

21世纪高等教育计算机规划教材

COMPUTER

数据库应用基础

(Access 2010)

DataBase System and Access 2010
Applications

■ 涂振宇 刘文军 主编
■ 熊志文 楼明珠 副主编

— 体现 CDIO 工程教育理念

— 理论适用、够用为度

— 用实例引导学生思维



 人民邮电出版社
POSTS & TELECOM PRESS

21世纪高等教育计算机规划教材

COMPUTER

数据库应用基础

(Access 2010)

DataBase System and Access 2010
Applications

■ 涂振宇 刘文军 主编
■ 熊志文 楼明珠 副主编



人民邮电出版社
北京

数据库应用基础 : Access 2010 / 涂振宇, 刘文军

主编. — 北京 : 人民邮电出版社, 2014.9

21世纪高等教育计算机规划教材

ISBN 978-7-115-36544-6

I. ①数… II. ①涂… ②刘… III. ①关系数据库系
统一高等学校—教材 IV. ①TP311.138

中国版本图书馆CIP数据核字(2014)第178061号

内 容 提 要

本书循序渐进地介绍了 Access 2010 数据库应用开发工具的详细内容。全书共分 10 章, 包括数据库系统概论、Access 2010 数据库对象、数据库的创建、表的创建与使用、创建查询与窗体、报表的创建与打印、宏与模块等 VBA 编程技巧、数据管理与安全保护等。全书内容丰富, 结构清晰, 语言简练, 图文并茂, 具有很强的实用性和可操作性, 是一本适合于大学本科院校、职业院校及相关各类社会培训机构的优秀教材, 也是广大初、中级电脑用户的自学参考书。

-
- ◆ 主 编 涂振宇 刘文军
 - 副 主 编 熊志文 楼明珠
 - 责任编辑 刘 博
 - 责任印制 彭志环 杨林杰

 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京鑫正大印刷有限公司印刷

 - ◆ 开本: 787×1092 1/16
印张: 12.5 2014 年 9 月第 1 版
字数: 324 千字 2014 年 9 月北京第 1 次印刷
-

定价: 29.80 元

读者服务热线: (010)81055256 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京崇工商广字第 0021 号

前言

阿尔文·托夫勒(Alvin Toffler)是一位美国的“未来学”大师,也是一位洞察到现代科技将深刻改变人类社会结构及生活形态的学者。他于20世纪80年代初所著的《第三次浪潮》一书中将人类发展史划分为第一次浪潮的“农业文明”,第二次浪潮的“工业文明”以及第三次浪潮的“信息社会”,从而给历史研究与未来思想带来了全新的视角。他特别强调“数据与数据库技术”将是“第三次浪潮的华彩乐章”。

随着移动互联网、智能手机、可穿戴设备的普及,人们每天都在和数据库打交道。不论是网上购物,网络聊天,还是发微博、微信、电子邮件等,这些信息技术的应用都离不开后台的数据库服务。当今的信息社会,每天产生巨大的结构化和非结构化数据,其中可能蕴含着极其重要的信息。我们已进入“大数据”时代,人们已经宣称“数据是信息时代的石油”、“得数据者得天下”。因此基本的数据库知识和技能已是大学生信息素养的重要组成部分。

为了帮助非计算机专业的大学生学习数据库技术,同时对“数据库技术及应用”这门课陈旧的教学内容进行一次大刀阔斧的教学改革,本书的编者在编写过程中,进行了大胆的尝试。适逢教育部在全国推广“卓越工程师教育培养计划”,通过多年的工程教育思想的学习与实践,我们在教材编写中体现了构思、设计实现和运作的“CDIO”工程教育理念。教材对基本的数据库知识做了选择性介绍,确保“够用,适用”,用实例来引导学生思维;安排的实验与课程设计中着重学生的个人能力和工程实施能力的培养。

本书采用 Access 2010 为蓝本来介绍数据库知识。Access 关系型数据库系统是 Microsoft Office 的一个成员,它与许多优秀的关系数据库管理系统一样,可以有效地组织、管理和共享数据库的信息,并将数据库与 Web 技术结合在一起。

本书重点介绍了 Access 2010 关系型数据库系统的各项功能和操作方法。

本书共分 10 章,从数据库的基础理论开始,由浅入深、循序渐进地介绍了 Access 2010 各种对象的功能及创建方法。

第 1 章主要介绍数据库的基础知识、数据模型、关系型数据库、关系运算及数据库系统设计的一般步骤。

第 2 章主要介绍数据库和表的一般创建方法,详细介绍了 Access 2010 数据库中表的基本操作。

第 3 章主要介绍利用查询向导和查询设计视图来创建选择查询。

第 4 章主要介绍数据库标准语言——结构化查询语言 SQL 的功能及用法。

第 5 章主要介绍 Access 2010 中窗体设计及窗体中控件的使用。

第 6 章主要介绍 Access 2010 中报表的创建与编辑等功能。

第 7 章主要介绍 Access 2010 中宏创建、调试和运行方法。

第 8 章主要介绍 Access 2010 中 VBA 的基础知识和基本操作。

第 9 章主要介绍 Access 2010 中 SharePoint 的基础知识。

第 10 章主要介绍 Access 2010 中的数据管理与安全保护。

涂振宇负责第 1 章、第 2 章的编写,刘文军负责第 3 章、第 4 章的编写,熊志文负责第 7 章、第 8 章的编写,楼明珠负责第 5 章、第 6 章、第 9 章、第 10 章的编写。在本书的编写和出版过程中,得到了校院领导的大力支持,在此表示衷心的感谢。

