

高等学校公共课计算机规划教材

Access 数据库 与程序设计

吴宏瑜 主编 夏 欣 副主编 伍良富 主审

<http://www.phei.com.cn>



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

高等学校公共课计算机规划教材

Access 数据库与程序设计

吴宏瑜 主编

夏欣 副主编

伍良富 主审



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书以示例驱动基本概念和基本操作,由浅入深、循序渐进地阐述 Access 2003 数据库与程序设计。内容包括:数据库系统、Access 2003 数据库管理系统、Access 数据表的建立与维护、创建 Access 查询、Access 的窗体设计、Access 的报表设计、创建数据访问页、宏、模块、上机实训。本书以“教学管理数据库”为主线,围绕数据应用,强调可阅读性,以求掌握操作方法和应用技能。全书在编写过程中力求通俗易懂、以图析文、直观生动。书中提供了大量的操作实例,第 1~9 章附有习题,能够帮助读者在较短的时间内迅速掌握 Access 2003 数据库技术。本书配有电子课件、源代码等教学资源。

本书可作为高等院校计算机基础课程“Access 数据库”的教材,也可供计算机技术培训人员和自学考试人员参考。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有,侵权必究。

图书在版编目(CIP)数据

Access 数据库与程序设计 / 吴宏瑜主编. —北京:电子工业出版社, 2011.2

高等学校公共课计算机规划教材

ISBN 978-7-121-12265-1

I. ①A… II. ①吴… III. ①关系数据库—数据库管理系统, Access 2003—高等学校—教材

IV. ①TP311.138

中国版本图书馆 CIP 数据核字(2010)第 221678 号

策划编辑:史鹏举

责任编辑:侯丽平

印 刷:北京市李史山胶印厂

装 订:

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本:787×1092 1/16 印张:17 字数:435.2 千字

印 次:2011 年 2 月第 1 次印刷

定 价:30.80 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线:(010)88258888。

前 言

本书以“教学管理数据库”为主线,以数据应用为基础,强调可阅读性,在编写过程中借鉴了一些先进教学理念和知识框架。本书在编写时所考虑的基本原则是“实例教学、抓住关键、解决问题”。具体内容安排如下。

第1章主要讲述关系数据库的基础知识。

第2章主要讲述 Access 2003 数据库的特点及安装、Access 2003 的启动与退出、操作界面设置、系统提供的联机帮助等内容。

第3章主要讲述如何创建数据库、创建新表、修改表的结构、设置表的主键与创建索引及排序、建立多表之间的关联关系、在数据表视图中的表操作和数据库窗口中的表操作等内容。

第4章主要讲述查询基础、查询设计器、各种相关的查询技术、查询向导、查询中的运算符与函数、使用 SQL 查询等内容。

第5章主要讲述窗体的组成、窗体的创建、窗体的属性、窗体中控件的使用和属性的设置、窗体的制作实例等内容。

第6章主要讲述报表的类型与视图、使用报表设计器和使用向导创建报表、在报表中进行分组与排序及汇总计算、报表的页面设置及打印等内容。

第7章主要讲述数据访问页的创建、数据访问页的设置等内容。

第8章主要讲述宏设计视图的使用,有关宏、宏组及条件宏的创建、调试与运行,设置宏的快捷方式等内容。

第9章主要讲述模块、面向对象的基本概念、VBA 程序设计基础、过程与函数、事件编程、VBE 编程环境、DoCmd 对象等内容。

第10章主要讲述上机实训。

本书根据项目驱动式教学的理念来组织教学内容,以理论联系实际的方法来分析要点难点,全书在编写过程中力求深入浅出、通俗易懂、以图析文、直观生动。书中提供了大量的操作实例,并配有丰富的实例图片,第1~9章附有习题,能够帮助读者在较短的时间内迅速掌握 Access 2003 数据库技术。本书配有电子课件、源代码等教学资源,有需要的读者可在华信教育资源网(<http://www.hxedu.com.cn>)免费下载。

本书主编为吴宏瑜,副主编为夏欣,主审为伍良富,编委有郭新明、何凯林、孙亚飞、向孟光,其中第1章、第2章、第7章、第10章由何凯林编写;第3章、第4章由夏欣编写;第5章、第6章由吴宏瑜编写;第8章、第9章由郭新民编写,孙亚飞、向孟光进行了校稿与上机验证。在此,向以上所有为本书出版做出贡献的同事表示感谢和敬意。

