



SQL Server 2005

数据库应用系统 开发技术

朱如龙 主编



增值回报
电子教案·练习素材



21 世纪高职高专规划教材系列

SQL Server 2005 数据库 应用系统开发技术

朱如龙 主编



机械工业出版社

本书全面系统地介绍了数据库的基本知识和基本概念、Microsoft SQL Server 2005 的安装配置、数据库设计、各种常用数据库对象的创建和管理、数据库的备份与恢复、数据的导入与导出、访问 SQL 数据库的常用方法。对数据库系统设计中较为常用的数据检索、视图、存储过程、触发器及数据完整性进行了详细的阐述。本书对数据库应用系统的各种开发技术也进行了详细的阐述,并给出了经过调试的示例。

本书内容涵盖了设计一个数据库管理系统要用到的主要知识。全书理论联系实际,内容翔实、可操作性强。本书既可以作为高职高专计算机及相关专业的数据库应用系统开发技术教材,也可作为各种 SQL Server 2005 培训班的培训教材,是一本理想的入门参考书。

图书在版编目(CIP)数据

SQL Server 2005 数据库应用系统开发技术/朱如龙主编. —北京:机械工业出版社,2006.6

(21 世纪高职高专规划教材系列)

ISBN 7-111-19293-1

I. S... II. 朱... III. 关系数据库—数据库管理系统,SQL Server 2005—高等学校:技术学校—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字 (2006) 第 060478 号

机械工业出版社(北京市百万庄大街 22 号 邮政编码 100037)

策 划:胡毓坚

责任编辑:李利健

责任印制:杨 曦

北京外文印刷厂印刷

2006 年 6 月第 1 版·第 1 次印刷

184mm×260mm·19 印张·470 千字

0001—5000 册

定价:27.00 元

凡购本图书,如有缺页、倒页、脱页,由本社发行部调换

本社购书热线电话:(010)68326294

编辑热线电话:(010)88379739

封面无防伪标均为盗版

出版说明

为了贯彻国务院发[2002]16号文件《国务院关于大力推进职业教育改革与发展的决定》的精神,进一步落实《中华人民共和国职业教育法》和《中华人民共和国劳动法》,实施科教兴国战略,大力推进高等职业教育改革与发展,我们组织力量,对实现高等职业教育培养目标和保证基本教学规格的文化基础课程、专业技术基础课程和重点建设专业主干课程的教材进行了规划和编写。

本套教材内容涵盖了高职高专院校计算机类、电子信息类、通信类、自动化类、市场营销类专业的专业基础课、专业课以及选修课,为配合高职教育关于“培养 21 世纪与我国现代化建设要求相适应的一线科技实用型人才”的最新理念,我们特为本系列教材配备了实践指导丛书,以利于老师的教学和学生的学习。

本套教材将理论教学和实践教学紧密结合,图文并茂、内容实用、层次分明、讲解清晰,其中融入了作者长期的教学经验和丰富的实践经验。可作为各类高职高专院校的教材,也可作为各类培训班的教材。

机械工业出版社

前 言

作为关系型数据库领域出类拔萃的软件,Microsoft SQL Server 2005 可以满足各种类型的用户和独立软件供应商的要求,它适用于从移动设备到 64 位系统的各种平台,可以说,Microsoft SQL Server 2005 是面向 21 世纪的关系型数据库,它在易用性、伸缩性、可靠性和数据仓库等方面的优异性能使其成为客户构建、管理和部署商业数据库的最佳选择方案之一。Microsoft SQL Server 2005 是管理程序方面最为友好的数据库之一,编程人员通过简单的操作就可以完成一个复杂的管理操作,已经成为业界领先的数据库管理系统。

本书以一个实际的数据库管理系统为范例,较为系统地阐述了数据库管理系统从基本需求分析到数据库逻辑设计、物理实现、数据库的创建与管理、各种数据库对象的创建与管理、数据完整性设计与实施、Transact-SQL 语言及数据库安全性。全书力求用浅显易懂的语言,深入浅出地阐述数据库管理系统的基本概念和基本理论,做到理论联系实际。

