

21世纪高等院校

财经类专业

计算机规划教材

数据库原理与应用教程 (Access 2010版)

唐小毅 吴靖 金鑫 编著

清华大学出版社



21世纪高等院校

财经类专业

计算机规划教材

数据库原理与应用教程

(Access 2010版)

唐小毅 吴靖 金鑫 编著

清华大学出版社
北京

内 容 简 介

本书从一个 Access 数据库应用系统实例——商贸公司的管理系统入手,系统地介绍数据库的基本原理及 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、查询、窗体、宏、报表、VBA 程序设计及数据库编程技术,数据库的安全和管理。

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

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库原理与应用教程: Access 2010 版/唐小毅,吴靖,金鑫编著. —北京:清华大学出版社,2018
(21 世纪高等院校财经类专业计算机规划教材)

ISBN 978-7-302-49546-8

I. ①数… II. ①唐… ②吴… ③金… 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(2018)第 029429 号

责任编辑:孟毅新

封面设计:傅瑞学

责任校对:刘 静

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市国英印务有限公司

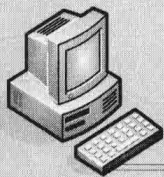
经 销:全国新华书店

开 本:185mm×260mm 印 张:22.25 字 数:509 千字

版 次:2018 年 8 月第 1 版 印 次:2018 年 8 月第 1 次印刷

定 价:56.00 元

产品编号:069794-01



前言

随着信息化进程的不断推进,社会工作和生活都离不开信息技术,对于非计算机专业类学生,掌握一定的计算机科学知识的也越来越高,对于信息管理技术的要求也日趋加深。在日常的工作和学习中,可以深深体会到,企业的信息化,数据组织和处理技术的好坏,往往关系到企业信息化的成败。因此,需要所有非计算机类专业的学生,均要掌握数据建模的基本知识,掌握数据库的设计方法和相关原则,以帮助他们具有一定的企业信息化和数据处理的能力。

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。

本书力求做到与实际教学紧密结合。在内容的组织方面,以学生认知规律作为主线,从数据库的基本知识学习开始,介绍数据库的发展及关系数据库的相关知识,并以 Access 作为工具,介绍从数据库的创建,到数据表操作、各种查询的应用,SQL 语句的撰写,窗体、报表、宏的学习等相关内容,并以一定篇幅介绍了 VBA 程序的编写及数据库编程方法,以及数据库安全与管理的相关技术,以期通过本书的学习,帮助读者掌握一个综合的数据库管理系统的开发技术。

本书除了知识结构的提炼外,还通过大量的图文介绍操作步骤,培养读者理论和实际动手能力。我们期望通过这门课程的学习使读者具备数据库方面的基本知识和良好的逻辑思维能力。

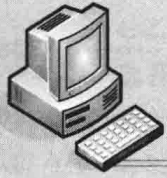
本书主要由中央财经大学教师唐小毅、吴靖和金鑫编写。在此,感谢中央财经大学教师江森和刘艳萍在本书编写过程中付出的劳动。

在本书编写过程中,编者参考了不少书籍,在此向这些书籍的作者表示衷心的感谢!

由于计算机技术日新月异,加之编者能力有限,书中难免有不足之处,恳请读者批评指正。

编者

2018年3月



目 录

| | |
|-------------------------------|----|
| 第 1 章 数据库系统概述 | 1 |
| 1.1 引言 | 1 |
| 1.2 数据库系统 | 2 |
| 1.2.1 数据库系统的构成 | 2 |
| 1.2.2 数据库系统的特点 | 3 |
| 1.2.3 数据管理技术的发展 | 5 |
| 1.3 数据库系统三级模式结构 | 6 |
| 1.3.1 模式结构概念 | 7 |
| 1.3.2 数据库系统三级模式与二级映像 | 7 |
| 1.4 数据库设计的基本步骤 | 8 |
| 1.5 实体—联系模型 | 10 |
| 1.5.1 实体—联系模型中的基本概念 | 10 |
| 1.5.2 实体集之间的联系形式 | 12 |
| 1.6 习题 | 17 |
| 第 2 章 关系模型和关系数据库 | 19 |
| 2.1 数据模型 | 19 |
| 2.2 关系模型的数据结构 | 20 |
| 2.3 关系数据库和关系数据库规范化 | 22 |
| 2.3.1 关系数据库 | 22 |
| 2.3.2 关系数据库规范化 | 22 |
| 2.3.3 关系数据完整性规则 | 25 |
| 2.4 E-R 模型向关系模型的转换 | 25 |
| 2.4.1 实体转换为关系模式 | 25 |
| 2.4.2 实体之间联系的转换 | 25 |
| 2.5 关系数据操作基础 | 29 |
| 2.5.1 集合运算 | 29 |
| 2.5.2 关系运算 | 30 |
| 2.6 习题 | 34 |