虽然我们希望为读者提供全面、准确、最新的计算机应用基础知识,但由于时间仓促,加之作者水平有限,书中难免会有不足和疏漏,欢迎读者来信提出意见和建议。

编 者

目 录

第1章 数据库系统	(1)	2.3.1 创建空数据库	(21)
1.1 数据库的基本概述	(1)	2.3.2 使用向导创建数据库	(21)
1.1.1 信息、数据与数据处理	(1)	2.3.3 数据库的打开与关闭	(23)
1.1.2 数据管理技术的发展	(1)	2.4 Access 的启动和退出	(24)
1.1.3 数据库	(2)	小结	(25)
1.2 数据库系统	(2)	习题2	(25)
1.2.1 数据库系统的组成	(2)	第3章 Access 数据表的建立与	
1.2.2 数据库系统的模式结构	(2)	维护	(27)
1.2.3 数据库系统的特点	(3)	3.1 数据表的构成	(27)
1.2.4 数据库管理系统	(4)	3.1.1 表的构成	(27)
1.3 数据模型	(5)	3.1.2 表的视图	(29)
1.3.1 概念模型	(5)	3.2 表的建立	(30)
1.3.2 数据模型	(7)	3.2.1 使用向导创建表	(30)
1.3.3 常用的数据模型	(7)	3.2.2 通过输入数据创建表	(31)
1.4 关系数据库	(9)	3.2.3 使用设计器创建表	(32)
1.4.1 关系模型的基本概念	(9)	3.2.4 字段属性的设置	(34)
1.4.2 关系代数	(10)	3.3 输入表中数据	(41)
1.4.3 关系范式	(13)	3.3.1 不同类型数据的输入	(41)
1.5 数据库的设计	(14)	3.3.2 数据的导入和导出	(42)
1.5.1 数据库设计的基本		3.4 建立表间关系	(44)
步骤	(14)	3.4.1 表间关系的建立	(44)
1.5.2 需求分析	(15)	3.4.2 编辑和删除关系	(46)
1.5.3 概念结构设计	(16)	3.5 数据表的维护	(47)
1.5.4 逻辑结构设计	(16)	3.5.1 表结构的修改	(47)
1.5.5 物理结构设计	(16)	3.5.2 编辑数据表中的记录	(48)
1.5.6 数据库的实施	(16)	3.5.3 数据表的修饰	(49)
1.5.7 数据库的运行与维护	(17)	3.6 数据表的操作、优化及	
小结	(17)	调整	(52)
习题1	(18)	3.6.1 查找或替换记录	(52)
第2章 Access 2003 数据库管理		3.6.2 排序记录	(54)
系统	(19)	3.6.3 筛选记录	(56)
2.1 Access 2003 的特点	(19)	小结	(58)
2.2 Access 2003 的数据库对象	(19)	习题3	(59)
2.3 创建 Access 数据库	(20)		

第4章 创建 Access 查询	(61)	5.2.3 在设计视图中创建窗体	(105)
4.1 查询概述	(61)	5.2.4 创建主窗体/子窗体	(109)
4.1.1 查询简介	(61)	5.2.5 窗体的基本操作	(113)
4.1.2 查询的类型	(61)	5.3 窗体的属性	(115)
4.1.3 查询准则的建立	(62)	5.3.1 设置窗体控件的属性	(116)
4.2 创建选择查询	(65)	5.3.2 “格式”属性	(117)
4.2.1 使用向导创建选择查询	(65)	5.3.3 “数据”属性	(118)
4.2.2 在设计视图中创建选择 查询	(67)	5.3.4 “事件”属性	(119)
4.2.3 在查询中进行计算	(70)	5.3.5 “其他”属性	(120)
4.3 创建交叉表查询	(74)	5.4 窗体的常用控件	(120)
4.3.1 使用向导创建交叉表 查询	(75)	5.4.1 控件的类型	(120)
4.3.2 在设计视图中创建交叉表 查询	(76)	5.4.2 常用控件的使用	(121)
4.4 创建参数查询	(78)	小结	(137)
4.5 创建操作查询	(80)	习题5	(137)
4.5.1 生成表查询	(80)	第6章 Access 的报表设计	(139)
4.5.2 更新查询	(81)	6.1 报表的基本概念	(139)
4.5.3 追加查询	(82)	6.1.1 报表的结构	(139)
4.5.4 删除查询	(85)	6.1.2 报表的类型	(140)
4.6 结构化查询语言 SQL	(86)	6.1.3 报表的视图	(140)
4.6.1 SQL 概述	(86)	6.2 报表的创建	(140)
4.6.2 SQL 语言的数据定义 功能(DDL)	(87)	6.2.1 自动创建报表	(140)
4.6.3 SQL 语言的数据操纵 功能(DML)	(89)	6.2.2 使用向导创建报表	(141)
4.6.4 SQL 语言的查询 功能(DQL)	(91)	6.2.3 使用图表向导创建报表	(142)
小结	(94)	6.2.4 使用标签向导创建报表	(143)
习题4	(95)	6.2.5 在设计视图中创建报表	(144)
第5章 Access 的窗体设计	(98)	6.2.6 建立主/子报表	(148)
5.1 窗体的基本概念	(98)	6.3 报表的常用控件	(149)
5.1.1 窗体的结构	(99)	6.4 报表的高级功能	(153)
5.1.2 窗体的类型	(99)	6.4.1 报表的排序与分组	(153)
5.1.3 窗体的视图	(100)	6.4.2 数据汇总	(155)
5.2 窗体的创建	(100)	6.5 报表的输出打印	(155)
5.2.1 使用向导创建窗体	(100)	小结	(156)
5.2.2 自动创建窗体	(104)	习题6	(158)
		第7章 创建数据访问页	(159)
		7.1 数据访问页的基本概念	(159)
		7.2 创建数据访问页	(159)
		7.2.1 自动创建数据访问页	(159)
		7.2.2 使用向导创建数据访 问页	(160)

