

“三五”普通高等教育规划教材



北京高等教育精品教材
BEIJING GAODENG JIAOYU JINGPIN JIAOCAI

数据库原理及应用 (Access版)

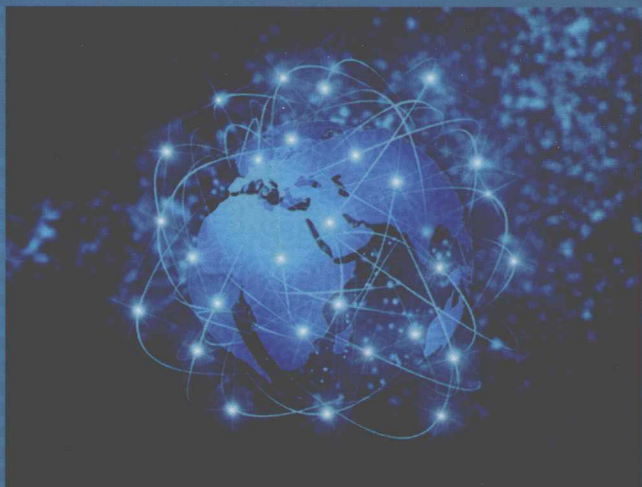
第4版

吴靖 主编
唐小毅 马燕林 金鑫 参编



提供电子教案

<http://www.cmpedu.com>



机械工业出版社
CHINA MACHINE PRESS

“十二五”普通高等教育规划教材

数据库原理及应用（Access版）

第4版

吴靖 主编

唐小毅 马燕林 金鑫 参编



机械工业出版社

本书从一个 Access 数据库应用系统实例——学生管理系统入手，系统地介绍数据库的基本原理与 Access 各种主要功能的使用方法，主要包括数据库的基本原理和相关概念、关系数据库的基本设计方法、数据库的建立、数据表、SQL 语言、查询、窗体、宏、报表、VBA 程序设计和数据库编程技术。

本书内容全面，结构完整，深入浅出，图文并茂，通俗易懂，可读性、可操作性强，既可以作为各类高校学生学习数据库原理及应用的教材，也可作为相关领域技术人员的参考用书或培训教材。

本书配有电子教案，需要的教师可登录 www.cmpedu.com 免费注册，审核通过后下载，或联系编辑索取（QQ：2966938356，电话：010-88379739）。

图书在版编目（CIP）数据

数据库原理及应用：Access 版 / 吴靖主编. —4 版. —北京：机械工业出版社，2019.2

“十三五”普通高等教育规划教材

ISBN 978-7-111-61792-1

I. ①数… II. ①吴… III. ①关系数据库系统—高等学校—教材
IV. ①TP311.132.3

中国版本图书馆 CIP 数据核字（2019）第 007860 号

机械工业出版社（北京市百万庄大街 22 号 邮政编码 100037）

策划编辑：和庆娣 责任编辑：胡静 和庆娣

责任校对：张艳霞 责任印制：孙炜

保定市中国画美凯印刷有限公司印刷

2019 年 2 月第 4 版·第 1 次印刷

184mm×260mm·17.25 印张·421 千字

0001—3000 册

标准书号：ISBN 978-7-111-61792-1

定价：52.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

电话服务

网络服务

服务咨询热线：（010）88379833

机工官网：www.cmpbook.com

读者购书热线：（010）88379649

机工官博：weibo.com/cmp1952

教育服务网：www.cmpedu.com

封面防伪标均为盗版

金书网：www.golden-book.com

前 言

随着社会信息化进程的不断推进和电子商务的蓬勃发展, 对非计算机类专业学生掌握一定计算机科学知识的要求也越来越高。在参与企业信息化建设的咨询和实施过程中, 我们深刻地体会到数据建模是企业信息化过程中的一个重要环节, 数据组织得好坏关系到企业信息化的成败。而企业信息化又与各专业领域相联系, 需要各专业的密切合作。因此, 需要非计算机类专业学生具有数据建模的基本知识, 掌握数据库的设计步骤和原则, 使学生具有解决实际问题的能力。

本书适用于非计算机类专业学生学习“数据库原理与应用”课程, 面向各类专业学生讲授数据库系统最基本的内容——数据库设计和数据库编程。通过对该课程的学习, 学生能够具备数据库方面的基本知识和良好的逻辑思维能力。本书在第3版的基础上, 根据 Access 2016 的特点和功能进行了修改。

本书的编著者为中央财经大学“Access 数据库原理及应用”课程的主讲教师, 在长期的教学中积累了丰富的教学经验。本书以数据库应用为重点, 以学生管理为主线, 主要讲解如何建立数据库、建立数据库应遵循的步骤以及每一步骤中应遵循的原则, 并以实例介绍在 Access 环境中如何实现数据库的建立和各种对象的创建等。

本书有配套的习题与实验教材——《数据库原理及应用实验指导及习题 (Access 版) (第2版)》, 以一个商品进销系统为例, 提供了丰富的练习题和题解, 对本书做了强有力的补充, 是教师教学和学生学习本书的重要资源。