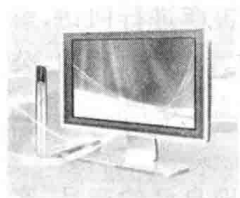
| | |
|---------------------------|----|
| 第 3 章 数据库和表 | 36 |
| 3.1 Access 概述 | 36 |
| 3.1.1 Access 的特点 | 37 |
| 3.1.2 Access 的启动与退出 | 38 |
| 3.1.3 Access 数据库的结构 | 38 |
| 3.1.4 Access 设置 | 41 |
| 3.1.5 帮助系统 | 44 |
| 3.2 创建 Access 数据库 | 45 |
| 3.2.1 创建数据库 | 45 |
| 3.2.2 数据库的简单操作 | 46 |
| 3.3 Access 数据类型 | 48 |
| 3.4 创建数据表 | 51 |
| 3.4.1 创建表 | 51 |
| 3.4.2 设置字段属性 | 56 |
| 3.4.3 修改表结构 | 63 |
| 3.4.4 输入数据 | 65 |
| 3.5 建立联系 | 68 |
| 3.5.1 创建索引与主键 | 68 |
| 3.5.2 建立表之间的联系 | 70 |
| 3.6 操作数据表 | 73 |
| 3.6.1 调整表的外观 | 73 |
| 3.6.2 数据的查找与替换 | 75 |
| 3.6.3 记录排序 | 78 |
| 3.6.4 记录筛选 | 79 |
| 3.6.5 更名、复制和删除 | 81 |
| 3.7 操作实例：商务管理数据库的创建 | 83 |
| 3.7.1 创建数据表 | 83 |
| 3.7.2 建立表间联系 | 88 |
| 3.8 习题 | 90 |
| 第 4 章 查询 | 92 |
| 4.1 查询的功能及类型 | 92 |
| 4.1.1 查询的功能 | 93 |
| 4.1.2 查询的类型 | 93 |
| 4.2 表达式 | 94 |
| 4.2.1 常量 | 95 |
| 4.2.2 Access 常用函数 | 95 |

| | | |
|------------|----------------|------------|
| 4.2.3 | 运算符与表达式 | 99 |
| 4.3 | 选择查询 | 102 |
| 4.3.1 | 利用向导创建查询 | 102 |
| 4.3.2 | 利用设计视图创建查询 | 103 |
| 4.3.3 | 查询属性 | 109 |
| 4.3.4 | 添加计算字段 | 110 |
| 4.3.5 | 总计查询 | 111 |
| 4.4 | 交叉表查询 | 116 |
| 4.4.1 | 利用向导创建交叉表查询 | 116 |
| 4.4.2 | 利用设计视图创建交叉表查询 | 118 |
| 4.5 | 动作查询 | 120 |
| 4.5.1 | 生成表查询 | 120 |
| 4.5.2 | 更新查询 | 121 |
| 4.5.3 | 追加查询 | 122 |
| 4.5.4 | 删除查询 | 124 |
| 4.6 | 参数查询 | 125 |
| 4.6.1 | 单参数查询 | 126 |
| 4.6.2 | 多参数查询 | 126 |
| 4.7 | 其他类型的查询 | 128 |
| 4.7.1 | 查找重复项查询 | 128 |
| 4.7.2 | 查找不匹配项查询 | 130 |
| 4.8 | 操作实例：数据查询 | 131 |
| 4.9 | 习题 | 134 |
| 第5章 | 结构化查询语言 | 136 |
| 5.1 | 数据查询语言 | 137 |
| 5.1.1 | SELECT 语句 | 137 |
| 5.1.2 | 简单查询 | 138 |
| 5.1.3 | 多表查询 | 141 |
| 5.1.4 | 排序 | 143 |
| 5.1.5 | 子查询 | 144 |
| 5.1.6 | 分组查询 | 146 |
| 5.1.7 | 联接查询 | 147 |
| 5.1.8 | 联合查询 | 148 |
| 5.2 | 数据定义语言 | 149 |
| 5.2.1 | 创建表 CREATE | 149 |
| 5.2.2 | 修改表 ALTER | 151 |
| 5.2.3 | 删除表 DROP | 152 |