7.2.3 利用已有的页创建数据访问页	(162)	9.1.3 将宏转换为模块	(180)
7.3 在设计视图中创建数据访问页	(162)	9.2 创建模块	(181)
7.3.1 创建数据访问页	(162)	9.3 VBA 程序设计基础	(181)
7.3.2 设置数据访问页控件	(164)	9.3.1 面向对象程序设计的基本概念	(182)
7.3.3 设置数据访问页主题	(166)	9.3.2 VBA 编程环境: VBE 界面	(186)
小结	(167)	9.3.3 VBA 编程基础: 常量、变量、运算符和表达式	(188)
习题 7	(167)	9.3.4 VBA 程序流程控制语句	(201)
第 8 章 宏	(169)	9.3.5 过程调用和参数传递	(212)
8.1 宏的概念	(169)	9.3.6 常用操作方法	(215)
8.1.1 宏的基本概念	(169)	9.3.7 VBA 的数据库编程	(223)
8.1.2 宏与 Visual Basic	(170)	9.4 程序调试	(236)
8.1.3 宏向 Visual Basic 程序代码转换	(171)	9.4.1 设置断点	(236)
8.2 宏的操作	(171)	9.4.2 调试工具使用	(236)
8.2.1 操作序列宏的创建	(171)	9.4.3 程序错误处理	(238)
8.2.2 宏组的创建	(171)	小结	(239)
8.2.3 条件操作宏	(172)	习题 9	(239)
8.2.4 宏的操作参数的设置	(173)	第 10 章 上机实训	(243)
8.2.5 宏的运行	(173)	实训 1 创建数据库	(243)
8.2.6 宏的调试	(174)	实训 2 创建数据表	(244)
8.2.7 常用宏操作	(175)	实训 3 创建查询	(248)
小结	(176)	实训 4 设计窗体	(251)
习题 8	(178)	实训 5 创建报表	(256)
第 9 章 模块	(180)	实训 6 宏的使用	(259)
9.1 模块的基本概念	(180)	实训 7 VBA 程序设计	(261)
9.1.1 类模块	(180)	附录 A 习题答案	(265)
9.1.2 标准模块	(180)		

第1章 数据库系统

本章主要介绍有关数据库、数据库管理系统与数据库系统的概念,数据模型与关系数据库的基本概念,数据库设计的过程与步骤。

1.1 数据库的基本概述

在信息时代,利用计算机对已收集的信息进行存储、加工处理已成为信息管理必不可少的手段,信息社会离不开信息管理。

1.1.1 信息、数据与数据处理

信息是客观事物状态和运动特征的一种普遍形式,客观世界中大量地存在、产生和传递着以这些方式表示出来的各种各样的信息。信息就是指以声音、语言、文字、图像、动画、气味等方式所表示的实际内容。

数据处理是对数据的加工和整理,即采集、存储、检索、加工、变换和传输。数据处理的基本目的是从大量的、可能是杂乱无章的、难以理解的数据中抽取并推导出对于某些特定的人们来说有价值、有意义的数据。数据处理技术的发展及其应用的广度和深度,极大地影响着人类社会发展的进程。

1.1.2 数据管理技术的发展

数据管理具体就是指人们对数据进行收集、组织、存储、加工、传播和利用的一系列活动的总和,经历了人工管理、文件管理、数据库管理三个阶段。每一阶段的发展以数据存储冗余不断减小、数据独立性不断增强、数据操作更加方便和简单为标志,各有各的特点。下面简单描述数据管理的三个阶段。

1. 人工管理阶段

在计算机出现之前,人们运用常规的手段从事记录、存储和对数据加工,也就是利用纸张来记录和利用计算工具(算盘、计算尺)来进行计算,并主要使用人的大脑来管理和利用这些数据。而早期的计算机主要用于数值计算,也无管理数据的软件,因此从计算机内记录的数据上看,数据量小,数据无结构。用户直接管理,且数据间缺乏逻辑组织,数据仅依赖特定的应用,缺乏独立性。

2. 文件系统阶段

这一阶段的数据管理技术得益于计算机的处理速度和存储能力的惊人提高,这一时期的数据处理系统是把计算机中的数据组织成相互独立的被命名的数据文件,并可按文件的名字来进行访问,对文件中的记录进行存取的数据管理技术。数据可以长期保存在计算机外存上,可以对数据进行反复处理,并支持文件的查询、修改、插入和删除等操作,这就是文件系统。文件系统实现了记录内的结构化,但从文件的整体来看却是无结构的。其数据面向特定的应用程序,因此数据共享性、独立性差,且冗余度大,管理和维护的代价也很大。

3. 数据库系统阶段

20 世纪 60 年代后期, 计算机性能得到进一步提高, 更重要的是出现了大容量磁盘, 存储容量大大增加且价格下降。在此基础上, 才有可能克服文件系统管理数据时的不足, 而满足和解决实际应用中多个用户、多个应用程序共享数据的要求, 从而使数据能为尽可能多的应用程序服务, 这就出现了数据库这样的数据管理技术。数据库的特点是数据不再只针对某一个特定的应用, 而是面向全组织, 具有整体的结构性, 共享性高, 冗余度低, 具有一定的程序与数据之间的独立性, 并且对数据进行统一的控制。

