

21世纪高等学校计算机规划教材

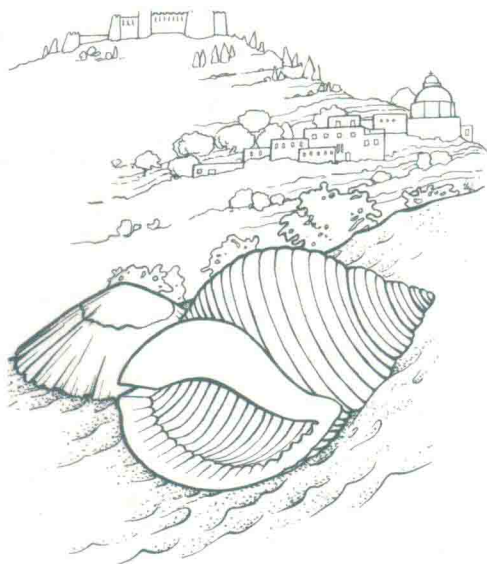
21st Century University Planned Textbooks of Computer Science

数据库技术与应用 ——Access 2010

Technology and Application of Database
——Access 2010

姜林枫 徐长滔 杨燕 曹锋 盛欣 等 编著

- 知行统一地搭建知识体系架构
- 由浅入深地讲解数据工程技术
- 日积月累地培养应用创新能力



高校系列



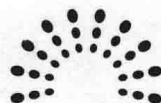
中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

21st Century University Planned

姜林枫 徐长滔 杨燕 曹锋 盛欣 等 编著



高校系列

人民邮电出版社

北京

图书在版编目 (CIP) 数据

数据库技术与应用 : Access 2010 / 姜林枫等 编
著. — 北京 : 人民邮电出版社, 2017. 1
21世纪高等学校计算机规划教材
ISBN 978-7-115-44475-2

I. ①数… II. ①姜… 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(2016)第321677号

内 容 提 要

本书基于 Microsoft Access 2010 关系数据库管理系统, 系统地介绍了数据库的基本理论、成熟技术和经典应用: 基于数据库技术产生的原因, 介绍数据库的概念、理论和技术, 建立学习数据库技术的学习框架; 基于典型案例讲解数据库的设计、创建及管理, 培养学生数据组织和管理的能力; 基于大量示例介绍 SQL 的应用, 培养学生数据处理和分析的能力; 基于学以致用原则, 介绍模块对象设计、查询对象设计、窗体对象设计、宏对象设计、报表对象设计和数据库应用系统的开发, 培养学生的应用创新能力。

本书可作为普通高等学校数据库技术与程序设计方面课程的教材, 还可作为全国计算机等级考试二级 Access 程序设计的培训教材, 或者作为广大信息管理用户和数据分析技术初学者的自学用书。

◆ 编 著 姜林枫 徐长滔 杨 燕 曹 锋 盛 欣 等

责任编辑 吴 婷

责任印制 沈 蓉 彭志环

◆ 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>

中国铁道出版社印刷厂印刷

▼ 开本: 787×1092 1/16

印张: 20

字数: 552 千字

2017 年 1 月第 1 版

2017 年 1 月北京第 1 次印刷

定价: 49.80 元

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

反盗版热线: (010)81055315

前 言

随着数据成为国民经济的三大资源之一,“数据库技术与应用”从一门重要的专业基础课演变为一门重要的公共课。近年来,越来越多的高校将数据库技术与应用课程列入了非计算机专业的计算机课程的教学范畴。与此同时,随着社会各个领域对信息管理和应用人才需求的迅猛增长,数据库技术与应用课程的普及呈现出进一步扩展和延伸的趋势。

数据库技术与应用课程的知识体系包括基础理论、成熟技术和经典应用三大部分。本书基于 Microsoft Access 2010,以销售订单管理和学生成绩管理为背景,深入浅出地介绍了数据组织、处理和分析等方面的理论、技术和应用。本书具有三大特点:一是基于案例分析进行启发式教学,二是基于数据建模进行数据工程能力的培养,三是基于对象设计进行创新能力的培养。

针对学生具体操作中可能出现的各种困难,本书提供了相应的实验素材,配套使用将使学习效果更佳。另外,为方便教学,本书还免费提供教学用的电子教案和全部习题解答,读者可到 <http://www.ryjiaoyu.com> 网站下载。

本书主要由齐鲁工业大学的姜林枫、徐长滔、杨燕、曹锋、盛欣、陈玲、代新利编著,另外参加本书编写工作的还有林冬梅和赵龙两位。其中:第 1 章、第 2 章、第 5 章、第 6 章、第 7 章、第 8 章的 8.5 节由姜林枫编写;第 3 章由盛欣编写;第 4 章、第 12 章由曹锋编写;第 8 章的前 4 节以及第 9 章由杨燕编写;第 10 章的前 5 节以及第 11 章由徐长滔编写;第 10 章的 10.6 节由赵龙编写;各章习题由林冬梅、陈玲和代新利编写。本书的读者分析由陈玲负责,代码验证由赵龙负责。全书由姜林枫负责框架结构的设计、初稿的修改和最后的统稿工作。