为了便于教学,我们将为选用本书的任课教师提供电子教案和数据库文件,可在人民邮电出版社教学服务与资源网(www.ptpedu.com.cn)上免费下载。

由于编者水平有限,书中难免有疏漏和欠缺之处,敬请广大读者提出宝贵意见。

编者
2014.夏

目 录

第 1 章 数据库绪论	1	2.3.5 建立有效性规则	35
1.1 数据库概述	1	2.4 数据表编辑	37
1.1.1 数据与数据库系统	1	2.4.1 修改表结构	37
1.1.2 数据库技术发展简史	4	2.4.2 设置表外观	37
1.1.3 数据模型	6	2.4.3 编辑表内容	39
1.1.4 结构模型	9	2.4.4 记录排序与筛选	39
1.2 关系数据库	10	2.4.5 记录筛选	40
1.2.1 关系模型	11	2.4.6 表的关联	40
1.2.2 关系运算	11	2.5 思考与练习	42
1.2.3 关系的完整性约束	13	第 3 章 查询	45
1.3 数据库设计方法	14	3.1 查询概述	45
1.3.1 需求分析阶段	15	3.1.1 查询的功能	45
1.3.2 概念结构设计阶段	15	3.1.2 查询的类型	46
1.3.3 逻辑结构设计阶段	17	3.2 建立查询	46
1.3.4 物理设计阶段	19	3.2.1 查询视图	46
1.3.5 数据库实施	19	3.2.2 查询设计器	48
1.3.6 数据库运行和维护	20	3.2.3 设置查询的组合条件	50
1.4 思考与练习	20	3.2.4 在查询中进行计算	54
第 2 章 Access 2010 初步	23	3.3 参数查询	56
2.1 Access 2010 简介	23	3.3.1 单参数查询	57
2.1.1 Access 2010 的特点	23	3.3.2 多参数查询	58
2.1.2 Access 2010 应用环境	25	3.4 交叉表查询	59
2.1.3 Access 2010 上下文命令和 快速访问	27	3.4.1 使用查询向导创建交叉表查询	59
2.1.4 数据库对象	28	3.4.2 使用“查找重复项查询向导” 创建查询	61
2.2 新建数据库	29	3.4.3 使用“查找不匹配项查询向导” 创建查询	62
2.2.1 使用模板创建数据库	29	3.5 操作查询	64
2.2.2 新建空白数据库	30	3.5.1 生成表查询	64
2.3 新建数据表	30	3.5.2 删除查询	65
2.3.1 通过数据表视图创建表	31	3.5.3 追加查询	66
2.3.2 通过设计视图新建表	32	3.5.4 更新查询	68
2.3.3 字段属性	32	3.6 思考与练习	69
2.3.4 创建主键和索引	33		

第 4 章 关系数据库标准语言 SQL72	6.2.4 创建子报表.....109
4.1 SQL 概述.....72	6.3 设计报表.....111
4.1.1 SQL 的特点.....72	6.3.1 控件的使用.....111
4.1.2 SQL 的功能.....73	6.3.2 控件的种类.....111
4.2 SQL 数据定义.....73	6.3.3 控件与报表属性的设置.....112
4.2.1 表的定义.....73	6.4 报表的计算、排序、分组与汇总.....112
4.2.2 建立索引.....76	6.4.1 数据分组.....112
4.3 SQL 数据操作.....77	6.4.2 数据筛选和排序.....112
4.3.1 插入数据.....77	6.4.3 报表中常用函数.....114
4.3.2 更新数据.....79	6.5 报表外观与打印设置.....116
4.3.3 删除数据.....80	6.5.1 设置报表外观.....116
4.4 SQL 数据查询.....80	6.5.2 报表打印.....117
4.4.1 SELECT 语句的格式.....80	6.6 思考与练习.....118
4.4.2 单表查询.....81	第 7 章 宏120
4.4.3 多表连接查询.....83	7.1 宏概述.....120
4.4.4 嵌套查询.....85	7.1.1 宏的概念.....120
4.4.5 SQL 特定查询.....88	7.1.2 宏设计窗口.....121
4.5 思考与练习.....89	7.1.3 常用的宏操作.....122
第 5 章 Access 2010 窗体设计92	7.1.4 宏组.....124
5.1 Access 2010 窗体对象.....92	7.2 宏的基本操作.....124
5.1.1 创建“纵栏式”员工信息浏览窗体.....93	7.2.1 创建宏.....124
5.1.2 创建“主—子”窗体.....94	7.2.2 运行调试宏.....126
5.1.3 分割窗体.....95	7.2.3 在宏中使用条件.....128
5.2 自定义窗体的创建.....97	7.3 思考与练习.....131
5.2.1 控件类型.....97	第 8 章 VBA 编程133
5.2.2 控件的使用.....97	8.1 模块的基本概念.....133
5.3 控件与窗体属性编辑.....99	8.1.1 模块的分类.....133
5.3.1 控件的编辑.....99	8.1.2 模块的组成.....134
5.3.2 设置窗体属性.....99	8.2 VBA 程序设计概述.....135
5.4 思考与练习.....101	8.2.1 数据类型和数据库对象.....140
第 6 章 报表103	8.2.2 常量.....142
6.1 报表的基本概念.....103	8.2.3 变量.....143
6.1.1 报表结构.....103	8.2.4 运算符和表达式.....148
6.1.2 报表的分类.....104	8.2.5 常用标准函数.....152
6.1.3 报表视图.....105	8.2.6 DoCmd 对象及常用方法.....154
6.2 创建报表.....105	8.3 VBA 程序结构.....157
6.2.1 快速报表生成.....105	8.3.1 顺序结构.....157
6.2.2 使用向导创建报表.....106	8.3.2 选择结构.....157
6.2.3 使用标签向导创建报表.....107	8.3.3 循环结构.....159

