



普通高等院校计算机基础教育“十三五”规划教材

Access 数据库与 VBA 面向对象程序设计

Access SHUJUKU YU VBA MIANXIANG DUIXIANG CHENGXU SHEJI

◆ 黎升洪 编著
◆ 万常选 主审

中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE

Access 数据库与 VBA 面向对象程序设计

黎升洪 编著

万常选 主审

内 容 简 介

本书根据教育部高等教育司组织制定的《高等学校文科类专业大学计算机教学基本要求(2011年版)》对数据库技术和程序设计方面的要求编写。

本书以 Microsoft Access 2010 中文版为平台,介绍了学习关系数据库的基础知识,阐述了基于数据库模式导航的多表 SQL 查询语句的工作原理,描述了结构化程序设计思想,讨论了面向对象程序设计的核心概念、编程过程,给出了应用实例。本书的特点是强化给定数据库模式的应用,弱化数据库模式设计。

本书内容丰富,层次清晰,讲解深入浅出。全书配有源码,并为教师提供电子课件。本书适合作为高等学校财经管理类专业和其他非计算机专业的数据库应用课程教材,也可作为全国计算机等级考试人员的参考资料,还可供从事办公软件开发的人员使用。

图书在版编目(CIP)数据

Access 数据库与 VBA 面向对象程序设计/黎升洪编著. —北京:中国铁道出版社, 2017. 2

普通高等院校计算机基础教育“十三五”规划教材

ISBN 978-7-113-22557-5

I. ①A… II. ①黎… III. ①关系数据库系统-程序设计-高等学校-教材 ②BASIC 语言-程序设计-高等学校-教材 IV. ①TP311.138 ②TP312.8

中国版本图书馆 CIP 数据核字 (2016) 第 286861 号

书 名: Access 数据库与 VBA 面向对象程序设计
作 者: 黎升洪 编著

策 划: 曹莉群

读者热线: (010) 63550836

责任编辑: 周海燕 徐盼欣

封面设计: 乔 楚

责任校对: 汤淑梅

责任印制: 郭向伟

出版发行: 中国铁道出版社 (100054, 北京市西城区右安门西街 8 号)

网 址: <http://www.51eds.com>

印 刷: 三河市宏盛印务有限公司

版 次: 2017 年 2 月第 1 版 2017 年 2 月第 1 次印刷

开 本: 787mm×1092mm 1/16 印张: 15.5 字数: 329 千

书 号: ISBN 978-7-113-22557-5

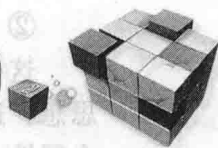
定 价: 36.50 元

版权所有 侵权必究

凡购买铁道版图书,如有印制质量问题,请与本社教材图书营销部联系调换。电话: (010) 63550836

打击盗版举报电话: (010) 51873659

前 言



大数据时代数据存储的核心技术是关系数据库。这对财经管理类专业学生掌握关系数据库的核心概念和工作原理以及结构化和面向对象编程等知识提出了更高要求。Access 作为桌面级数据库是一种易用的关系数据库系统,它具有实体完整性和参照完整性等大型数据库具备的特性。VBA 作为 Access 的自动化编程语言具有简单易学的特性,同时具备当前面向对象程序设计语言的图形化界面、事件驱动等面向对象特性,非常适合财经管理类专业与非计算机专业学生学习关系数据库知识。Access 2010 将关系数据库管理知识、结构化程序设计和面向对象程序设计有机结合,提供了开发简单信息系统所需要的所有功能。Access 的这些特点使其具有使用简单、起点低等特性,是目前流行的关系数据库,也是全国计算机等级考试的可选内容之一。

《Access 数据库与 VBA 面向对象程序设计》的编写围绕教育部高等学校文科计算机基础教学指导委员会制定的《高等学校文科类专业大学计算机教学基本要求(2011 年版)》,吸纳当今计算机学科发展中出现的新技术、新成果,强调理论与应用相结合,注意合理取材和安排,力求重点突出、概念清晰、讲究实效,注重培养学生掌握计算机的基础知识、基本概念和基本操作技能,使学生具有能够应用计算机进行学习、工作以及解决实际问题的能力。本书的主要特色如下:

(1) 以学生熟悉的教学管理系统为背景介绍数据库各种相关知识,同时以图书馆管理系统为背景来设置相应的练习,好处是易于理解。

(2) 数据库概念部分详细介绍数据库完整性约束思想,通过数据库完整性约束的主键和外键约束,使数据库模式具备数据判定功能,避免垃圾数据进入数据库。

(3) 如同地图使用和地图绘制知识是分离的一样,数据库关系模式设计与关系数据库记录的查询、更新、插入、删除等知识也是分离的。数据库关系模式设计需要数据库范式理论知识,通过数据库模式设计,可以定义一个关系数据表。而关系数据库记录的查询、更新、插入、删除等操作是最常见的关系数据库操作。本书将弱化前者,强化后者。这样安排的目的是:使非计算机类专业的学生在具备正确数据库概念的基础上,集中精力更好地掌握关系数据库的操作和使用。

(4) 抽象了面向对象程序设计思想框架,以便于学生自主学习面向对象程序设计示例。基于面向对象系统观认为系统由对象和对象间交互构成。具体说,描述和构造一个系统过程为三步:

① 刻画该系统中的各个核心概念——类和它具有的属性和方法(或事件)。

② 将类实例化为对象。对于初次涉及面向对象编程的用户,为简化问题,通常不涉及编写自己的类代码,仅需将系统提供的类实例化为对象。

③ 在所有控件已经命名的基础上,描述这些对象间的交互,即这些对象间的消息关系。

因此,基于这三步,编写图形化窗体程序时的步骤可概括为:

① 设计窗体界面的过程，就是实例化各种控件对象。

② 编写事件过程，就是实现对象间交互。

这样做的好处是：使读者能从较高层面理解基于图形化界面的面向对象程序设计思想，避免只见树木、不见森林的情况，也不会被 Access 对象模型中多个对象、多个属性、多个事件（或方法）困扰。面向对象编程的重点是理解 VBA 对象属性和方法（或事件）的引用方式，而某个属性、事件或方法的具体功能可通过翻阅书籍或联机帮助了解。这样有助于学生的模仿和理解，进而编写自己的图形化界面代码。总之，在事件触发的图形化界面程序设计中，其宏观层面是面向对象程序设计思想，但微观层面是结构化程序设计思想。

为便于学生自学，本书附录中给出了 VBA 函数一览表和 VBA 语言简明手册等实用内容，并附有索引。本书给出了每章例子的（程序）源码，并为教师提供电子课件，可从中国铁道出版社网站下载。

本书内容涵盖以下三部分：

（1）关系数据库基础知识，包括数据库基本概念、数据库角色、主键（外键）和数据库实体（参照）完整性等内容。

（2）结构化编程知识，包括数据类型、常量、变量、函数、表达式和三种基本控制结构等内容。

（3）面向对象程序设计知识，包括面向对象核心概念、面向对象编程的基本步骤、Access 对象模型、Access 控件（窗体）属性和方法（或事件）调用方式、DAO 数据库编程对象模型等内容。

32 课时授课，32 课时上机的参考课时安排如下：第 1 章（2 授课，0 上机）；第 2 章（0 授课，2 上机）；第 3 章（8 授课，8 上机）；第 4 章（0 授课，2 上机）；第 5 章（选讲）；第 6 章（2 授课，2 上机）；第 7 章（8 授课，8 上机）；第 8 章（4 授课，2 上机）；第 9 章（6 授课，8 上机）；机动课时（2 授课，0 上机）。

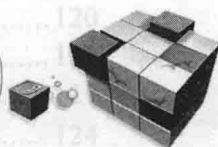
本教程由黎升洪编著，万常选主审。万常选教授对本书初稿进行了认真审阅，提出了许多宝贵意见和建议。此外，郭勇博士和刘谦副教授为本教程提供了资料，徐升华、吴京慧教授和毛小兵院长对本书编写给予了许多帮助，在此表示衷心的感谢。

虽然本书是关于 Access 和 VBA 的教材，但在编写过程中，力图反映大型数据库和面向对象语言具备的特性，力求反映数据库新技术，以保持其先进性和实用性。由于编者水平有限，疏漏和不足之处在所难免，恳请同行专家和广大读者批评指正。联系邮件 lee.shenghong@gmail.com。

