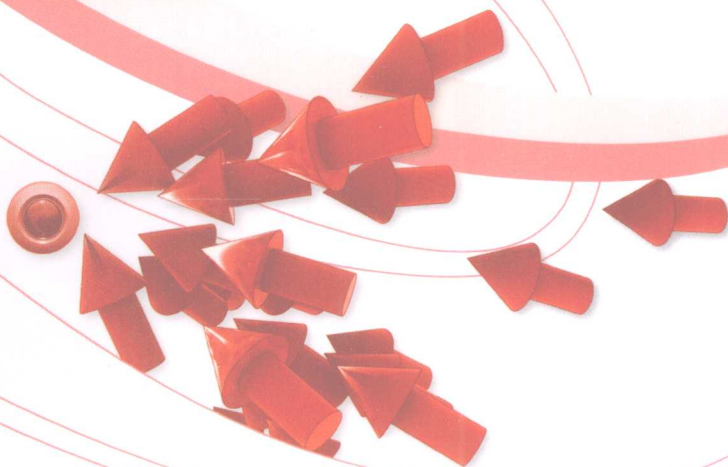




北京市高等教育精品教材立项项目

为教师
提供实验
指导书



高等学校计算机规划教材

网络数据库技术 (第2版)

■ 逯燕玲 戴 红 李志明 主编



电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY

<http://www.phei.com.cn>

高等学校计算机规划教材

网络数据库技术

(第2版)

逯燕玲 戴 红 李志明 主编

电子工业出版社

Publishing House of Electronics Industry

北京 • BEIJING

内 容 简 介

本书全面、系统地介绍了网络数据库的基本概念、Microsoft SQL Server 2000 的安装、管理、应用、数据维护的方法以及网络数据库设计和应用系统开发技术。本书共分 12 章, 主要包括: 数据库系统的概念, 关系数据库基本理论, 网络数据库技术, SQL 语言基础, 数据库的创建与管理, 查询分析器的使用与 T-SQL 高级查询, SQL Server 的视图、索引、游标、事务、存储过程与触发器, SQL Server 的安全性管理, 数据库的备份与恢复, ASP/ADO 网络数据库技术和基于 C/S 模式的应用系统开发实例等。本书为教师提供习题解答和实验指导书。

本书重点突出了技术应用性, 章后附有大量的习题和实训练习。本书适合作为高等院校本科生(或专科生)相关课程的教材, 也可供数据库管理员参考使用。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

图书在版编目(CIP)数据

网络数据库技术 / 逯燕玲, 戴红, 李志明主编. —2 版. —北京: 电子工业出版社, 2009.8

高等学校计算机规划教材

ISBN 978-7-121-09150-6

I. 网… II. ①逯… ②戴… ③李… III. 关系数据库—数据库管理系统—高等学校—教材 IV. TP311.138

中国版本图书馆 CIP 数据核字(2009)第 106113 号

策划编辑: 章海涛

责任编辑: 章海涛

特约编辑: 何 雄/王 纲

印 刷: 北京市天竺颖华印刷厂

装 订: 三河市鑫金马印装有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本: 787×1092 1/16 印张: 19.5 字数: 500 千字

印 次: 2009 年 8 月第 1 次印刷

印 数: 4 000 册 定价: 29.8 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

第 2 版前言

随着计算机技术、通信技术和网络技术的飞速发展,信息系统渗透到社会的各个领域,作为其核心和基础的数据库技术也得到了越来越广泛的发展和应用,数据库系统的建设规模、信息量的大小和使用频度已成为衡量一个国家信息化程度的重要标志。

20 世纪 90 年代以来,Internet 日益普及,Web 成为最流行、最大的网络系统,并以惊人的速度继续发展。Web 技术和数据库技术的结合,产生了网络数据库这一新兴的数据库应用领域。在高等院校计算机相关专业的课程中,网络数据库技术已经成为一门重要的课程。

本书第 1 版是“北京市高等教育精品教材立项项目”之一——“计算机网络技术与应用系列教材”中的一本,第 2 版重点突出了技术应用性。

相比于第 1 版,本书第 2 版修订了第 1、5、6、7、8、9、11 章的相关内容,主要修订思路如下:

- ◎ 调整章节结构,以更加适应案例化教学。
- ◎ 增加实例,围绕贯穿全书的综合性实例给出实例。
- ◎ 整理、完善章后习题和实验题。

同时,本书第 2 版增加了第 12 章数据库应用系统开发实例,用于加强技术应用。

本书的主要内容包括:第 1 章数据库系统概述,第 2 章关系数据库基本理论,第 3 章网络数据库技术概论,第 4 章 SQL 语言基础,第 5 章数据库的创建与管理,第 6 章数据库查询,第 7 章视图、索引、游标及事务,第 8 章存储过程与触发器,第 9 章 SQL Server 的安全性管理,第 10 章数据库备份与恢复,第 11 章 ASP/ADO 网络数据库技术,第 12 章基于 C/S 模式的应用系统开发。

为了配合教学,本书在每章后附有大量的习题和实训练习,以巩固学习效果。

另外,我们编写了配套的实验指导书,提供给使用本书作为教材的教师。

本书由逯燕玲修订编写了第 1 至 8 章,由戴红修订编写了第 9 章、第 10 章和第 12 章,由李志明编写了第 11 章,全书由逯燕玲统稿。在本书的编写和修订过程中,我们得到了北京市高等教育精品教材指导委员会和北京联合大学的大力支持,在此表示衷心感谢。

限于作者的水平和学识,书中难免会有错误和不妥之处,恳请读者指正。

本书为任课教师提供配套的教学资源(包含电子教案),需要者可[登录华信教育资源网](http://www.huaxin.edu.cn)[站](http://www.hxedu.com.cn)(<http://www.huaxin.edu.cn> 或 <http://www.hxedu.com.cn>),注册之后进行免费[下载](#),或发邮件到 unicode@phei.com.cn 进行[咨询](#)。

作 者

目 录

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 数据处理的三个发展阶段	1
1.1.3 数据库技术的发展	3
1.2 数据模型	3
1.2.1 信息的三个世界	3
1.2.2 概念模型	5
1.2.3 逻辑模型	8
1.3 数据库系统的基本概念	13
1.3.1 数据库系统的组成	13
1.3.2 数据库管理系统的功能	13
1.3.3 数据库系统的三级模式结构	15
1.4 数据库设计	16
1.4.1 需求分析	16
1.4.2 概念结构设计	21
1.4.3 逻辑结构设计	23
1.4.4 物理结构设计	23
1.4.5 数据库实施	24
1.4.6 数据库运行与维护	25
习题 1	25
第 2 章 关系数据库基本理论	28
2.1 关系模型的概念	28
2.1.1 关系的数学定义	28
2.1.2 关系操作	30
2.1.3 关系的完整性	30
2.2 关系代数	31
2.2.1 传统的集合运算	31
2.2.2 专门的关系运算	32
2.3 关系模型的规范化	35
2.3.1 函数依赖	35
2.3.2 范式	36
2.3.3 E-R 模型向关系数据库转换	38
2.4 关系数据库系统的查询优化	40
2.4.1 关系系统	40
2.4.2 全关系型系统的基本准则	40

2.4.3	查询优化的准则	42
2.5	常用的关系数据库管理系统	43
2.5.1	Access 数据库管理系统	44
2.5.2	SQL Server 数据库管理系统	44
2.5.3	Oracle 数据库管理系统	45
2.5.4	DB2 数据库管理系统	46
2.5.5	Sybase 系列	47
习题 2		48
第 3 章	网络数据库技术概论	52
3.1	网络数据库的基本概念	52
3.1.1	网络数据库	52
3.1.2	SQL Server 2000 的系统结构	55
3.1.3	SQL Server 2000 的新特性	57
3.2	SQL Server 2000 的安装	58
3.2.1	SQL Server 2000 系统要求	58
3.2.2	设置 Windows 服务账户	59
3.2.3	SQL Server 2000 的安装过程	59
3.3	SQL Server 服务器管理	64
3.3.1	启动和停止 SQL Server	64
3.3.2	SQL Server 服务器组的管理	67
3.3.3	连接与注册服务器	68
3.3.4	配置 SQL Server 服务器	74
3.4	Web 数据库技术概述	75
3.4.1	CGI 技术	76
3.4.2	Web API 技术	76
3.4.3	ODBC 技术	77
3.4.4	JDBC 技术	77
3.4.5	ASP 技术	78
3.5	IIS 简介	79
3.5.1	IIS 的特性	79
3.5.2	IIS 的 Web 服务器	79
3.5.3	IIS 的 FTP 服务器	82
3.5.4	IIS 的 SMTP 服务器	82
习题 3		83
实训 1		83
第 4 章	SQL 语言基础	85
4.1	SQL 语言的基本概念	85
4.1.1	标准 SQL 语言及其特点	85
4.1.2	T-SQL 语言	87
4.1.3	T-SQL 语言的语法结构	88
4.1.4	数据类型	95