8.3.4 过程和模块.....	163	9.3.1 发布 Access 2010 数据库.....	181
8.4 数据库编程.....	165	9.3.2 导出或链接 Access 2010 数据到 SharePoint 列表.....	182
8.4.1 数据库引擎及访问的数据类型.....	165	9.4 思考与练习.....	182
8.4.2 数据访问对象 (DAO).....	165	第 10 章 Access 2010 的数据 管理与安全保护	184
8.4.3 ActiveX 数据对象 (ADO).....	167	10.1 Access 2010 数据库安全.....	184
8.5 程序的调试.....	169	10.1.1 使用数据库密码加密 Access 数据库.....	184
8.5.1 VBA 程序错误类型.....	169	10.1.2 解密并打开数据库.....	185
8.5.2 VBA 程序的调试方法.....	169	10.1.3 改进的安全措施.....	186
8.6 思考与练习.....	172	10.1.4 生成 ACCDE 文件.....	186
第 9 章 SharePoint 与 Access 数据集成	177	10.2 数据库的安全管理.....	187
9.1 SharePoint 网站概述.....	177	10.2.1 拆分数据库.....	187
9.2 SharePoint 与 Access 2010 的 数据关联.....	179	10.2.2 压缩与修复数据库.....	188
9.2.1 迁移 Access 2010 数据库.....	179	10.3 数据库打包、签名和分发.....	188
9.2.2 SharePoint 列表的导入与链接.....	179	10.3.1 什么是数字签名.....	188
9.2.3 脱机使用 SharePoint 列表.....	180	10.3.2 使用数字签名发布数据包.....	189
9.3 SharePoint 与 Access 2010 的 数据共享.....	181	10.4 思考与练习.....	190

第 1 章

数据库绪论

数据库是当今社会应用最为广泛的信息技术之一。社会各行各业都依赖于长期积累的、海量的数据信息资源。有效地管理、应用、挖掘这些数据资源是信息社会的重要应用。数据库技术广泛应用于企业 ERP 系统、电子商务、电子政务、制造业和各类与人们日常生活相关的信息服务产业。现在人们在即时通信软件上聊天、在网上购物、社交网络上交友、发微博,以及包括学生在校内的各种学习、生活都依赖于后台数据库的服务。“数据库应用基础”课程的开设也是提高大学生计算机素养的必要环节。

在本章中读者将达到以下学习目标:

1. 理解数据、数据库、数据库管理系统、数据库系统的概念。
2. 了解层次模型、网状模型和关系模型的概念。
3. 掌握关系代数的并、差、交和选择、投影、连接运算。
4. 掌握实体完整性约束、参照完整性约束和用户定义的完整性约束的含义。
5. 了解数据管理技术的发展中三个阶段的特点,了解分布式数据库系统和并行数据库系统。
6. 掌握数据库结构设计的步骤。

1.1 数据库概述

本节将介绍几个入门概念:数据、信息、数据库、数据库管理系统。

1.1.1 数据与数据库系统

1. 数据与信息

人们将描述各类客观事物的文字、数字、音频、视频进行数字化处理(通过一定的编码,形成二进制数字)存入计算机形成数据。相关数据通过一定的组织结构形成“数据表”,相关数据表又被纳入同一个数据仓库,构成“数据库”。例如在“学生信息管理系统”这个数据库应用程序中,对于指定的某个学生有如下数据描述:

(2012001014, 王勇, 男, 1994-07-12, 水利与生态学院, 照片, 视频介绍)

每个班的学生记录可以组成“学生基本情况表”。对于学生要选修的课程,可以有如下描述:

(C001012 数据库应用基础 48 学时 3 学分 信息工程学院)

