



普通高等教育“十二五”规划教材

Access 基础教程 (第四版)



主 编 于繁华 李 民
副主编 贾学婷 赵 东



中国水利水电出版社
www.waterpub.com.cn

普通高等教育“十二五”规划教材

Access 基础教程（第四版）

主 编 于繁华 李 民

副主编 贾学婷 赵 东



中国水利水电出版社
www.waterpub.com.cn

内 容 提 要

本书是按照教育部文科专业大学计算机教学基本要求编写的教材,同时适用于教育部提出的非计算机专业计算机基础教学三层次的要求。

本书以 Access 2010 为环境,介绍了包括数据库基础知识、表、查询、窗体、报表、宏、VBA 和 Web 数据库等对象的功能及使用方法。为培养学生的数据库应用开发能力,还给出了一个利用 Access2010 开发数据库应用系统的综合实例。另外本书配套有相应的习题与实验指导书,以详尽细致的实验内容辅助读者对有关操作进行系统训练。

本书配套有《Access 基础教程(第四版)习题与实验指导》,以详尽细致的实验内容辅助读者对有关操作进行系统训练。

本书结构严谨、可操作性和实用性强,既可以作为高等学校非计算机专业的教材,也可以作为全国计算机等级考试考生的培训辅导参考书。

本书配有电子教案,读者可以从中国水利水电出版社网站和万水书苑上下载,网址为:
<http://www.waterpub.com.cn/softdown/>和 <http://www.wsbookshow.com>。

图书在版编目(CIP)数据

Access基础教程 / 于繁华, 李民主编. -- 4版. --
北京: 中国水利水电出版社, 2013. 8
普通高等教育“十二五”规划教材
ISBN 978-7-5170-1072-2

I. ①A… II. ①于… ②李… III. ①关系数据库系统
—高等学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2013)第168608号

策划编辑: 石永峰 责任编辑: 李 炎 加工编辑: 于杰琼 封面设计: 李 佳

书 名	普通高等教育“十二五”规划教材 Access 基础教程(第四版)
作 者	主 编 于繁华 李 民 副主编 贾学婷 赵 东
出版发行	中国水利水电出版社 (北京市海淀区玉渊潭南路1号D座 100038) 网址: www.waterpub.com.cn E-mail: mchannel@263.net (万水) sales@waterpub.com.cn 电话: (010) 68367658 (发行部)、82562819 (万水)
经 售	北京科水图书销售中心(零售) 电话: (010) 88383994、63202643、68545874 全国各地新华书店和相关出版物销售网点
排 版	北京万水电子信息有限公司
印 刷	三河市铭浩彩色印装有限公司
规 格	184mm×260mm 16开本 18.75印张 460千字
版 次	2004年8月第1版 2004年8月第1次印刷 2013年8月第4版 2013年8月第1次印刷
印 数	0001—4000册
定 价	36.00元

凡购买我社图书,如有缺页、倒页、脱页的,本社发行部负责调换

版权所有·侵权必究

前 言

在《Access 基础教程（第三版）》出版的五年时间里，本书入选了普通高等教育“十一五”国家级规划教材，很多高校选用此系列丛书作为非计算机专业学生的计算机应用基础课程方面的教材，并对教材体系提出了很多建设性意见。在充分考虑这些意见的基础上，并结合 Access 版本的升级与变化，我们对第三版教材做了较大篇幅的调整。

第四版主要的修改为：

（1）将本书的实例环境全面升级为 Access 2010。

（2）调整第 1 章和第 2 章的顺序，有助于读者从数据库最基础的知识全面掌握 Access 2010 的功能与操作。

（3）第 2 章 Access 2010 基础中，介绍了 Access 2010 新特性及数据库的操作环境；在 Web 数据库中突出了 Access 对象变化及其创建过程，

（4）第 3 章中对新增数据类型做出了简要说明，同时对 Access 2010 新增的汇总行及其功能给出了实例，方便读者对新版本 Access 2010 的理解和掌握。

（5）第 5 章增加了 Access 新增布局视图的介绍，并对新版本 Access 2010 关于创建窗体的方法和过程给出了实例，供读者参考。

（6）第 6 章关于图表报表的创建方法，Access 2010 做出了较大调整，本书以图表控件的形式来实例说明。

（7）第 7 章宏是本书变动相对较大的一章，在 Access 2010 中 Microsoft 对宏做出了重要变更，使得宏的设计更接近于代码的设计过程，同时新增了嵌入宏和数据宏，这是读者需要注意的。

（8）本书删去了《Access 基础教程（第三版）》中的第 7 章“数据访问页”的内容，因为 MicroSoft Office 对 Access 基本对象设置及功能做出了调整，本书在第 10 章中对 Web 数据库进行了介绍。

本书配套有《Access 基础教程（第四版）习题与实验指导》，以详尽细致的实验内容辅助读者对有关操作进行系统训练。

本书由于繁华、李民担任主编，贾学婷、赵东担任副主编。在编写过程中，王春艳、陈然、晏愈光、杨鑫、刘研、姜艳、孙英娟、孙晔、翟雷、张未名、孔祥营、朱海泉、赵宇、李晓宁、张志勇、肖明亮等也参与了书稿内容的讨论，并提出了许多宝贵的意见和建议，在此一并表示感谢！