4.1.5 运算符与表达式	96
4.2 流程控制语言	99
4.3 函数	104
习题 4	111
实训 2	112
第 5 章 数据库的创建与管理	114
5.1 数据库的创建	114
5.1.1 数据库的存储结构	114
5.1.2 系统数据库	115
5.1.3 估算数据库的空间需求	117
5.1.4 使用 SQL Server Management Studio 创建数据库	118
5.1.5 使用向导创建数据库	121
5.2 数据库的管理	124
5.2.1 数据库属性的查看与修改	124
5.2.2 设置数据库的选项	125
5.2.3 删除数据库	127
5.3 表的创建	128
5.3.1 设计表的结构	128
5.3.2 使用 SQL Server Management Studio 创建表	130
5.4 表结构的修改	132
5.5 表的数据操作	135
5.6 删除表	135
5.7 数据库的完整性设计	135
5.7.1 数据完整性实施方法	136
5.7.2 约束	137
5.7.3 使用默认值	145
5.7.4 规则	146
5.7.5 使用标识列	148
5.8 创建数据库关系图	148
习题 5	152
实训 3	152
第 6 章 数据库查询	154
6.1 使用 SELECT 语句进行简单查询	154
6.2 T-SQL 高级查询	156
6.2.1 多表查询	156
6.2.2 分组查询	161
6.2.3 统计查询	163
6.2.4 嵌套查询	166
6.3 存储查询结果	169
6.3.1 将查询结果存储到新表中	169
6.3.2 将查询结果添加到已有表中	169

6.3.3 对查询结果进行修改	169
习题 6	170
实训 4	170
第 7 章 视图、索引、游标及事务	172
7.1 视图	172
7.1.1 视图概述	172
7.1.2 创建视图	173
7.1.3 视图信息的查看和修改	175
7.1.4 视图数据的更新	176
7.1.5 删除视图	176
7.2 索引	177
7.2.1 索引的分类	177
7.2.2 创建索引	177
7.2.3 创建索引视图	181
7.2.4 删除索引	183
7.3 游标	183
7.3.1 游标概述	183
7.3.2 游标的定义及使用过程	183
7.3.3 使用游标修改数据	184
7.4 事务的使用	185
7.4.1 事务概述	185
7.4.2 事务处理语句	186
7.4.3 分布式事务	189
习题 7	191
实训 5	191
第 8 章 存储过程和触发器	192
8.1 存储过程概述	192
8.2 存储过程的设计与执行创建存储过程	193
8.2.1 使用企业管理器创建存储过程	193
8.2.2 使用“创建存储过程向导”创建存储过程	193
8.2.3 使用 T-SQL 创建存储过程	193
8.3 存储过程调用	197
8.3.1 创建执行存储过程	197
8.3.2 查看存储过程	198
8.3.3 删除存储过程	199
8.3.4 修改和重命名存储过程	199
8.4 创建触发器	200
8.4.1 使用企业管理器创建触发器	200
8.4.2 使用 CREATE TRIGGER 创建触发器	200
8.5 触发器的管理与应用	203
8.5.1 查看触发器	203