从文件系统到数据库系统, 标志着数据管理技术质的飞跃。20 世纪 80 年代后, 不仅在大、中型计算机上实现并应用了数据管理的数据库技术, 如 Oracle、Sybase、Informix 等, 在微型计算机上也可使用数据库管理软件, 如常见的 Access、FoxPro 等软件, 使数据库技术得到广泛应用和普及。

1.1.3 数据库

数据库(Database, DB)是依照某种数据模型组织起来并存放在二级存储器中的数据集合。这种数据集合具有如下特点: 尽可能不要重复, 以最优方式为某个特定组织的多种应用服务, 其数据结构独立于使用它的应用程序, 对数据的增、删、改和检索由统一软件进行管理和控制。简单来说, 数据库是“按照数据结构来组织、存储和管理数据的仓库”。在经济管理的日常工作中, 常常需要把某些相关的数据放进这样的“仓库”, 并根据管理的需要进行相应的处理。

1.2 数据库系统

数据库系统(Data Base System, DBS), 是由数据库及其管理软件组成的系统。它是为适应数据处理的需要而发展起来的一种较为理想的数据处理的核心机构。它是一个软件系统, 是存储介质、处理对象和管理系统的集合体, 一般是由数据库、数据库管理系统、数据库管理员和用户构成的。

1.2.1 数据库系统的组成

数据库系统一般由 4 个部分组成: ① 数据库, 即存储在磁带、磁盘、光盘或其他外存储介质上、按一定结构组织在一起的相关数据的集合。② 数据库管理系统(DBMS)。它是一组能完成描述、管理、维护数据库的程序系统。它按照一种公用的和可控制的方法完成插入新数据、修改和检索原有数据的操作。③ 数据库管理员(DBA)。④ 用户和应用程序。

1.2.2 数据库系统的模式结构

数据库系统分为三级模式结构, 由外模式、模式、内模式三级构成, 如图 1-1 所示。

1. 外模式

外模式又称子模式, 对应于用户级。它是某个或某几个用户所看到的数据库的数据视图, 是与某一应用有关的数据的逻辑表示。外模式是从模式导出的一个子集, 包含模式中允许特定用户使用的那部分数据。用户可以通过外模式描述语言来描述、定义对应于用户的数据记录(外模式)。

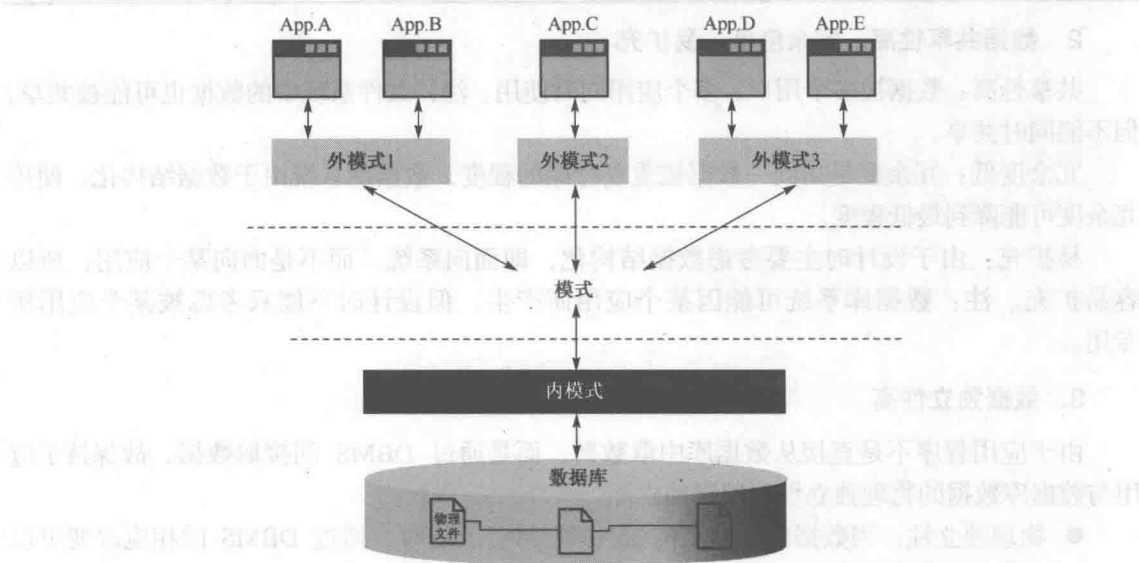


图 1-1 数据库系统的三级模式结构

外模式是保证数据库安全性的一种方式，每个用户只能看见和访问所对应的外模式中的数据，数据库中的其余数据是不可见的。

2. 模式

模式又称概念模式或逻辑模式。它是由数据库设计者综合所有用户的数据，按照统一的观点构造的全局逻辑结构，是对数据库中全部数据的逻辑结构和特征的总体描述，是所有用户的公共数据视图。

一个数据库只有一个模式，其中概念模式可以用实体-联系模型来描述，逻辑模式以某种数据模型为基础，综合考虑所有用户的需求，将其形成全局逻辑结构。

3. 内模式

内模式又称存储模式，它是数据库中全体数据的内部表示或底层描述，是数据库最低一级的逻辑描述，它描述了数据在存储介质上的存储方式和物理结构，对应着实际存储在存储介质上的数据库。内模式由内模式描述语言来描述、定义。