由于作者水平有限，书中难免有不足之处，敬请读者批评指正。

编 者

2013 年 6 月

目 录

前言

第1章 数据库基础知识	1
1.1 数据库系统的基本概念	1
1.1.1 数据、数据库、数据库管理系统	1
1.1.2 数据库技术的发展	2
1.1.3 数据库系统的特点	5
1.1.4 数据库系统的体系结构	5
1.2 数据模型	6
1.2.1 E-R 模型	7
1.2.2 层次模型	8
1.2.3 网状模型	9
1.2.4 关系模型	10
1.2.5 面向对象数据模型	11
1.3 关系数据库系统	11
1.3.1 关系模型的组成	12
1.3.2 关系运算的基本概念	12
1.3.3 关系数据库管理系统的功能	15
1.3.4 常见的关系数据库管理系统及分类	15
1.3.5 关系数据库管理系统——Access	15
1.4 关系数据库标准语言 SQL	15
1.4.1 SQL 的特点	16
1.4.2 SQL 的数据查询和数据操纵功能	16
1.5 关系数据库设计	18
1.5.1 数据库关系完整性设计	18
1.5.2 数据库规范化设计	19
1.5.3 Access 数据库应用系统设计实例	22
本章小结	25
习题	25
第2章 Access 2010 基础	27
2.1 Access 2010 简介	27
2.2 Access 2010 基本对象	27
2.3 Access 2010 的新特点	29
2.4 Access 2010 的操作环境	31
2.5 创建数据库	36

本章小结	38
习题	38
第3章 表	39
3.1 创建表	39
3.2 数据类型和字段属性	43
3.2.1 数据类型	43
3.2.2 字段属性	45
3.3 字段编辑操作	53
3.4 主键和索引	54
3.4.1 主键	54
3.4.2 索引	55
3.5 表的联接	56
3.5.1 定义表之间的关系	56
3.5.2 编辑关系	59
3.5.3 参照完整性定义	60
3.6 输入和编辑数据	61
3.6.1 数据的输入	61
3.6.2 编辑记录	65
3.6.3 调整表的外观	72
3.6.4 数据表中的汇总行	77
本章小结	78
习题	78
第4章 查询	79
4.1 查询简介	79
4.2 查询视图	80
4.2.1 数据表视图	80
4.2.2 设计视图	80
4.2.3 SQL 视图	82
4.3 使用查询向导创建查询	82
4.3.1 简单查询向导	82
4.3.2 交叉表查询向导	84
4.3.3 查找重复项查询向导	87
4.3.4 查找不匹配项查询向导	89

4.4 查询条件	91	5.5.2 创建多页(屏)窗体	134
4.5 对查询进行编辑	94	5.6 窗体中的常用控件	135
4.5.1 编辑查询中的字段	94	5.6.1 标签	135
4.5.2 运行查询	96	5.6.2 文本框	136
4.5.3 排序查询的结果	96	5.6.3 组合框和列表框	136
4.6 选择查询	96	5.6.4 命令按钮	138
4.6.1 创建选择查询	96	5.6.5 选项组	140
4.6.2 查询中的计算功能	98	5.7 窗体和控件的属性	140
4.7 参数查询	102	5.7.1 常用的格式属性	140
4.8 操作查询	103	5.7.2 常用的数据属性	143
4.8.1 保护数据	104	5.7.3 常用的事件属性	144
4.8.2 更新查询	104	5.7.4 常用的其他属性	146
4.8.3 追加查询	106	5.8 在窗体上放置控件	146
4.8.4 删除查询	108	5.8.1 使用“控件”在窗体上放置控件	146
4.8.5 生成表查询	109	5.8.2 使用字段列表放置控件	147
4.9 SQL 查询	110	5.9 为控件命名	147
4.9.1 SQL 查询视图的切换	111	5.10 控件的尺寸统一与对齐	148
4.9.2 联合查询	111	本章小结	149
4.9.3 传递查询	113	习题	149
4.9.4 数据定义查询	113	第6章 报表	150
4.9.5 子查询	114	6.1 认识报表	150
4.9.6 用 SQL 语句实现各种查询	115	6.1.1 报表的作用	150
本章小结	116	6.1.2 报表的类型	150
习题	116	6.1.3 报表的节	150
第5章 窗体	117	6.1.4 “报表设计工具”选项卡	151
5.1 窗体简介	117	6.2 创建和修改报表	151
5.1.1 窗体的种类	117	6.2.1 使用“报表”创建报表	152
5.1.2 窗体的视图	118	6.2.2 使用“空报表”创建报表	152
5.2 创建窗体	120	6.2.3 使用“报表向导”创建报表	153
5.2.1 使用“窗体”创建窗体	121	6.2.4 在设计视图中创建和修改报表	155
5.2.2 使用“多个项目”创建窗体	121	6.3 报表的排序、分组和计算	157
5.2.3 创建“分割窗体”	122	6.3.1 报表的排序与分组	157
5.2.4 创建数据透视图窗体	122	6.3.2 报表的计算	160
5.2.5 使用窗体向导创建窗体	123	6.4 创建图表报表	162
5.2.6 使用“空白窗体”创建窗体	126	6.5 创建标签报表	165
5.3 窗体的设计视图	127	6.6 打印报表和创建多列报表	167
5.4 子窗体	129	6.7 创建和链接子报表	168
5.5 创建多页或多选项卡窗体	132	本章小结	169
5.5.1 创建多选项卡窗体	132	习题	170

