

DB2

从入门到精通

17小时语音视频讲解

明日科技 编著



- ✓ 实例资源库
- ✓ 模块资源库
- ✓ 项目资源库
- ✓ 面试资源库
- ✓ 测试题库系统
- ✓ PPT电子课件

循序渐进，实战讲述

基础知识 ⇨ 核心技术 ⇨ 高级应用 ⇨ 项目实战
204个学习实例，57个练习实例

海量资源，可查可练

除本书配套的17小时视频讲解外，根据学习顺序，光盘还额外配备如下海量开发资源库：

实例资源库(1010个实例) ⇨ 模块资源库(15个典型模块) ⇨ 项目资源库(15个项目案例) ⇨ 测试题库系统(596道测试题) ⇨ 面试资源库(369个面试真题)

在线解答，高效学习

QQ: 400 675 1066(可容纳10万人在线)



软件开发视频大讲堂

DB2 从入门到精通

明日科技 编著

清华大学出版社

北京

内 容 简 介

《DB2 从入门到精通》一书从初学者角度出发，通过通俗易懂的语言、丰富多彩的实例，详细介绍了使用 DB2 进行数据管理的各种技术。全书共分 3 篇 19 章，包括初识 DB2，DB2 的体系结构与 DB2 工具，实例和服务器管理，DB2 数据库基础，管理表空间和表，表数据的基本操作，数据库对象，DB2 数据查询技术，视图的使用，约束、索引和别名，SQL 过程语言，DB2 中的函数，存储过程和触发器，DB2 数据库安全处理，数据库的备份与恢复，PureXML 技术应用，XQuery 技术的使用，数据移动，事务与锁等内容。书中所有知识都结合具体实例进行介绍，涉及的程序代码给出了详细的注释，可以使读者轻松领会 DB2 管理数据库的精髓，快速提高数据库管理技能。

本书列举了大量的实例，所附 DVD 光盘内容有同步视频讲解、实例源程序、“实践与练习”答案等；本书的服务网站提供了模块库、案例库、题库、素材库、答疑服务。

本书内容详尽，实例丰富，非常适合作为 DB2 初学者的学习用书，也适合作为开发人员的查阅、参考资料。

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

DB2 从入门到精通/明日科技编著. —北京：清华大学出版社，2017

(软件开发视频大讲堂)

ISBN 978-7-302-45800-5

I. ①D… II. ①明… III. ①关系数据库系统 IV. ①TP311.138

中国版本图书馆 CIP 数据核字 (2016) 第 291072 号

责任编辑：赵洛育

封面设计：刘洪利

版式设计：魏 远

责任校对：王 云

责任印制：李红英

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：清华大学印刷厂

经 销：全国新华书店

开 本：203mm×260mm 印 张：29 字 数：793 千字

(附海量开发资源库 DVD1 张)

版 次：2017 年 6 月第 1 版

印 次：2017 年 6 月第 1 次印刷

印 数：1~5000

定 价：79.80 元

产品编号：058899-01

目 录

Contents

第 1 篇 基础知识

第 1 章 初识 DB2.....	2	2.4.4 “命令编辑器”“命令窗口”“命令行处理器”用法的区别.....	27
📺 视频讲解: 31 分钟		2.5 DB2 常用工具.....	28
1.1 DB2 的发展历史.....	3	2.5.1 配置助手.....	28
1.2 关系型数据库的基本理论.....	4	2.5.2 控制中心.....	29
1.2.1 数据模型的概念.....	4	2.5.3 工具设置.....	31
1.2.2 常见的数据模型.....	4	2.5.4 向导.....	32
1.2.3 关系型数据库与数据库管理系统.....	5	2.5.5 信息中心.....	32
1.2.4 关系型数据库的 E-R 模型.....	6	2.5.6 运行状况中心.....	33
1.2.5 关系型数据库的设计范式.....	8	2.5.7 任务中心.....	33
1.2.6 关系数据库的设计原则.....	10	2.5.8 复制中心.....	34
1.3 其他常见关系数据库.....	10	2.5.9 日志.....	35
1.3.1 Access 数据库.....	11	2.6 小结.....	36
1.3.2 SQL Server 数据库.....	11	2.7 实践与练习.....	36
1.3.3 Oracle 数据库.....	11	第 3 章 实例和服务器管理.....	37
1.4 DB2 的安装与配置.....	11	📺 视频讲解: 26 分钟	
1.4.1 DB2 的安装.....	11	3.1 实例.....	38
1.4.2 验证 DB2 的安装.....	19	3.1.1 什么是 DB2 实例.....	38
1.5 小结.....	21	3.1.2 创建实例.....	38
1.6 实践与练习.....	21	3.1.3 实例目录.....	39
第 2 章 DB2 的体系结构与 DB2 工具.....	22	3.1.4 列出实例.....	40
📺 视频讲解: 66 分钟		3.1.5 启动/停止实例.....	40
2.1 DB2 体系结构.....	23	3.1.6 更新实例.....	40
2.2 对象层次关系.....	24	3.1.7 删除实例.....	41
2.3 数据访问过程.....	24	3.1.8 配置实例.....	41
2.4 DB2 命令工具的使用.....	25	3.2 管理服务器.....	42
2.4.1 命令编辑器.....	25	3.2.1 管理服务器的概念.....	42
2.4.2 命令窗口.....	26	3.2.2 创建管理服务器.....	42
2.4.3 命令行处理器.....	26		