在一个数据库系统中，只有唯一的数据库，因而作为定义、描述数据库存储结构的内模式和定义、描述数据库逻辑结构的模式，也是唯一的，但建立在数据库系统之上的应用则是非常广泛、多样的，所以对应的外模式不是唯一的，也不可能是唯一的。

在三层模式结构中，数据库模式是数据库的核心和关键，外模式通常是模式的子集。数据按外模式的描述提供给用户，按内模式的描述存储在硬盘上，而模式介于外、内模式之间，既不涉及外部的访问，也不涉及内部的存储，有利于保持数据的独立性。

1.2.3 数据库系统的特点

1. 数据结构化

在文件系统中，尽管其记录内部已有了某些结构，但记录之间没有联系。而数据库系统则实现了整体数据的结构化，这是数据库系统的主要特征之一，也是数据库系统与文件系统的本质区别。

2. 数据共享性高、冗余度低、易扩充

共享性高：数据被多个用户、多个应用同时使用。注：文件系统中的数据也可能被共享，但不能同时共享。

冗余度低：冗余度是指同一数据被重复存储的程度，数据库系统由于数据结构化，使得冗余度可能降到最低程度。

易扩充：由于设计时主要考虑数据结构化，即面向系统，而不是面向某个应用，所以容易扩充。注：数据库系统可能因某个应用而产生，但设计时不能只考虑被某个应用所专用。

3. 数据独立性高

由于应用程序不是直接从数据库中取数据，而是通过 DBMS 间接取数据，故保持了应用与数据库数据的物理独立性和逻辑独立性。

- 物理独立性：当数据的存储结构(或物理结构)改变时，通过 DBMS 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。
- 逻辑独立性：当数据的总体逻辑结构改变时，通过 DBMS 的相应改变可以保持数据的局部逻辑结构不变，应用程序是依据数据的局部逻辑结构编写的，所以应用程序不必改变。

4. 数据由 DBMS 统一管理和控制

由于数据库的共享是并发的共享，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。为此，DBMS 必须提供以下几方面的数据控制功能。

- 数据的安全性保护：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对某些数据以某种方式进行访问和处理。
- 数据的完整性检查：指数据的正确性、有效性和相容性，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。

正确性：如输入成绩时，应该输入数值，而实际输入了字符，即不正确。

有效性：如输入年龄时，应该输入 0~150 之间的数据，而实际输入了-8，即无效。

相容性：指表示同一个事实的两个数据应该一致。

- 并发控制：指控制多个用户同时存取、修改数据库中的数据，以保证数据库的完整性。例如，多个用户可以同时读，但同一时间只能允许一个用户写数据。注：由于计算机的高速且分时操作，用户可能感觉到多个用户可以同时写数据。
- 数据库恢复：指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)的功能。计算机系统的硬件故障、软件故障、操作员的失误及故意的破坏均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。

1.2.4 数据库管理系统

数据库管理系统(Database Management System, DBMS)是一种操纵和管理数据库的大型软件，用于建立、使用和维护数据库。它对数据库进行统一的管理和控制，以保证数据库的安全性和完整性。用户通过 DBMS 访问数据库中的数据，数据库管理员也通过 DBMS 进行数据库的维护工作。它提供多种功能，可使多个应用程序和用户用不同的方法在同一时刻或

不同时刻去建立、修改和询问数据库。它使用户能方便地定义和操纵数据,维护数据的安全性和完整性,以及进行多用户下的并发控制和恢复数据库。

DBMS 的主要目标是使数据作为一种可管理的资源来处理,其主要功能如下。

1. 数据定义

DBMS 提供数据定义语言,供用户定义数据库的数据对象。

2. 数据操作

DBMS 提供数据操作语言,供用户实现对数据的操作。

3. 数据库的运行管理

数据库的运行管理功能是 DBMS 的运行控制、管理功能,包括多用户环境下的并发控制、安全性检查和存取限制控制、完整性检查和执行、运行日志的组织管理、事务的管理和自动恢复,即保证事务的原子性。这些功能保证了数据库系统的正常运行。

4. 数据库的保护

因为数据库中的数据是信息社会的战略资源,所以数据库的保护至关重要。DBMS 对数据库的保护通过 4 个方面来实现:数据库的恢复、数据库的并发控制、数据库的完整性控制、数据库的安全性控制。

5. 数据库的维护

这一部分包括数据库的数据载入、转换、转储、数据库的重构及性能监控等功能,这些功能分别由各个使用程序来完成。

1.3 数据模型

数据(Data)是描述事物的符号记录。模型(Model)是现实世界的抽象。数据模型(Data Model)是数据特征的抽象,是数据库管理的教学形式框架。数据模型是数据库系统中用以提供信息表示和操作手段的形式构架。人们常常将现实社会抽象为信息世界,然后再将信息世界转化为机器世界。换言之,先将现实世界中的客观对象抽象为某一种信息结构,这种信息结构常常叫做概念模型,图 1-2 所示为现实世界中客观对象的转换过程。