编 者

2016 年 7 月

目 录



第 1 章 数据库系统概述	1
1.1 数据库技术	1
1.1.1 信息与数据	1
1.1.2 数据处理	2
1.1.3 数据库系统	2
1.2 数据模型	4
1.2.1 数据模型的三要素	5
1.2.2 概念模型与实体—联系方法	6
1.2.3 数据模型	8
1.3 关系数据库	10
1.3.1 关系模型	10
1.3.2 关系完整性约束	17
1.4 数据库系统应用模式	18
1.5 E-R 图与关系数据库设计	20
习题	24
第 2 章 Access 2010 中对象与操作基础	26
2.1 Access 2010 的运行及其工作界面	26
2.1.1 运行 Access 2010	26
2.1.2 Access 2010 工作界面	26
2.2 Access 中的对象及其视图	29
2.2.1 表对象	30
2.2.2 查询对象	30
2.2.3 窗体对象	31
2.2.4 报表对象	31
2.2.5 宏和模块对象	32
2.3 Access 2010 帮助使用	32
习题	33
第 3 章 SQL 查询与操纵语句的使用	34
3.1 SQL 历史和优点	34
3.1.1 SQL 的历史和标准	34
3.1.2 SQL 的优点	35
3.2 数据查询	36



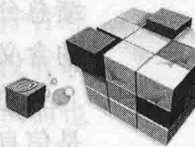
3.2.1	数据查询建立与保存的方法	36
3.2.2	单表查询	41
3.2.3	多表自然连接查询	50
3.2.4	参数查询	57
3.2.5	SQL 子查询与合并查询	57
3.2.6	交叉表查询	62
3.3	数据操纵	63
3.4	数据库与数据表建立	64
3.4.1	Access 数据库和数据表建立	64
3.4.2	使用 SQL 语句建立、删除和修改表结构	69
习题		71
第 4 章	窗体设计	74
4.1	窗体概述	74
4.1.1	窗体的功能	74
4.1.2	窗体的类型	75
4.2	窗体创建	76
4.2.1	创建平面、分割和多项目窗体	76
4.2.2	主/子窗体创建和使用	78
4.2.3	创建数据透视表和数据透视图	79
4.3	个性化窗体设计	80
4.3.1	个性化窗体示例	81
4.3.2	个性化窗体设计界面	83
4.3.3	个性化窗体设计中的方法	84
习题		90
第 5 章	报表设计	91
5.1	报表概述	91
5.2	报表创建	92
5.3	个性化报表设计	95
习题		102
第 6 章	VBA 程序语言基础	103
6.1	VBA 的开发环境 VBE	103
6.2	数据类型概述	109
6.3	常量	110
6.4	变量	111
6.5	函数	114
6.5.1	函数概述	114
6.5.2	数值处理函数	116



6.5.3	字符函数	117
6.5.4	日期和时间处理函数	120
6.5.5	金融函数	121
6.5.6	输入与输出函数	121
6.6	表达式	124
6.7	VBA 注释符和续行符	129
	习题	129
第 7 章	结构化程序设计	131
7.1	结构化程序设计概述	131
7.1.1	程序设计过程	131
7.1.2	算法与三种控制结构	132
7.1.3	程序流程图及示例	133
7.2	VBA 程序的控制语句结构	136
7.2.1	顺序结构程序	136
7.2.2	分支结构程序	137
7.2.3	循环结构程序	142
7.2.4	循环嵌套程序设计举例	147
7.3	数组及其应用	150
7.4	过程与自定义函数	154
7.5	VBA 程序调试方法	159
7.5.1	程序错误的种类	159
7.5.2	良好编程风格	160
7.5.3	使用 VBA 调试器调试程序	160
	习题	161
第 8 章	面向对象程序设计基础	165
8.1	面向对象的基本概念与示例	165
8.1.1	面向对象核心概念	165
8.1.2	类或对象的特性	167
8.1.3	面向对象编程示例	170
8.2	Access VBA 面向对象方法	177
8.2.1	Access 对象模型	177
8.2.2	Access 对象模型常见属性、方法和事件	181
8.2.3	VBA 面向对象编程常见操作方法	190
8.2.4	Access VBA 面向对象编程方法	191
	习题	196
第 9 章	VBA 面向对象程序设计	197
9.1	Access VBA 面向对象编程控件示例	197