第7章 宏	171	8.5.1 Sub 子过程	211
7.1 宏的概念	171	8.5.2 Function 函数过程	212
7.2 宏的结构	172	8.5.3 Property 过程	212
7.3 宏选项卡和设计器	173	8.6 文件	213
7.4 创建宏与宏操作	174	8.7 API 函数与 ActiveX 数据对象	214
7.4.1 创建独立宏	174	8.7.1 API 函数	214
7.4.2 创建子宏	175	8.7.2 ActiveX 数据对象 (ADO)	215
7.4.3 设置宏操作	177	8.7.3 数据访问和处理的特殊函数	216
7.4.4 设置宏操作参数	179	8.8 常用操作方法	217
7.4.5 在宏中使用条件	181	8.8.1 消息框	217
7.5 创建嵌入宏	182	8.8.2 输入框	217
7.6 创建数据宏	185	8.8.3 计时器	218
7.7 宏的调试和运行	186	本章小结	220
7.7.1 宏的调试	186	习题	220
7.7.2 宏的运行	187	第9章 VBA 应用实例	221
本章小结	188	9.1 程序流程控制	221
习题	188	9.1.1 选择结构	221
第8章 VBA 程序设计基础	189	9.1.2 循环结构	225
8.1 模块	189	9.2 文件	231
8.1.1 标准模块	189	9.3 过程调用	231
8.1.2 类模块	189	9.4 计时器 Timer	235
8.1.3 宏和模块	190	9.5 ADO 数据库编程实例	236
8.2 面向对象程序设计基础	191	本章小结	238
8.2.1 面向对象程序设计的基本概念	191	习题	238
8.2.2 VBA 的开发环境 VBE	193	第10章 Web 数据库	239
8.2.3 VBA 程序的调试	193	10.1 创建 Web 数据库	239
8.2.4 VBA 程序运行错误处理	194	10.2 设计 Web 表	243
8.3 VBA 编程基础	195	10.3 创建 Web 查询	247
8.3.1 常量	195	10.4 创建 Web 窗体	249
8.3.2 变量	196	10.5 Web 数据库设计说明	252
8.3.3 表达式	198	本章小结	253
8.3.4 用户定义的数据类型	204	习题	253
8.3.5 数组	205	第11章 Access 应用程序设计	254
8.3.6 数据库对象变量	205	11.1 系统分析与设计	254
8.4 VBA 程序流程控制	206	11.2 数据库的设计	255
8.4.1 程序语句书写	206	11.2.1 数据库的需求分析	255
8.4.2 选择结构	206	11.2.2 数据库的结构设计	255
8.4.3 循环结构	209	11.3 系统功能概述	257
8.5 子过程与函数过程	211	11.4 创建数据表和索引	258

11.4.1	创建表格	258
11.4.2	创建主键和关系	260
11.5	设计窗体	262
11.5.1	创建切换面板	262
11.5.2	创建登录窗体	268
11.5.3	创建数据录入窗体	269
11.5.4	创建查询修改窗体	277
11.6	编码实现	280
11.6.1	公用模块	280
11.6.2	登录窗体的代码	281

11.6.3	数据录入窗体代码	283
11.6.4	查询修改窗体的代码	284
11.7	系统的调试及发布	286
11.7.1	系统分析与调试	286
11.7.2	数据库启动选项设置	287
11.7.3	生成 ACCDE 文件	289
	本章小结	290
	习题	290
	参考文献	291

第1章 数据库基础知识

数据库技术是计算机领域的一个重要分支。在计算机应用的三大领域（科学计算、数据处理和过程控制）中，数据处理约占其中的 70%，而数据库技术就是作为一门数据处理技术发展起来的。在信息技术日益普及的今天，作为信息系统的核心技术和基础，数据库系统几乎触及到人类社会生活的各个方面，从企事业内部的信息管理到各行业的业务处理。随着计算机应用的普及和深入，数据库技术更是不断发展，应用范围不断扩大，如多媒体系统、企业管理、工程、统计、汽车工业等领域都在利用数据库技术。了解并掌握数据库系统的基本概念和基本技术是应用数据库技术的前提。本章首先介绍数据库系统的基础知识，并对基本数据模型进行讨论，特别是关系模型；然后介绍关系代数及其在关系数据库中的应用，并对关系规范化理论做出简要说明；最后较为详细地介绍了数据库的设计过程。

1.1 数据库系统的基本概念

1.1.1 数据、数据库、数据库管理系统

1. 数据（Data）

数据是数据库系统研究和处理的对象，从本质上讲是描述事物的符号记录。符号不仅仅是指数字、字母和文字，而且包括图形、图像、声音等。因此数据有多种表现形式，都是经过数字化处理后存入计算机能够反映或描述事物的特性。

2. 数据库（DataBase，简称 DB）

数据库是数据的集合，它具有一定的组织形式并存储于计算机存储器上，具有多种表现形式并可被各种用户所共享。数据库中的数据具有较小的冗余度、较高的数据独立性和扩展性。信息社会中，人们收集各种各样的数据后，对它们进行加工，借助计算机和数据库技术科学地保存和管理大量的复杂数据，以便充分利用这些数据资源。例如，为了方便地管理学生档案，可以把学生的有关信息，如学号、姓名、性别、出生日期、民族、政治面貌等存储在一个数据库的表对象中。