各门课程可以组成“课程表”。

以上各类表都可以收入同一个数据库中。

上述信息经过数字化变成 0、1 构成的二进制数字存入计算机中, 通过某个数据库管理软件, 人们可以把该学生上述数据包涵的语义通过计算机表示出来, 从而获得对该生基本信息的认知。

因此, 数据 (Data) 是数据库中存储的基本单位, 是客观事物的数字化表示。信息是数据所表达的语义。对于同一个数据, 人们可以有不同的解读, 因此可以得到不同的信息。“数据”如何表达信息, 取决于软件开发者不同的解读。

2. 数据库

数据库 (DataBase) 是按特定的组织方式存放的各类相关数据的数据集合。例如对于“学生信息管理”数据库来说, 其中存放有“学生基本情况”、“课程开设情况”、“任课老师情况”、“学生成绩”等各种数据表。这种科学地组织在一起的数据具备较低的冗余度, 较高的数据独立性且可以供各种应用管理程序共享。

例如“学生基本情况表”里的数据或其携带的信息可以被用在“学生成绩管理子系统”中, 也可被用在“学生社团管理子系统”中。数据库中的各种表可以以一个“查询”组合在一起, 不需要在各个表中重复存储同样的数据, 这样在一定程度上降低了数据的冗余度, 提高了数据的相关性与完整性约束。

3. 数据库管理系统

为了能方便地定义各种数据, 规定其数据结构、物理地创建数据对象、提高各类数据操作方法和和管理数据的安全性、解决并发访问冲突, 人们研究了“数据库管理系统”(DBMS)。它位于用户与操作系统之间。著名的数据库管理系统包括 Oracle、DB2、SQL Server、SYBASE 等。本书介绍的 Access 2010 是一种小型桌面关系型数据库管理系统。

数据库管理系统的功能包括以下几种。

(1) 数据定义语言 (Data Definition Language, DDL)。该语言负责数据的模式定义与数据的物理存取构建。

(2) 数据操纵语言 (Data Manipulation Language, DML)。该语言负责数据的操纵, 包括查询及增加、删除、修改等。

(3) 数据控制语言 (Data Control Language, DCL)。该语言负责数据完整性、安全性的定义与检查以及并发控制、故障恢复等功能等。

DBMS 通常还提供一些实用程序, 包括数据初始装入程序、数据转储程序、数据库恢复程序、性能监测程序、数据库再组织程序、数据转换程序、通信程序等。数据库用户可以利用这些实用程序完成数据库的建立与维护, 以及数据格式的转换与通信。通过 DBMS 的支持, 用户可以抽象地处理数据, 而不必关心这些数据在计算机中的存放方式以及计算机处理数据的过程细节, 把处理数据具体而繁杂的工作交给 DBMS 完成。常用的小型数据库管理软件只具备数据库管理系统的一些简单功能, 如 FoxPro 和 Access 等。严格意义上的 DBMS 应具备其全部功能, 包括数据组织、数据操纵、数据维护、数据控制及保护、数据服务等, 如 Oracle、MySQL 和 SQL Server 等。

4. 数据库系统人员

数据库系统人员包括数据库管理员、应用程序员和最终用户。

(1) 数据库管理员

数据库管理员 (DBA) 是负责数据库的建立、使用和维护的专门人员。数据库管理员的主要职责包括: 设计与定义数据库系统; 指导最终用户使用数据库系统; 监督与控制数据库系统的使用和运行; 对数据库中的数据安全性、完整性、并发控制及系统恢复、数据定期转存等进行实施与维护; 提高系统效率。数据库管理员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。当效率下降时,数据库管理员需采取适当的措施,如进行数据库的重组、重构等。每个数据库系统,如学校的成绩查询系统、网上购物系统等,都需要数据库管理员对数据库进行日常的管理与维护,数据库管理员的素质在一定程度上决定了数据库应用的水平。

(2) 应用程序员

应用程序员是指开发数据库应用程序的人员。数据库系统一般需要应用程序员在开发周期内完成数据库结构设计、应用程序开发等任务,在后期管理应用程序,保证使用周期中对应用程序在功能及性能方面的维护、修改工作。应用程序员是指负责设计和编写应用程序的人员,应用程序员使用高级语言编写应用程序,对数据库进行开发、存取、维护等操作。

(3) 最终用户

最终用户是应用程序的使用者,通过应用程序与数据库进行交互。例如,我们通过网上购物系统购物、在QQ上聊天或者通过学校的学生成绩管理系统进行选课和查询时,我们就是最终用户。最终用户通过计算机联机终端存取数据库的数据,具体操作应用程序,通过应用程序的用户界面,操作数据库来完成其业务活动。

5. 数据库系统

数据库系统是由数据库、数据库管理系统、数据库管理员、数据库应用系统构成的完成一个具体应用的大系统,如12306火车票购买系统就是一个典型的数据库系统。

数据库系统(DataBase System, DBS)是指计算机引进数据库技术后的整个系统构成,包括系统硬件平台(硬件)、系统软件平台(软件)、数据库管理系统(DBMS)、数据库(数据)、数据库系统用户。这五个部分构成了一个以数据库为核心的完整的运行实体,称为数据库系统。数据库系统的层次结构如图1-1所示。