9.2 数据库编程中的 DAO 对象模型和 DoCmd 对象	202
9.2.1 DAO 对象模型	202
9.2.2 DoCmd 对象	206
9.3 面向对象编程综合示例	217
习题	228
附录	230
附录 A VBA 函数一览表	230
附录 B VBA 语言简明手册	231
索引	236
参考文献	240
第 5 章 报表设计	246
5.1 报表概述	246
5.2 报表创建	246
5.3 个性化报表设计	246
5.4 报表打印	246
5.5 习题	246
第 6 章 VBA 编程语言基础	246
6.1 VBA 的开发环境	246
6.2 数据类型概述	246
6.3 变量	246
6.4 函数	246
6.5 习题	246
6.5.1 函数概述	246
6.5.2 函数处理函数	246



计算机已成为信息社会人们日常工作中处理数据的得力助手和工具,渗透到人们生活和工作中的各个领域。数据处理、科学计算、过程控制和辅助设计是计算机四大应用。数据处理的主要技术是数据库技术。本章讲解数据库技术的基本知识和概念,重点是关系数据库、数据完整性和数据库模式的概念。



1.1 数据库技术

信息社会中,信息是一种资源。对企业来说,各种必需的信息是其赖以生存和发展的基石;对一个国家来说,信息决定其如何建设和发展;对一个人来说,信息是其决定如何发展才能适应社会的基本要素。信息是维持生产活动、经济活动和社会活动必不可少的基本资源,它是有价值的,是构成客观世界的三大要素(信息、能源和材料)之一。因此,人们为了获取有价值的信息用于决策,就需要对信息和用于表示信息的数据进行处理和管理。人们用计算机对数据进行处理的应用系统称为计算机信息系统,而计算机信息系统的核心是数据库。

1.1.1 信息与数据

信息和数据是数据处理中的两个核心概念。在一些不是很严格的场合下,对它们没有做严格的区分,甚至当作同义词来使用。这里,数据是记录现实世界中各种信息并可以识别的物理符号,是信息的载体,是信息的具体表现形式。数据的表示形式不仅仅只是数字,还包括字符(文字和符号)、图表(图形、图像和表格)及声音(视频)等形式。数据以某种特定格式来表示事实和概念,这种形式有助于通信、解释和处理。数据有两方面的特征:一是客体属性的反映,这是数据的内容;二是记录信息的符号,这是数据的形式。

信息是数据所包含的意义。信息具有如下重要特征:

- (1) 信息具有表征性。它能够表达事物的属性、运动特性及状态。
- (2) 信息具有可传播性。信息可以进行获取、存储、传递、共享。
- (3) 信息具有可处理性。信息可以进行压缩、加工、再生。
- (4) 信息具有可用性、可增值性、可替代性。

数据与信息是密切关联的。信息是向人们提供有关现实事物的知识,数据则是承载信息的物理符号,二者是不可分离而又有一定区别的两个相关的概念。信息可以用



不同形式的数据来表示，也不随它的数据形式不同而改变。例如，张平同学的高考成绩总分为 630 分。这里符号 630 就是数据；630 解释为高考成绩总分，表示的是 630 的含义，即信息。

总之，数据形式是信息内容的表现方式，信息内容是数据形式的实质，即“数据是信息的载体，信息是数据的内涵”。

1.1.2 数据处理

要使获得的信息能够充分地发挥作用，就必须对其进行处理，这种处理称为数据处理。数据处理是指利用计算机对各种形式的数据进行一系列的存储、加工、计算、分类、检索、传输等处理。如果稍加扩展就包括数据的采集、整理、编码、输入和输出等数据组织，数据组织过程也应属于数据处理的内容，只不过数据组织过程主要是由人对其进行有效的处理，并把数据组织到计算机中。

1.1.3 数据库系统

下面介绍数据库系统的相关知识。

1. 数据库的概念

在日常工作中，需要处理的数据量往往很大，为便于计算机对其进行有效的处理，可以将采集的数据存放在建立在磁盘、光盘等外存媒介的“仓库”中，这个“仓库”就是数据库（DataBase，DB）。数据集中存放在数据库中，便于对其进行处理，提炼出对决策有用的数据和信息。这就如同一个学校采购大量的图书存放在图书馆（书库），供学生借阅。因此，数据库就是在计算机外部存储器中存储的数据仓库。