1.3.1 概念模型

概念模型是对客观事物及其联系的抽象,是数据库设计者进行数据库设计的很重要的工具。概念模型是现实世界到机器世界的一个中间层次。要了解它就先要了解几个常见术语。

1. 实体

客观上实实在在存在并可以相互区分的事物称为实体。例如,一位老师、一门功课等。

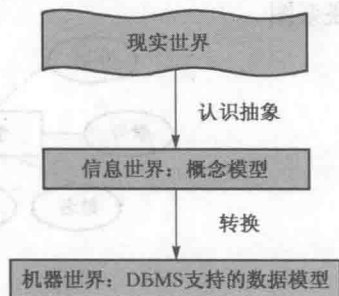


图 1-2 现实世界中客观对象的转换过程

2. 属性

实体的特征。属性的差异能使我们区分出同类的实体。比如公司员工可以由编号、姓名、性别、年龄等属性组成，这些属性组合起来就表示一名员工。

3. 域

属性的取值集合。比如性别的域为男、女。

4. 关键字

能在实体集中唯一标识一个实体的属性，如身份证号码、学号、借书证号等。

现实世界中的客观事物往往不是孤立存在的，我们还需要将现实世界中相关事物之间错综复杂的关系抽象为概念世界中实体之间的关联。从参与关联的两个实体集的数量关系来说，一般情况下实体之间的关联可以分为三种：“一对一”、“一对多”和“多对多”。

“一对一”联系(1:1)表示为假设有 A、B 两个实体集，如果 A 中的每个实体至多和 B 中的一个实体有联系，反之，B 中的每个实体至多和 A 中的一个实体有联系，则称 A 对 B 或者 B 对 A 是一对一联系。例如，一所学校有一个校长，而一个校长又只能在一所学校里面任职，这样学校和校长就是一对一的联系。

“一对多”联系(1:N)表示为假设有 A、B 两个实体集，如果 A 中的每个实体可以和 B 中的几个实体有联系，而 B 中的每个实体至多和 A 中的一个实体有联系，则称 A 对 B 是一对多的联系。例如，一所学校有若干学生，而每个学生只能在一所学校里面学习，那么学校和学生之间就是一对多的关系。

“多对多”联系(M:N)表示为假设有 A、B 两个实体集，如果 A 中的每个实体可以和 B 中的多个实体有联系，而 B 中的每个实体也可以和 A 中的多个实体有联系，那么就称 A 对 B 或 B 对 A 是多对多联系。例如，一名学生可以选修多门课程，一门课程有多名学生选修，学生和课程之间就是多对多的联系。

概念模型的表示方法中，最著名的是 E-R(实体-联系)方法，它是用来描述实体集、属性和联系的图形。图中每种元素都用结点表示。我们用特殊形状的结点来表示特定的元素类别：矩形表示实体集；椭圆表示属性；菱形表示联系，同时要在无向边旁边标上联系的类型(1:1, 1:N, M:N)；用实线来连接实体集与它的属性以及联系与它的实体集，如图 1-3 所示的实体-联系图。

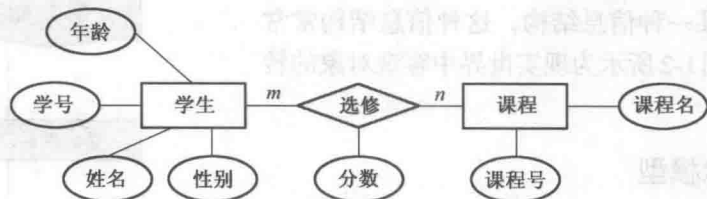


图 1-3 实体-联系图

从图 1-3 中可知一共有两个实体，学生实体的属性分别是年龄、姓名、学号、性别，课程实体的属性有课程名、课程号。

其中学生实体和课程实体之间有“选修”联系，是 M:N 联系，并且选修也有自己的属性，那就是分数。

1.3.2 数据模型

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。通俗来讲，数据模型就是对现实世界的模拟、描述或表示。通常一个数据模型由三部分组成：数据结构、数据操作、数据的约束条件。

1. 数据结构

数据结构用于描述系统的静态特性。数据结构不仅要描述数据本身，还要描述数据之间的联系。数据结构是刻画一个数据模型性质最重要的方面。

2. 数据操作

数据操作用于描述系统的动态特性，包括操作及有关的操作规则。数据库的主要操作有插入、删除、修改和查询。

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。完整性规则是数据模型中的数据及其联系所具有的约束规则，用来限定数据库状态及状态的变化，以保证数据的正确。

1.3.3 常用的数据模型

目前，数据库领域最常用的数据库模型有三种，它们是层次模型、网状模型、关系模型，前两种为非关系模型。关系模型是用二维表来表示数据之间联系的模型。

1. 层次模型

层次模型是数据库系统中最早出现的数据模型。20世纪60年代后期，IBM开发出IMS DBMS，是层次数据模型的基础。

层次数据模型用树形结构表示各类实体及实体之间的联系。现实世界中许多实体之间的联系就呈现出一种很自然的层次关系，如行政机构、家庭关系等。层次模型是以记录为结点的有向树。