在本书的出版过程中, 得到了各有关部门同事的大力支持和帮助, 在此表示衷心的感谢。参加本书编写的有吴靖、唐小毅、马燕林、金鑫。

由于计算机技术日新月异, 加之时间仓促及编者能力所限, 书中难免有疏漏和不足之处, 恳请读者批评指正。

编 者

目 录

前言

第 1 章 数据库系统概述 1	2.5.1 集合运算..... 25
1.1 引言..... 1	2.5.2 关系运算..... 26
1.2 数据库系统..... 2	2.6 习题..... 29
1.2.1 数据库系统的构成..... 2	第 3 章 数据库和表 31
1.2.2 数据库系统的特点..... 3	3.1 Access 概述..... 31
1.2.3 数据库的发展过程..... 4	3.1.1 Access 的特点..... 31
1.2.4 数据管理技术的发展..... 5	3.1.2 Access 的启动与退出..... 32
1.3 数据库系统三级模式结构..... 6	3.1.3 Access 数据库的结构..... 32
1.3.1 模式结构的概念..... 6	3.1.4 Access 设置..... 34
1.3.2 数据库系统三级模式与二级映像..... 6	3.1.5 帮助系统..... 37
1.4 数据库设计的基本步骤..... 7	3.2 创建数据库..... 38
1.5 实体-联系模型..... 8	3.2.1 新建数据库..... 38
1.5.1 实体-联系模型中的基本概念..... 9	3.2.2 数据库的简单操作..... 40
1.5.2 实体集之间的联系形式..... 10	3.3 创建数据表..... 40
1.6 习题..... 15	3.3.1 Access 数据类型..... 40
第 2 章 关系模型和关系数据库 16	3.3.2 创建表..... 42
2.1 数据模型..... 16	3.3.3 设置字段属性..... 49
2.2 关系模型的数据结构..... 16	3.3.4 修改表结构..... 54
2.3 关系数据库和关系数据库 规范化..... 18	3.3.5 输入数据..... 55
2.3.1 关系数据库..... 18	3.3.6 创建索引与主键..... 57
2.3.2 关系数据库规范化..... 18	3.3.7 建立表之间的关系..... 59
2.3.3 关系数据完整性规则..... 21	3.4 表操作..... 62
2.4 E-R 模型向关系模型的转换..... 21	3.4.1 调整表的外观..... 63
2.4.1 实体转换为关系模式..... 21	3.4.2 数据的查找与替换..... 66
2.4.2 实体之间联系的转换..... 21	3.4.3 记录排序..... 67
2.5 关系数据操作基础..... 24	3.4.4 记录筛选..... 69
	3.4.5 数据表的更名、复制和删除..... 72

3.5 习题	72	5.4.1 利用向导创建交叉表查询	113
第4章 结构化查询语言 SQL	74	5.4.2 利用“设计视图”创建交叉表 查询	115
4.1 数据查询语言	74	5.5 操作查询	115
4.1.1 简单查询	76	5.5.1 生成表查询	116
4.1.2 多表查询	78	5.5.2 更新查询	117
4.1.3 排序	80	5.5.3 追加查询	119
4.1.4 子查询	81	5.5.4 删除查询	120
4.1.5 分组查询	83	5.6 参数查询	121
4.1.6 连接查询	85	5.6.1 单参数查询	121
4.1.7 联合查询	86	5.6.2 多参数查询	122
4.2 数据操纵语言	86	5.7 其他类型的查询	123
4.2.1 追加	87	5.7.1 查找重复项查询	124
4.2.2 更新	87	5.7.2 查找不匹配项查询	125
4.2.3 删除	88	5.8 查询的其他操作	127
4.3 数据定义语言	88	5.8.1 查询的更名	127
4.3.1 创建表	88	5.8.2 查询的复制	127
4.3.2 删除表	89	5.8.3 查询的删除	127
4.3.3 修改表	90	5.9 习题	127
4.4 习题	90	第6章 窗体	129
第5章 查询	92	6.1 窗体概述	129
5.1 查询概述	92	6.1.1 窗体的功能	129
5.1.1 查询的功能	92	6.1.2 窗体的结构	129
5.1.2 查询的类型	93	6.1.3 窗体的类型	130
5.2 表达式	94	6.1.4 窗体的视图	131
5.2.1 常量	94	6.2 创建窗体	132
5.2.2 Access 常用函数	95	6.2.1 自动窗体	132
5.2.3 Access 表达式	97	6.2.2 使用向导创建窗体	134
5.3 选择查询	99	6.2.3 创建图表窗体	136
5.3.1 利用向导创建查询	99	6.3 利用设计视图创建窗体	137
5.3.2 利用“设计视图”创建查询	102	6.3.1 窗体设计视图	137
5.3.3 查询属性	107	6.3.2 常用控件的功能	138
5.3.4 添加计算字段	108	6.3.3 常用控件的使用	140
5.3.5 总计查询	109	6.3.4 窗体中控件的常用操作	145
5.4 交叉表查询	113		