第 3 章 数据库的创建与 编者 52

2016 年 9 月

第 2 章 Access 基础 26

3.1 数据库的创建 52

3.1.1 空数据库的创建 52

3.1.2 简单数据库的创建 53

3.2 数据库对象的分类 55

3.2.1 数据库的打开 55

3.2.2 数据库对象的视图 56

3.2.3 数据库对象的打印 57

3.2.4 数据库的属性管理 58

3.2.5 数据库的保存与关闭 59

3.2.6 数据库的删除 60

3.3 数据库的安全管理 61

3.3.1 Access 数据库的安全体系 61

目 录

第 1 章 数据库技术基础.....1

1.1 数据库技术的产生	1
1.1.1 文件技术组织数据的弊端	1
1.1.2 数据库技术组织数据的优势	4
1.2 数据库技术的相关概念	5
1.2.1 数据和信息	5
1.2.2 数据库	6
1.2.3 数据库管理系统	6
1.2.4 数据库应用程序	7
1.2.5 数据库系统	7
1.2.6 数据模型	8
1.2.7 数据库模式	10
1.2.8 数据库语言	10
1.3 关系数据库的数据模型	11
1.3.1 关系数据库的数据结构	11
1.3.2 关系数据库的数据操作	13
1.3.3 关系数据库的数据约束	15
1.4 关系数据库的数据库模式	16
1.5 关系数据库的设计	17
1.5.1 关系数据库设计的内容	17
1.5.2 关系数据库设计的步骤	17
1.5.3 关系数据库概念设计的方法	18
1.5.4 E-R 模型转换为关系数据库 模式	21
习题	24

第 2 章 Access 基础..... 26

2.1 Access 的用户界面	26
2.1.1 Access 的 Backstage 界面	26
2.1.2 Access 的主界面	27
2.2 Access 的工作环境	28
2.2.1 数据库默认文件夹的设置	29
2.2.2 功能区的设置	30
2.2.3 快速访问工具栏的设置	31

2.3 Access 的数据库对象	32
2.3.1 表	32
2.3.2 查询	33
2.3.3 窗体	33
2.3.4 报表	33
2.3.5 宏	34
2.3.6 模块	34
2.4 Access 中的数据	34
2.4.1 Access 中的数据类型	35
2.4.2 Access 中的函数	38
2.4.3 Access 中的表达式	42
2.4.4 表达式的计算输出	45
2.5 Access 的设计工具	46
2.5.1 模板	46
2.5.2 向导	48
2.5.3 设计器	48
2.5.4 生成器	49
2.6 Access 的操作方式	49
2.6.1 交互方式	49
2.6.2 批处理方式	50
习题	50

第 3 章 数据库的创建与管理 52

3.1 数据库的创建	52
3.1.1 空数据库的创建	52
3.1.2 简单数据表的创建	53
3.2 数据库的日常管理	55
3.2.1 数据库的打开	55
3.2.2 数据库对象的视图	56
3.2.3 数据库对象的组织	57
3.2.4 数据库的属性管理	58
3.2.5 数据库的保存与关闭	59
3.2.6 数据库的删除	60
3.3 数据库的安全管理	61
3.3.1 Access 数据库的安全体系	61

3.3.2 信任中心的设置.....	61	5.5.1 数据分析概述.....	116
3.3.3 数据库的打包、签名和分发.....	63	5.5.2 分析型查询对象设计的案例分析.....	117
3.3.4 密码的设置与撤销.....	65	5.6 操作型查询对象的设计和应用.....	120
3.3.5 数据库的备份与还原.....	67	5.6.1 生成表查询.....	121
3.3.6 数据库的压缩与修复.....	68	5.6.2 追加查询.....	122
习题.....	69	5.6.3 更新查询.....	123
第4章 表对象的创建与维护.....	71	5.6.4 删除查询.....	123
4.1 表的创建.....	71	习题.....	124
4.1.1 表结构的创建.....	71	第6章 数据库语言 SQL.....	126
4.1.2 表数据的追加.....	75	6.1 SQL 概述.....	126
4.1.3 表索引的创建.....	76	6.1.1 SQL 的功能.....	126
4.1.4 表约束的创建.....	77	6.1.2 SQL 的特点.....	127
4.2 表的维护.....	80	6.1.3 SQL 语句.....	127
4.2.1 表结构的维护.....	80	6.1.4 Access 支持的 SQL.....	127
4.2.2 表数据的维护.....	81	6.2 SQL 的定义功能.....	129
4.2.3 表约束的维护.....	82	6.2.1 数据库的创建.....	129
4.2.4 表的应用维护.....	83	6.2.2 数据表的定义.....	129
习题.....	89	6.2.3 索引的定义.....	135
第5章 查询对象的设计与应用.....	92	6.2.4 视图的定义.....	137
5.1 查询对象概述.....	92	6.3 SQL 的更新功能.....	138
5.1.1 查询对象的概念.....	92	6.3.1 插入数据.....	138
5.1.2 查询对象和表对象的关系.....	92	6.3.2 更新数据.....	139
5.1.3 查询对象的应用.....	93	6.3.3 删除数据.....	139
5.1.4 查询的类型.....	93	6.4 SQL 的查询功能.....	139
5.1.5 查询对象的视图.....	94	6.4.1 简单查询.....	141
5.2 查询对象的设计工具.....	95	6.4.2 嵌套查询.....	144
5.2.1 查询向导.....	95	6.4.3 连接查询.....	145
5.2.2 查询设计视图.....	104	6.4.4 统计查询.....	147
5.3 检索型查询对象的设计和应用.....	106	6.4.5 集合查询.....	149
5.3.1 无条件检索查询.....	106	6.5 SQL 的综合应用.....	150
5.3.2 静态条件检索查询.....	108	习题.....	155
5.3.3 动态条件检索查询.....	110	第7章 宏对象的设计及应用.....	158
5.4 计算型查询对象的设计和应用.....	113	7.1 宏对象概述.....	158
5.4.1 行计算型查询对象设计的案例分析.....	113	7.1.1 宏对象的概念.....	158
5.4.2 列计算型查询对象设计的案例分析.....	115	7.1.2 宏对象的功能.....	158
5.5 分析型查询对象的设计和应用.....	116	7.1.3 宏对象的分类.....	159
		7.1.4 宏的组成.....	159
		7.1.5 宏的操作命令.....	160