与书库需要管理员和一套管理制度一样，数据库的管理也需要一个管理系统，这个管理系统就称为数据库管理系统（DataBase Management System，DBMS）。以数据库为核心，并对其进行管理的计算机系统称为数据库系统（DataBase System，DBS）。那么，什么是数据库呢？数据库是一个复杂的系统，给它下一个确切的定义是困难的，目前还没有一个公认的、统一的定义。

但对一个特定数据库来说，它是集中、统一地保存、管理着某一单位或某一领域内所有有用信息的系统，这个系统根据数据间的自然联系结构而成，数据较少冗余，且具有较高的数据独立性，能为多种应用服务。

2. 数据库的发展

数据管理的发展经历了人工管理、文件系统到数据库系统三个阶段。

在人工管理阶段，由于没有软件系统对数据进行管理和计算机硬件的限制，数据的管理是靠人工进行的，而计算机只能对数据进行计算。当时对数据处理的过程是：先将程序和数据输入计算机，计算机运行结束后，将结果再输出，由人工保存，计算机并不存储数据。

20 世纪 50 年代后期到 60 年代中期，由于计算机外存得到发展，软件又有了操作系统，对数据管理产生了文件系统。在文件系统阶段，是按照数据文件的形式来存



放数据的, 在一个文件中包含若干“记录”, 一个记录又包含若干“数据项”, 用户通过对文件的访问实现对记录的存取。这种数据管理方式称为文件管理系统。文件管理系统的一个致命的不足是: 数据的管理没有实现结构化组织, 数据与数据之间没有联系, 文件与文件之间没有有机的联系, 数据不能脱离建立其数据文件的程序, 从而也使文件管理系统中的数据独立性和一致性差, 冗余度大, 限制了大量数据的共享和有放的应用。

20 世纪 60 年代末期, 随着计算机技术的发展, 为了克服文件管理系统的缺点, 人们对文件系统进行了扩充, 研制了一种结构化的数据组织和处理方式, 即数据库系统。数据库系统建立了数据与数据之间的有机联系, 实现了统一、集中、独立地管理数据, 使数据的存取独立于使用数据的程序, 实现了数据的共享。从 20 世纪 90 年代至今, 数据库技术得到飞速的发展。

3. 数据库的特征

作为信息管理中的核心技术, 数据库技术在计算机应用中得到迅速的发展, 目前已经成为信息管理的最新、最重要的技术。数据库有以下明显特点:

(1) 数据结构化。数据库中的数据不再像文件系统中的数据那样从属特定的应用, 而是按照某种数据模型组织成为一个结构化的数据整体。它不仅描述了数据本身的特性, 而且描述了数据与数据之间的种种联系, 这使数据库具备了复杂的内部组织结构。

(2) 实现数据共享。这是数据库技术先进性的重要体现。由于数据库中的数据实现了按某种数据模型组织为一个结构化的数据, 实现了多个应用程序、多种语言及多个用户能够共享一个库中的数据, 甚至在一个单位或更大的范围内共享, 大大提高了数据的利用率, 提高了工作效率。

(3) 减少数据冗余度。在数据库技术之前, 许多应用系统都需要建立各自的数据文件, 即使相同的数据也都需要在各自的系统中保留, 造成大量的数据重复存储, 这一现象称为数据的冗余。由于数据库实现了数据共享, 减少了存储数据的重复, 节省了存储空间, 减少了数据冗余。

(4) 数据独立性。数据库技术中的数据与操作这些数据的应用程序相互独立, 互不依赖, 不因一方的改变而改变另一方, 这大大简化了应用程序设计与维护的工作量, 同时数据也不会随应用程序的结束而消失, 可长期保留在计算机系统中。

(5) 统一的数据安全保护。数据共享在提供了多个用户共享数据资源的同时, 还需解决数据的安全性、一致性和并发性问题。这里安全性是指只有合法授权的用户才能对数据进行操作; 一致性是指当多个用户对同一数据操作时不能互相干扰, 从而出现操作结果不确定或不一致的情况; 在保证一致性的前提下, 数据库系统提供并发功能, 使多用户同时对数据库的操作有一致的正确结果。