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

数据库管理系统是位于用户与操作系统之间的一层数据管理软件，属于系统软件。它是数据库系统的一个重要组成部分，是使数据库系统具有数据共享、并发访问、数据独立等特性的根本保证，主要提供以下功能：

- （1）数据定义功能。
- （2）数据操纵及查询优化。
- （3）数据库的运行管理。
- （4）数据库的建立和维护。

Microsoft Access 就是一个关系型数据库管理系统（简称 RDBMS），它提供一个软件环境，利用它用户可以方便快捷地建立数据库，并对数据库中的数据实现查询、编辑、打印等操作。

4. 数据库管理员（DataBase Administrator，简称 DBA）

由于数据库的共享性，因此对数据库的规划、设计、维护、监视等需要有专人管理，从事这方面工作的人员称为数据库管理员。其主要工作如下：

（1）数据库设计。DBA 主要任务之一是进行数据库设计，具体地说就是进行数据模式的设计。由于数据库的集成与共享性，因此需要有专门人员对多个应用的数据需求进行全面的规划、设计与集成。

（2）数据库维护。DBA 必须对数据库中的数据安全性、完整性、并发控制及系统恢复、数据定期转储等进行实施与维护。

（3）改善系统性能，提高系统效率。DBA 必须随时监视数据库运行状态，不断调整内部结构，使系统保持最佳状态与最高效率。当数据库运行效率下降时，DBA 需采取适当的措施，如进行数据库的重组、重构等。

5. 数据库系统（DataBase System，简称 DBS）

数据库系统通常是指带有数据库的计算机应用系统。它一般由数据库、数据库管理系统（及其开发工具）、应用系统、硬件系统、数据库管理员和用户组成。在不引起混淆的情况下常把数据库系统简称为数据库。

6. 数据库应用系统（DataBase Application System，简称 DBAS）

利用数据库系统进行应用开发可构成一个数据库应用系统，数据库应用系统是数据库系统再加上应用软件和界面组成，具体包括数据库、数据库管理系统、数据库管理员、硬件平台、软件平台、应用软件、应用界面，其结构如图 1.1 所示。

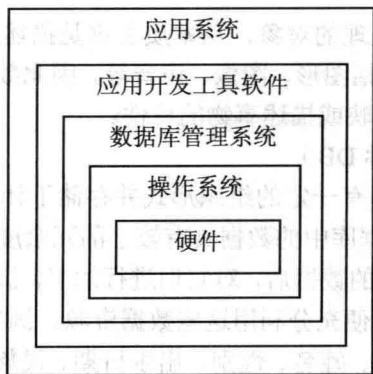


图 1.1 数据库系统软硬件层次结构图

1.1.2 数据库技术的发展

数据库技术产生于 20 世纪 60 年代后期，是随着数据管理的需要而产生的。在此之前，数据管理经历了人工管理阶段和文件系统阶段。20 世纪 60 年代，计算机技术迅速发展，其主要应用领域从科学计算转移到数据事务处理，从而出现了数据库技术，它是数据管理的最新技术，是计算机科学中发展最快、应用最广泛的重要分支之一。在短短的三十几年里，数据库技术的发展经历了三代：第一代为层次、网状数据库系统，第二代为关系数据库系统，第三代为以面向对象模型为主要特征的数据库系统。目前，数据库技术与网络通信技术、人工智能技术、面向对象程序设计技术、并行计算机技术等相互渗透，成为数据库技术发展的主要特征。

1. 第一代数据库系统——层次、网状数据库系统

数据库发展阶段的划分是以数据模型的发展为主要依据的。数据模型的发展经历了格式化数据模型（包括层次数据模型和网状数据模型）、关系数据模型两个阶段，并正向面向对象的数据模型等非传统数据模型阶段发展。实际上层次数据模型是网状数据模型的特例，层次数据库系统和网状数据库系统在体系结构、数据库语言和数据存储管理上均具有相同特征，并且都是在 20 世纪 60 年代后期研究和开发的，属于第一代数据库系统。

第一代数据库系统具有如下特点：

(1) 支持三级模式的体系结构。三级模式通常指外模式、模式、内模式，模式之间具有转换功能。

(2) 用存取路径来表示数据之间的联系。数据库系统不仅存储数据，而且存储数据之间的联系。在层次和网状数据库系统中，数据之间的联系是用存取路径来表示和实现的。

(3) 独立的数据定义语言。层次数据库系统和网状数据库系统有独立的数据定义语言，用以描述数据库的外模式、模式、内模式以及相互映像。三种模式一经定义，就很难修改。这就要求数据库设计人员在建立数据库应用系统时，不仅要充分考虑用户的当前需求，还要充分了解可能的需求变化和发展。

(4) 导航的数据操纵语言。层次数据库和网状数据库的数据查询和数据操纵语言是一个一个记录的导航式的过程化语言。这类语言通常嵌入某一种高级语言如 COBOL、Fortran、PL/1 中，其优点是存取效率高。缺点是编程繁琐，应用程序的可移植性较差，数据的逻辑独立性也较差。

第一代数据库系统的代表是：

(1) 1969 年，IBM 公司开发的层次模型的数据库系统 IMS (Information Management System)，它可以让多个程序共享数据库。