8.5.2 修改和重命名触发器	204
8.5.3 删除触发器	205
习题 8.....	206
实训 6.....	206
第 9 章 SQL Server 的安全性管理	207
9.1 SQL Server 的安全性机制	207
9.2 SQL Server 的登录模式	208
9.2.1 登录认证模式概述	208
9.2.2 SQL Server 认证模式的设置与管理	208
9.3 SQL Server 用户标识和鉴别	211
9.3.1 数据库用户概述	211
9.3.2 管理数据库用户	211
9.4 角色与权限	213
9.4.1 角色	213
9.4.2 权限	217
9.5 数据库的并发控制	222
9.5.1 并发事务	222
9.5.2 封锁	223
9.5.3 不加锁的并发控制	225
习题 9.....	226
实训 7.....	226
第 10 章 备份与恢复	227
10.1 数据备份概述.....	227
10.2 数据库备份策略.....	227
10.2.1 备份策略.....	227
10.2.2 恢复策略.....	228
10.3 数据库备份与恢复的实施.....	228
10.3.1 备份设备.....	228
10.3.2 数据库备份	229
10.3.3 数据库恢复	232
10.3.4 备份和恢复系统数据库	235
10.4 数据库表数据的导入导出.....	236
10.4.1 数据转换服务概述	236
10.4.2 数据转换服务的导入/导出向导	236
习题 10	247
实训 8.....	248
第 11 章 ASP/ADO 网络数据库技术	249
11.1 Web 环境操作 SQL Server 数据库	249
11.2 ASP/ADO 技术概述	250
11.2.1 ASP 技术概述.....	250
11.2.2 ADO 技术概述	250

11.2.3	IIS/Web 服务器的配置	251
11.3	ASP 的内建对象	254
11.3.1	Response 对象	254
11.3.2	Request 对象	256
11.3.3	Server 对象	264
11.3.4	Session 对象	266
11.3.5	Application 对象	266
11.4	使用 ASP/ADO 操作 SQL Server 数据库	267
11.4.1	ADO 数据库接口简介	267
11.4.2	在 ASP 中使用 ADO 对象访问数据库	268
11.5	通过 ODBC 访问数据库	269
11.5.1	ODBC 的作用	269
11.5.2	DSN 数据源的创建	270
11.5.3	利用 ODBC 访问数据库	273
11.6	对数据库中数据的基本操作	275
11.6.1	通过视图查询数据库中的数据	275
11.6.2	利用表单向数据表添加数据	277
11.6.3	数据更新修改	281
11.6.4	数据删除	284
第 12 章	基于 C/S 模式的应用系统开发	288
12.1	数据库设计和实施	288
12.1.1	学生信息管理系统需求分析	288
12.1.2	概念结构设计	289
12.1.3	数据库逻辑结构设计	289
12.1.4	创建“数据库创建”脚本文件并建库	290
12.1.5	创建“表创建”脚本文件并创建表	290
12.1.6	使用视图和存储过程实现一些查询和操作要求	290
12.1.7	使用触发器、规则等加入完整性控制机制	292
12.1.8	制定数据库备份和恢复策略	292
12.2	系统设计	294
12.2.1	系统功能及处理流程	294
12.2.2	创建工程、添加模块	295
12.2.3	创建数据源	295
12.2.4	设计系统主界面	295
12.2.5	设计系统登录模块	296
12.2.6	设计各功能模块	296
小结		301
实训 9		301
参考文献		302

第 1 章 数据库系统概述

本章将介绍数据处理的三个发展阶段，讲述数据模型、数据库系统的基本概念，介绍关系数据库的特点和数据库设计的步骤。

数据库技术是随着信息社会对数据处理任务的需要而产生的，所研究的问题是如何科学地组织和存储数据，如何高效地获取和处理数据。随着社会对数据处理任务的要求不断提高，数据库技术也随之不断向前发展，从问世到现在的半个多世纪的时间，形成了坚实的理论基础、成熟的商业产品和广泛的应用领域。数据库技术是一门综合性的软件技术，已经成为计算机科学中的一个重要分支。

数据库技术是使用计算机对各种信息进行收集、管理的必备知识。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，数据库的数量和规模越来越大，数据库的研究领域也不断拓广和深化。

1.1 计算机数据处理

在信息社会中，信息成为比物质和能源更重要的资源，在国民经济中占据主导地位。实际上，信息就是新的、有用的事实和知识，其表现形式是数据。

1.1.1 数据及数据处理

数据（data）是数据库中存储的基本对象，是指对客观存在的事物的一种描述，其表现形式可以是数字、文字、图像、声音等，这些数据经过数字化才能为计算机识别和处理。例如，学生李伟的学号是“2005013045”，性别是“男”，所在院系为“计算机学院”，高等数学考试成绩为 85 分等。“2005013045”、“李伟”、“男”、“计算机学院”、“85”就是数据，用来分别描述学生李伟的学号、姓名、性别、所在院系和高等数学考试成绩等信息。