本书共分为 15 章。第 1 章讲述数据库的基本概念和基本理论;第 2~15 章分别介绍 SQL Server 2005 的安装、配置,数据库系统的设计方法,各种数据库基本对象的设计,数据库安全性及几种常用的访问 SQL Server 数据库的方法等。附录部分是作者经过精心设计的,在本书授课和实验中用到的全部表的结构和数据。本书的重点放在数据库的分析与设计、数据的完整性设计与实施、数据库的安全性及数据的查询与检索及各种常用数据库对象的使用上以及数据库应用系统的开发方法和技术。本书主要由朱如龙编写完成,参与编写的人员有刘焰。

本书内容基本上是按数据库管理系统开发的流程组织的,章与章之间是循序渐进的关系,建议读者在学习时按照本书的章节顺序进行,以确保学习时知识点不脱节。另外,联机帮助是微软资深程序员多年编程工作的结晶,建议大家在遇到问题时首先看联机帮助,这是一个优秀程序员必备的良好素养。

本书结构清晰,内容翔实、丰富,注重实际操作,具有很强的知识性、实用性和可操作性。书中所有的程序实例均经过测试并能编译、执行。通过本书的学习,相信读者一定能熟练掌握并能灵活运用 SQL Server 2005 软件,开发出具有一定实用价值的 SQL Server 数据库管理系统。

和本书配套的还有一本《SQL Server 2005 数据库应用系统开发技术实验指导及习题解答》,和本书配合起来一起学习,可以达到更好的效果。另外,本书提供了与教学配套的电子教案及部分程序源代码,读者可到机械工业出版社网站(www.cmpbook.com)上下载。

由于作者水平所限,书中错漏之处,恳请读者批评指正。

编 者

目 录

出版说明

前言

第 1 章 数据库系统概述	1
1.1 计算机数据处理	1
1.1.1 数据及数据处理	1
1.1.2 数据处理的三个阶段	1
1.2 数据模型	2
1.2.1 信息的三个世界	2
1.2.2 概念模型	5
1.2.3 逻辑模型	6
1.3 数据库应用系统概述	9
1.3.1 数据库应用系统的组成	9
1.3.2 数据库管理系统	10
1.3.3 数据库系统的结构	14
1.3.4 数据库系统的规范化设计	15
1.4 习题	19
第 2 章 安装和配置 SQL Server 2005	21
2.1 安装前的准备	21
2.1.1 最小硬件和软件需求	21
2.1.2 有关服务器的几点考虑	22
2.1.3 安装 SQL Server 2005 的配置选项	23
2.1.4 安装前的准备	24
2.2 SQL Server 2005 的安装	25
2.3 认识 Microsoft SQL Server 2005	28
2.3.1 Microsoft SQL Server 2005 的目录结构	28
2.3.2 系统数据库和样本数据库	29
2.3.3 Microsoft SQL Server 2005 程序组	31
2.3.4 Microsoft SQL Server Management Studio 界面及组件介绍	32
2.4 SQL Server 数据库的几个基本概念	35
2.4.1 数据库对象及其引用	35
2.4.2 系统表、系统存储过程、系统函数	36
2.5 Microsoft SQL Server 2005 的特点	36
2.6 习题	40
第 3 章 Transact-SQL 语言	41
3.1 SQL 语言简介	41
3.2 数据类型	42
3.2.1 二进制数据类型	42

3.2.2 数字数据类型	42
3.2.3 字符数据类型	43
3.2.4 unicode 数据类型	43
3.2.5 日期和时间数据类型	43
3.2.6 货币数据类型	44
3.2.7 特殊数据类型	44
3.2.8 用户自定义数据类型	44
3.3 局部变量与全局变量	46
3.4 运算符及优先级	48
3.4.1 一元运算符	48
3.4.2 赋值运算符	48
3.4.3 算术运算符	48
3.4.4 字符串连接运算符	49
3.4.5 比较运算符	49
3.4.6 逻辑运算符	49
3.4.7 运算符优先级	49
3.5 函数	50
3.5.1 数学函数	50
3.5.2 字符串函数	51
3.5.3 转换函数	53
3.5.4 日期时间函数	54
3.5.5 系统函数	55
3.5.6 集合函数	56
3.6 简单的 SELECT 语句	56
3.7 批处理和流程控制语言	57
3.7.1 批和脚本	57
3.7.2 流程控制语言	58
3.8 习题	63
第 4 章 网络数据库应用系统设计	64
4.1 网络数据库应用系统设计概述	64
4.1.1 网络数据库应用的特点	64
4.1.2 网络数据库应用系统体系结构	64
4.1.3 B/S、C/S 体系数据库应用系统的比较	66
4.2 数据库应用系统的开发方法和一般步骤	67
4.2.1 数据库应用系统的开发方法	67
4.2.2 数据库应用系统开发的一般步骤	68
4.3 教务管理数据库系统的设计概述	72
4.4 教务管理系统的基本需求	72
4.5 数据库系统的逻辑设计	72
4.5.1 教务管理系统的 E-R 模型	73
4.5.2 从 E-R 模型到逻辑数据库的转化	74
4.5.3 逻辑设计的其他考虑	75