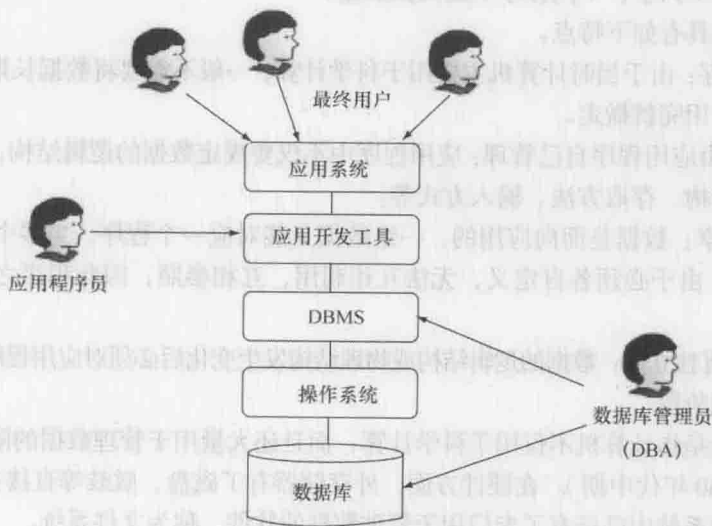


图 1-1 数据库系统

在数据库系统中,硬件包括如下两方面。

(1) 计算机:是系统中硬件的基础平台。目前常用的有微型机、小型机、中型机、大型机及巨型机。

(2) 网络: 过去数据库系统一般建立在单机上, 但是近年来它较多地建立在网络上, 从目前发展看, 数据库系统今后将以建立在网络上为主, 而其结构形式又以客户机/服务器 (Client/Server, C/S) 方式与浏览器/服务器 (Browser/Server, B/S) 方式为主。

在数据库系统中, 软件包括三方面内容。

(1) 操作系统: 是系统的基础软件平台, 目前常用的有 UNIX、Linux 和 Windows。

(2) 数据库系统开发工具: 为开发数据库应用程序所提供的计算机语言工具, 它包括面向对象程序如 Java、C# 等, 也包括可视化开发工具 Visual Basic、PowerBuilder、Delphi 等, 还包括与 Internet 有关的编程语言 HTML 及 XML 等以及一些专用开发工具。

(3) 后台数据库: 包括 Access、SQL Server 等。

1.1.2 数据库技术发展简史

1. 数据库技术的历史和发展

数据库技术是 20 世纪 60 年代开始兴起的一门信息管理自动化的新兴学科, 是计算机科学中的一个重要分支。随着计算机应用的不断发展, 在计算机应用领域中, 数据处理越来越占主导地位, 数据库技术的应用也越来越广泛。

数据库是数据管理技术发展的产物。数据管理是数据库的核心任务, 内容包括对数据的分类、组织、编码、储存、检索和维护。随着计算机硬件和软件的发展, 数据库技术也不断地发展。从数据管理的角度看, 数据库技术经历了人工管理阶段、文件系统阶段和数据库系统阶段。

(1) 人工管理阶段

在 20 世纪 50 年代前期, 计算机主要用于科学计算, 当时没有大容量的存储设备。人们把程序和要计算的数据通过打孔的纸带送入计算机中, 计算的结果由用户自己手工保存。处理方式只能是批处理, 数据不共享, 不同程序不能交换数据。

人工管理数据具有如下特点。

① 数据不保存: 由于当时计算机主要用于科学计算, 一般不需要将数据长期保存, 只是在计算时将数据输入, 用完就撤走。

② 数据需要由应用程序自己管理: 应用程序中不仅要规定数据的逻辑结构, 而且要设计物理结构, 包括存储结构、存取方法、输入方式等。

③ 数据不共享: 数据是面向应用的, 一组数据只能对应一个程序。当多个应用程序涉及某些相同的数据时, 由于必须各自定义, 无法互相利用, 互相参照, 因此程序之间有大量的冗余数据。

④ 数据不具有独立性: 数据的逻辑结构或物理结构发生变化后必须对应用程序作相应的修改。

(2) 文件系统阶段

文件系统阶段是指计算机不仅用于科学计算, 而且还大量用于管理数据的阶段 (从 20 世纪 50 年代中后期到 60 年代中期)。在硬件方面, 外存储器有了磁盘、磁鼓等直接存取的存储设备。在软件方面, 操作系统中已经有了专门用于管理数据的软件, 称为文件系统。

这个时期数据管理的特点有以下几个。

① 数据需要长期保存在外存上供反复使用。由于计算机大量用于数据处理, 经常对文件进行查询、修改、插入和删除等操作, 所以数据需要长期保留, 以便于反复操作。

② 程序之间有了一定的独立性。操作系统提供了文件管理功能和访问文件的存取方法, 程序和数据之间有了数据存取的接口, 程序可以通过文件名和数据打交道, 不必再寻找数据