3.2.3 管理服务器的相关命令	43	6.2 使用界面操作表数据	96
3.2.4 删除管理服务器	44	6.3 插入记录	98
3.2.5 配置管理服务器	44	6.3.1 用界面方式插入记录	98
3.3 小结	45	6.3.2 用命令方式插入记录	99
3.4 实践与练习	45	6.4 修改记录	103
第4章 DB2 数据库基础	46	6.4.1 用界面方式修改记录	103
 视频讲解: 30 分钟		6.4.2 用命令方式修改记录	104
4.1 数据库结构	47	6.5 删除记录	107
4.2 数据库基本概念	48	6.5.1 用界面方式删除记录	108
4.3 数据库的创建、编目和删除	49	6.5.2 用命令方式删除记录	108
4.3.1 创建数据库	49	6.6 小结	109
4.3.2 SQL 语句中的注释符	54	6.7 实践与练习	109
4.3.3 编目数据库	54	第7章 数据库对象	110
4.3.4 删除数据库	64	 视频讲解: 47 分钟	
4.4 小结	65	7.1 数据库对象概述	111
4.5 实践与练习	65	7.2 模式	111
第5章 管理表空间和表	66	7.2.1 使用控制中心的方式创建模式	112
 视频讲解: 53 分钟		7.2.2 使用命令的方式创建模式	114
5.1 表空间的创建、修改和删除	67	7.3 表	115
5.1.1 创建表空间	68	7.3.1 表的维护	115
5.1.2 修改表空间	70	7.3.2 表约束	116
5.1.3 删除表空间	72	7.3.3 表状态	118
5.2 表的创建、修改和删除	73	7.3.4 表压缩	119
5.2.1 DB2 数据类型	74	7.3.5 表分区	120
5.2.2 创建表	77	7.4 索引	121
5.2.3 修改表	84	7.4.1 B+树结构	121
5.2.4 删除表	91	7.4.2 索引的好处	122
5.3 小结	92	7.4.3 合适的索引	124
5.4 实践与练习	92	7.4.4 复合索引	124
第6章 表数据的基本操作	93	7.5 其他常用数据库对象	126
 视频讲解: 30 分钟		7.5.1 视图	127
6.1 SQL 语言的概述	94	7.5.2 昵称	127
6.1.1 SQL 语言的特点	94	7.5.3 序列	127
6.1.2 SQL 语言的分类	95	7.5.4 自增字段	129
6.1.3 SQL 语言的编写规则	95	7.6 小结	131
		7.7 实践与练习	132