7.2 宏对象的设计界面	161
7.2.1 “宏设计工具”选项卡	162
7.2.2 “宏设计”窗格	163
7.2.3 “操作目录”窗格	163
7.3 宏对象的创建与编辑	165
7.3.1 宏对象的创建	165
7.3.2 宏对象的编辑	170
7.4 宏对象的执行与调试	170
7.4.1 宏对象的执行	170
7.4.2 宏对象的调试	171
7.5 宏对象的应用示例	172
7.5.1 序列宏	172
7.5.2 条件宏	173
7.5.3 自动运行宏	175
习题	175

第 8 章 VBA 程序设计语言 177

8.1 程序设计语言概述	177
8.1.1 程序	177
8.1.2 程序设计方法	177
8.1.3 程序设计语言	177
8.1.4 VBA 简介	178
8.2 VBA 语法知识	180
8.2.1 关键字和标识符	180
8.2.2 常量和变量	180
8.2.3 数据类型	182
8.2.4 表达式	183
8.2.5 函数	185
8.2.6 声明语句	189
8.2.7 赋值语句	190
8.2.8 注释语句	191
8.3 VBA 流程控制语句	191
8.3.1 选择控制语句	191
8.3.2 循环控制语句	194
8.4 面向过程的程序设计	197
8.4.1 Sub 过程	197
8.4.2 Function 过程	199
8.5 面向对象的程序设计	200
8.5.1 面向对象程序设计基础	200
8.5.2 Access VBA 中的对象类	204

8.5.3 Access VBA 中的对象模型	204
8.5.4 Access VBA 中的集合对象	206
8.5.5 Access VBA 对象的属性、方法与事件	206
8.5.6 Access VBA 对象的引用	209
8.5.7 Access VBA 对象的设计	210
8.5.8 基于 DAO 接口的 Access 数据库的访问	211
习题	220

第 9 章 模块对象的设计及应用 ... 224

9.1 模块对象概述	224
9.1.1 模块对象的概念	224
9.1.2 模块对象的分类	224
9.1.3 模块对象的组成	225
9.2 模块对象的建立	226
9.2.1 类模块的建立	226
9.2.2 标准模块的建立	228
9.3 模块对象的执行与调试	229
9.3.1 模块对象的执行	229
9.3.2 模块的调试	230
9.4 模块对象的应用案例	232
9.4.1 模块中过程的协作	232
9.4.2 将宏转换为模块	234
习题	236

第 10 章 窗体对象的设计及应用 237

10.1 窗体对象概述	237
10.1.1 窗体对象的概念	237
10.1.2 窗体对象的功能	237
10.1.3 窗体的类型	237
10.1.4 窗体对象的视图	238
10.1.5 创建窗体对象的功能按钮	239
10.1.6 创建窗体对象的主要方法	240
10.2 使用向导创建窗体对象	240
10.2.1 自动创建窗体	240
10.2.2 创建简单窗体	241
10.2.3 创建分割窗体	242
10.2.4 创建数据透视表窗体	242

10.2.5	创建数据透视图窗体	244
10.3	使用设计视图创建窗体对象	245
10.3.1	窗体对象的设计视图	245
10.3.2	窗体对象的设计工具	245
10.3.3	窗体对象的常用控件	246
10.3.4	窗体和控件的常用属性	247
10.3.5	窗体和控件的常用事件	248
10.3.6	控件的常用操作	249
10.4	常用窗体控件的设计及应用	252
10.4.1	标签与文本框控件	252
10.4.2	命令按钮控件	253
10.4.3	选项按钮与复选框	255
10.4.4	列表框与组合框	256
10.4.5	选项卡控件	258
10.4.6	子窗体控件	259
10.5	窗体对象的完善和美化	261
10.5.1	设置窗体的背景图片	261
10.5.2	设置条件格式	262
10.5.3	设置窗体的主题	262
10.6	窗体对象的应用示例	263
	习题	271

第 11 章 报表对象的设计及应用 274

11.1	报表概述	274
11.1.1	报表的类型	274
11.1.2	报表对象的视图	274
11.1.3	创建报表对象的方法	275
11.2	使用向导创建报表对象	275
11.2.1	自动创建报表	275
11.2.2	使用报表向导创建报表	276
11.2.3	使用标签向导创建报表	278

11.3	使用设计视图创建报表对象	279
11.3.1	报表对象的设计视图	279
11.3.2	报表对象的设计工具	280
11.3.3	报表对象的创建起点——页面设置	280
11.4	常用报表控件的设计及应用	281
11.4.1	在报表中添加简单控件	281
11.4.2	在报表中添加计算控件	282
11.4.3	在报表中添加排序和分组控件	284
11.4.4	在报表中添加子报表控件	286
11.5	报表对象的完善和美化	287
11.5.1	修饰报表的常用方法	288
11.5.2	报表的预览与打印	289
	习题	290

第 12 章 数据库应用系统的开发 292

12.1	Access 数据库的访问技术	292
12.1.1	常用的数据库访问接口	292
12.1.2	ADO 对象模型	293
12.1.3	使用 ADO 访问 Access 数据库的基本步骤	294
12.2	数据库应用系统的开发概述	294
12.2.1	数据库应用系统开发的一般过程	295
12.2.2	数据库设计的步骤	296
12.3	案例分析——订单管理系统的开发	298
12.3.1	需求分析	298
12.3.2	系统设计	299
12.3.3	系统实现	301
	习题	312

第 1 章

数据库技术基础

数据库技术是随着信息技术的发展和人们对信息需求的增加发展起来的，它主要研究如何科学地组织和管理数据，以提供可共享、安全和可靠的数据。数据库的建设规模、信息容量和使用频度已成为衡量一个国家信息化程度的重要标志。