(2) 1969 年 10 月，美国数据库系统语言协会 CODASYL (Conference On Data System Language) 的数据库研制者提出了网状模型数据库系统规范报告，称为 DBTG (Data Base Task Group) 报告，使数据库系统开始走向规范化和标准化。它是数据库网状模型的典型代表。

2. 第二代数据库系统——关系数据库系统

1970 年美国 IBM 公司 San Jose 研究室的高级研究员埃德加·考特 (E. F. Codd) 发表了论文《大型共享数据库数据的关系模型》，提出了数据库的关系模型，开创了数据库关系方法和关系数据理论的研究，奠定了关系数据库技术的理论基础，为数据库技术开辟了一个新时代。

20 世纪 70 年代，关系方法的理论研究和软件系统的研制均取得了很大成果。IBM 公司的 San Jose 实验室研制出关系数据库实验系统 System R。与 System R 同期，美国 Berkeley 大学也研制了 INGRES 数据库实验系统，并发展成为 INGRES 数据库产品，使关系方法从实验室走向了市场。

关系数据库产品一问世，就以其简单清晰的概念、易懂易学的数据库语言，使得用户不需了解复杂的存取路径细节，不需说明“怎么干”，只需指出“干什么”，就能操作数据库，从而深受广大用户喜爱。

20 世纪 80 年代以来，大多数厂商推出的数据库管理系统的产品都是关系型的，如 FoxPro、Access、DB2、Oracle 及 Sybase 等都是关系型数据库管理系统（简称 RDBMS），使数据库技术日益广泛地应用到企业管理、情报检索、辅助决策等各个方面，成为实现和优化信息系统的基本技术。

关系数据库是以关系模型为基础的，具有以下特点：

(1) 关系数据库对实体及实体之间的联系均采用关系来描述，对各种用户提供统一的单一数据结构形式，使用户容易掌握和应用。

(2) 关系数据库语言具有非过程化特性，将用户从数据库记录的导航式检索编程中解脱出来，降低了编程难度，可面向非专业用户。

(3) 数据独立性强，用户的应用程序、数据的逻辑结构与数据的物理存储方式无关。

(4) 以关系代数为基础，数据库的研究更加科学化，尤其在关系操作的完备性、规范化及查询优化等方面，为数据库技术的成熟奠定了很好的基础。

3. 第三代数据库系统

第一代和第二代数据库技术基本上是处理面向记录、以字符表示为主的数据，能较好地满足商业事务处理的需求，但远远不能满足多种多样的信息类型处理需求。新的数据库应用领域如计算机辅助设计/制造 (CAD/CAM)、计算机集成制造 (CIM)、办公信息系统 (OIS) 等需要数据库系统能支持各种静态和动态的数据，如图形、图像、语音、文本、视频、动画、音乐等，并且还需要数据库系统具备处理复杂对象、实现程序设计语言和数据库语言无缝集成等能力。这种情况下，原有的数据库系统就暴露出了多种局限性。正是在这种新应用的推动下，数据库技术得到进一步发展。

1990 年高级 DBMS 功能委员会发表了《第三代数据库系统宣言》，提出了第三代数据库应具有的三个基本特征，并从三个基本特征导出了 13 个具体特征和功能。

经过多年的研究和讨论，对第三代数据库系统的基本特征已有了如下共识：

(1) 第三代数据库系统应支持数据管理、对象管理和知识管理。以支持面向对象数据模型为主要特征，并集数据管理、对象管理和知识管理为一体。

(2) 第三代数据库系统必须保持或继承第二代数据库系统的技术，如非过程化特性、数据独立性等。

(3) 第三代数据库系统必须对其他系统开放，如支持数据库语言标准、在网络上支持标准网络协议等。

4. 数据库技术的新进展

20 世纪 80 年代以来，数据库技术经历了从简单应用到复杂应用的巨大变化，数据库系统的发展呈现出百花齐放的局面，目前在新技术内容、应用领域和数据模型三个方面都取得了很大进展。

数据库技术与其他学科的有机结合，是新一代数据库技术的一个显著特征，从而出现了各种新型的数据库，例如：

- 数据库技术与分布处理技术相结合，出现了分布式数据库。
- 数据库技术与并行处理技术相结合，出现了并行数据库。
- 数据库技术与人工智能技术相结合，出现了知识库和主动数据库系统。
- 数据库技术与多媒体处理技术相结合，出现了多媒体数据库。
- 数据库技术与模糊技术相结合，出现了模糊数据库等。
- 数据库技术应用到其他领域中，出现了数据仓库、工程数据库、统计数据库、空间数据库及科学数据库等多种数据库技术，扩大了数据库应用领域。

数据库技术发展的核心是数据模型的发展。数据模型应满足三方面的要求：一是能比较真实地模拟现实世界；二是容易为人们所理解；三是便于在计算机上实现。目前，一种数据模

型要很好地满足这三方面的要求是很困难的。新一代数据库技术则采用多种数据模型,例如面向对象数据模型、对象关系数据模型、基于逻辑的数据模型等。

1.1.3 数据库系统的特点

数据管理技术经历了人工管理、文件系统和数据库系统三个阶段,数据库技术是在文件系统的基础上发展起来的,以数据文件来组织数据,并在文件系统之上加入了 DBMS 对数据进行管理,其特点如下:

1. 数据的集成性

数据库系统的数据集成性主要表现在如下几个方面:

(1) 在数据库系统中采用统一的数据结构,如在关系数据库中采用关系(在用户角度来看是二维表)作为统一结构方式。

(2) 在数据库系统中按照多个应用的需要组织全局的、统一的数据结构(即数据模式、全局结构)。

(3) 数据库系统中的数据模式是多个应用共同的、全局的数据结构,而每个应用的数据则是全局结构中的一部分,称为局部结构(即视图),这种全局与局部的结构模式构成了数据库系统数据集成性的主要特征。

2. 数据的高共享性与低冗余性

数据的集成性使得数据可为多个应用所共享,而数据共享又可极大地减少数据冗余,不仅减少了不必要的存储空间,更为重要的是可以避免数据的不一致性。所谓数据的一致性是指在系统中同一数据的不同出现场合应保持相同的值,减少数据冗余是保证系统一致性的基础。

3. 数据独立性

数据独立性是数据与程序间的互不依赖性,即数据库中的数据独立于应用程序而不依赖于应用程序。也就是说,数据的逻辑结构、存储结构与存取方式的改变不会影响应用程序。数据独立性包括:

(1) 物理独立性。简单地说,物理独立性就是指数据的物理结构(包括存储结构、存取方式)的改变不影响数据库的逻辑结构,从而不会引起应用程序的变化。

(2) 逻辑独立性。简单地说,逻辑独立性就是指数据的全局逻辑结构的改变不会引起应用程序的变化。

当然,数据独立性的实现需要模式间的映射关系作为保障。

1.1.4 数据库系统的体系结构

数据库系统的体系结构包括三级模式和两级映射,三级模式分别为外模式、概念模式和内模式;两级映射分别为外模式与概念模式间的映射以及概念模式与内模式间的映射,其抽象结构关系如图 1.2 所示。

1. 数据库系统的三级模式

数据模式是数据库系统中数据结构的一种表示形式,它具有不同的层次与结构方式。

(1) 外模式(External Schema)。外模式又称为用户模式或子模式,是某个或某几个数据库用户所看到的数据库的数据视图。外模式是与某一应用有关的数据的逻辑结构和特征描述,也就是前面所介绍的局部结构,它由概念模式推导而来。概念模式给出了系统全局的数据描述,而外模式则给出每个用户的局部数据描述。对于不同的数据库用户,由于需求的不同,外模式

的描述也互不相同。一个概念模式可以有若干个外模式, 每个用户只关心与其有关的外模式, 这样有利于数据保护, 对数据所有者和用户都极为方便。

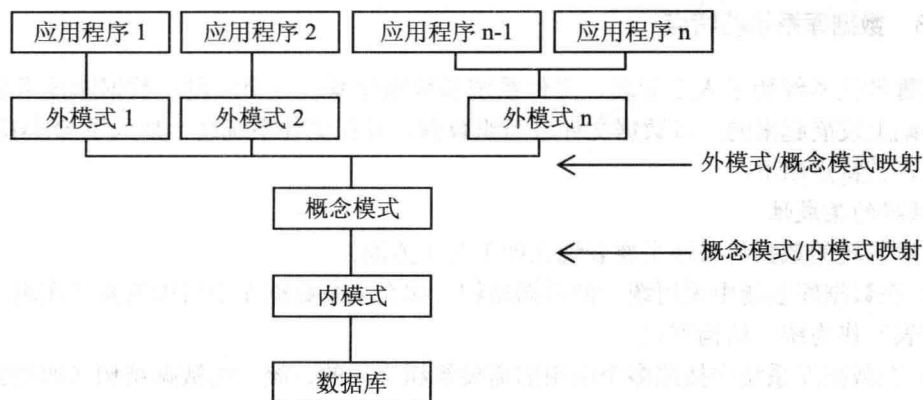


图 1.2 数据库三级模式结构

(2) 概念模式 (Conceptual Schema)。概念模式又称为模式或逻辑模式, 它介于内模式与外模式之间, 是数据库设计者综合各用户的数据, 按照统一的需求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。外模式涉及的仅是局部的逻辑结构, 通常是概念模式的子集。概念模式是用模式描述语言来描述的, 在一个数据库中只有一个概念模式。

(3) 内模式 (Internal Schema)。内模式又称为存储模式或物理模式, 是数据库中全体数据的内部表示, 它描述了数据的存储方式和物理结构, 即数据库的“内部视图”。它是数据库的底层描述, 定义了数据库中各种存储记录的物理表示、存储结构与物理存取方式, 如数据存储文件的结构、索引、集簇等存取方式和存取路径等。内模式是用模式描述语言严格定义的, 在一个数据库中只有一个内模式。

在数据库系统体系结构中, 三级模式是根据所描述的三层体系结构的三个抽象层次定义的, 外模式处于最外层, 它反映了用户对数据库的实际要求; 概念模式处于中间层, 它反映了设计者的数据全局的逻辑要求; 内模式处于最内层, 它反映了数据的物理结构和存取方式。

2. 数据库系统的两级映射

数据库系统的三级模式是数据的三个级别的抽象, 使用户能逻辑地、抽象地处理数据而不必关心数据在计算机中的表示和存储。为实现三个抽象层次间的联系和转换, 数据库系统在三个模式间提供了两级映射: 外模式与概念模式间的映射、概念模式与内模式间的映射。