4.6	逻辑设计的物理实现	75
4.7	设计数据完整性	75
4.8	数据库服务器编程	76
4.9	设计完成后的表及表中的列	76
4.10	习题	78
第5章	数据库的创建与管理	79
5.1	数据库简介	79
5.1.1	数据库的存储结构	79
5.1.2	事务日志	81
5.2	估算数据库的空间需求	82
5.3	数据库的创建	83
5.3.1	创建数据库的注意事项	83
5.3.2	使用 T-SQL 语言创建数据库	84
5.3.3	使用 SQL Server Management Studio 创建数据库	87
5.3.4	由已有库生成创建数据库的脚本	88
5.4	数据库的管理	89
5.4.1	数据库属性的查看	89
5.4.2	修改数据库大小	90
5.4.3	数据库的收缩	90
5.4.4	数据库的更名	91
5.4.5	数据库的删除	92
5.5	设置数据库的选项	92
5.5.1	使用 Transact-SQL 语句设置数据库的选项	92
5.5.2	使用 SQL Server Management Studio 设置数据库的选项	93
5.6	习题	94
第6章	表的创建与管理	95
6.1	表的创建	95
6.1.1	设计表的结构	95
6.1.2	使用 T-SQL 语言创建表	97
6.1.3	使用 SQL Server Management Studio 创建表	97
6.2	表结构的修改	98
6.2.1	使用 ALTER TABLE 语句修改表结构	98
6.2.2	使用 SQL Server Management Studio 修改表的结构	99
6.3	表的重命名与删除	100
6.4	向表中添加、更新、删除数据	100
6.4.1	使用 T-SQL 语言进行	101
6.4.2	使用 SQL Server Management Studio 进行	103
6.5	数据完整性的概念与实施方法	103
6.5.1	数据完整性概念	104
6.5.2	数据完整性实施方法	104
6.5.3	约束	105

6.5.4 使用默认值	111
6.5.5 规则	113
6.5.6 使用标识列	115
6.6 习题	116
第 7 章 数据库安全管理	117
7.1 SQL Server 2005 的验证模式	117
7.1.1 Windows 身份验证模式	117
7.1.2 SQL Server 身份验证模式	118
7.1.3 设置验证模式	118
7.1.4 SQL Server 系统登录验证过程	119
7.2 登录管理	120
7.2.1 系统管理员登录账户	120
7.2.2 用 T-SQL 语句创建、查看、删除 SQL Server 登录账户	120
7.2.3 用 SQL Server Management Studio 创建、查看、删除 SQL Server 登录账户	122
7.3 用户管理	124
7.3.1 数据库用户名和登录名的关系	124
7.3.2 用 T-SQL 语句创建、查看、删除数据库用户	124
7.3.3 使用 SQL Server Management Studio 创建、查看、删除数据库用户	126
7.3.4 改变数据库所有权	126
7.4 角色管理	126
7.4.1 固定服务器角色	126
7.4.2 数据库角色	128
7.5 许可管理	131
7.5.1 许可类型	131
7.5.2 许可的验证	132
7.5.3 管理许可	133
7.6 习题	135
第 8 章 数据检索	136
8.1 SELECT 语句	136
8.2 使用 SELECT 语句进行简单查询	137
8.2.1 最基本的 SQL 查询语句	138
8.2.2 改变列标题的显示	138
8.2.3 使用 WHERE 子句的查询	138
8.2.4 TOP 和 DISTINCT 关键字	141
8.2.5 使用 ORDER BY 子句对结果进行排序	141
8.2.6 计算列的使用	142
8.2.7 基于多个检索条件的查询	142
8.3 使用 T-SQL 语句进行高级查询	142
8.3.1 多表查询	143
8.3.2 使用 UNION 子句	144
8.3.3 使用 GROUP BY 子句	145
8.3.4 使用 COMPUTE 和 COMPUTE BY 子句	147