按照树的定义，层次模型有以下两个限制：

- (1) 只有一个结点没有双亲结点，即根结点。
- (2) 根以外的其他结点有且只有一个双亲结点。

因此，层次数据库系统只能处理一对多的实体关系。

层次模型中，每个结点表示一个记录类型，结点之间的连线表示记录类型间的联系。这种联系只能是父子联系。

层次模型的另一个最基本的特点是：任何一个给定的记录值只有按其路径查看时，才能显出它的全部意义，没有一个子女记录值能够脱离双亲记录值而独立存在。

层次模型的优点：层次模型的数据结构比较简单清晰，查询效率高，性能优于关系模型，不低于网状模型，层次数据模型提供了良好的完整性支持。

层次模型的缺点：无法直观地表示多对多联系，对插入和删除操作的限制多，应用程序的编写比较复杂，查询子女结点必须通过双亲结点，由于结构严密，层次命令趋于程序化。

2. 网状模型

用层次模型表示非树形结构很不直接,网状模型则可克服这一弊病。网状数据模型的典型代表是 DBTG 系统,也称 CODASYL 系统。

网状模型比层次模型更具普遍性。它去掉了层次模型的两个限制,允许结点有多个双亲结点,比层次模型更能直接地描述现实世界。

网状数据库系统对数据操纵加了一些限制,提供了一定的完整性约束。网状数据模型的操纵主要包括查询、插入、删除和更新数据。允许插入尚未确定双亲结点值的子女结点值。删除操纵允许只删除双亲结点值而不删除子女结点值。

网状模型的优点:能够更为直接地描述现实世界,如一个结点可以有多个双亲、允许结点之间为多对多的联系。具有良好的性能,存取效率较高。

网状模型的缺点:结构比较复杂,而且随着应用环境的扩大,数据库的结构就变得越来越复杂,不利于最终的用户掌握。DDL、DML 语言复杂,用户不容易使用。

3. 关系模型

将概念世界中的概念模型(E-R 模型)进一步转化为机器世界中便于数据库管理系统处理的数据模型,转换原则如下:

(1) 每一个实体型转换为一个关系模式,实体的属性就是关系的属性,实体的关键字就是关系的关键字。

(2) 一个 1:1 的联系转换为关系模式,可以在两个实体型转换成的两个关系模式中的任意一个关系模式的属性中加入另一个关系模式的关键字和联系类型的属性。

(3) 一个 1:N 的联系转换为关系模式,可以在 N 端实体型转换成的关系模式中加入 1 端实体型的关键字和联系类型的属性。

(4) 一个 M:N 的联系转换为一个关系模式,其属性为两端实体型的关键字加上联系类型的属性。

将学生成绩管理系统的 E-R 模型转换为学生表、课程表和成绩表三张表。学生表和课程表之间多对多的关联需要通过成绩表建立两个一对多关系来实现,它的主键包括两个字段:

“课程编号”字段与“学号”字段,分别来源于学生表和课程表的外部关键字。这样一来,数据之间“多对多”的关系就一目了然,如图 1-4 所示。

学生表: 表				
学号	姓名	性别	民族	
110	田野牧歌	男	汉族	
111	黎明	男	汉族	
112	王敏	女	汉族	
113	张平	男	汉族	
114	李小明	男	汉族	
115	王艳梅	女	汉族	
116	马力	男	回族	
117	王红	女	汉族	
118	王敏	男	羌族	
119	李建平	男	汉族	
120	段阿鹏	男	白族	
121	王燕茹	女	满族	
122	关健	男	汉族	
123	李阿林	男	汉族	
124	斯琴高娃	女	蒙古族	
125	央金	女	藏族	
126	才旦卓玛	女	藏族	

课程表: 表	
课程编号	课程名称
101	英语
102	高等数学
103	大学语文
104	数据库技术
105	体育

成绩表: 表				
选课ID	学号	课程编号	课程成绩	
1	110	101	90	
2	110	102	88	
3	110	103	80	
4	110	104	95	
5	110	105	85	
6	111	101	90	

图 1-4 学生成绩管理系统的关系模型

关系数据模型有下列优点:

(1) 建立在严格的数学概念的基础上。

(2) 概念单一, 实体和实体间的各类联系都用关系来表示, 对数据的检索结果也是关系。

(3) 关系模型的存取路径对用户透明, 具有更高的数据独立性、更好的安全保密性, 简化了程序员的工作和数据库开发建立的工作。

关系数据模型的缺点: 存取路径对用户透明导致查询效率往往不如非关系数据模型。为了提高性能, 必须对用户的查询请求进行优化, 增加了开发 DBMS 的难度。

1.4 关系数据库

关系数据库系统是支持关系模型的数据库系统。关系模型的数据结构单一, 现实世界的实体及实体之间的各种联系均用关系来表示。在关系模型中, 实体及实体之间的联系也是用关系来表示的。例如, 导师实体、研究生实体、导师与研究生之间的一对多联系都可以分别用一个关系来表示。在一个给定的应用领域中, 所有实体及实体之间联系的关系的集合构成一个关系数据库。

关系数据库也有型和值之分。关系数据库的型也称为关系数据库模式, 是对关系数据库的描述, 它包括若干域的定义以及在这些域上定义的若干关系模式。关系数据库的值是这些关系模式在某一时刻对应的关系的集合, 通常称为关系数据库。