(1) 外模式与概念模式间的映射。该映射定义了外模式与概念模式之间的对应关系, 保证了逻辑数据的独立性, 即外模式不受概念模式变化的影响。

(2) 概念模式与内模式间的映射。该映射定义了内模式与概念模式之间的对应关系, 保证了物理数据的独立性, 即概念模式不受内模式变化的影响。

1.2 数据模型

数据库中组织数据应从全局出发, 不仅考虑到事物内部的联系, 还要考虑到事物之间的联系。表示事物以及事物之间联系的模型就是数据模型。数据模型是用来抽象、表示和处理现

实世界的数据和信息的工具,也就是现实世界数据特征的抽象。数据模型是数据库系统的核心和基础,现有的数据库系统均是基于某种数据模型的。

数据模型有三个基本组成要素:数据结构、数据操作和完整性约束。

(1) 数据结构:用于描述系统的静态特性,研究的对象包括两类,一类是与数据类型、内容、性质有关的对象;另一类是与数据之间的联系有关的对象。

(2) 数据操作:是指对数据库中各种对象(型)的实例(值)允许执行的所有操作,即操作的集合,包括操作及有关的操作规则。数据库主要有检索和更新两类操作。

(3) 完整性约束:是给定的数据模型中数据及其联系所具有的制约和依存规则,用以限定数据库的状态及状态的变化,以保证数据的正确、有效和相容。

数据模型按不同的应用层次分成三种类型:概念数据模型、逻辑数据模型、物理数据模型。

概念数据模型简称概念模型,它是一种面向客观世界、面向用户的模型,与具体的数据库管理系统和计算机平台无关。概念模型着重于对客观世界复杂事物的结构及它们之间的内在联系的描述。概念模型是整个数据模型的基础,设计概念模型常用的方法是ER方法,也就是E-R模型(实体-联系模型)。

逻辑数据模型又称为数据模型,它是一种面向数据库系统的模型,该模型着重于数据库系统级别的实现。概念模型只有在转换成数据模型后才能在数据库中得以表示。数据库领域中过去和现在最常见的数据模型有四种:层次模型(Hierarchical Model)、网状模型(Network Model)、关系模型(Relational Model)和面向对象模型(Object Oriented Model),其中层次模型和网状模型统称为非关系模型。在关系模型出现以前,它们是非常流行的数据模型。非关系模型中数据结构的单位是基本层次联系。所谓基本层次联系是指两个记录以及它们之间的一对多(包括一对一)的联系,如图1.3所示。图中 R_i 位于联系 L_{ij} 的始点,称为双亲结点, R_j 位于联系 L_{ij} 的终点,称为子女结点。每个结点表示一个记录类型(实体),结点间的连线表示记录类型之间一对多的联系。

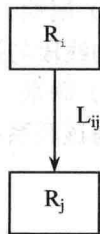


图1.3 基本层次联系

物理数据模型又称为物理模型,它是一种面向计算机物理表示的模型,此模型给出了数据模型在计算机上物理结构的表示。

1.2.1 E-R模型

概念模型是面向现实世界的,它的出发点是便于有效和自然地模拟现实世界给出数据的概念化结构。长期以来被广泛使用的概念模型是E-R模型,它于1976年由Peter Chen首先提出。该模型将现实世界的要求转化成实体、联系、属性等几个基本概念,以及它们之间的两种基本联接关系,E-R模型可用图形直观地表示。

1. E-R模型的基本概念

(1) 实体。实体是客观存在并且可以相互区分的事物。实体可以是具体的人、事、物,如一个学生、一门课程、一本书;也可以是抽象的概念与联系,如学生与课程之间的联系,即学生选课的情况。凡是有相同属性的实体可组成一个集合,称为实体集,如“王立”、“李美”是学生实体,那么这个集合就是一个实体集。

(2) 属性。实体有若干个特性,每个特性称为实体的一个属性。也可以说,属性是实体某一方面特征的描述,如学生实体包括学号、姓名、性别、院系等若干属性。每个属性可以有

属性值，如（“200103001”，“王立”，“男”，“计算机”）。一个属性的取值范围称为该属性的值域，如“性别”的值为{“男”，“女”}。

（3）联系。联系是两个或两个以上的实体集间的关联关系的描述，如学生与课程实体间的选课关系。实体集间的联系类型有如下三种：

①一对一联系。假设有实体集 A 与实体集 B，如果 A 中的一个实体至多与 B 中的一个实体关联，反过来，B 中的一个实体至多与 A 中的一个实体关联，则称 A 与 B 是一对一联系类型，记为 1:1，如班级与班长。

②一对多联系。假设有实体集 A 与实体集 B，如果 A 中的一个实体可与 B 中多个实体关联，反过来，B 中的一个实体至多与 A 中的一个实体关联，则称 A 与 B 是一对多联系类型，记为 1:N，如班级与学生。

③多对多联系。假设有实体集 A 与实体集 B，如果 A 中的一个实体可与 B 中多个实体关联，反过来，B 中的一个实体也可与 A 中多个实体关联，则称 A 与 B 是多对多联系类型，记为 M:N，如学生与课程。