6.4 修饰窗体	146	8.1.3 报表的结构	188
6.4.1 使用主题	147	8.1.4 报表的类型	190
6.4.2 使用条件格式	148	8.2 创建报表	191
6.4.3 提示信息的添加	148	8.2.1 使用“报表”工具自动创建 报表	192
6.5 定制系统控制窗体	149	8.2.2 使用“报表向导”创建报表	193
6.5.1 创建切换窗体	149	8.2.3 使用“标签”创建报表	194
6.5.2 创建导航窗体	153	8.2.4 使用“报表设计”创建报表	195
6.5.3 设置启动窗体	154	8.2.5 使用“空报表”工具创建报表	197
6.6 对象与属性	155	8.3 编辑报表	199
6.6.1 面向对象的基本概念	155	8.3.1 设置报表格式	200
6.6.2 对象属性	156	8.3.2 修饰报表	201
6.6.3 对象事件和方法	158	8.3.3 创建多列报表	203
6.7 窗体设计实例	161	8.4 报表的高级应用	203
6.8 习题	171	8.4.1 报表的排序和分组	204
第7章 宏	172	8.4.2 使用计算控件	206
7.1 Access 宏对象的概念	172	8.4.3 创建子报表	208
7.2 宏的创建与编辑	172	8.5 报表的预览和打印	211
7.2.1 操作序列宏的创建	173	8.6 习题	212
7.2.2 宏操作分组	175	第9章 VBA 程序设计	214
7.2.3 子宏的创建	176	9.1 VBA 概述	214
7.2.4 条件宏的创建	177	9.1.1 VBA 的概念	214
7.2.5 宏的编辑	179	9.1.2 VBA 的编程步骤	214
7.3 宏的运行和调试	180	9.1.3 VBA 的编程环境	215
7.3.1 宏的运行	180	9.1.4 模块的基础知识	218
7.3.2 宏的调试	181	9.1.5 模块与宏	220
7.4 利用宏创建菜单	181	9.2 面向对象程序设计基础	220
7.4.1 创建窗口下拉菜单	181	9.2.1 面向对象程序设计的基本概念	221
7.4.2 创建窗口多级菜单	182	9.2.2 Access 中的对象	221
7.5 常用宏操作	185	9.3 VBA 程序设计基础	223
7.6 习题	187	9.3.1 VBA 中的数据和表达式	223
第8章 报表	188	9.3.2 VBA 基本语句	232
8.1 概述	188	9.4 VBA 程序流程控制	235
8.1.1 报表的功能	188	9.4.1 选择结构	235
8.1.2 报表的视图	188		

9.4.2	循环结构	240	10.1	VBA 数据库编程技术简介	256
9.4.3	GoTo 控制语句	247	10.1.1	数据库引擎与数据库访问 接口	256
9.4.4	过程调用与参数传递	248	10.1.2	数据访问对象 (DAO)	257
9.5	VBA 程序调试	251	10.1.3	ActiveX 数据对象 (ADO)	259
9.5.1	错误类型	251	10.2	VBA 数据库编程技术	259
9.5.2	错误处理	251	10.2.1	DAO 编程	259
9.5.3	调试程序	252	10.2.2	ADO 编程	266
9.6	习题	254	10.3	习题	267
第 10 章	VBA 数据库编程	256			

第1章 数据库系统概述

20世纪80年代,美国信息资源管理学家霍顿(F.W.Horton)和马钱德(D.A.Marchand)等人指出:信息资源(Information Resources)与人力、物力、财力及自然资源一样,都是企业的重要资源,因此,应该像管理其他资源一样管理信息资源。

随着社会的进步和经济的发展,各类组织和机构都意识到当今社会已经跨入信息时代,数据是信息时代的重要资源之一。通常,大多数数据是经过一段时间积累下来的重要资源。人们对大量的数据进行管理,从数据中获取信息和知识,从而帮助人们进行决策,于是就有了数据库的蓬勃发展。数据库技术是计算机科学中一门重要的技术,在管理和财经等领域得到广泛的应用。特别是Internet技术的发展,为数据库技术开辟了更广泛的应用舞台。

本章的重点是介绍数据库系统的基本概念和数据库设计的步骤。

1.1 引言

首先,下面通过几个事例,讨论为什么需要数据库。

A公司的业务之一是销售一种科技含量较高的日常生活用品,为适应不同客户群的需求,这种商品有9个型号;产品通过分布在全市的3000多家不同类型零售商(如各类超市、便利店等)销售;同时,公司在全国各主要城市都设有办事处,通过当地的代理商销售这种商品。

如果是你在管理这家公司,你需要什么信息?