1.4.1 关系模型的基本概念

数据库的数据模型先后经历了层次模型、网状模型、关系模型和面向对象模型等阶段。在现实生活中, 表达数据之间关联性最常用、最直观的方法就是制作表格, 这些表通俗易懂。比如图 1-5 就是一个学生信息二维表。

关系模型就是将数据组织成二维表的形式, 通过一张二维表来描述实体的属性、描述实体间联系的数据模型。

下面就来了解一些关系模型的常见术语。

关系: 以属性分类的、包含若干行的二维表称为“关系模型的关系”。一个关系就是一张二维表, 每一个关系可选取一个关系名。

记录: 二维表中的行称为“记录”, 或者称为“元组”。一行为一个记录。记录用来表示现实世界中某一个实体集合中不同的、独立的个体事物。

属性、值、域: 每一个记录由若干列组成, 横向排列一个事物实体的多种属性, 在二维表中列对应实体的属性(或称为字段), 列名就是属性名。例如, (学号, 姓名, 性别, 民族)描述了学生实体的 4 个属性。值为记录中各属性的具体数值, 例如, (计算机学院, 计算机科学与工程系, 张东, 男, 20), 它描述了学生张东的信息。域为属性的取值范围, 也就是说不同记录对应同一个属性的取值所限定的范围, 例如“性别”的域为(男, 女)。

关键字(主码、主键): 关键字由二维表的属性或属性组合组成, 其值能够唯一地标识一个记录。例如, 学生关系中的学号, 课程表中的课程号。

外码: 如果一个关系中的属性或属性组并非该关系的主码, 但它们是另一个关系的主码, 则称其为该关系的外码。

全码: 关系模型的所有属性组是这个关系模式的候选码, 称为全码。

学号	姓名	性别	民族
110	田野牧歌	男	汉族
111	黎明	男	汉族
112	王敏	女	汉族
113	张平	男	汉族
114	李刘强	男	汉族
115	王艳梅	女	汉族
116	马力	男	回族
117	王红	女	汉族
118	王敏	男	羌族
119	李建平	男	汉族
120	段阿鹏	男	白族
121	王燕茹	女	满族
122	关健	男	汉族
123	李阿林	男	汉族
124	斯琴高娃	女	蒙族
125	央金	女	藏族
126	才旦卓玛	女	藏族

图 1-5 学生信息二维表

1.4.2 关系代数

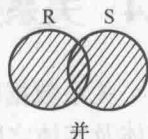
数据库中关系的基本运算有两类：一类是传统的集合运算(并、差、交等)，另一类是专门的关系运算(选择、投影、连接等)。

1. 传统的集合运算

传统的集合运算是二目运算，包括并、差、交、广义笛卡儿积四种运算。设关系 R 和关系 S 具有相同的目 n (即两个关系都有 n 个属性)，且相应的属性取自同一个域，则可以定义并、差、交运算如下。

1) 并(Union)

关系 R 与关系 S 的并记做： $R \cup S = \{t | t \in R \vee t \in S\}$ ，其结果仍为 n 目关系，由属于 R 或属于 S 的元组组成，如图1-6所示为并运算。



并

图 1-6 并运算

【例 1-1】有两个学生关系表 R (见表 1-1)、 S (见表 1-2)，并运算结果见表 1-3。

表 1-1 关系 R

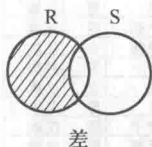
学生	姓名	性别
110	田野牧歌	男
112	王敏	女

表 1-2 关系 S

学生	姓名	性别
112	王敏	女
113	张平	男

表 1-3 并运算结果

学生	姓名	性别
110	田野牧歌	男
112	王敏	女
113	张平	男



差

2) 差(Difference)

关系 R 与关系 S 的差记做： $R - S = \{t | t \in R \wedge t \notin S\}$

其结果关系仍为 n 目关系，由属于 R 而不属于 S 的所有元组组成，如图1-7所示。

图 1-7 差运算

【例 1-2】有两个学生关系表 R (见表 1-4)、 S (见表 1-5)，差运算结果见表 1-6。

表 1-4 关系 R

学生	姓名	性别
110	田野牧歌	男
112	王敏	女

表 1-5 关系 S

学生	姓名	性别
112	王敏	女
113	张平	男

表 1-6 差运算结果

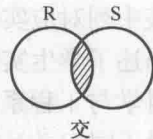
学生	姓名	性别
110	田野牧歌	男

3) 交(Intersection)

关系 R 与关系 S 的交记做： $R \cap S = \{t | t \in R \wedge t \in S\}$

其结果关系仍为 n 目关系，由既属于 R 又属于 S 的元组组成，如图1-8所示。关系的交可以用差来表示，即 $R \cap S = R - (R - S)$ 。

【例 1-3】有两个学生关系表 R (见表 1-7)、 S (见表 1-8)，交运算结果见表 1-9。



交

图 1-8 交运算

4) 广义笛卡儿积

两个分别为 n 目和 m 目的关系 R 和 S 的广义笛卡儿积是一个 $(n + m)$ 列的元组的集合。元组的前 n 列是关系 R 的一个元组，后 m 列是关系 S 的一个元组。若 R 有 k_1 个元组， S 有 k_2 个元组，则关系 R 和关系 S 的广义笛卡儿积有 $k_1 \times k_2$ 个元组。