基于微软发布的关系数据库管理系统 Microsoft Access 2010，本书对关系数据库的技术和应用进行了全面的讲解。本章首先介绍数据库技术的产生和相关概念；然后介绍主流数据库——关系数据库的数据模型，这个内容是数据库技术的基础，将贯穿本书始终；最后介绍关系数据库的设计技术。请读者注意的是，由于关系数据库是市场主流，本书只介绍关系数据库技术，如无特别明示，本书提到的数据库都是关系数据库。

1.1 数据库技术的产生

随着智能手机和计算机的普及，信息技术已经融入了我们的生活，人们除了用智能终端进行网购、理财、支付、学习、聊天和娱乐外，很多人开始使用文件等信息技术来组织、保存和管理生活中的数据，例如，使用 Excel 工作表记录家庭财务信息，又如使用 Word 表格记录通讯信息，还有利用电子文档记日记等。

然而，使用工作表之类的文件技术组织和管理数据常常出现一些问题，如数据管理异常问题、数据冗余问题、数据独立性问题、数据的共享问题等，这些问题经常困扰着人们，于是数据库技术应时而生。

本节下文将以数据管理异常问题为视角，首先通过分析 Excel 工作表文件组织数据所导致的数据操作异常，说明文件技术组织和管理数据的弊端，然后通过几个示例说明如何利用数据库技术来解决这些问题，从而将数据库技术和文件技术对数据组织和管理差异揭示出来，自然而然地告诉人们数据库技术产生的必然性。

1.1.1 文件技术组织数据的弊端

有一定计算机文化基础的人，可能觉得数据的组织和管理好像用文件技术就可以了，例如，Microsoft 公司的 Excel 可以将数据组织成一个个的工作表，通过对工作表的有效管理，就可以将数据提供给用户分析和使用了。

这种观点，对于一些简单结构的数据来说，应该是正确的。但当数据的数据结构比较复杂的时候，就会发现文件技术组织和管理数据存在数据不一致或数据管理困难等方面的问题。下面以 Microsoft 公司的 Excel 为例，分析文件技术组织数据的一些弊端。

表 1-1 是一个关于销售员 E-mail 的简单工作表。由于这个工作表的主题很简单，所以管理也很轻松，也就是查询销售员 E-mail、添加销售员 E-mail、修改销售员 E-mail 或者删除销售员 E-mail。对于这样的工作表，使用 Excel 之类的电子表格文件技术绰绰有余。

即使工作表中销售员 E-mail 行很多，也可以按“销售员姓名”这一列或按“销售员 E-mail”列排序，以提高检索速度，降低管理难度。总之，使用 Excel 之类的文件技术组织和管理表 1-1 所示的销售员 E-mail 信息没有任何问题，不需要麻烦数据库技术。

表 1-1 电子表格格式的销售员 E-mail 表

销售员姓名	销售员 E-mail
姜刘敏	547948328@qq.com
徐莉莉	lxul127@163.com
宋苏娟	276960500@qq.com
李晓东	Lidong91928@163.com
张大猛	774568142@qq.com
耿小丽	1570818754@qq.com

如果在表 1-1 中增加两列，存储销售员所服务的顾客手机号码，形成表 1-2，虽然仍然可以使用文件技术组织和管理，但有些操作会出现问题。

例如，假设要删除销售员张大猛的数据（如表 1-2 所示），那么就需要删除工作表的第 5 行，这时，我们会发现不仅删除了销售员张大猛的数据，也删除了顾客的姓名和电话（杨燕燕，17788816961）。上面看到的这个删除异常问题，是 Excel 之类的文件技术组织数据不可避免的。

表 1-2 销售员/顾客工作表的删除问题

销售员姓名	销售员 E-mail	顾客姓名	顾客电话
姜刘敏	547948328@qq.com	姜笑枫	17788816965
徐莉莉	lxul127@163.com	徐涛	17788816967
宋苏娟	276960500@qq.com	姜笑枫	17788816965
李晓东	Lidong91928@163.com	徐涛	17788816967
张大猛	774568142@qq.com	杨燕燕	17788816961
耿小丽	1570818754@qq.com	徐涛	17788816967

删除行
丢失了过多的数据

同样，更新工作表中的值也会导致一些意外结果。例如，如果改动了表 1-3 中第 1 行的手机号码，数据就会不一致。改动后，第 1 行显示了顾客姜笑枫的一个手机号码，第 3 行却显示该顾客的另一个手机号码，这就导致了数据的不一致性。

表 1-3 销售员/顾客工作表中的修改问题

销售员姓名	销售员 E-mail	顾客姓名	顾客电话
姜刘敏	547948328@qq.com	姜笑枫	17788816966
徐莉莉	lxul127@163.com	徐涛	17788816967
宋苏娟	276960500@qq.com	姜笑枫	17788816965
李晓东	Lidong91928@163.com	徐涛	17788816967
张大猛	774568142@qq.com	杨燕燕	17788816961
耿小丽	1570818754@qq.com	徐涛	17788816967

修改行
不一致的数据

执行修改操作后，工作表会导致这样的困惑：是顾客姜笑枫有两个不同的手机号码，还是两个同名顾客各有一个手机号码？这就是说，如果使用文件技术对表 1-3 执行更新操作，工作表中的数据可能会产生更新不一致的问题，这会让用户产生困惑，导致数据语义的不确定性。

最后，如果要给没有关联销售员的顾客添加数据，该如何做？例如顾客孙叶青没有自己的关联销售员，但是仍需要存储她的手机号码，此时就必须在工作表的销售员姓名和 E-mail 字段中插入空值（待定的值，不知道的值），这样就出现了值不完全的行，如表 1-4 所示。值不完全的行在管理、维护和使用时会带来很多问题，应尽量避免使用。