2. E-R 模型的表示方法

E-R 模型可以用一种非常直观的图来表示，称为 E-R 图。

（1）实体（集）。在 E-R 图中，实体用矩形来表示，在矩形内写上该实体（集）的名字，如图 1.4 中的学生、课程实体（集）。

（2）属性。在 E-R 图中，属性用椭圆形来表示，在椭圆形内写上该属性的名字，并用没有方向的线段与该属性所关联的实体（集）连接，如图 1.4 中的学号、姓名等属性。

（3）联系。在 E-R 图中，联系用菱形来表示，在菱形内写上联系的名字，并用没有方向的线段与该联系相关的实体（集）连接，同时在线段上表明联系的类型，如图 1.4 中的选课联系。



图 1.4 学生选课 E-R 图

1.2.2 层次模型

层次模型（Hierarchical Model）用树形结构来表示数据间的从属关系结构，其主要特征如下：

- 仅有一个无双亲的结点，这个结点称为根结点。
- 其他结点向上仅有一个双亲结点，向下有若干子女结点。

如图 1.5 所示的层次模型就像一棵倒置的树，根结点在上，层次最高；子女结点在下，逐层排列。同一双亲的子女结点称为兄弟结点，没有子女结点的结点称为叶结点。一所学校的人员数据库可以采用层次模型来表示，如图 1.6 所示。记录项学校是根结点，它有院系和行政部门两个子女结点。记录项院系是学校的子女结点，同时又是教师的双亲结点。记录项行政部门是学校的另一个子女结点，同时是工作人员的双亲结点。教师和工作人员是叶结点，它们没有子女结点。由学校到院系、院系到教师、学校到行政部门、行政部门到工作人员均是一对多的

联系。

层次数据模型比较简单，结构清晰，容易理解。但由于现实世界中很多联系是非层次的，采用层次模型表示这种非层次的联系很不直接，只能通过冗余数据或创建非自然的数据组织来解决。

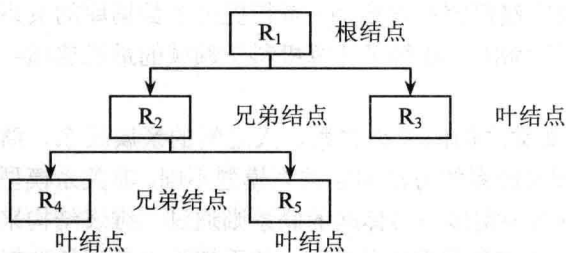


图 1.5 层次模型示例

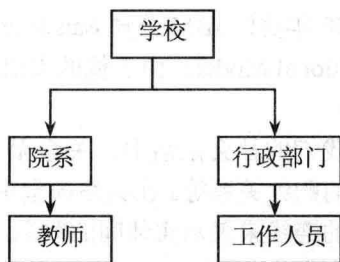


图 1.6 学校人员数据库模型

1.2.3 网状模型

网状模型 (Network Model) 是层次模型的扩展，呈现一种交叉关系的网络结构，可以表示较复杂的数据结构。其主要特征如下：

- 可以有一个以上的结点无双亲结点。
- 一个结点可以有多个双亲结点。

在网状模型中，子女结点与双亲结点的联系可以不唯一。因此，要为每个联系命名，并指出与该联系有关的双亲记录和子女记录。如图 1.7 (a) 中， R_3 有两个双亲记录 R_1 和 R_2 ，把 R_1 和 R_3 之间的联系称为 L_1 ，把 R_2 和 R_3 之间的联系称为 L_2 ；图 1.7 (b) 中 R_1 和 R_3 均无双亲， R_4 和 R_5 有两个双亲。

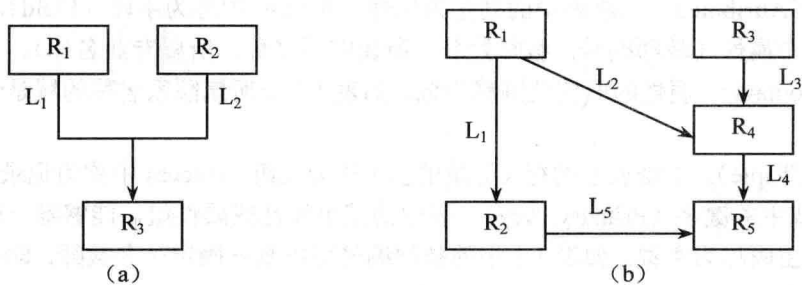


图 1.7 网状模型示例

学生选课数据库可以采用网状模型。学生选课时一个学生可以选修多门课程，一门课程也可以被多个学生选修，学生与课程是多对多的联系。尽管网状模型不支持多对多联系，但由于一个多对多联系可以转化为两个一对多联系，所以网状模型可以间接地描述多对多联系。可以在学生与课程之间建立一个联接记录“学生—课程”，把原来的多对多联系转化为“学生”与“学生—课程”、“课程”与“学生—课程”这两个一对多联系。图 1.8 是学生选课系统数据库网状模型。

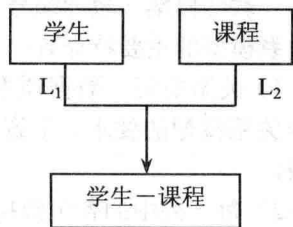


图 1.8 学生/学生—课程/课程网状模型