4. 数据库的角色

数据库的系统结构及角色如图 1-1 所示。

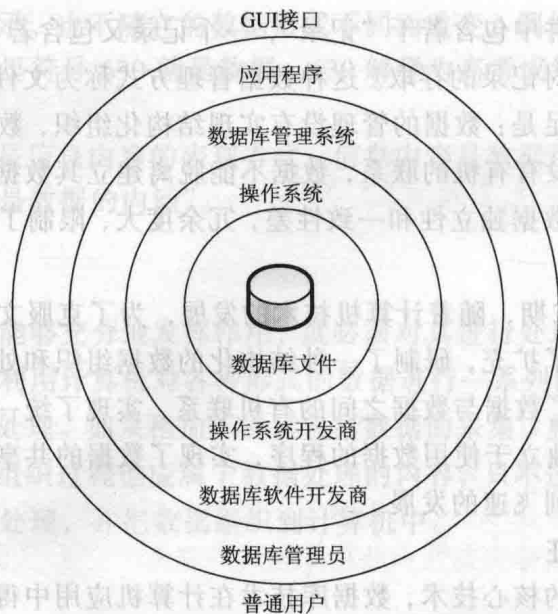


图 1-1 数据库系统结构及角色

由内至外，一个数据库系统角色及其负责功能如下：

(1) 数据库软件开发商负责数据库软件开发工作，其编程基础是操作系统提供的各种功能，开发的软件称为数据库管理系统。数据库管理系统功能是完成对数据定义、描述、操作、维护等功能（例如，SQL 语句功能的实现），接受并完成用户程序及终端命令对数据库的不同请求，并负责保护数据免受各种干扰和破坏。数据库软件的核心是数据库管理系统。

(2) 数据库管理员（DataBase Administrator, DBA）负责应用程序的开发工作。应用程序通常是图形化界面的程序，它封装了数据库操作知识，使得没有数据库知识的普通用户也能够使用数据库操作、存储数据。数据库管理员必须理解数据库管理系统提供的各种功能（如 SQL 语句编写和图形化面向对象编程知识）。数据库管理员负责数据库的设计、建立、执行和维护，以及应用程序开发。

(3) 普通用户负责日常事务处理软件的操作。普通用户只要会操作图形化用户界面（Graphic User Interface, GUI）的程序即可，不需要理解数据库的 SQL 和编程等知识。因为图形化用户接口将普通用户对界面的操作最终转换为相应的 SQL 语句等有数据库管理系统提供的功能。

本书介绍成为关系数据库管理员所必须具备的基础知识。



1.2 数据模型

提到模型人们自然会联想到建筑模型、飞机模型等事物。广义地说，模型是现实世界特征的模拟和抽象。在数据库中，用数据模型（Data Model）这个工具来对现实世界进行抽象。数据模型是数据库系统中用于提供信息表示和操作手段的形式构架。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。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。不同的数据模型是提供给人们模型化数据和信息的不同工具。根据模型应用的目的,可以将数据模型分为两种类型:第一类模型是概念模型,也称信息模型,它是独立于计算机之外的模型,如实体—联系模型,这种模型不涉及信息在计算机中如何表示,而是用来描述某一特定范围内人们所关心的信息结构,它是按用户的观点来对数据和信息建模,主要用于数据库设计;另一类模型是数据模型,它是直接面向计算机的,是按计算机系统的观点对数据进行建模,主要用于 DBMS 的实现,常称为基本数据模型,数据库中常用的基本数据模型有网状模型、层次模型和关系模型。

数据模型是数据库系统的核心和基础。各种机器上实现的 DBMS 软件都是基于某种数据模型的。

图 1-2 显示了把现实世界中的具体事物抽象、组织为某一 DBMS 支持的数据模型的过程。过程包括:

(1) 将现实世界经过信息抽象变成信息模型(也称概念模型)。

(2) 信息模型(概念模型)经过数据抽象编程计算机实现。

这里,信息模型使用实体—联系模型描述,计算机实现使用数据模型表示。

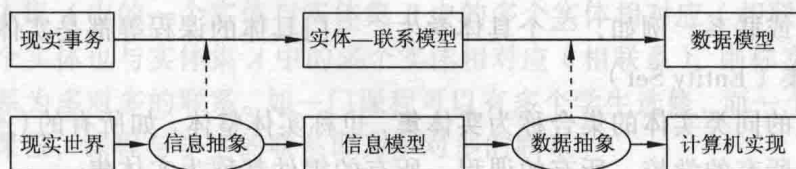


图 1-2 数据抽象过程

1.2.1 数据模型的三要素

数据模型是现实世界中的各种事物及其间的联系用数据和数据间的联系来表示的一种方法。一般地讲,数据模型是严格定义的概念的集合,这些概念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,数据模型通常由数据结构、数据操作和完整性约束三部分组成。