表 1-4 销售员/顾客工作表中的插入问题

销售员姓名	销售员 E-mail	顾客姓名	顾客电话
姜刘敏	547948328@qq.com	姜笑枫	17788816966
徐莉莉	Ixul127@163.com	徐涛	17788816967
宋苏娟	276960500@qq.com	姜笑枫	17788816965
李晓东	Lidong91928@163.com	徐涛	17788816967
张大猛	774568142@qq.com	杨燕燕	17788816961
耿小丽	1570818754@qq.com	徐涛	17788816967
NULL	NULL	孙叶青	17788816962

不完全的数据插入行

为什么对表 1-1 这样一个简单的工作表，再添加两列就会带来上述的删除异常、更新不一致和插入空值等问题呢？这难道是工作表的列数问题吗？

带着这个问题，我们又设计了一个同样具有 4 列的销售员/宿舍工作表，如表 1-5 所示。然后再分析一下对表 1-5 中的数据进行插入、删除和修改会不会出现操作异常。

在表 1-5 所示的销售员/宿舍工作表中，如果删除销售员张大猛的数据，仅会丢失与该销售员相关的数据，没有删除其他实体的数据。同样，修改销售员姜刘敏的字段值也不会带来任何更新不一致问题。最后，添加销售员马晓秀的数据也不会导致空值行的出现。

看来不是工作表列数的问题。那么又是什么原因呢？仔细观察表 1-3 和表 1-5 你会发现，表 1-3 组织的数据和表 1-5 组织的数据有一个本质区别：表 1-5 中的销售员/宿舍工作表中的数据是关于一个实体的，所有数据都和销售员有关，所添加的两列都是销售员这个实体的手机号码和宿舍信息。而表 1-3 的销售员/顾客工作表是关于两个实体的，有些数据和销售员有关，有些数据和顾客有关。

表 1-5 销售员/宿舍工作表

销售员姓名	销售员 E-mail	销售员电话	销售员宿舍
姜刘敏	547948328@qq.com	15999916912	公寓 2#501
徐莉莉	Ixul127@163.com	15999916916	公寓 2#501
宋苏娟	276960500@qq.com	15999916915	公寓 2#501
李晓东	Lidong91928@163.com	15999916919	公寓 1#201
张大猛	774568142@qq.com	15999916917	公寓 1#201
耿小丽	1570818754@qq.com	15999916915	公寓 2#109

通常情况下，只要工作表中的数据是关于两个或多个不同的实体的，修改、删除以及添加行就会出现插入、删除和修改的异常问题。

原因找到了，有的读者会说，这好办，把数据分别组织在 Excel 工作簿的不同工作表中，每

一个工作表只保存一个实体的数据，问题不就解决了吗？好像这种解决方案是正确的，但很遗憾地告诉你：这种解决方案会导致新问题的出现，当把不同实体的数据分别放在工作簿的不同工作表之中后，不同实体就被割裂开来，实体之间的固有联系被人为切断，很难基于实体之间的固有联系，对不同实体的数据进行集成分析和使用。

1.1.2 数据库技术组织数据的优势

早在 20 世纪 60 年代，运用工作表之类的文件技术组织数据的弊端就被发现了，因此业界一直在寻找一种技术来组织数据以克服这些弊端，不少相关技术应运而生。随着时间的流逝，基于关系模型的数据库技术成为计算机人的选择。现在，主流的商用数据库都是基于关系模型的。基于关系模型的数据库，称为关系数据库，它的基本特征是使用关系表来组织和管理数据，本章 1.3 节将深入介绍关系模型的相关内容，这里只是用关系数据表来组织和管理表 1-2 中的数据，看看是否可以解决用工作表文件管理数据时所产生的问题。

图 1-1 包括两个表：seller 表和 Customer 表。下面分析一下对图 1-1 中的数据进行删除、修改和插入操作是否会出现上面提到的删除异常、更新不一致以及插入空值行的问题。



销售员姓名	销售员 E-mail	顾客姓名
姜刘敏	547948328@qq.com	姜笑枫
徐莉莉	lxu1127@163.com	徐涛
宋苏娟	276960500@qq.com	姜笑枫
李晓东	Lidong91928@163.com	徐涛
张大猛	774568142@qq.com	杨燕燕
耿小丽	1570818754@qq.com	徐涛

顾客姓名	顾客电话
姜笑枫	17788816965
徐涛	17788816967
杨燕燕	17788816962

图 1-1 seller 表和 Customer 表

(1) 删除操作。例如从 seller 表中删除销售员张大猛的数据，只是删除了销售员张大猛的数据，他的顾客杨燕燕的数据信息仍然保存在 Customer 表中。

(2) 修改操作。如果将顾客杨燕燕的手机号码改为 13188896888，显然不会出现数据行不一致的数据，因为顾客杨燕燕的电话信息仅在 Customer 表中存储一次。

(3) 插入操作。如果需要添加顾客孙叶青的信息，只需将她的数据添加到 Customer 表中就可以了。因为现在没有销售员关联顾客孙叶青，因此在 seller 表中不会出现空值行。

通过上面的分析，得到一个结论，使用数据库技术组织和管理数据可以解决文件技术所遇到的操作异常问题，关键的原因在于两种技术的数据组织不同：数据库技术将同一个应用系统的不同实体的数据组织在不同的表中，这些表不是孤立的，而是通过联系组织成一个整体；而 Excel 文件技术只能将同一个应用系统的不同实体的数据组织在同一个电子表中，如果将不同实体的数据组织在 Excel 的不同工作表中，实体之间的固有联系就被切断了，无法对不同实体的数据进行集成使用。

读者会提出这样的问题：将同一个应用系统中所有实体的数据分割到不同的表中时，如果用

户需要访问多个表的相关信息,到底应该怎么办?还有,如果删除了 Customer 表中顾客杨燕燕的信息,那么 seller 表中的销售员张大猛的信息就会不完整。这又怎么办?这些问题,数据库技术都有相应的方法和机制来解决,第3章和第5章会详细讨论这些问题的解决方法和机制。