A公司的管理层需要随时掌握各代理商和零售商的进货情况、销货情况和库存情况;需要掌握各销售渠道的销售情况;需要了解不同型号产品在不同地域的销售情况,以便及时调整销售策略等。A公司的工作人员需要定期对代理商和零售商进行回访,解决销售过程中的各种问题,并对自己的客户(代理商和零售商)进行维护。在此过程中,公司还需要对自己市场部门的工作业绩进行考核。

随着市场范围的不断扩大,业务量迅速增长,A公司需要有效地管理自己的产品、客户和员工等数据,并且这类数据正在不断地积累、增大。

这样大量的数据靠人工管理已经不再可能,比较好的方法之一是用数据库系统来管理。那么,应该如何去抽象数据、组织数据并能够有效地使用数据,从中得到有价值的信息呢?这正是我们要讨论的问题。

另一个例子是银行,每个人都有在银行接受服务的经历。人们首先在银行开户,向银行提供个人基本信息(如姓名和身份证号码等),然后不断地存款、取款、消费;而银行需要及时地记录这些数据,并实时地更新账户余额。

解决上述问题的最佳方案之一就是使用数据库。产生数据库的动因和使用数据库的目的正是为了及时地采集数据、合理地存储数据、有效地使用数据,保证数据的准确性、一致性和安全性,在需要的时间和地点获得有价值的信息。

1.2 数据库系统

本节讨论数据库系统的构成、数据库系统的特点、数据库的发展过程以及数据管理技术的发展。

1.2.1 数据库系统的构成

数据库技术所要解决的基本问题如下。

- 1) 如何抽象现实世界中的对象，如何表达数据以及数据之间的联系。
- 2) 如何方便、有效地维护和利用数据。

通常意义下，数据库是数据的集合。一个数据库系统的主要组成部分是数据、数据库、数据库管理系统、应用程序及用户。数据存储于数据库中，用户和用户应用程序通过数据库管理系统对数据库中的数据进行管理和操作。

1. 数据

数据 (Data) 是对客观事物的抽象描述。数据是信息的具体表现形式，信息包含在数据之中。数据的形式或者说数据的载体是多种多样的，它们可以是数值、文字、图形、图像、声音等。例如，用会计分录描述企业的经济业务，会计分录反映了经济业务的来龙去脉。会计分录就是其所描述的经济业务的抽象，并且是以文字和数值的形式表现的。

数据的形式还不能完全表达数据的内容，数据是有含义的，即数据的语义或数据解释。所以数据和数据的解释是不可分的。例如，(983501011, 张捷, 女, 1978, 北京, 信息学院) 就仅仅是一组数据，如果没有数据解释，读者就无法知道这是一名学生还是一名教师的数据，1978 应该是一个年份，但它是出生年份还是参加工作或入学的年份就无法了解了。在关系数据库中，上述数据是一组属性值，属性是它们的语义。

通过对数据进行加工和处理，从数据中获取信息。数据处理通常包括数据采集、数据存储、数据加工、数据检索和数据传输 (输出) 等环节。

数据的 3 个范畴分为现实世界、信息世界和计算机世界。数据库设计的过程，就是将数据的表示从现实世界抽象到信息世界 (概念模型)，再从信息世界转换到计算机世界 (数据世界)。

2. 数据库

数据库 (DataBase) 是存储数据的容器。通常，数据库中存储的是一组逻辑相关的数据的集合，并且是企业或组织经过长期积累保存下来的数据集合，是组织的重要资源之一。数据库中的数据按一定的数据模型描述、组织和存储。人们从数据中提取有用信息，信息的积累成为知识，丰富的知识创造出智慧。

3. 数据库管理系统

数据库管理系统 (DataBase Management System, DBMS) 是一类系统软件，提供能够科学地组织和存储数据、高效地获取和维护数据的环境。其主要功能包括数据定义、数据查询、数据操纵、数据控制、数据库运行管理、数据库的建立和维护等。DBMS 一般由软件厂商提供，如 Microsoft 公司的 SQL Server 和 Access 等。

4. 数据库系统

一个完整的数据库系统 (DataBase System, DBS) 由保存数据的数据库、数据库管理系

统、用户应用程序和用户组成，如图 1-1 所示。DBMS 是数据库系统的核心。用户以及应用程序都是通过数据库管理系统对数据库中的数据进行访问的。

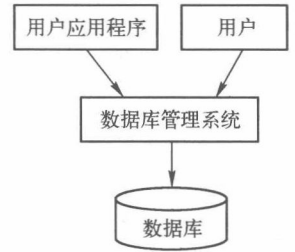


图 1-1 数据库系统组成

通常一个数据库系统应该具备如下功能。

- 1) 提供数据定义语言，允许使用者建立新的数据库并建立数据的逻辑结构 (Logical Structure)。
- 2) 提供数据查询语言。
- 3) 提供数据操纵语言。
- 4) 支持大量数据存储。
- 5) 控制并发送访问。

1.2.2 数据库系统的特点

数据库系统具有如下特点。

1. 数据结构化

数据库中的数据是结构化的。这种结构化就是数据库管理系统所支持的数据模型。使用数据模型描述数据时，不仅描述了数据本身，同时描述了数据之间的联系。按照应用的需要，建立一种全局的数据结构，从而构成了一个内部紧密联系的数据整体。关系数据库管理系统支持关系数据模型，关系模型的数据结构是关系——满足一定条件的二维表格。