| | | |
|------------|-----------------------|------------|
| 5.3 | 数据操纵语言 | 152 |
| 5.3.1 | 追加 INSERT | 152 |
| 5.3.2 | 更新 UPDATE | 153 |
| 5.3.3 | 删除 DELETE | 153 |
| 5.4 | 习题 | 154 |
| 第6章 | 窗体 | 156 |
| 6.1 | 窗体概述 | 157 |
| 6.1.1 | 窗体的功能 | 157 |
| 6.1.2 | 窗体的结构 | 157 |
| 6.1.3 | 窗体的类型 | 159 |
| 6.1.4 | 窗体的视图 | 160 |
| 6.2 | 创建快速窗体 | 160 |
| 6.2.1 | 自动窗体 | 160 |
| 6.2.2 | 利用向导创建窗体 | 161 |
| 6.2.3 | 创建数据透视表和数据透视图窗体 | 163 |
| 6.2.4 | 创建图表窗体 | 165 |
| 6.3 | 设计视图创建窗体 | 168 |
| 6.3.1 | 窗体设计视图 | 168 |
| 6.3.2 | 常用控件的功能 | 169 |
| 6.3.3 | 常用控件的使用 | 171 |
| 6.3.4 | 窗体中控件的常用操作 | 176 |
| 6.4 | 修饰窗体 | 179 |
| 6.4.1 | 利用主题 | 179 |
| 6.4.2 | 利用属性 | 179 |
| 6.4.3 | 利用条件格式 | 180 |
| 6.4.4 | 提示信息的添加 | 181 |
| 6.5 | 定制系统控制窗体 | 183 |
| 6.5.1 | 创建切换窗体 | 183 |
| 6.5.2 | 创建导航窗体 | 187 |
| 6.5.3 | 设置启动窗体 | 188 |
| 6.6 | 对象与属性 | 188 |
| 6.6.1 | 面向对象的基本概念 | 189 |
| 6.6.2 | 对象属性 | 189 |
| 6.6.3 | 对象事件和方法 | 192 |
| 6.7 | 窗体设计实例 | 194 |
| 6.8 | 习题 | 204 |
| 第7章 | 报表 | 206 |
| 7.1 | 概述 | 206 |

| | | |
|--------------|------------------|------------|
| 7.1.1 | 报表的功能 | 206 |
| 7.1.2 | 报表的视图 | 207 |
| 7.1.3 | 报表的结构 | 208 |
| 7.1.4 | 报表的类型 | 210 |
| 7.2 | 创建报表 | 212 |
| 7.2.1 | 使用“报表”工具自动创建报表 | 212 |
| 7.2.2 | 使用“报表向导”工具创建报表 | 213 |
| 7.2.3 | 使用“标签向导”工具创建标签报表 | 219 |
| 7.2.4 | 使用“报表设计”工具创建报表 | 219 |
| 7.2.5 | 使用“空报表”工具创建报表 | 223 |
| 7.3 | 编辑报表 | 226 |
| 7.3.1 | 设置报表格式 | 226 |
| 7.3.2 | 修饰报表 | 228 |
| 7.3.3 | 创建多列报表 | 230 |
| 7.4 | 报表的高级应用 | 231 |
| 7.4.1 | 报表的排序和分组 | 231 |
| 7.4.2 | 使用计算控件 | 236 |
| 7.4.3 | 创建子报表 | 237 |
| 7.5 | 报表的预览和打印 | 241 |
| 7.6 | 操作实例 | 242 |
| 7.7 | 习题 | 247 |
| 第 8 章 | 宏 | 249 |
| 8.1 | 宏的概念 | 249 |
| 8.2 | 宏的创建与编辑 | 250 |
| 8.2.1 | 操作序列宏的创建 | 251 |
| 8.2.2 | 宏操作分组 | 253 |
| 8.2.3 | 子宏的创建 | 254 |
| 8.2.4 | 条件宏的创建 | 255 |
| 8.2.5 | 宏的编辑 | 257 |
| 8.3 | 宏的运行和调试 | 258 |
| 8.3.1 | 宏的运行 | 258 |
| 8.3.2 | 宏的调试 | 261 |
| 8.4 | 操作实例 | 261 |
| 8.5 | 常用宏操作 | 263 |
| 8.6 | 习题 | 265 |
| 第 9 章 | VBA 与模块 | 267 |
| 9.1 | VBA 简介 | 267 |