第2篇 核心技术

第8章 DB2 数据查询技术.....	134	9.4.2 使用命令方式操作视图数据	191
 视频讲解: 78 分钟		9.5 修改视图的注释	193
8.1 选择、投影、连接和集合运算	135	9.5.1 使用界面方式为视图 CJB_VIEW1 添加注释	193
8.1.1 选择	135	9.5.2 使用命令方式为视图添加注释	194
8.1.2 投影	136	9.6 删除视图	194
8.1.3 连接	136	9.6.1 使用界面方式删除视图	194
8.1.4 集合运算.....	137	9.6.2 使用命令方式删除视图	195
8.2 检索数据	139	9.7 视图的维护	196
8.2.1 简单查询.....	140	9.8 小结	196
8.2.2 筛选查询.....	145	9.9 实践与练习	196
8.2.3 分组查询.....	150		
8.2.4 排序查询.....	154	第10章 约束、索引和别名.....	197
8.2.5 多表关联查询.....	156	 视频讲解: 46 分钟	
8.2.6 使用聚集函数计算列值.....	162	10.1 默认值和约束	198
8.3 子查询的用法	164	10.1.1 默认值.....	198
8.3.1 子查询.....	164	10.1.2 约束.....	199
8.3.2 单行子查询.....	166	10.2 索引	205
8.3.3 多行子查询.....	166	10.2.1 索引的概念	205
8.3.4 关联子查询.....	168	10.2.2 索引的分类.....	206
8.4 小结	169	10.2.3 创建索引	206
8.5 实践与练习	169	10.2.4 查看已创建的索引	212
第9章 视图的使用	170	10.2.5 修改索引	214
 视频讲解: 37 分钟		10.2.6 重命名索引	214
9.1 视图概述	171	10.2.7 删除索引	215
9.1.1 视图与表的关系.....	171	10.3 别名	216
9.1.2 视图的优点.....	172	10.3.1 创建别名.....	217
9.1.3 使用视图的注意事项.....	172	10.3.2 使用别名.....	218
9.2 创建视图	172	10.3.3 删除别名.....	218
9.2.1 使用界面方式创建视图.....	172	10.4 小结	218
9.2.2 使用命令方式创建视图.....	177	10.5 实践与练习	219
9.2.3 创建特殊类型视图.....	178		
9.3 查询视图	185	第11章 SQL 过程语言.....	220
9.4 更新视图	188	 视频讲解: 120 分钟	
9.4.1 使用界面方式操作视图数据.....	188	11.1 SQL PL 语言	221

11.1.1 SQL PL 语言元素.....	221	11.10 实践与练习	253
11.1.2 VALUES 语句.....	222	第 12 章 DB2 中的函数	254
11.2 数据类型	223	 📺 视频讲解: 38 分钟	
11.2.1 系统数据类型.....	223	12.1 系统内置函数	255
11.2.2 创建单值数据类型.....	223	12.1.1 系统内置函数介绍	255
11.2.3 创建结构数据类型.....	227	12.1.2 常用系统内置函数	256
11.2.4 创建数组数据类型.....	230	12.2 用户定义函数	261
11.3 声明变量	230	12.2.1 创建和调用用户定义函数	261
11.4 赋值	232	12.2.2 用户定义函数的删除	267
11.4.1 赋值语句的语法.....	232	12.3 小结	267
11.4.2 专用寄存器.....	232	12.4 实践与练习	267
11.5 游标	233	第 13 章 存储过程和触发器.....	268
11.5.1 游标的概念.....	234	 📺 视频讲解: 71 分钟	
11.5.2 声明游标.....	234	13.1 存储过程	269
11.5.3 打开游标.....	236	13.1.1 存储过程的概念	269
11.5.4 读取游标.....	237	13.1.2 存储过程的类型	270
11.5.5 关闭游标.....	237	13.1.3 SQL 存储过程的创建与执行.....	270
11.5.6 游标和结果集.....	238	13.1.4 存储过程的重载	284
11.6 流程控制语句	239	13.1.5 存储过程的查询、修改和删除	285
11.6.1 条件语句.....	239	13.2 触发器	286
11.6.2 迭代语句与退出循环语句.....	243	13.2.1 触发器的概念	287
11.6.3 RETURN 语句.....	248	13.2.2 触发器的类型	287
11.7 异常处理机制	248	13.2.3 用 CREATE TRIGGER 语句创建	
11.7.1 声明有名称的条件.....	248	触发器	288
11.7.2 声明条件处理程序.....	249	13.2.4 查询、修改和删除触发器	296
11.7.3 强制发出异常.....	249	13.3 小结	300
11.8 MERGE 语句	250	13.4 实践与练习	300
11.9 小结	253		

第 3 篇 高级应用

第 14 章 DB2 数据库安全处理	302	14.4.1 DB2 身份验证方式.....	305
 📺 视频讲解: 72 分钟		14.4.2 其他身份验证参数	306
14.1 DB2 数据库安全概述.....	303	14.4.3 在客户机-服务器环境中设置身份	
14.2 认证机制	303	验证	307
14.3 身份验证和授权	304	14.5 管理权限	314
14.4 DB2 身份验证.....	305	14.5.1 管理权限级别	314