2. 数据高度共享、低冗余度、易扩充

数据的共享度直接关系到数据的冗余度。数据库系统从整体角度看待和描述数据，数据不再面向某个应用，而是面向整个系统。因此，数据库中的数据可以高度共享。数据的高度共享本身就减少了数据的冗余，同时确保了数据的一致性，同一数据在系统中的多处引用是一致的。

3. 数据独立

数据的独立性是指数据库系统中的数据与应用程序之间是互不依赖的。数据库系统提供了两方面的映像功能，从而使数据既具有物理独立性，又具有逻辑独立性。

数据库系统的一个映像功能是数据的总体逻辑结构与某类应用所涉及的局部逻辑结构之间的映像或转换功能。这一映像功能保证了当数据的总体逻辑结构改变时，通过对映像的相应改变可以保持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，所以应用程序不必修改。这就是数据与程序的逻辑独立性，简称数据的逻辑独立性。

数据库系统的另一映像功能是数据的存储结构与逻辑结构之间的映像或转换功能。这一映像功能保证了当数据的存储结构（或物理结构）改变时，通过对映像的相应改变可以保持数据的逻辑结构不变，从而应用程序也不必改变。这就是数据与程序的物理独立性，简称数据的物理独立性。

4. 数据由数据库管理系统统一管理和控制

DBMS 提供以下几方面的数据管理与控制功能。

(1) 数据安全性

数据安全性 (Security) 是指保护数据，防止不合法使用数据造成数据的泄密和破坏，使每个用户只能按规定权限对某些数据以某种方式进行访问和处理。例如，部分用户对学生成绩只能查阅不能修改。

(2) 数据完整性

数据完整性 (Integrity) 是指数据的正确性、有效性、相容性和一致性, 即将数据控制在有效的范围内, 或要求数据之间满足一定的关系。

(3) 并发控制

当多用户的并发 (Concurrency) 进程同时存取、修改数据库时, 可能会发生相互干扰而得到错误的结果, 并使得数据库的完整性遭到破坏, 因此必须对多用户的并发操作加以控制和协调。

(4) 数据库恢复

计算机系统的硬件故障、软件故障、操作员的失误以及故意的破坏都会影响数据库中数据的正确性, 甚至造成数据库部分或全部数据的丢失。DBMS 必须具有将数据库从错误状态恢复到某一已知的正确状态 (也称为完整状态或一致状态) 的能力, 这就是数据库的恢复 (Recovery) 功能。

1.2.3 数据库的发展过程

美国学者詹姆斯·马丁在其《信息工程》和《总体数据规划方法论》中, 将数据环境分为 4 种类型, 阐述了数据管理即数据库的发展过程。

1. 数据文件

在数据库管理系统出现之前, 程序员根据应用的需要, 用程序语言分散地设计应用所需要的各种数据文件。数据组织技术相对简单, 但是随着应用程序的增加, 数据文件的数量也在不断增加, 最终会导致很高的维护成本。数据文件阶段, 会为每一个应用程序建立各自的数据文件, 数据是分离的、孤立的, 并且随着应用的增加, 数据被不断地重复, 数据不能被应用程序所共享。

2. 应用数据库

意识到数据文件带来的各种各样的问题, 于是就有了数据库管理系统。但是各个应用系统的建立依然是“各自为政”, 每个应用系统建立自己的数据库文件。随着应用系统的建立, 孤立的数据库文件也在增加, “数据孤岛”产生, 数据仍然在被不断地重复, 数据不能共享, 并且导致了数据的不一致和不准确。

3. 主题数据库

主题数据库是面向业务主题的数据组织存储方式, 即按照业务主题重组有关数据, 而不是按照原来的各种登记表和统计报表来建立数据库。它强调信息共享 (不是信息私有或部门所有)。主题数据库是对各个应用系统“自建自用”数据库的彻底否定, 强调各个应用系统“共建共用”的共享数据库; 所有源数据一次、一处输入系统 (不是多次、多处输入系统)。同一数据必须一次、一处进入系统, 保证其准确性、及时性和完整性, 经由网络-计算机-数据库系统, 可以多次、多处使用; 主题数据库由基础表组成, 基础表具有如下特性: 原子性 (表中的数据项是数据元素)、演绎性 (可由表中的数据生成全部输出数据) 和规范性 (表中数据结构满足三范式要求)。

4. 数据仓库

数据仓库是将从多个数据源收集的信息进行存储, 存放在一个一致的模式下。数据仓库通过数据清理、数据变换、数据集成、数据装入和定期数据刷新来构造。建立数据仓库的目的是进行数据挖掘。

数据挖掘是从海量数据中提取出知识。数据挖掘是以数据仓库中的数据为对象, 以数据挖掘算法为手段, 最终以获得的模式或规则为结果, 并通过展示环节表示出来。

1.2.4 数据管理技术的发展

计算机数据管理随着计算机硬件、软件技术和计算机应用范围的发展而不断发展,经历了人工管理阶段、文件系统阶段和数据库技术阶段 3 个阶段。对数据管理是为了对数据进行处理,数据处理的过程包括数据收集、存储、加工和检索等过程。

