与专家系统和知识库一起被称为智能数据库。其关键是逻辑推理，如果推理模式出现问题，便可能导致出现荒诞的结果。

1.2 数据模型

数据模型是工具，是用来抽象表示和处理现实世界中的数据和信息的工具。数据模型应满足三方面的要求：一是能够比较真实地模拟现实世界；二是容易被人理解；三是便于在计算机系统中实现。

1.2.1 基本概念

数据模型是客观事物及其联系的数据描述，它应具有描述数据和数据联系两方面的功能。数据模型由数据结构、数据操作和数据的约束条件三部分组成。其中数据结构是所研究对象类型的集合，是对系统静态特征的描述；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；数据的约束条件是一组完整性规则的集合。所谓完整性规则是指数据模型中数据及其联系所具有的制约和依存规则，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、相容。

数据模型是模型的一种，它是现实世界数据特征的抽象，是数据和数据之间相互联系的形式，是数据和信息的处理工具，现实世界中的具体事务必须用数据模型这个工具来抽象和表示。

根据模型应用的目的不同，可以将这些模型划分为两类，它们分属于两个不同的层次。一类模型是概念模型，它是按用户的观点对数据和信息建模，是用户和数据库设计人员之间进行交流的工具，主要用于数据库设计，该类模型中最主要的是实体关系模型。另一类模型是数据模型，主要包括网状模型、层次模型和关系模型等，它是按计算机系统的观点对数据建模，主要用于 DBMS 的实现。

数据模型是数据库系统的核心和基础，DBMS 软件都是基于某种数据模型的。

1.2.2 概念模型与 E-R 图

在数据库技术中，用数据的概念模型描述数据库的结构和语义，表示实体及实体之间的联系。概念模型是对客观事物及其联系的一种抽象的描述。

1. 几个相关的基本概念

1）实体（Entity）

客观存在并且可以相互区别的事物称为实体。实体可以是具体的事物，如一名教师、一门课程、一本教材；也可以是抽象的事件，如一次授课、一场考试等。

2）属性（Attribute）

实体所拥有的各类性质称为属性。实体有很多性质，如教师实体有教师号、姓名、性别、出生日期、学历等方面的属性。

属性有“型”和“值”之分。“型”即为属性名，如教师号、姓名、性别、出生日期是属性的型；“值”即为属性的具体内容，如“20180936”“张大鹏”“男”“1999 年 10 月 01 日”

“大学本科”“航天学院”等，这些属性值的集合表示了一名教师实体。

3）实体集（Entity Set）和实体型（Entity Type）

具有相同类型及相同性质的实体的集合称为实体集。而属性的集合表示一种实体的类型，称为实体型。例如，学生（学号，姓名，性别，年龄，政治面貌，简历，照片）就是一个实体型，对于学生来说，全体学生就是一个实体集。

4）联系（Connection）

实体之间的对应关系称为联系，它反映现实世界事物之间的相互联系。实体内部的联系通常是指组成实体的各属性之间的联系，实体之间的联系通常是指不同实体集之间的联系。

2. 实体间的联系

两个实体集（设 A，B）的联系可分为3种类型，即一对一的联系、一对多的联系和多对多的联系。

1）一对一的联系（1∶1）

实体集 A 中的一个实体只能与实体集 B 中的一个实体相对应，反之亦然，则称实体集 A 与实体集 B 为一对一的联系，记为1∶1。例如，一个学校有一名校长，校长和学校之间存在一对一的联系。

2）一对多的联系（1∶n）

实体集 A 中的一个实体与实体集 B 中的多个实体相对应，反之，实体集 B 中的一个实体只能与实体集 A 中的一个实体相对应，记为1∶n。例如，班级与学生两个实体集之间存在一对多的联系，一个班有多名学生，一名学生只能属于一个班，班级和学生之间存在一对多的联系。

3）多对多的联系（m∶n）

实体集 A 中的一个实体与实体集 B 中的多个实体相对应，反之，实体集 B 中的一个实体与实体集 A 中的多个实体相对应，记为 m∶n。例如，学生与课程两个实体集之间存在多对多的联系，因为一名学生可以选修多门课程，而一门课程又可以被多名学生所选修，所以学生和课程之间存在多对多的联系。

一对多的联系是最普遍的联系，我们可以把一对一的联系看作一对多联系的一个特例。

3. E-R 模型与 E-R 图

E-R 模型（Entity Relationship Model）是人们描述数据及其联系的概念数据模型，是数据库应用系统设计人员和普通非计算机专业用户进行数据建模和沟通交流的有力工具，使用起来非常直观易懂、简单易行。进行数据库应用系统设计时，首先要根据用户需求建立合乎需要的E-R模型，然后建立与计算机数据库管理系统相适应的逻辑数据模型和物理数据模型，最后才能在计算机系统上安装、调试和运行数据库。