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。数据处理是现代计算机应用中的一个重要组成部分。

1.1.2 数据处理的三个发展阶段

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，半个多世纪以来，数据处理经历了三个发展阶段。

1. 人工发展阶段

20 世纪 50 年代中期以前，计算机主要用于科学计算。从计算机硬件来看，没有磁盘等直接存取的存储设备；从计算机软件来看，没有操作系统和各种工具软件，计算机的程序和数据是合为一体的，数据的组织方式必须由程序员自行设计与安排。这个时期的数据管理具有以下 4 个特点：

① 数据不保存，只是在计算某一课题时将数据输入，用完后既不保存原始数据，也不

保存计算结果。

② 没有对数据进行管理的软件系统。程序员不仅要规定数据的逻辑结构，还要在程序中设计物理结构，程序中存取数据的子程序随着存储的改变而改变，数据与程序不具有一致性。

③ 没有文件的概念。数据是面向应用的，一组数据只能对应一个程序，即使两个程序用到相同的数据，也必须各自定义、各自组织，数据无法共享，从而导致程序与程序之间有大量的冗余数据。

④ 数据与程序不具有独立性。当数据的逻辑结构或物理结构发生变化后，必须对应用程序做相应的修改，这就加重了程序员的负担。

2. 文件系统阶段

20 世纪 50 年代后期至 60 年代中期，计算机在硬件方面有了磁盘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——文件系统。这个时期的数据管理具有以下 4 个特点：

① 数据可以以文件的形式长期保存在外存储器中，可以经常对文件进行查询、修改、插入和删除等数据处理，便于反复操作。

② 由文件系统对数据进行管理。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与数据打交道，数据有了物理结构和逻辑结构的区别，但此时程序与数据之间的独立性尚不分明。

③ 文件的形式已经多样化，文件的物理存储结构有顺序文件、链接文件、索引文件等，而对文件的访问既可以是顺序的，也可以是随机访问。

④ 数据具有一定的独立性，如果两个程序用到相同的数据，可以共享同一个数据文件；但是，一般情况下，一个数据文件是为某一特定应用服务的，数据共享性、独立性仍然较差。

3. 数据库系统阶段

20 世纪 60 年代后期，计算机主要用于管理的规模越来越大，人们对数据的共享要求越来越高，传统的文件系统已经不能满足人们的需要，能够统一管理和共享数据的数据库管理系统便应运而生。在这一阶段中，数据库中的数据不再面向某个应用或某个程序，而是面向整个企业或整个实际应用，数据库的数量急剧增长。这个时期的数据管理具有以下特点：

① 采用复杂的结构化的数据模型，数据库系统不仅描述数据本身，还要描述数据之间的联系。

② 较高的数据独立性。数据存储结构的变化尽量不影响用户程序的使用，数据与程序的独立把数据的定义从程序中分离出去，加上数据是由数据库管理系统来管理，简化了应用程序的编制和程序员的负担。

③ 最低的数据冗余度。数据共享性高，易于扩充，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 具有数据控制功能。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、有效和相容；具有数据的并发控制功能，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。

1.1.3 数据库技术的发展

20 世纪 70 年代以前，数据库系统并没有一个完整和坚固的数学理论基础，数据库产品直接按实际应用的要求来设计。这段时期的数据库产品以层次型数据库系统为主，业界称为“前关系型数据库系统（Pre-Relational Database）”。1970 年，IBM 公司的研究员 Codd 发表了题为“大型数据库的数据关系模型”的文章，从此数据库系统的发展进入到“关系型数据库系统（Relational Database）”时期。关系型数据库系统以关系代数为理论基础，经过几十年的发展和实际应用，技术越来越成熟和完善。