1. 人工管理阶段

20 世纪 50 年代中期以前,计算机主要用于数值计算。从硬件系统看,当时的外存储设备只有纸带、卡片、磁带,没有直接存取设备;从软件系统看,没有操作系统以及管理数据的软件;从数据看,数据量小,数据无结构,由用户直接管理,且数据间缺乏逻辑组织,数据依赖于特定的应用程序,缺乏独立性。人工管理阶段数据管理的特点如下。

- 数据不能保存:一个目标计算完成后,程序和数据都不能被保存。
- 应用程序管理数据:应用程序与数据之间缺少独立性。
- 数据不能共享:数据是面向应用的,一组数据只能对应一个程序。
- 数据不具有独立性:数据结构改变后,应用程序必须修改。

2. 文件系统阶段

20 世纪 50 年代后期到 60 年代中后期,计算机应用从科学计算发展到科学计算和数据处理。1954 年出现了第一台商业数据处理的电子计算机 UNIVAC1,标志着计算机开始应用于以加工数据为主的事务处理阶段。基于计算机的数据处理系统也就从此迅速发展起来。这个阶段,硬件系统出现了磁鼓、磁盘等直接存取数据的存储设备;软件系统有了文件系统,处理方式也从批处理发展到了联机实时处理。文件系统阶段的数据管理特点如下。

- 数据可以长期保存:数据能够保存在存储设备上,可以对数据进行各种数据处理操作,包括查询、修改、增加、删除操作等。
- 文件系统管理数据:数据以文件形式存储在存储设备上,有专门的文件系统软件对数据文件进行管理,应用程序按文件名访问数据文件,按记录进行存取,可以对数据文件进行数据操作。应用程序通过文件系统访问数据文件,使得程序与数据之间具有一定的独立性。
- 数据共享差、数据冗余大:仍然是一个应用程序对应一个数据文件(集),即便是多个应用程序需要处理部分相同的数据时,也必须访问各自的数据文件,由此造成数据的冗余,并可能导致数据不一致;数据不能共享。
- 数据独立性不好:数据文件与应用程序一一对应,数据文件改变时,应用程序就需要改变;同样,应用程序改变时,数据文件也需要改变。

3. 数据库技术阶段

20 世纪 70 年代开始有了专门进行数据组织和管理的软件——数据库管理系统,特别在 20 世纪 80 年代后期到 20 世纪 90 年代,由于金融和商业的需求,数据库管理系统得到了迅猛的发展。数据库技术阶段数据管理具有如下特点。

- 数据结构化。
- 数据共享性高,冗余度低,易扩充。
- 数据独立性高。
- 数据由 DBMS 统一管理,DBMS 具有完备的数据管理和控制功能。

1.3 数据库系统三级模式结构

从数据库管理系统的内部体系结构角度看，数据库管理系统对数据库数据的存储和管理采用三级模式结构。数据库系统的三级模式结构是指数据库系统由模式、外模式和内模式三级构成。数据库系统的三级模式结构如图 1-2 所示。

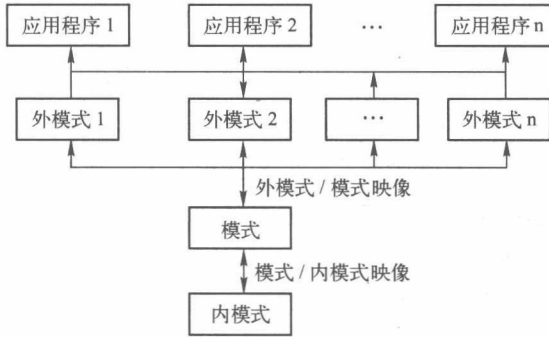


图 1-2 数据库系统的三级模式结构

1.3.1 模式结构的概念

1. 模式

模式 (Schema) 又称逻辑模式，是数据库中全部数据的逻辑结构和特征的描述，是对数据的结构和属性的描述。

关系数据库用关系数据模型来描述数据的逻辑结构 (数据项、数据类型、取值范围等) 和数据之间的联系，以及数据的完整性规则。

在关系数据模型中，对学生数据的一组描述 (学号，姓名，性别，所在学院) 就是一个模式，这个模式可以有多个不同的值与其对应，每一组对应的值称为模式的实例。例如，(2008350222, 钟红, 女, 信息学院) 就是上述模式的一个实例。

数据库设计的主要任务之一就是数据库的模式设计。

2. 外模式

外模式 (External Schema) 又称子模式或用户视图，是用户能够看到和使用的逻辑数据模型描述的数据。外模式通常是从模式得到的子集；用户的需求不一样，用户视图就不一样，因此，一个模式可以有很多个外模式。