1）E-R 模型中的基本构件符号

E-R 模型是一种用图形表示数据及其联系的方法，所使用的图形构件（元件）包括矩形、菱形、椭圆形和连接线。

矩形表示实体，矩形框内写上实体名。

菱形表示联系，菱形框内写上联系名。

椭圆形表示属性，椭圆形框内写上属性名。

连接线表示实体、联系与属性之间的所属关系或实体与联系之间的相连关系。

2）各种联系的 E-R 图表示

对于一对一、一对多和多对多3种实体联系，可以用 E-R 图来表示，如图1.5所示。

若每种联系来自同一个实体，则它们的 E-R 图如图1.6所示。

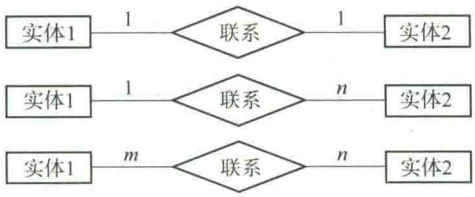


图 1.5　3 种联系的 E-R 图

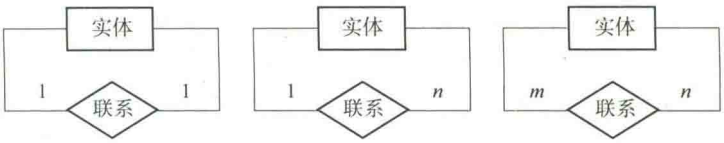


图 1.6　3 种联系的单实体的 E-R 图

两个实体的联系是基本联系，在现实世界中经常出现 3 个或更多的实体相互联系的情况，有时 3 个实体之间，两两存在着不同类型的联系。例如，学生、课程和教师这 3 个实体，多名学生可以选多门课程，而每一门课程又可以被多名学生选，每门课程唯一对应一名教师，一名教师又可以教授多门课程，这样学生与课程之间就是多对多的联系，课程和教师之间就是多对一的关系，学生和教师之间无须直接给出，它可以从两个联系中推导出来，即学生和教师是多对多的关系。学生、课程和教师这 3 个实体之间的联系所对应的 E-R 图如图 1.7 所示。

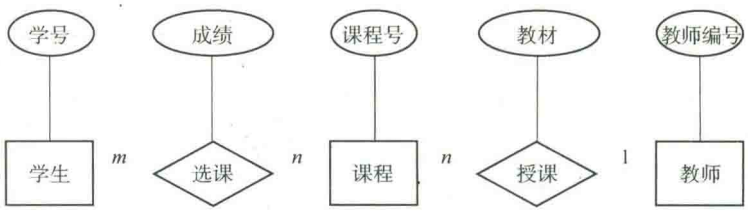


图 1.7　学生、课程和教师这 3 个实体之间的联系所对应的 E-R 图

在图 1.7 中，每一个实体和联系上，只给出了一个代表属性，其他属性没有给出。“选课”联系的属性“成绩”表示某名学生选修某门课程的考试成绩，“授课”联系的属性“教材”表示教师教授某门课程时所选用的教材。若要找出某名学生所选课程的任课教师，则首先通过选课联系查出相应的课程，再通过授课联系找出对应的任课教师即可。

在实际的 E-R 图设计中，除了设计各实体和联系，还要确定每个实体和联系所包含的属性，不必是全部属性，但要取其相关和必要的属性。任何一个应用系统设计的 E-R 图都不是唯一的，与对应的设计思路和设计方法及对系统的分析程度相关联。合理的、符合实际和贴近运营管理要求的 E-R 图，对于应用系统的使用者和应用系统的设计开发者都是有好处的。

1.2.3　3 种数据模型

数据模型是数据库系统的基础，任何一个数据库管理系统都是基于某种数据模型的。数据库管理系统所支持的传统数据模型分为 3 种：层次模型、网状模型和关系模型。

3 种数据模型之间的根本区别在于实体之间联系的表示方式不同。层次模型用树形结构来表示实体之间的联系；网状模型用网状结构来表示实体之间的联系；关系模型用二维表（或称关系）来表示实体之间的联系。

1. 层次模型

层次模型表示数据间的从属关系结构，是一种以记录某一事物的类型为根节点的有向树结构。层次模型像一棵倒立的树，根节点在上，层次最高，子节点在下，逐层逐级排列。上级节点与下级节点之间为一对多的联系。图 1.8 给出了一个层次模型的例子，其中，“工业大学”为根节点，“工业大学”以下为各级子节点。