的物理存放位置,至此,数据有了物理结构和逻辑结构的区别,但此时程序和数据之间的独立性尚还不充分。

③ 文件的形式已经多样化。由于有了直接存取的存储设备,文件也就不再局限于顺序文件,还有了索引文件、链表文件等,因而,对文件的访问可以是顺序访问,也可以是直接访问。

④ 数据的存取基本上以记录为单位。

(3) 数据库系统阶段

数据库系统阶段是从20世纪60年代后期开始的。在这一阶段中,数据库中的数据不再是面向某个应用或某个程序,而是面向整个企业(组织)或整个应用的。

数据库系统阶段的特点如下。

① 采用复杂的结构化的数据模型。数据库系统不仅要描述数据本身,还要描述数据之间的联系。这种联系是通过存取路径来实现的。

② 较高的数据独立性。数据和程序彼此独立,数据存储结构的变化尽量不影响用户程序的使用。

③ 最低的冗余度。数据库系统中的重复数据被减少到最低程度,这样,在有限的存储空间内可以存放更多的数据并减少存取时间。

④ 数据控制功能。数据库系统具有数据的安全性,以防止数据的丢失和被非法使用;具有数据的完整性,以保护数据的正确、有效和兼容;具有数据的并发控制,避免并发程序之间的相互干扰;具有数据的恢复功能,在数据库被破坏或数据不可靠时,系统有能力把数据库恢复到最近某个时刻的正确状态。

2. 数据库系统发展的三个时代

数据模型是数据库系统的核心。按照数据模型发展的主线,数据库技术的形成过程和发展可从以下三个方面反映。

(1) 第一代数据库系统:层次和网状数据库管理系统

层次和网状数据库的代表产品是IBM公司在1969年研制出的层次模型数据库管理系统。层次数据库是数据库系统的先驱,而网状数据库则是数据库概念、方法、技术的基础。

(2) 第二代数据库系统:关系数据库管理系统(RDBMS)

1970年,IBM公司的研究员E.F.Codd在题为《大型共享数据库数据的关系模型》的论文中提出了数据库的关系模型,为关系数据库技术奠定了理论基础。到了20世纪80年代,几乎所有新开发的数据库系统都是关系型的。

真正使得关系数据库技术实用化的关键人物是James Gray。Gray在解决如何保障数据的完整性、安全性、并发性以及数据库的故障恢复能力等重大技术问题方面发挥了关键作用。

关系数据库系统的出现,促进了数据库的小型化和普及化,使得在微型机上配置数据库系统成为可能。

(3) 新一代数据库技术的研究和发展

目前现行的数据库系统技术发展成熟。我们可以从数据模型、新技术内容、应用领域三个方面概括新一代数据库系统的发展。

① 面向对象的方法和技术对数据库发展的影响最为深远

20世纪80年代,面向对象的方法和技术的出现,对计算机各个领域,包括程序设计语言、软件工程、信息系统设计以及计算机硬件设备等都产生了深远的影响,也给面临新挑战的数据库

技术带来了新的机遇和希望。数据库研究人员借鉴和吸收了面向对象的方法和技术,提出了面向对象的数据库模型(简称对象模型)。当前有许多研究是建立在数据库已有的成果和技术上的,针对不同的应用,对传统的 DBMS,主要是 RDBMS 进行不同层次上的扩充,例如建立对象关系(OR)模型和建立对象关系数据库(ORDB)。

② 数据库技术与多学科技术的有机结合

数据库技术与多学科技术的有机结合是当前数据库发展的重要特征。计算机领域中其他新兴技术的发展对数据库技术产生了重大影响。传统的数据库技术和其他计算机技术的结合、互相渗透,使数据库中新的技术内容层出不穷。数据库的许多概念、技术内容、应用领域,甚至某些原理都有了重大的发展和变化。建立和实现了一系列新型的数据库,如分布式数据库、并行数据库、演绎数据库、知识库、多媒体库、移动数据库等,它们共同构成了数据库大家族。

③ 面向专门应用领域的数据库技术的研究

为了适应数据库应用多元化的要求,在传统数据库基础上,结合各个专门应用领域的特点,研究适合该应用领域的数据库技术,如工程数据库、统计数据库、科学数据库、空间数据库、地理数据库、Web 数据库等,这是当前数据库技术发展的又一重要特征。

同时,数据库系统结构也由主机/终端的集中式结构发展到网络环境的分布式结构,随后又发展成两层、三层或多层客户/服务器结构以及 Internet 环境下的浏览器/服务器和移动环境下的动态结构。多种数据库结构满足了不同应用的需求,适应了不同的应用环境。

1.1.3 数据模型

1. 数据模型三要素

人们将现实世界的实体属性抽象到信息世界并在计算机软硬件系统中用物理数据来描述的过程,称为建立数据模型。

现实世界是指客观存在的事物及其相互之间的联系。现实世界中的事物有着众多的特征和千丝万缕的联系,但人们只选择感兴趣的一部分来描述,如学生,人们通常用学号、姓名、班级、成绩等特征来描述和区分,而对身高、体重、长相不太关心。事物可以是具体的、可见的实物(如学生),也可以是抽象的事物(如课程)。