1.2 数据模型

信息（Information）就是通过各种方式传播的能被感受的声音、文字、图像、符号等。例如，一辆小轿车的价格信息是 15 万元，颜色信息是黑色。“15”是表示价格的数据，“黑色”是表示颜色的数据。信息可以告诉人们有用的事实和知识，数据可以有效地表示、存储和提取信息。信息是现实世界事物状况在人们头脑中的反映，通过加工可以用一系列数据来表示。但是，在不同的领域，数据的描述有所不同。

1.2.1 信息的三个世界

客观存在的事物要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、概念化并逐步抽象存储到计算机中去，即需要经历现实世界—信息世界—数据世界三个领域。

现实世界（Real World）就是客观存在的事物及其相互联系。现实世界存在无数事物，每一个客观存在的事物可以看作一个个体，每个个体都有属于自己的特征。现实世界中的事物之间既有共性又有个性，事物之间的联系是由事物本身的性质决定的。现实世界的主要概念如下。

1. 实体（Entity）

现实世界中存在的可以相互区分的客观事物或概念称为实体。例如，一个学生、一门课程、一名教师、一台计算机、一头大象等都是实体。

2. 实体的特征（Entity Characteristic）

每个实体都有自己的特征，用于区别不同的实体。例如，课程有课程名称、学时、学分等特征，教师有姓名、性别、职务和职称等特征。现实世界就是通过每个实体所特有的特征来相互区分的。

3. 实体集（Entity Set）

具有相同特征的实体的集合称为实体集。例如，全体学生的集合构成学生实体集，全部课程的集合构成课程实体集。

信息世界（Information World）是现实世界在人们头脑中的反映，也称为概念世界。人们以现实世界为基础，用思维对客观事物及其联系进行选择、命名、分类和抽象，并用数据来描述，就形成了信息世界。信息世界是对现实世界的第一次抽象，抽象的重点在于数据的结构。信息世界的主要概念如下。

4. 属性 (Attribute)

实体的特征在人们思想意识中形成的知识称为属性。一个实体可以有多个属性，如学生的学号、姓名、性别、出生年月、入学时间、所学专业等均为学生实体的属性；课程的课程号、课程名称、学时、学分等均为课程实体的属性。

5. 码 (Key)

在实体的多个属性中，能够唯一标识实体的属性或属性集合称为码。例如，学号是学生的唯一标识，课程号唯一标识一门课程。

6. 域 (Domain)

域就是属性的取值范围。每个属性都有一个值域，例如，性别的域为(男,女)；学时是整数类型的，值域为 0~500。

7. 联系 (Relation)

联系就是实体与实体之间的关系。联系也可以有属性，例如，一个学生选修一门课程取得了成绩，“选修”就是学生实体与课程实体之间的联系，而“成绩”就是“选修”联系的属性。实体之间的联系比较复杂，可以抽象为 3 类：

① 一对一联系 (1:1)。对于实体集 A 中的每一个实体，实体集 B 中有且仅有一个实体和它有联系，反之亦然。例如，如果一个班级只有一个班主任，每个班主任又只负责管理一个班级，那么班主任和班级之间的联系就是一对一联系。

② 一对多联系 (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与之有联系，反之，对于实体集 B 中的每一个实体，实体集 A 中有且仅有一个实体与之有联系。例如，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就是一对多联系。

③ 多对多联系 ($M:N$)。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与之有联系，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与之有联系。例如，一个学生可以选修多门课程，而每门课程可以被多个学生选修，那么学生和课程之间的关系就是多对多联系。

数据世界 (Data World) 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。在现代计算机系统中，数据组织的层次结构一般有 4 个基本逻辑元素：数据项、记录、文件和数据库，形成以数据库为最高层的层次结构。一个记录包含若干数据项，相关记录的集合组成文件，逻辑相关的文件组成数据库。

8. 数据项 (Item)

数据项用来标记实体的一个属性，也称为字段 (Field)，是可以命名的最小信息单位，与信息世界的属性相对应。例如，课程有课程号、课程名称、学时、学分等字段，教师有姓名、性别、职务和职称等字段。

9. 记录 (Record)