数据库技术不仅从组织结构上解决了数据管理的操作异常问题,另外也解决了文件技术不能完全实现的数据共享、数据独立性以及数据冗余等问题。有兴趣的读者可查阅相关文献,这里就不展开了。

1.2 数据库技术的相关概念

数据库技术涉及的基本概念有:数据、信息、数据库、数据库管理系统、数据库应用程序、数据库系统、数据模型、数据库模式以及数据库语言等。下面分别介绍这几个概念。

1.2.1 数据和信息

1. 数据和信息的概念

日常生活中,大家将数据和信息混为一谈,认为数据就是信息,信息就是数据。这个观点是错误的,信息和数据根本不是一回事。数据和信息的区别,可以用一句话来概括:数据是信息的形式,信息是数据的内容。

(1) 数据的概念

如果将客观存在并且可以相互区分的事物称为实体,那么数据是对实体特性的一种记载,这种记载通常表现为符号的记录。

纯粹的数据没有任何意义,需要经过解释才能明确其表达的含义。例如 21,当解释其代表人的年龄时就是 21 岁;当解释其代表商品价格时,就是 21 元。

(2) 信息的概念

将从数据中获得的有意义的内容称为信息。信息和解释不可分,数据的解释是对数据含义的说明,数据的含义称为数据的语义,也就是数据的信息。

例如,对于(姜笑枫, 197101, 1989, 计算机系)这样一个数据集合,其语义可以解释为:姜笑枫, 1971 年 1 月生, 1989 年考入计算机系;还可以解释为:姜笑枫,工号 197101, 1989 年任职于计算机系。

2. 数据的静态特征

数据的静态特征指的是数据的基本结构、数据间的联系以及数据的约束。对于学生成绩而言,可以用这三个特征来描述学生成绩信息。

(1) 数据的基本结构

对于每一个学生的成绩,既可以用{学号、姓名、性别、专业、班级、数学、外语、计算机}这样的集合结构来描述;也可以用{学号、姓名、性别、专业、班级}以及{学号、数学、外语、计算机}这样的两个集合结构来描述。

(2) 数据间的联系

如果用{学号、姓名、性别、专业、班级}以及{学号、数学、外语、计算机}这样的两个集合结构来描述学生的成绩,那么第一个集合结构的学号与第二个集合结构的学号必然存在着某种联系。

(3) 数据的约束

数据反映的是客观对象的信息,它必然要遵循某些约束。例如对于百分制的课程成绩,学生的成绩必然在 0~100 分之间;又如,学生的性别只能取“男”或“女”这两个值。

3. 数据的动态特征

数据的动态特征包括对数据可以进行的操作以及操作规则。对数据库技术而言,数据的操作主要有查询和更新两大类。查询最常用,例如查询成绩不及格的学生名单;更新又包括插入、删除和修改三项操作,对于一般的数据库应用而言,也是必不可少的。

1.2.2 数据库

数据库(database),顾名思义,就是存放数据的仓库,只是这个仓库是在计算机存储器上建立的,而且仓库中的数据按一定的格式组织并长期存储,提供给用户共同使用。

严格地讲,数据库是长期存储在计算机外部存储器上的有组织的、可共享的数据集合。有组织的数据集合意味着数据的结构化和关联化:结构化表现为数据打包成一个个既定结构的数据对象;关联化表现在数据对象之间的互连性。

图 1-2 描述了数据库“销售订单”的数据对象及其联系:订单数据库包含 sllers、product、Customer、order 和 orderdetail 5 个数据对象;这 5 个数据对象都有很多属性,例如数据对象 product 包括商品编号、商品名称、商品价格、商品库存、商品简介和畅销否 6 个属性;这 5 个数据对象之间都通过公共属性相互关联,例如数据对象 sllers 和数据对象 order 通过销售员编号这个公共属性相互关联。

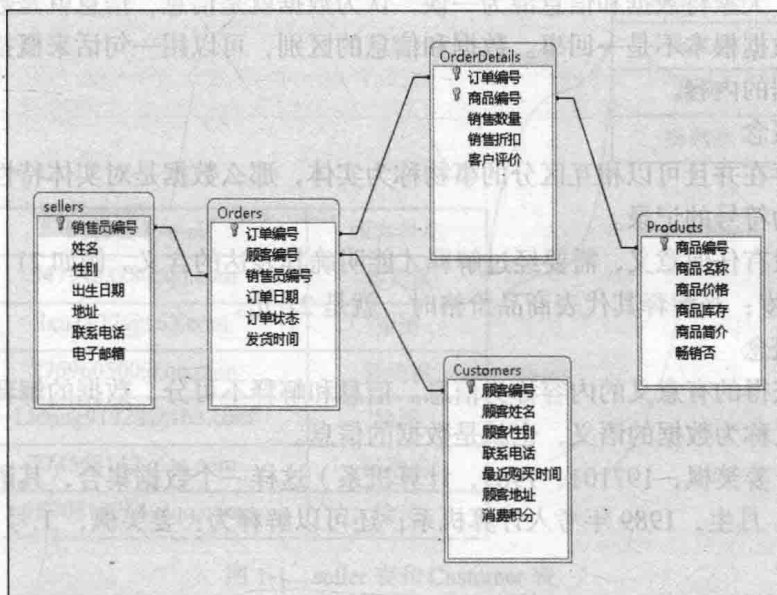


图 1-2 数据库的数据结构

数据的组织总是基于某种数据模型的,科学组织的数据可以极大地提高数据的共享度、检索速度和管理效率。数据模型这一内容将在 1.2.6 小节中介绍。

1.2.3 数据库管理系统

1. 数据库管理系统的概念

数据库管理系统(database management system)是一种管理和操纵数据库的软件,英文缩写为 DBMS,其主要功能是科学地组织和存储数据、高效地获取和维护数据,在数据库系统中起核心的作用,是用户程序与数据库中数据的接口。