层次模型具有以下特征：

①有且仅有一个根节点而且无双亲；

②根节点以下的子节点，向上层仅有一个父节点，向下层有若干子节点；

③最下层为叶节点且无子节点。

支持层次模型的数据库管理系统称为层次数据库管理系统，其中的数据库称为层次数据库。

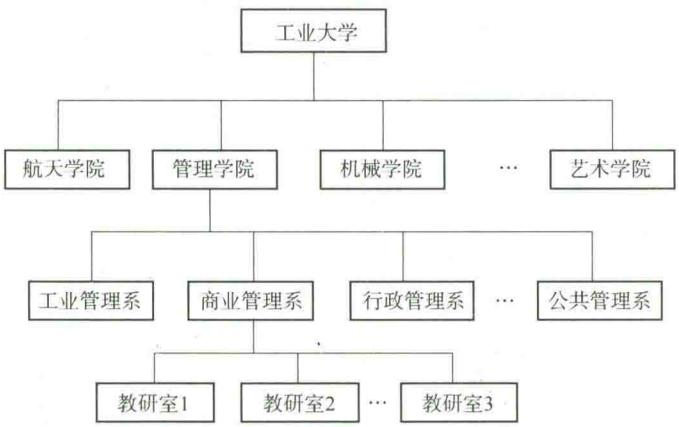


图 1.8 层次模型示意

2. 网状模型

现实世界中事物之间的联系更多是非层次关系的，用层次模型表示这种关系很不直观，网状模型克服了这一弊病，可以清晰地表示非层次关系。

网状模型是用网状结构表示实体与实体之间联系的模型。网状模型是层次模型的扩展，它表示多个从属关系的层次结构，可以允许两个节点之间有多种联系。网状模型表现为一种交叉关系的网络结构。

网状模型具有以下特征：

①有一个以上的节点无双亲；

②至少有一个节点有多双亲。

网状模型可以表示较复杂的数据结构，它不但可以表示数据间的纵向关系，而且可以表示数据间的横向关系。

网状模型中每个节点表示一个记录（实体），每个记录可包含若干个字段（实体的属性），节点间的连线表示记录（实体）间的父子关系。

网状模型的典型代表是数据库任务组（DataBase Task Group，DBTG）系统，也称 CODASYL 系统，它是 20 世纪 70 年代数据系统语言协会（Conference on Data System Language，CODASYL）下属的数据库任务组提出的一个系统方案。图 1.9 给出了一个抽象的简单的网状模型。

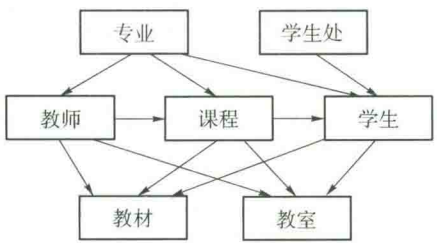


图 1.9 网状模型示意

3. 关系模型

美国 IBM 公司的研究员 E. F. Codd 于 1970 年发表了题为《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关系模型。20 世纪 80 年代以来，计算机厂商推出的数据库管理系统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。IBM 公司的 DB2 和微软公司的 SQL Server 数据库是关系数据库的代表。

用二维表结构来表示实体与实体之间联系的模型称为关系模型，由 IBM 公司的 E. F. Codd 于 1970 年首次提出，其特点是，理论基础完备、模型简单、具有说明性的查询语言和使用方便。

关系模型是发展较晚，也是最常用、最重要的一种数据模型。

在关系模型中，操作的对象和结果都是二维表，这种二维表就是关系。关系模型的主要特征是用二维表来表示实体集。例如，表 1.2 的“学生”表就是一个关系。

表 1.2 “学生”表

学号	姓名	性别	民族	出生日期	政治面貌	专业	班级	兴趣爱好
20220101	王琳	女	藏族	2003/1/30	团员	经济	2205	游泳，旅游
20220203	赵正	男	壮族	2003/7/9	中共党员	统计	2206	游泳，摄影
20220305	陈瑞	女	汉族	2002/12/3	团员	财务	2206	游泳，电影
20220407	崔婷	女	白族	2001/2/5	团员	软件	2203	电影，旅游
20220509	马福良	男	汉族	2002/8/24	中共党员	物流	2202	电影，体育
20220610	徐舒怡	女	汉族	2001/9/3	团员	经济	2201	电影，旅游
20220112	蔡泓	男	回族	2003/1/10	团员	统计	2205	电影，网球
20220214	张楠	女	汉族	2002/7/19	团员	财务	2206	摄影，旅游
20220316	冯佳	女	回族	2001/1/24	中共党员	经济	2027	看书，唱歌
20220418	赵阳	男	藏族	2003/12/9	团员	软件	2208	游泳，电影
20220506	兰云	女	回族	2003/7/9	团员	财务	2209	电影，体育
20220606	吴艳	女	汉族	2001/5/26	团员	软件	2202	摄影，唱歌
20220700	张悦	女	汉族	2002/12/19	团员	经济	2205	游泳，体育
20200205	李一博	男	汉族	2002/11/26	团员	软件	2201	摄影，旅游
20200208	施正	男	回族	2001/9/10	团员	财务	2203	电影，体育
20201201	胡龙	男	白族	2002/11/3	团员	统计	2206	电影，旅游

（1）二维表的特点如下。

①表有表名。

②表由两部分构成，即一个表头和若干行数据。

③从垂直方向看，表由若干列组成，每列都有列名，如“学号”“姓名”等。

④同一列的值取自同一个定义域。例如，“性别”的定义域是（男，女）。

⑤每一行的数据代表一名学生的信息，同样每一名学生在表中也有一行。

（2）对一个二维表可以进行以下操作。

①填表：将每名同学的数据填写进表格。

②修改：改正表中的错误数据。

③删除：去掉一名学生的数据（如：某名同学已退学或出国等）。

④查询：在表中按某些条件查找满足条件的学生。