信息世界是人们把现实世界的信息和联系,通过“符号”记录下来,然后用规范化的数据库定义语言来定义描述而构成的一个抽象世界。信息世界实际上是对现实世界的一种抽象描述。在信息世界中,不是简单地对现实世界进行符号化,而是要通过筛选、归纳、总结、命名等抽象过程产生出概念模型,用以表示对现实世界的抽象与描述。

计算机世界是将信息世界的内容数据化后的产物。将信息世界中的概念模型,进一步地转换成数据模型,形成便于计算机处理的数据表现形式,如图 1-2 所示。

数据模型三要素:数据结构、数据操作与数据约束。

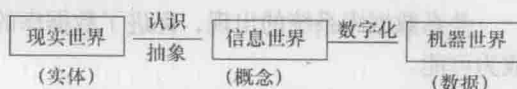


图 1-2 数据模型的建立

(1) 数据结构

数据结构描述的是计算机存储、组织数据的方式。数据结构是指相互之间存在一种或多种特定关系的数据元素的集合。通常情况下,精心选择的数据结构可以带来更高的运行或者存储效率。数据结构往往同高效的检索算法和索引技术有关。在数据库系统中也按照数据模型中数据结构的类型来区分、命名各种不同的数据模型,如层次结构、网状结构、关系结构的数据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

数据操作是指对各种对象类型的实例(或值)所允许执行的操作的集合,包括操作及有关的操作规则。在数据库中,主要的操作有检索和更新(包括插入、删除、修改)两大类。数据模型定义了这些操作的定义、语法(即使用这些操作时所用的语言)。数据结构是对系统静态特性的描述,而数据操作是对系统动态特性的描述。两者既有联系,又有区别。

(3) 数据约束

数据的约束条件是完整性规则的集合。完整性规则是指在给定的数据模型中,数据及其联系所具有的制约条件和依存条件,用限制符合数据模型的数据库的状态以及状态的变化来确保数据的正确性、有效性和一致性。

数据模型的三要素紧密依赖、相互作用、形成一个整体才能全面正确地抽象、描述现实世界数据的特征。例如在学生数据表中规定大学生年龄不得超过29岁,学生毕业需修满175个学分等。

2. 概念模型

概念数据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,概念模型的一些基本概念包括实体、属性、域、实体型、实体集、联系等。

(1) 实体。客观存在并可相互区别的事物称为实体。

一个实体是现实世界客观存在的一个事物,可以是一个具体的事物,如一个学生;也可以是抽象的事物,如一门课程。实体由它们自己的属性值表示其特征。例如,学生具有“学号,姓名,性别,出生年份,系,入学时间”等属性。

(2) 实体集。同类型实体的集合称为实体集。例如,全体学生就是一个实体集。

(3) 属性。用来描述实体或联系的特性。实体的每个特性称为一个属性。属性有属性名、属性类型、属性定义域和属性值之分。一个实体可以由若干个属性来刻画。

(4) 码。其值能唯一地标识每个实体的属性集称为码或键。例如,学号是学生实体的码。

(5) 域。属性的取值范围称为该属性的域。例如,学号的域为9位整数,姓名的域为字符串集合,年龄的域为小于35的整数,性别的域为(男,女)。

(6) 联系。在现实世界中,事务内部以及事务之间是有联系的,这些联系在信息世界中反映为实体内部的联系和实体之间的联系。

实体内部的联系通常是指组成实体的各属性之间的联系。

例如,每个教师隶属一个院系,每个教师和其隶属的一个院系之间有一个隶属联系。

3. 实体的联系

现实世界中,实体之间的联系是多种多样的。实体间的各种联系按照联系涉及的群体可以划分为两类:一类是不同的实体集中实体之间的联系,另一类是相同的实体集内实体之间的联系。这两种联系都有三种不同的联系情况。

(1) 一对一联系。如果实体集EA中每个实体至多和实体集EB中的一个实体有联系,反之亦然,就称实体集EA和实体集EB的联系为“一对一联系”,记为“1:1”。例如,学生和学号之间的管理联系,一个学生只有一个学号,一个学号只能唯一地对应一个学生。

(2) 一对多联系。如果实体集EA中每个实体与实体集EB中的任意多个(零个或多个)实体有联系,而EB中每个实体至多与实体集EA中的一个实体有联系,就称实体集EA对EB的联系为“一对多联系”,记为“1:n”。例如,院系与其多个教师之间的“属于”联系、工厂里的车间和车间内多个工人之间的联系等,都是1:n的联系。一个院系内有多个教师,每个教师只能属于

一个院系。同样,一个车间内有多个人,每个工人只能属于一个车间。

(3) 多对多联系。如果实体集 EA 中的每个实体与实体集 EB 中的任意个(零个或多个)实体有联系;反之,实体集 EB 中的每个实体与实体集 EA 中的任意个(零个或多个)实体有联系,就称实体集 EA 和 EB 的联系为“多对多联系”,记为“ $m:n$ ”联系。例如,学生和课程之间的联系,一个学生可以选修多门课程,每门课程有多个学生选修。零件供应商和零件之间的联系,一个供应商可供应多种零件,每种零件可以由多个供应商提供等,都是多对多的联系。