记录是有一定逻辑关系的数据项的组合，与信息世界的实体相对应。一个记录可以描述一个实体，构成该记录的数据项表示实体的若干属性。例如，一个课程实体的一组数据

(0101105, 数据库原理及应用, 68, 4)就是一条记录。

10. 文件（File）

文件是被命名了的同一类记录的集合，与信息世界的实体集相对应，又可以称为描述一个实体集的所有记录集。例如，所有的学生记录可以组成一个学生文件。

11. 数据库（DataBase, DB）

数据库是具有特定联系的数据集合，也可以看成是具有特定联系的多种类型的记录集合。数据库是用来存放数据的仓库，是一个按数据结构来存储和管理数据的计算机软件系统，是数据管理的方法和技术，能够合理地组织数据、方便地维护数据，严密地控制数据，有效地利用数据。数据库是比文件更大的数据组织形式，如一个图书管理数据库中可以包含所有图书记录、所有读者记录和所有借阅记录等。

现实世界、信息世界和数据世界三者之间术语的对应关系如表 1.1 所示。

表 1.1 信息的三个世界术语的对应关系

现实世界	信息世界	数据世界
特征	属性	数据项
事物个体	实体	记录
事物总体	实体集	文件
事物间的联系	实体间的联系	数据间的联系
	概念模型	逻辑模型

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存储，而是要设计数据的逻辑结构和物理结构。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。

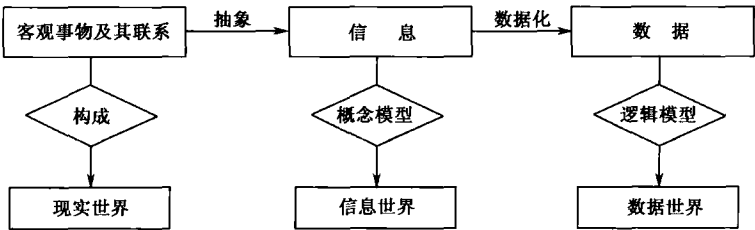


图 1-1 信息的三个世界

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，需要使用两个不同层次的数据模型。数据模型（Data Model）是用于表示现实问题的数据概念、数据关系以及数据约束的集合，是数据库系统中用于提供信息表示和操作手段的形式构架。

1.2.2 概念模型

所谓概念模型，就是为正确、直观地反映客观事物及其联系，对所研究的信息世界建立的一个抽象的模型。该模型是按照用户的观点来对现实世界建模，完全不涉及信息在计算机

中的表示，只是用来描述某个特定组织所关心的信息结构。实际上，概念模型是现实世界到数据世界的第一次抽象，是用户与数据库设计人员之间进行交流的语言。而逻辑模型直接面向数据库的逻辑结构，是现实世界到数据世界的第二次抽象，是按计算机系统的观点对数据建模的。

概念模型具有以下特征：

- ⊙ 具有较强的语义表达能力，能够方便、直接地表达应用中的各种语义知识。
- ⊙ 简单、清晰、易于理解。
- ⊙ 独立与计算机系统，完全不涉及信息在计算机中的表示。
- ⊙ 独立于任何数据库管理系统，可以方便地转换为计算机上任意数据库管理系统所支持的特定数据模型。

目前，能够对现实世界的信息结构进行描述的最常用的方法是 1976 年由 RS.Chen 提出的实体-联系方法（Entity-Relation Approach），简称 E-R 方法。E-R 方法是按用户的观点来描述现实世界的信息结构的，应用 E-R 方法建立的概念模型称为实体-联系模型，简称 E-R 模型。由于 E-R 模型独立于具体的数据库管理系统所支持的数据模型，是各种数据模型的共同基础，因此比数据模型更一般、更抽象、更接近于现实世界。

从本质上讲，E-R 方法是一种信息结构的图示法，简称 E-R 图，是直接表示概念模型的有效工具。在 E-R 图中有 4 个基本成分：

- ⊙ 矩形框——表示实体类型。
- ⊙ 菱形框——表示实体间联系的类型。
- ⊙ 椭圆框——表示实体类型和联系类型的属性。
- ⊙ 连线——实体与属性间、联系与属性间用直线连接。