外模式可以起到很好地保护数据安全的作用，是数据库数据安全的一个有力措施。外模式使得每个用户只能访问到与其相关的数据，不能看到模式中的其他数据。

3. 内模式

内模式 (Internal Schema) 又称存储模式，是数据物理结构和存储方式的描述，一个模式只有一个内模式。

1.3.2 数据库系统三级模式与二级映像

数据库系统的三级模式对应数据的三个抽象级别，数据的具体组织由 DBMS 管理，用户

可以逻辑地抽象处理数据，而无须关心数据在计算机内部的具体表示方式和存储方式。

数据库系统的三级模式提供了二级映像，从而保证了数据库系统中数据的逻辑独立性和物理独立性。

1. 外模式/模式映像

模式描述了数据的全局逻辑结构，外模式是根据用户需求描述的数据局部逻辑结构。

对应一个模式可以有任意多个外模式，如图 1-3 所示。对应于每一个外模式，都有一个外模式/模式映像，它定义了该外模式与模式之间的对应关系。

应用程序是依据数据的外模式编写的，因此当模式改变时，应用程序不必改变，从而实现了数据与程序之间的逻辑独立性，简称数据的逻辑独立性。

2. 模式/内模式映像

数据库中，模式是唯一的，内模式也是唯一的，模式与内模式是一一对应的，模式/内模式映像也是唯一的，如图 1-4 所示。模式/内模式映像定义了数据全局逻辑结构与存储结构之间的对应关系，并且实现了数据的物理独立性。

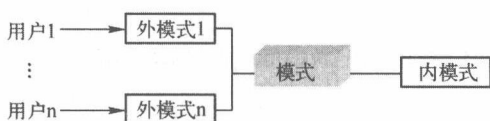


图 1-3 外模式/模式映像

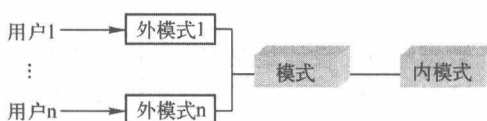


图 1-4 模式/内模式映像

1.4 数据库设计的基本步骤

数据是一个组织机构的重要资源之一，是组织积累的宝贵财富，通过对数据的分析，可以了解过去，把握今天，预测未来。但这些数据通常是大量的、甚至是杂乱无章的，如何合理、有效地组织这些数据，是数据库设计的重要任务之一。

正如前面所述，数据库是企业或组织所积累数据的聚集，除了每一个具体数据以外，这些数据是逻辑相关的，即数据之间是有联系的。数据库是组织和管理这些数据的常用工具。数据库设计讨论的问题是：根据业务管理和决策的需要，应该在数据库中保存什么数据；这些数据之间有什么联系；如何将所需要的数据划分到表格的列，并且建立表之间的关系。

数据的三个范畴分为：现实世界、信息世界和计算机世界。数据库设计的过程，就是将数据的表示从现实世界抽象到信息世界（概念模型），再从信息世界转换到计算机世界（数据模型）。

数据库设计的目的在于提供实际问题的计算机表示，在于获得支持高效存取数据的数据结构。数据库中用数据模型这个工具来抽象和描述现实世界中的对象（人或事物）。数据库设计分为 4 个步骤，如图 1-5 所示。

1. 需求分析

对需要使用数据库系统来进行管理的现实世界中对象（人或事物）的业务流程、业务规则和所涉及的数据进行调查、分析和研究，充分理解现实世界中的实际问题 and 需求。需求分析的策略一般有两种：自下向上的方法和自上向下的方法。

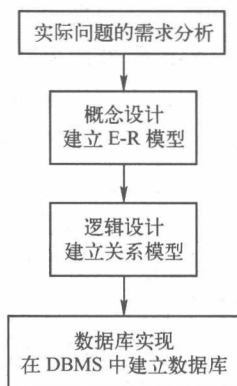


图 1-5 数据库设计步骤

(1) 自下向上的方法

对事物进行了解，理解实际问题的业务规则和业务流程。在此基础上，归纳出该事物处理过程中需要存放在数据库中的数据。

例如，一个产品销售数据库，需要保存客户的哪些数据？可以做出一个二维表格，每一列是一个数据项，每一行是一个客户信息，可能包括：客户姓名、地址、邮政编码、手机号码等。

(2) 自上向下的方法

从为描述事物最终提供的各种报表和经常需要查询的信息着手，分析出应包含在数据库中的数据。

例如，上述产品销售数据库的客户信息，是否需要按客户性别进行统计分析，如果需要，就应该增加一列“性别”数据项。

进行需求分析时，通常会同时使用上述两种方法。自下向上的方法反映了实际问题的信息需求，是对数据及其结构的需求，是一种静态需求；自上向下的方法侧重于对数据处理的需求，即实际问题的动态需求。

2. 数据库概念设计

数据库概念设计是在需求分析的基础上，建立概念数据模型（Conceptual Data Model）；用概念模型描述实际问题所涉及的数据以及数据之间的联系；这种描述的内容和描述的详细程度取决于期望得到的信息。一种较常用的概念模型是实体-联系模型，又称 E-R (Entity-Relationship) 模型。E-R 模型用实体和实体之间的联系来表达数据以及数据之间的联系。