| | | |
|---------------|-----------------------|------------|
| 9.1.1 | VBA 的概念 | 268 |
| 9.1.2 | VBA 编辑环境介绍 | 268 |
| 9.1.3 | 模块简介 | 271 |
| 9.2 | VBA 程序设计基础 | 274 |
| 9.2.1 | 数据类型 | 274 |
| 9.2.2 | 常量与变量 | 276 |
| 9.2.3 | 数组 | 279 |
| 9.2.4 | VBA 表达式 | 281 |
| 9.3 | VBA 基本语句 | 282 |
| 9.3.1 | 基本语法规则 | 282 |
| 9.3.2 | 赋值语句 | 283 |
| 9.3.3 | 交互式输入 | 284 |
| 9.3.4 | 输出语句 | 284 |
| 9.4 | 流程控制语句 | 286 |
| 9.4.1 | 分支结构 | 287 |
| 9.4.2 | 循环结构 | 293 |
| 9.4.3 | 过程调用与参数传递 | 300 |
| 9.4.4 | 变量的作用域与生存期 | 304 |
| 9.5 | VBA 常用操作 | 307 |
| 9.5.1 | DoCmd 命令 | 307 |
| 9.5.2 | 打开和关闭操作 | 308 |
| 9.5.3 | 操作实例 | 310 |
| 9.6 | VBA 数据库编程 | 312 |
| 9.6.1 | VBA 数据库编程技术简介 | 312 |
| 9.6.2 | 数据库编程示例 | 316 |
| 9.7 | 习题 | 323 |
| 第 10 章 | 数据库安全与管理 | 327 |
| 10.1 | 数据库的安全性 | 327 |
| 10.1.1 | 数据库加密与解密 | 327 |
| 10.1.2 | 通过备份和还原保护数据 | 329 |
| 10.2 | 数据库的管理 | 333 |
| 10.2.1 | 压缩和修复数据库 | 333 |
| 10.2.2 | 数据导入与导出 | 335 |
| 10.2.3 | 数据库文件格式转换 | 339 |
| 10.3 | 习题 | 342 |
| 参考文献 | | 344 |



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)等指出,信息资源(Information Resources)与人力、物力、财力和自然资源一样,都是企业的重要资源,因此,应该像管理其他资源那样管理信息资源。

数据是信息时代的重要资源之一。商业的自动化和智能化,使得企业收集到了大量的数据,积累下来许多重要资源。政府、企业等各类组织需要对大量的数据进行管理,从数据中获取信息和知识,从而进行决策,于是就有了数据库蓬勃发展的今天。数据库技术是计算机科学中一门重要的技术,数据库技术在政府、企业等机构得到广泛的应用。特别是 Internet 技术的发展,为数据库技术得以迅速发展奠定了重要基础。

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。

知识体系:

- ☞ 数据库系统的基本概念
- ☞ 数据库设计的基本步骤
- ☞ E-R 模型

学习目标:

- ☞ 了解数据库系统的发展历程
- ☞ 掌握数据库系统的三级模式结构
- ☞ 掌握数据库设计的基本步骤
- ☞ 学会利用 E-R 图进行数据库系统的设计

1.1 引言

首先,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,为分别适应不同客户群的需求,这种商品有九个型号;产品通过分布在全市的 3000 多个各种类型零售商销售(例如,各类超市、便利店等);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；等等。A公司的工作人员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，公司还需要对自己的市场部门工作业绩进行考核。这个例子涉及了产品、客户、员工和订单。

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。

这样大量且相互关联的数据，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其数据。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？这正是讨论的问题。