由于 DBMS 功能复杂,一般由软件供应商开发并授权用户使用。例如 Microsoft 公司的 Access 就是微软公司开发的一个 DBMS,其他商用的 DBMS 产品还有 Microsoft 公司的 Visual FoxPro、SQL Server、Oracle 公司的 Oracle、IBM 公司的 DB2 等。尽管还有其他 DBMS 产品,但这 5 种

DBMS 几乎囊括了所有的市场份额。

2. 数据库管理系统的功能

DBMS 的功能主要有：创建数据库、创建表、创建其他支持对象（如索引等），读取数据库数据，插入、更新或删除数据库数据，维护数据库结构，定义和执行约束规则，并发控制，提供数据安全保障等。

(1) DBMS 的首要功能是创建数据库、创建数据库中的表和其他辅助结构。例如，创建图 1-3 所示的数据库“订单”，创建数据库中的数据表 order 和 product。为了提高检索速度，还可以给数据表创建索引等支持结构。

(2) DBMS 的第二个功能是读取和修改数据库中的数据。为此，DBMS 接收用户或应用程序的请求，并将这些请求转化为对数据库文件的操作。DBMS 的第二个功能还包括数据库结构的维护。例如，根据业务变化修改表的结构或改变相关辅助对象的属性等。

(3) DBMS 的第三个功能是约束规则的定义和检查。例如，在图 1-2 所示的订单数据库中，如果用户在 order 表中提交一张订单，订单中有一个商品的商品编号在 product 表中没有相应的数据，就会导致错误。为了防止这种错误的发生，用户可以用 DBMS 制定如下的数据约束规则：order 表中商品编号的值必须引用 product 表中商品编号的值，对于 product 表中不存在的商品编号，DBMS 应该拒绝含有这样商品编号的订单的插入或更新请求。

(4) DBMS 的第四个功能是并发控制，它可以保证一个用户的工作不会干扰另一个用户的工作。另外 DBMS 具有安全保证功能，它可以保证只有授权用户能对数据库完成得到授权的活动，例如防止用户查看特定数据，又如限制用户操作在指定的范围内。

一般来说，DBMS 还应该具有备份数据库和恢复数据库的功能。数据库是数据的集中仓库，是具有相当价值的重要资源，必须采取有效步骤，确保在软硬件故障或自然灾害等事件中没有数据丢失。

1.2.4 数据库应用程序

为了提高数据库的易用性，普通用户对数据库的管理和访问一般通过数据库应用程序作为媒介，而不是直接使用 DBMS 命令。现在的 DBMS 都给用户提供了很多应用程序开发工具，如报表设计器、窗体设计器、查询设计器等，它们为数据库应用程序的开发和使用提供了良好的环境和帮助，可提高生产率 20~100 倍。

数据库应用程序的功能一般包括：窗体的创建和处理、查询的创建和处理、报表的创建和处理等。当然，数据库应用程序的上述功能都是围绕着特定应用的业务逻辑展开的。

1.2.5 数据库系统

数据库系统是在计算机系统中引入数据库后的系统，它是以计算机平台为基础，动态地组织、存储、管理和分析处理数据库数据的软硬件系统。数据库系统的组成如图 1-3 所示，它包括五个部分：计算机平台、用户、数据库应用程序、数据库管理系统和数据库。

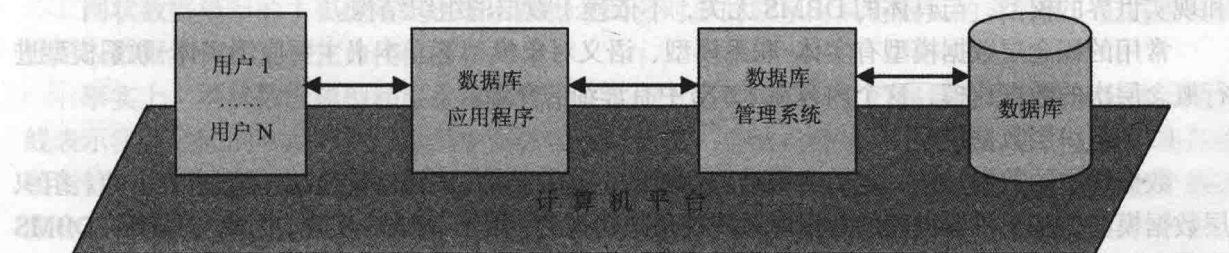


图 1-3 数据库系统的组成

在图 1-3 中,最右边的组成部分是数据库,它是描述实体的数据对象的集合。右数第二个组成部分是数据库管理系统,一般使用它的英文缩写 DBMS,这是一个计算机系统软件,一般由软件巨头开发并授权用户使用,它的主要功能是数据库的创建、管理和维护。

数据库应用程序是用户和 DBMS 间的媒介程序,它通过向 DBMS 发送请求命令来更新数据库中的数据,也可以通过 DBMS 检索数据库中的数据,并以友好的形式向用户显示结果。数据库应用程序可以由软件供应商提供,也可以由数据库用户编写。

用户是数据库系统的第四个组成部分,他们一般通过数据库应用程序进行事务管理,当然高级用户也可以直接通过 DBMS 操纵和管理数据库。

1.2.6 数据模型

对于模型,人们并不陌生。日常生活中所说的模型通常是指某个真实事物按比例缩小的版本,例如航模飞机、地图等,它们与所模拟的真实事物在结构上是相似的,因此,模型是对现实世界中研究对象的模拟和抽象。在日常生活中,模型的一个重要的作用是在制造真实事物之前,花费最少的代价,利用模型对研究对象的结构、性能等进行实验和评估,以降低真实对象的制造风险。