8.3.5 嵌套查询	148
8.3.6 在查询的基础上创建新表	149
8.4 用户定义函数在数据检索中的应用	150
8.4.1 用户自定义函数的定义及分类	150
8.4.2 标量值型用户自定义函数的定义、使用及删除	150
8.4.3 表值用户自定义函数的定义、使用及删除	151
8.5 数据查询的优化	153
8.6 分布式查询简介	156
8.7 习题	157
第9章 视图	158
9.1 视图概述	158
9.2 创建视图	159
9.2.1 使用 Transact-SQL 语句创建视图	159
9.2.2 使用 SQL Server Management Studio 创建视图	160
9.3 使用视图的优点和缺点	161
9.4 创建视图示例	162
9.4.1 创建水平视图	162
9.4.2 创建投影视图	162
9.4.3 创建联合视图	163
9.4.4 创建包含集合函数的视图	163
9.4.5 创建视图的视图	164
9.5 视图信息的查看	164
9.6 视图的修改、删除与重命名	165
9.6.1 视图的修改	165
9.6.2 视图的删除	166
9.6.3 视图的重命名	166
9.7 视图数据的查询、插入、修改与删除	167
9.7.1 用 T-SQL 语句进行视图数据的查询、插入、修改与删除	167
9.7.2 用 SQL Server Management Studio 进行视图数据的插入、修改与删除	168
9.7.3 修改视图数据的限制	169
9.8 视图的安全性	169
9.9 习题	170
第10章 创建索引文件	171
10.1 索引概述	171
10.1.1 索引的概念	171
10.1.2 创建索引的优、缺点	171
10.1.3 考虑建索引的列和不考虑建索引的列	172
10.2 聚簇索引与非聚簇索引	173
10.2.1 索引的分类	173
10.2.2 聚簇索引	173
10.2.3 非聚簇索引	174

10.3 索引的创建与管理	174
10.3.1 使用 T-SQL 语句创建索引	174
10.3.2 使用 T-SQL 语句管理索引	177
10.3.3 使用 SQL Server Management Studio 创建、查看、重命名及删除索引	178
10.4 索引的维护	179
10.5 习题	180
第 11 章 存储过程	181
11.1 存储过程的定义	181
11.2 存储过程的优点	182
11.3 存储过程的创建	182
11.3.1 使用 T-SQL 语句创建存储过程	183
11.3.2 创建带输入参数的存储过程	184
11.3.3 创建带输出参数的存储过程	187
11.4 重新编译存储过程	188
11.5 自动执行存储过程	189
11.6 查看、修改和删除存储过程	189
11.6.1 查看存储过程	189
11.6.2 修改存储过程	191
11.6.3 删除存储过程	191
11.7 扩展存储过程	192
11.7.1 扩展存储过程概述	192
11.7.2 注册扩展存储过程	192
11.7.3 如何使用扩展存储过程	193
11.7.4 删除扩展存储过程	193
11.8 习题	194
第 12 章 触发器	195
12.1 概述	195
12.1.1 触发器的概念及分类	195
12.1.2 触发器的工作原理	196
12.2 触发器的创建	196
12.2.1 使用 T-SQL 语句创建触发器	196
12.2.2 使用 SQL Server Management Studio 创建触发器	203
12.3 触发器实施数据完整性实例	204
12.3.1 实现参照完整性	204
12.3.2 实施特殊业务规则	205
12.4 查看、修改和删除触发器	205
12.4.1 查看触发器信息	205
12.4.2 修改触发器	207
12.4.3 删除触发器	208
12.5 使用触发器的注意事项	208
12.6 习题	210