例如，产品销售数据库，供应商是实体，客户是另一个实体，产品是实体，订单是实体，并且它们之间是有联系的；使用 E-R 模型描述这些实体以及它们之间的联系。

3. 数据库逻辑设计

数据库逻辑设计是根据概念数据模型建立逻辑数据模型（Logic Data Model）。逻辑数据模型是一种面向数据库系统的数据模型，本节使用目前应用广泛的关系数据模型来描述数据库逻辑设计：根据概念模型建立数据的关系模型（Relational Model）；用关系模型描述实际问题在计算机中的表示。关系模型是一种数据模型，用二维表格的形式来表示数据以及数据之间的联系。数据库的逻辑设计实际是把 E-R 模型转换为关系模型的过程。

E-R 模型和关系模型分属两个不同的层次，E-R 模型更接近于用户，不需要用户具有计算机知识，属于现实世界范畴；而关系模型是从计算机的角度描述数据及数据之间的联系，需要使用者具有一定计算机知识，属于计算机范畴。

4. 数据库实现（数据库物理设计）

依据关系模型，在数据库管理系统环境中建立数据库。如在 Access 中，Access 把数据组织到表格，表格由行和列组成。简单的数据库可能只包含一个表格，但是大多数数据库是包含多个表的，并且表之间有联系。

例如产品销售数据库，就应该至少包含供应商表、客户表、产品表、订单表等，这些表通过主键建立联系。

1.5 实体-联系模型

数据库设计的过程就是利用数据模型来表达数据和数据之间联系的过程。数据模型是一

种工具，用来描述数据 (Data)、数据的语义 (Data Semantics)、数据之间的联系 (Relationship) 以及数据的约束 (Constraints) 等。数据建模过程是一个抽象的过程，其目的是把一个现实世界中的实际问题用一种数据模型来表示，用计算机能够识别、存储和处理的数据形式进行描述。在本节中，将讨论一种用于数据库概念设计的数据模型——E-R 模型。一般地讲，任何一种数据模型都是经过严格定义的。

理解实际需求之后，需要用一种方法来表达这种需求，现实世界中使用概念数据模型来描述数据以及数据之间的联系，即数据库概念设计。概念模型的表示方法之一是 E-R 模型，即用 E-R 模型表达实际需求的需求。E-R 模型具有足够的表达能力且简明易懂，不需要使用者具有计算机知识。E-R 模型以图形的方式表示模型中各元素及其之间的联系，所以又称 E-R 图 (Entity-Relationship Diagram)。E-R 图便于理解且易于交流，因此，E-R 模型得到了相当广泛的应用。

1.5.1 实体 - 联系模型中的基本概念

下面介绍 E-R 模型中使用的基本元素。

1. 实体

实际问题中客观存在并可相互区别的事物称为实体 (Entity)。实体是现实世界中的对象，实体可以是具体的人、事、物。例如，实体可以是一名学生、一位教师或图书馆中的一部书籍。

2. 属性

实体所具有的某一特性称为属性 (Attribute)。在 E-R 模型中用属性来描述实体，例如通常用“姓名”“性别”“出生日期”等属性来描述人，用“图书名称”“出版商”“出版日期”等属性描述书籍。一个实体可以由若干个属性来描述。例如学生实体可以用学号、姓名、性别、出生日期等属性来描述。这些属性的集合 (学号, 姓名, 性别, 出生日期) 表征了一名学生的部分特性。一个实体通常具有多种属性，应该使用哪些属性描述实体，取决于实际问题的需要或者说取决于最终期望得到哪些信息。例如，教务处关心学生各门功课的成绩，而学生处可能会更关心学生的各项基本情况，如学生来自哪里、监护人是谁、如何联系等问题。

确定属性的两条原则。

- 1) 属性必须是不可分的最小数据项，属性中不能包含其他属性，不能再具有需要描述的性质。
- 2) 属性不能与其他实体具有联系，E-R 图中所表示的联系是实体集之间的联系。

属性的取值范围称为该属性的域 (Domain)。例如，“学号”的域可以是 9 位数字组成的字符串，“性别”的域是“男”或“女”，“工资”的域是大于零的数值等。但域不是 E-R 模型中的概念，E-R 模型不需要描述属性的取值范围。

3. 实体集

具有相同属性的实体的集合称为实体集 (Entity Set/Entity Class)。例如，全体学生就是一个实体集。实体属性的每一组取值代表一个具体的实体。例如，(983501011, 张捷, 女, 1978 年 12 月) 是学生实体集中的一个实体，而 (993520200, 李纲, 男, 1978 年 8 月) 是另一个实体。在 E-R 模型中，一个实体集中的所有实体具有相同的属性。

4. 键

在描述实体集的所有属性中，可以唯一地标识每个实体的属性称为键 (Key) 或标识 (Identifier)。首先，键是实体的属性；其次，这个属性可以唯一地标识实体集中每个实体。因此，作为键的属性取值必须唯一且不能“空置”。例如，在学生实体集中，学号属性唯一地标