1. 数据结构

数据结构是所研究对象和对象具有的特性、对象间的联系的集合,它是对数据静态特性的描述。这些对象是数据库的组成部分。如网状模型中的数据项、记录、系型,关系模型中的域、属性、关系等。

在数据库系统中,通常按照数据结构类型来命名数据模型,如层次结构、网状结构和关系结构的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。通常对数据库的操作有检索和更新(包括插入、删除和



修改)两大类,这些操作反映了数据的动态特性。现实世界中的实体及实体间的联系是在不断变化的,数据模型应能反映出这种变化。

3. 数据的约束条件

数据的约束条件是完整性规则的集合。完整性规则是给定的数据模型中数据及其联系必须满足给定要求。例如,年龄的数据取值一般不能大于 150 岁。

1.2.2 概念模型与实体—联系方法

由图 1-2 可以看出,信息模型(概念模型)实际上是现实世界到机器世界的一个中间层次。

概念模型用于信息世界的建模,是现实世界到信息世界的第一层抽象,是数据库设计人员进行数据库设计的有力工具,也是数据库设计人员和用户之间进行交流的语言,因此,概念模型一方面应该具有较强的表达能力,能够方便、直接地表达不同应用中的各种实际知识,另一方面它还应该简单、清晰,易于用户理解。

1. 信息世界中的基本概念

信息世界涉及的概念主要有:

1) 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,一个具体学生、一门具体的课程等都是实体。

2) 实体集(Entity Set)

性质相同的同类实体的集合称为**实体集**,也称**实体整体**,如所有的(全体)学生、所有的汽车、所有的学校、所有的课程、所有的零件都称为实体集。

3) 实体型(Entity Type)

具有相同属性的实体必然具有共同的特征和性质。用实体名及其属性名的集合来抽象和刻画同类实体,称为**实体型**。例如,学生(学号,姓名,性别,出生年份,系,入学时间)就是一个实体型。

4) 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。属性有“型”和“值”的区分,如学生实体属性的名称(姓名、性别、年龄等)是属性的型,而属性的值是其型的具体内容,如王源、男、18 分别是姓名、性别、年龄的值。由此可以看到,事物的若干属性值的集合可表征一个实体,而若干属性型所组成的集合可表征一个实体的类型,简称“**实体型**”。同类型的实体集合组成**实体集**。

5) 关键字(Key)

能唯一标识实体的属性或属性集称为**关键字**(或**码**)。例如,学号是学生实体的关键字(码)。本书将混用关键字和码这两个概念。

6) 域(Domain)

属性的取值范围称为该属性的**域**。例如,学号的域为 8 位数字符号,年龄的域为

小于 128 的整数, 性别的域为 (男, 女)。

7) 联系 (Relationship)

在现实世界中, 事物内部以及事物之间是有联系的, 这些联系在信息世界中反映为实体 (型) 内部的联系和实体 (型) 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。实体之间的联系通常是指不同实体集之间的联系。

两个实体型之间的联系可以分为三类:

1) 一对一的联系 (1:1)

如果实体集 A 中的一个实体至多与实体集 B 中的一个实体相对应 (相联系), 反之亦然, 则称实体集 A 与实体集 B 的联系为一对一的联系。如一个学校只能有一个校长, 一个校长也只能在一个学校任职, 则学校与校长的联系即为一对一的联系, 还有班长与班、学生与座位之间也都是一对一的联系。

2) 一对多联系 (1:n)

如果实体集 A 中的一个实体与实体集 B 中的多个实体相对应 (相联系), 反之, 实体集 B 中的一个实体至多与实体集 A 中的一个实体相对应 (相联系), 则称实体集 A 与实体集 B 的联系为一对多的联系。如一个班级可以有多个学生, 而一个学生只会拥有一个班级, 班级与学生的联系即为一对多的联系。

3) 多对多联系 ($m:n$)

如果实体集 A 中的一个实体与实体集 B 中的多个实体相对应 (相联系), 而实体集 B 中的一个实体也与实体集 A 中的多个实体相对应 (相联系), 则称实体集 A 与实体集 B 的联系为多对多的联系。如一门课程可以有多个学生选修, 而一个学生同时可以选修多门课程, 课程与学生的联系即为多对多的联系。