第 13 章 游标及事务	211
13.1 游标的定义及其优点	211
13.1.1 游标的概念	211
13.1.2 使用游标的优点	211
13.2 游标的使用	212
13.2.1 使用游标的步骤	212
13.2.2 游标的定义及使用过程	212
13.3 使用游标修改数据	215
13.3.1 更新数据	215
13.3.2 删除数据	216
13.4 事务的使用	217
13.4.1 事务概述	217
13.4.2 事务的特性及事务的管理	217
13.4.3 事务控制语句	218
13.5 习题	220
第 14 章 数据库的备份与恢复	221
14.1 备份概述	221
14.1.1 数据库备份的定义及其重要性	221
14.1.2 备份的分类	221
14.1.3 何时备份	222
14.2 备份设备	222
14.2.1 备份设备的类型	223
14.2.2 建立备份设备	223
14.2.3 管理备份设备	225
14.3 备份类型	227
14.3.1 完整数据库备份	227
14.3.2 差异数据库备份	228
14.3.3 事务日志备份	228
14.3.4 数据库文件和文件组备份	228
14.4 执行备份	229
14.4.1 备份前的准备	229
14.4.2 使用 SQL Server Management Studio 进行备份	229
14.4.3 使用 T-SQL 语句进行备份	232
14.5 数据库恢复概述	235
14.5.1 数据库恢复的定义、原则和方法	235
14.5.2 故障类型和恢复模式	236
14.6 从不同备份中恢复数据库	237
14.6.1 使用 T-SQL 语句进行数据库的恢复	237
14.6.2 使用 SQL Server Management Studio 恢复数据库备份	240
14.7 恢复系统数据库	242
14.8 习题	242

第 15 章 SQL Server 2005 数据库开发技术	244
15.1 Microsoft SQL Server 2005 的数据库体系结构	244
15.1.1 SQL Server 的通信	244
15.1.2 SQL Server 应用程序的开发	244
15.1.3 SQL Server 结构	245
15.2 应用程序访问 SQL Server 数据库	247
15.2.1 SQL Server 数据库访问应用程序接口(API)	247
15.2.2 ADO 简介	247
15.2.3 使用 ODBC 连接 SQL Server 数据库	248
15.3 基于 C/S 体系结构的数据库访问技术	253
15.3.1 使用 ADO 数据控件访问 SQL Server 数据库数据示例	253
15.3.2 使用 ODBC 接口访问 SQL Server 数据库数据示例	256
15.4 基于 B/S 体系结构的数据库访问技术	260
15.4.1 B/S 体系结构的数据库接口技术	260
15.4.2 B/S 体系结构的组成及各部分的功能	262
15.4.3 数据库访问组件 ADO	266
15.5 数据导入与导出	281
15.5.1 数据导入与导出的意义	281
15.5.2 SQL Server 数据库表数据导出	282
15.5.3 导入数据到 SQL Server 表中	284
15.6 习题	285
附录	286
附录 A JWGL 数据库各数据表的结构	286
附录 B JWGL 数据库各数据表数据实例	288

第 1 章 数据库系统概述

Microsoft SQL Server 2005 是微软公司一个最新的数据库产品,为叙述方便,本书简称其为 SQL Server。为更好地学习和理解 SQL Server,本书先介绍数据库有关的基础知识。学习和理解 SQL Server 的最好方法是:学习理论的同时多做练习。在阐述完每章的理论和要点后,我们将尽可能多地给出与之相关的、能够直接在 SQL Server 运行的实例,帮助大家理解和掌握其中的要领。

本章主要介绍以下几方面内容:

- 数据及数据处理;
- 数据模型;
- 数据库、数据库管理系统、数据库应用系统。

1.1 计算机数据处理

1.1.1 数据及数据处理

数据是指对客观存在的事物的一种描述,它的表现形式可以是数字、文字、图像、声音等,这些形式的数据经过数字化以后才能为计算机识别和处理。数据是数据库中存储的基本对象。

数据处理是现代计算机应用中的一个重要组成部分。数据处理是指对各种形式的数据进行分类、组织、编码、存储、检索、更新和维护等一系列活动的总和。其目的是从大量的、杂乱无章的原始数据中推导出对人们有价值的信息,用来作为行动和决策的依据。

1.1.2 数据处理的三个阶段

数据处理是随着计算机硬件和软件技术的发展而不断发展的。30 多年来数据处理的发展经历了三个阶段:

20 世纪 50 年代中期以前,计算机的程序和数据是合为一体的,数据与程序不具有独立性,各个程序使用自己的数据,数据的组织方式必须由程序员自行设计与安排。当数据的存储结构发生改变时,应用程序必须相应地随之改变。由于数据面向程序,即一组数据对应一个程序,因此,数据是不能够共享的,这就造成了数据的冗余和难于管理。不过这个时期的计算机主要用于科学计算,并没有大规模地用于数据处理。这个阶段称为“人工管理阶段”。