另一个例子是银行，大概每个人都有在银行接受服务的经历。首先在银行开户，向银行提供个人的基本数据，例如，姓名和身份证号码，之后作为银行客户就会不断地存款、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，数据库系统的特点以及数据库技术的发展历史。

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有两个：一是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；二是如何方便、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、应用程序以及用户。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,1970,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。1970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。例如，这组数据描述

的是学生,描述学生的属性包括学号、姓名、性别、出生日期、籍贯、所属学院,则上述数据就是这一组属性的值。

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)是一种系统软件,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,例如,Microsoft的SQL Server、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。DBMS是数据库系统的核心,其关系如图1.1所示。用户以及应用程序都是通过数据库管理系统对数据库中数据进行访问的。

通常,一个数据库系统应该具备以下功能。

(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) 控制并发访问。

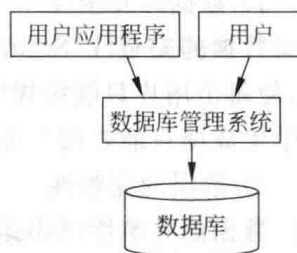


图 1.1 数据库系统组成

1.2.2 数据库系统的特点

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,关系数据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,从而使数据既具有逻辑独立性,又具有物理独立性。

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。

数据库系统的另一个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) 数据的安全性

数据的安全性(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,部分用户对学习成绩只能查阅不能修改。

2) 数据的完整性

数据的完整性(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)的功能,这就是数据库的恢复(Recovery)功能。

5. 数据库发展过程

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

为四种类型,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;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,基础表具有以下特性: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。数据仓库是从多个数据源收集的信息存储,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。

1.2.3 数据管理技术的发展

随着计算机应用范围的不断扩大,也伴随着各领域对数据处理的需求不断增强,数据管理技术在不断地发展。

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,经历了以下三个阶段:人工管理阶段、文件系统阶段和数据库技术阶段。对数据有效地管理,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。人工管理阶段的数据管理特点如下。

(1) 数据不保存。一个目标计算完成后,程序和数据都不被保留。

(2) 应用程序管理数据。应用程序与所要处理的数据集是一一对应的,应用程序与

数据之间缺少独立性。

(3) 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。

(4) 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了科学计算和数据处理。1954年出现了第一台商业数据处理的计算机UNIVAC I,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。这种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。

(1) 数据可以长期保存。数据能够被保存在存储设备上,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。

(2) 由文件系统管理数据。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。

(3) 程序与数据相互独立。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。

(4) 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,由此造成数据冗余,并可能导致数据不一致;数据不能共享。

(5) 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,应该程序改变时,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,特别是在20世纪80年代后期到90年代,由于金融和商业的需求,数据库技术得到了迅猛的发展。数据库管理系统管理数据具有以下特点。

(1) 数据结构化。

(2) 数据共享性高,冗余度低,易扩充。

(3) 数据独立性高。

(4) 数据由DBMS统一管理,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。数据库系统三级模式结构如图1.2所示。

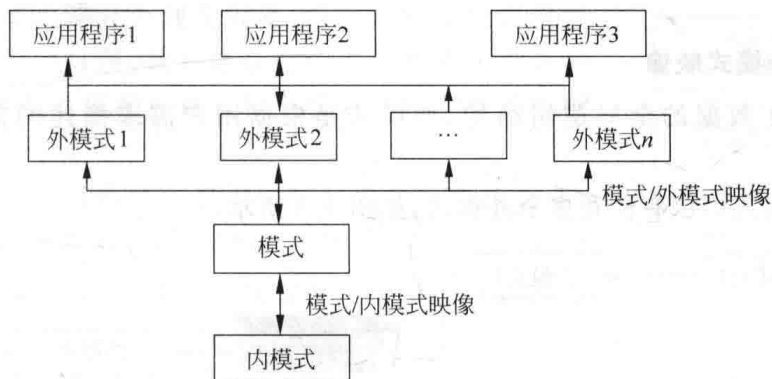


图 1.2 数据库系统三级模式结构

1.3.1 模式结构概念

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,这个模式可以有多组不同的值与其对应,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,一个模式可以有多个外模式。

外模式可以很好地起到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,不能看到模式中的全部数据。

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