两个实体之间的联系可以用 E-R 图表示，如图 1-2 所示。

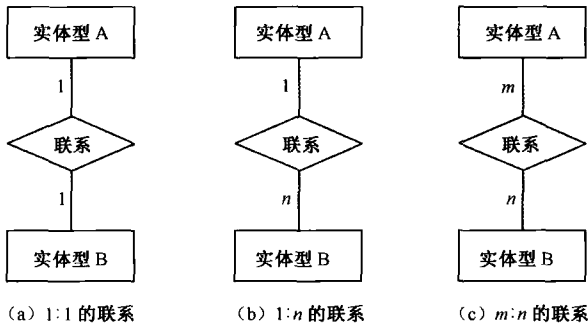


图 1-2 两个实体之间的联系

不但两个实体之间存在联系，多个实体之间也可以存在以上三种联系。例如，课程、教师、参考书三个实体之间存在一对多联系，一门课程可以由多名教师讲授，可以有多本参考书；供应商、项目、零件三个实体之间存在多对多联系，每个供应商可以给多个项目供应各种零件，每个项目所使用的多个零件可以由不同的供应商供应。上述两个联系用 E-R 图表示，如图 1-3 所示。

E-R 图的设计步骤如下。

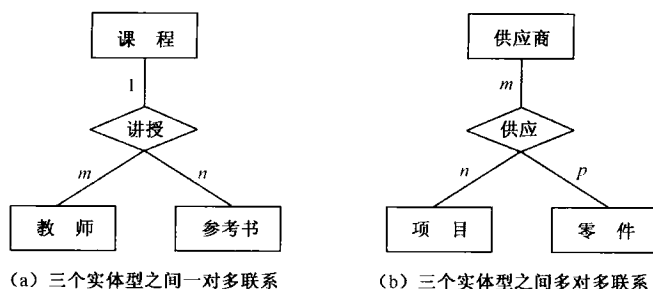


图 1-3 三个实体之间的联系

(1) 设计局部 E-R 图

① 采用分类、聚集和概括的方法对应用系统中的实体、关系抽取其共性，忽略其非本质的细节，确定实体和属性。划分实体和属性时一般遵循以下原则：属性是不可再分的数据项，不能再具有需要描述的性质；属性不能与其他实体发生联系，联系只能发生在实体之间。

② 对上面产生的数据再抽象，形成实体、实体的属性，标识实体的码，确定实体之间的联系，设计局部 E-R 图，如图 1-4 所示。

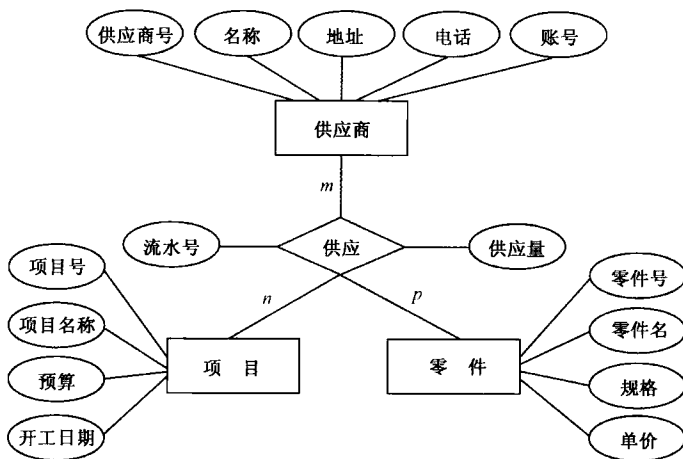


图 1-4 零件供应局部 E-R 图

(2) 综合成全局 E-R 图

① 解决各局部 E-R 图之间的冲突，如命名冲突、属性冲突、结构冲突等，将各局部 E-R 图合并起来生成初步 E-R 图。

② 对初步 E-R 图进行修改、合并，消除不必要的冗余，形成一个整体的概念结构图，即全局 E-R 图。

【例 1-1】用完整的 E-R 图来表示某工程公司数据库管理系统的概念模型，如图 1-5 所示。

如果应用系统的概念模型中涉及的实体和实体属性较多，可以将实体及其属性在另一幅图中画出来，全局 E-R 图中应着重清楚地刻画实体与实体之间的联系。

E-R 模型的优点为：一是简单、容易理解，真实地反映用户的需求；二是与计算机无关，用户容易接受。但由于 E-R 模型只能说明实体间语义的联系，因此在数据库设计时，遇到实际问题总是先设计 E-R 模型，然后将 E-R 模型转换成计算机能够实现的数据模型。