4. E-R 图

概念模型的表示方法很多,其中最为常用的是 P.P.S.Chen 于 1976 年提出的实体-联系方法。该方法用 E-R 图来描述现实世界的概念模型。

E-R 图提供了表示实体型、属性和联系的方法。

实体型:用矩形表示,矩形框内写明实体名。

属性:用椭圆形表示,并用无向边将其与相应的实体连接起来。

联系:用菱形表示,菱形框内写明联系名,并用无向边分别与有关实体连接起来,同时在无向边旁标上联系的类型(1:1,1:n 或 $m:n$)。

需要注意的是,联系本身也是一种实体型,也可以有属性。如果一个联系具有属性,则这些属性也要用无向边与该联系连接起来。

【例 1-1】学生、班级、课程 E-R 图。

需要注意的是,一个联系若具有属性,则这些属性也要用无向边与该联系连接起来,如图 1-3 所示。



图 1-3 实体及其联系

【例 1-2】用 E-R 图表示学生成绩管理系统中学生与课程的联系。绘图步骤如下。

(1) 分析:每个学生选修若干课程,每门课可由若干学生学习,联系为多对多,学生和课程的联系是“学习”产生属性“成绩”。

(2) 确定实体型及其属性。学生成绩管理系统中包含学生与课程两个实体型。其中学生的属性有学号、姓名、性别、所获助学金等。课程的属性有课程号、课程名、学时数。分别用矩形和圆角矩形表示两个实体型及其属性。

(3) 确定这两个实体型之间的联系为“学习”。学习后产生成绩,分别用矩形和圆角矩形表示联系及其属性。

(4) 对实体型和联系用连线组合,并标上联系的方式 $m:n$, 得到学生学习课程的 E-R 图,如图 1-4 所示。

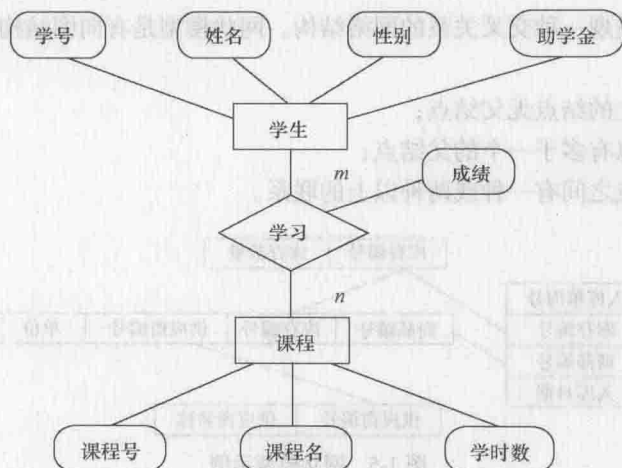


图 1-4 学生选课 E-R 图

1.1.4 结构模型

不同的数据模型具有不同的逻辑结构形式。目前最常用的逻辑模型有层次模型、网状模型和关系模型。其中层次模型和网状模型统称为非关系模型。非关系模型的数据库系统在 20 世纪 70 年代与 80 年代初非常流行，在数据库系统产品中占据了主导地位，现在已逐渐被关系模型的数据库系统取代。

1. 层次模型

层次模型 (Hierarchical Model) 是数据库系统中最早采用的数据模型，它是通过从属关系结构表示数据间的联系。层次模型是有向树结构。现实世界中许多实体之间的联系本来就呈现出一种自然的层次关系，如行政机构、家族关系等。因此层次模型可自然地表达数据间具有层次规律的分类关系、概括关系、部分关系等，但在结构上有一定的局限性。在数据库中，满足以下两个条件的数据模型称为层次模型。

(1) 有且仅有一个无父结点的根结点。

(2) 根结点以外的子结点，向上有且仅有一个父结点，向下可有若干子结点。

若用图形来表示，层次模型是一棵倒立的树，“学校”有且仅有一个结点，没有父结点，称为根结点。有的结点（如“资产”）具有子结点“房产”、“办公用品”和“实验设备”，这样的结点称为父结点，而子结点相互称为兄弟结点，如“学院”、“资产”。“学生”、“教师”等没有子结点，将这样的结点称为叶子结点。

在层次模型中，每一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。任何一个给定的记录值，只有按其路径查看时才能显示出它的全部内容，没有一个子女记录值能够脱离双亲记录值而独立存在。

层次模型的优点是数据结构比较简单，操作也比较简单。对于实体间联系是固定的、且预先定义好的应用系统，层次模型比较适用；层次模型提供了良好的完整性支持。由于层次模型受文件系统影响大，模型受限制多，物理成分复杂，不适用于表示非层次性的联系。

2. 网状模型

网状模型 (Network Model) 是层次模型的扩展，是一种更具有普遍性的结构，它表示多个从