1. 数据模型的概念

数据模型是一种用来表达数据的工具模型。在数据库领域,数据模型用于表达现实世界中的对象及其联系,即将现实世界中杂乱的信息用一种规范的、形象化的方式表达出来。

前面说过,数据呈现结构、约束和操作三类特征,相应地,在计算机中表示数据的数据模型应该能够全面地描述数据的数据结构、数据约束和数据操作。尽管数据模型具有结构、约束和操作三方面的要素,但数据模型的结构是最核心的,很多情况下提到的数据模型,指的就是数据对象的数据结构。

数据模型既要面向现实世界,又要面向机器世界,因此需满足三个要求:①能够真实地模拟现实世界;②容易被人们理解;③能够方便地在计算机上实现。

在数据库领域建立数据模型的目的有两个:一个是提高数据库的可研究性,另一个是规避风险,将真实数据的组织、管理和分析处理风险掌握在可控范围中。

2. 数据模型的层次

数据描述涉及三个世界:一是现实世界,这是存在于人们头脑之外的客观世界;二是信息世界,这是现实世界在人们头脑中的反映形式;三是机器世界,这是信息世界的信息在机器世界中的数据组织形式。

数据模型是对数据的抽象和模拟,根据不同应用目的,可以将数据模型分为三大类:第一类是概念层数据模型,第二类是组织层数据模型,第三类是物理层数据模型。

(1) 概念层数据模型

概念层数据模型,即概念模型,它按用户的观点对现实世界的论域建立模型。概念模型更关注数据的语义,是现实世界的论域在人脑中的模型,属于信息世界的建模。概念模型是面向用户和现实世界的模型,与具体的 DBMS 无关,不依赖于数据的组织结构。

常用的概念层数据模型有实体-联系模型、语义对象模型等。本书主要应用实体-联系模型进行概念层次的数据建模,这个内容在 1.5 节中有详细的介绍。

(2) 组织层数据模型

数据库中的数据是按一定的逻辑结构存放的,这种结构是用组织层数据模型来表示的。组织层数据模型是基于机器世界的数据库系统视角来建模的,用于 DBMS 实现,因此与具体的 DBMS 有很大关系。组织层数据模型,在数据库的设计中,有着非常重要的作用,直接影响数据库中数据的质量和效率。请注意:如果没有特别声明,下文提到数据模型,指的都是组织层数据模型。

(3) 物理层数据模型

物理层数据模型是对数据最底层的抽象,描述数据在计算机系统内部的表示方式和存取方法,在计算机外部存储设备上的存储方式和存取方法。

3. 组织层数据模型

任何一个数据库管理系统都是基于组织层次的某种数据模型的,数据库技术的发展就是沿着组织层数据模型的主线展开的。迄今为止,比较流行的组织层数据模型有三种:层次数据模型、网状数据模型以及关系数据模型。

(1) 层次数据模型

在层次数据模型(hierarchical data model)的数据集合中,各数据对象之间是一种依次的一对一的或一对多的联系。在这种模型中,层次清楚,可沿层次路径存取和访问各个数据对象。层次结构犹如一棵倒置的树,因而也称其为树形结构。图1-4所示即为层次数据模型的数据集合的一个例子。

层次数据模型的特点如下。

- 有且仅有一个根结点,它是一个无父结点的结点。
- 除根结点以外的所有其他结点有且仅有一个父结点。
- 同层次的结点之间没有联系。

层次数据模型的优点是结构简单、层次清晰,并且易于实现。适宜描述类似于行政编制、家族关系及书目章节等信息载体的数据结构。但用层次模型不能直接表示多对多的联系,因而难以实现对复杂数据关系的描述。

(2) 网状数据模型

在网状数据模型(network data model)中,各数据对象之间建立的往往是一种层次不清楚的一对一、一对多或多对多的联系,此种结构可用来表示数据间复杂的逻辑关系。图1-5即是一个网状数据模型的例子。

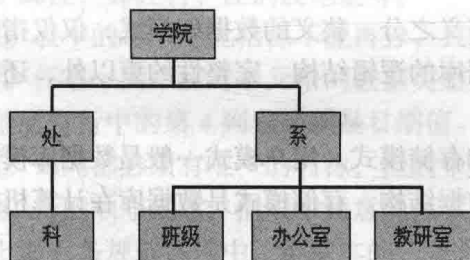


图 1-4 层次数据模型举例

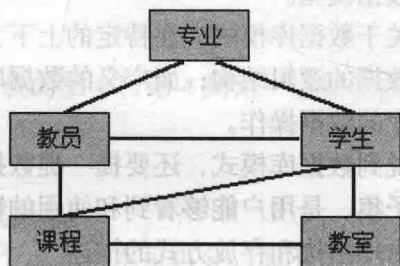


图 1-5 网状数据模型举例

网状数据模型的特点如下。

- 一个结点可以有多个父结点。
- 可以有一个以上的结点无父结点。
- 两个结点之间可以有多个联系。

网状数据模型的主要优点是在表示数据之间多对多的联系时具有很大的灵活性,但这种灵活性是以数据结构的复杂化为代价的。

事实上,网状数据模型和层次数据模型在本质上是类似的,它们都是用结点表示实体,用连线表示实体之间的联系。在计算机中具体实现时,每一个结点都是一个存储的数据记录,而用链接指针来实现数据记录之间的联系。这种用指针将数据记录联系在一起的方法,很难对整个数据集合进行修改和扩充。

(3) 关系数据模型

关系数据模型(relational data model)用二维表表示实体以及实体之间的联系,每一个二维表