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。可以用图形来表示两个实体型之间的这三类联系。

2. 概念模型的表示方法

为了在信息世界中简洁、清晰地描述现实世界的实体模型, 通常使用实体—联系图 (E-R 图) 描述。E-R 图是 P. P. S. Chen 于 1976 年提出的实体—联系模型 (E-R 模型) (Entity-Relationship Model)。E-R 图提供了实体、属性与联系的方法, 其图元符号如图 1-3 所示。在 E-R 图中, 实体集用矩形框表示, 并在矩形框里写上实体名。属性用椭圆框表示, 并在椭圆框里写上属性名。联系用菱形框表示, 并在菱形框里写上联系方式。

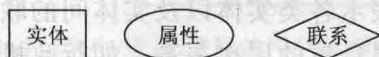


图 1-3 E-R 图图元符号

在图 1-4 中, 分别是学校与校长 (一对一)、班级与学生 (一对多)、学生与课程 (多对多) 的 E-R 实体模型图。班级、学生和课程对应的 E-R 图如图 1-5 所示, 由于实体班级、学生和课程属性太多, 我们这里只画出部分属性, 其中带下划线的属性表示对应实体的关键字。

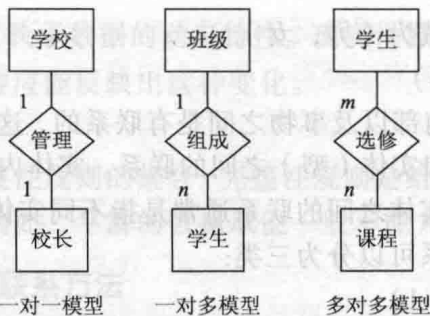


图 1-4 实体间的联系

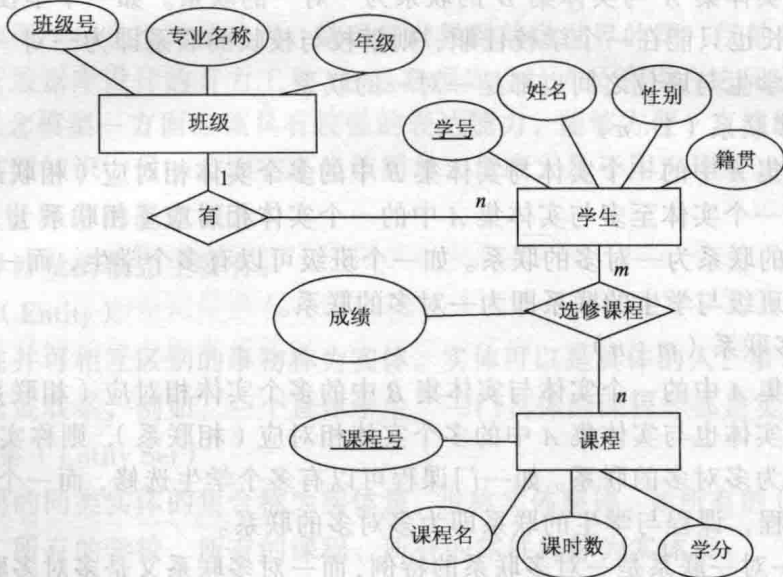


图 1-5 班级、学生和课程对应的 E-R 图（省略了部分属性）

1.2.3 数据模型

目前，数据库领域中最常用的数据模型有四种，它们是层次模型（Hierarchical Model）、网状模型（Network Model）、关系模型（Relational Model）、面向对象模型（Object Oriented Model）。

关系模型是本章的重点，将在 1.3 节中加以介绍。

1. 层次模型

层次模型用树形结构来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。

在层次模型中，每个结点表示一个记录类型，记录（类型）之间的联系用结点之间的连线（有向边）表示，这种联系是父子之间的一对多的联系。这就使得层次数据库系统只能处理一对多的实体联系。

每个记录类型可包含若干字段，这里，记录类型描述的是实体，字段描述的是实体的属性。各个记录类型及其字段都必须命名。各个记录类型、同一记录类型中各个字段不能同名。每个记录类型可以定义一个排序字段，也称码字段，如果定义该排序