20 世纪 50 年代后期到 60 年代中期,人们开始应用文件系统来保存数据,文件组织已多样化,有索引文件、直接存取文件等。程序和数据之间已经有一定的独立性。数据不再属于某个特定程序,可以重复使用。但数据还只是存放在数据文件中,数据文件之间相互独立、缺乏联系,数据文件的建立、存取、查询、插入、删除、修改等操作要通过程序来实现。这个阶段称为“文件系统阶段”。文件系统阶段对数据的操作以记录为单位。

20 世纪 60 年代后期开始,人们对数据处理和数据共享的要求越来越高。这个阶段计算

机硬件技术和软件研究水平的快速提高使得数据处理这一领域取得了长足的进步。伴随着大容量、高速度、低价格的存储设备的出现,用来存储和管理大量信息的“数据库管理系统”应运而生。这个阶段称为“数据库系统阶段”。

数据库系统的特点是数据共享性高、冗余低,采用数据模型表示复杂的数据结构。有较高的数据独立性。提供了方便的用户接口。用户可以使用查询语言或终端命令操作数据库,也可以用程序方式操作数据库。有统一的数据控制功能,主要包括数据安全性控制、数据完整性控制、并发控制、数据备份和恢复等。数据库系统阶段对数据的操作既可以记录为单位,也可以数据项为单位。

与文件系统阶段的数据处理相比,数据库系统的先进性表现在以下几个方面:

- 数据更容易被共享;
- 数据由面向程序变为面向系统,程序和数据相互独立;
- 数据库系统能将数据组织成一个结构化的整体,能描述复杂的数据结构;
- 数据存储有效地避免了冗余;
- 增加了数据控制功能,如安全性控制、完整性控制、并发处理的控制和出错时应急恢复的控制等。

数据库技术是当今计算机领域发展最快的技术之一,它的实用性很强,应用范围很广。目前,数据库技术已经和网络通信技术、面向对象技术、人工智能技术等交叉融合在一起,其应用领域几乎覆盖社会生活的各个方面。建设一个以数据库管理系统为核心的信息管理系统对于提高企业管理效率、增强企业竞争力具有极其重要的意义。

1.2 数据模型

数据库是一个企业或公司运营的相关数据的提取和集合,数据库中的数据是高度结构化的,它不仅反映数据本身,而且反映数据之间的联系,数据模型就是描述这些联系的数据结构形式。数据库中采用数据模型对现实世界进行抽象描述,所有的数据库系统均是基于某种数据模型的,数据模型的好坏,直接影响数据库的性能。为什么不直接进行数据库设计而要开发一个逻辑模型?因为首先开发数据库应用系统的逻辑数据模型将迫使用户把精力集中于一个系统的数据和数据的内在联系上,而不必一开始就为一些诸如列的数据类型等实际细节费心。与数据库设计相比,用户可以在更高层次上抽象地考虑一个数据模型。因此,数据模型的选择是设计数据库的一项首要任务。

1.2.1 信息的三个世界

将客观存在的事物以数据的形式存储到计算机中,需要对现实生活中的事物认识、概念化并逐级抽象存储到计算机中去,即需要经历现实世界-概念世界-数据世界三个领域。

1. 现实世界

现实世界就是客观存在的事物及其相互联系。事物之间的联系是由事物本身的性质决定的。现实世界中的事物之间既有“共性”,又具有“个性”。要求解现实问题,就要从中找出反映实际问题的对象,研究它们的性质及其内在规律,从而找到求解方法。

例如,学校的教学管理中涉及学生管理、教师管理及课程管理等。制作学生选修课程情况

表时,应包括学号、姓名、课程名、选修课程类别等;制作教师授课安排表时,应包括教师号、课程名、授课类别、学时数等。这就是现实世界,是数据库设计者接触到的最原始的数据。

2. 信息世界

信息世界也称为概念世界,是现实世界在人们头脑中的反映,是对客观事物及其联系的一种抽象描述,即用数据来描述现实世界中的事物及其联系,是对现实世界的第一次抽象。其中,经过抽象描述的现实世界中的客观事物在信息世界中称为“实体”。

例如,教师授课安排表是现实世界中存在的一个客观事物,可以用一组数据(教师号、课程名、授课类别、学时数)来描述,通过这样一组数据就可以了解该教师授课的基本情况。可以看出,信息世界仅涉及对某个特定组织所关心的特定信息结构的描述,而不涉及信息在系统中的表示。

信息世界中的主要概念如下:

(1) 实体

实体(Entity)就是指现实世界中客观存在并且可以相互区别的对象。实体可以是具体的事物,例如,一个学生、一门课程。

(2) 实体集

同型实体的集合称为实体集(Entity Set)。例如,全体职工、全体学生均是实体集。

(3) 属性

对实体或联系特性的描述称为属性(Attribute)。一个实体可以由若干个属性来描述,联系也有自己特有的属性。例如,学生的学号、姓名、性别、出生年份、系、入学时间等均为学生实体的属性;一个学生学习一门课程的成绩就是“学习”联系的属性。

(4) 域

域(Domain)就是属性的取值范围。每个属性有一个值域,其类型可以是整数型、实数型、字符串型或其他类型。例如,性别的域为(男,女)。

(5) 联系

联系(Relation)就是指实体与实体之间的关系。通过联系,可以用一个实体的信息来查找另一个实体的信息。实体之间的联系比较复杂,但抽象以后,可以分为3类:

1) 一对一关系(1:1)。对于实体集A中的每一个实体,实体集B中有并且至多有一个实体和它有联系,反之亦然。例如,如果一个班级有一个班主任,每个班主任只负责一个班级,那么,班主任和班级之间的关系是一对一关系。

2) 一对多关系(1:N)。对于实体集A中的每一个实体,实体集B中有多于一个的实体和它有联系,反之,对于实体集B中的每一个实体,实体集A中有并且至多有一个实体和它有联系。例如,如果一个班级有多个学生,而每个学生只能属于一个班级,那么,班级和学生之间的关系是一对多的关系。

3) 多对多关系(M:N)。对于实体集A中的每一个实体,实体集B中有多于一个的实体和它有联系,反之对于实体集B中的每一个实体,实体集A中也有多于一个的实体和它有联系。如:一个书店可以销售多本书,而同一种书可以在多个书店销售。那么,书店和书之间的关系就是多对多的关系。

3. 数据世界

数据世界也称为计算机世界、存储世界或机器世界,是对信息世界中的有关信息进行加工、编码及格式化等具体处理后以一定的形式存储在计算机内的数据。可以说,数据世界中的

数据既能展示信息世界,同时又向计算机数据处理前进了一步。

数据世界中的主要概念如下:

(1) 数据项

数据项(Item)就是标记实体属性的命名单位,也称为元素或字段。它是可以命名的最小信息单位。数据项的取值范围称为域。域以外的任何值对该数据项都是无意义的。例如,表示月份的数据项的域是1~12,则13就是无意义的值。数据项的值可以是数值、字母、汉字等形式。数据项的物理特点在于它具有确定的物理长度,一般用字节数表示。

(2) 记录

若干相关联的数据项的集合构成记录(Record)。记录是应用程序输入/输出的逻辑单位。对于大多数数据库系统来讲,记录是处理和存储信息的基本单位。通常用一条记录描述一个实体。构成该记录的数据项表示实体的若干属性。如实体“学生”的一组数据(9940201,黎明,男,计算机)就是一条记录。其数据项包括学号、姓名、性别及专业。

(3) 文件

同一给定类型记录的集合构成一个文件(File)。文件对应于信息世界的实体集,所以文件又可定义为描述一个实体集的所有记录集。如所有的学生记录组成一个学生文件。文件用文件名称标识。文件根据记录的组织方式和存取方法可以分为:顺序文件、索引文件、直接存取文件等。

(4) 数据库

数据库是比文件更大的数据组织形式。数据库是具有特定联系的数据的集合,也可以看成是具有特定联系的多种类型的记录的集合。

(5) 关键字

关键字(Key)是能够惟一标识文件中的每一条记录的字段或字段集,如学生的“学号”可以作为学生记录的关键字。

从现实世界到信息世界不是简单的数据描述,而是从现实世界中抽象出适合数据库技术处理的数据即实体,同时要求这些实体能够很好地反映客观世界中的事物;从信息世界到数据世界也不是简单的数据对应存储,而是要设计数据的逻辑结构和物理存储结构。所谓数据的逻辑结构,是指程序员或用户用以操作的数据形式,是抽象的概念化数据;所谓数据的物理结构,是指实际存储在存储设备上的数据。

4. 三个世界之间的关系

现实世界、信息世界和数据世界三者之间的关系如图1-1所示。从图1-1可知,将现实世界中的事物及其联系逐步抽象为数据世界中具有一定结构的便于计算机处理的数据形式,需要使用以下两个不同层次的数据模型。