14.5.2 授予或撤销实例级权限	316	第 16 章 PureXML 技术应用	359
14.5.3 授予或撤销数据库级权限	319	📺 视频讲解: 55 分钟	
14.6 特权	322	16.1 pureXML 技术简介	360
14.6.1 特权级别	322	16.2 创建可以存储 XML 数据的 DB2	
14.6.2 授予和撤销特权	323	数据库	361
14.6.3 隐式特权	328	16.3 创建具有 XML 列的表	361
14.7 LBAC 凭证	329	16.4 样本数据	362
14.8 GET AUTHORIZATIONS 命令	330	16.5 XML 模式存储库	366
14.9 使用模式控制对数据库对象的		16.5.1 XML 模式存储库概念	366
访问	331	16.5.2 增大应用程序堆栈大小配置参数	366
14.10 小结	331	16.5.3 增大代理程序堆栈大小配置参数	367
14.11 实践与练习	331	16.5.4 编写 XML 模式文档	369
第 15 章 数据库的备份与恢复	332	16.5.5 向 XSR 注册 XML 模式	370
📺 视频讲解: 53 分钟		16.5.6 删除 XSR 中的 XML 模式	376
15.1 备份与恢复	333	16.6 将 XML 文档插入到 XML 类型的	
15.2 恢复操作的种类	333	列中	378
15.3 事务日志记录	334	16.6.1 插入时针对 XML 模式验证 XML 文档	378
15.3.1 为什么需要日志	334	16.6.2 插入时不验证 XML 文档	379
15.3.2 事务日志记录的概念	335	16.7 小结	379
15.3.3 主日志文件和辅助日志文件	335	16.8 实践与练习	380
15.3.4 循环日志记录和归档日志记录	336	第 17 章 XQuery 技术的使用	381
15.3.5 修改事务日志模式	337	📺 视频讲解: 49 分钟	
15.4 备份	338	17.1 XQuery 基础	382
15.4.1 离线备份	339	17.1.1 XQuery 查询的组成部分	382
15.4.2 在线备份	339	17.1.2 基本规定	383
15.4.3 表空间备份	340	17.1.3 XML 名称空间和 QName	383
15.4.4 增量备份	341	17.1.4 序言	384
15.5 恢复	342	17.1.5 表达式	386
15.5.1 版本恢复	342	17.2 查询 XML 数据	394
15.5.2 前滚恢复	347	17.2.1 界面查询表中 XML 数据	394
15.5.3 崩溃恢复	348	17.2.2 如何选择查询 XML 数据的语言	395
15.6 常见的恢复场景	349	17.2.3 用普通 SQL 查询 XML 数据	396
15.6.1 整个数据库意外删除和损坏	349	17.2.4 XQuery 和嵌入了 SQL 的 XQuery	398
15.6.2 表空间容器意外丢弃或损坏	353	17.2.5 SQL/XML	399
15.6.3 恢复到时间点	356	17.3 更新 XML 数据	402
15.7 小结	358	17.3.1 使用命令更新 XML 数据	402
15.8 实践与练习	358	17.3.2 变换表达式	403

17.4 小结	405	18.7.4 导入/导出大字段 (LOB)	432
17.5 实践与练习	405	18.7.5 导入 identity 数据	433
第 18 章 数据移动	406	18.7.6 数据移动出现乱码	433
视频讲解: 68 分钟		18.7.7 表数据从一个表空间迁移到另一个 表空间	434
18.1 数据移动概述	407	18.8 小结	435
18.2 文件格式	407	18.9 实践与练习	435
18.2.1 DEL 格式	408	第 19 章 事务与锁	436
18.2.2 ASC 格式	408	视频讲解: 27 分钟	
18.2.3 PC/IXF 格式	408	19.1 数据库事务	437
18.2.4 WSF 格式	408	19.2 事务日志记录	437
18.2.5 CURSOR 格式	409	19.2.1 控制事务日志记录的参数	438
18.3 数据的导出	409	19.2.2 日志文件存储的位置	441
18.3.1 使用 EXPORT 导出数据	409	19.3 并发性控制	442
18.3.2 EXPORT 命令格式及选项	413	19.4 锁	443
18.4 数据的导入	416	19.4.1 锁的基本概念	443
18.4.1 使用 IMPORT 导入数据	416	19.4.2 行级锁和表级锁的模式	444
18.4.2 IMPORT 命令格式及选项	420	19.4.3 获取不同级别的锁的方法	446
18.5 数据的装入	422	19.5 隔离级别	447
18.5.1 装入操作的 4 个阶段	422	19.5.1 隔离级别的种类	447
18.5.2 使用 LOAD 清空表数据	423	19.5.2 设置数据库的事务隔离级别	448
18.5.3 LOAD 实用程序	424	19.5.3 重写隔离级别 (WITH 从句)	448
18.6 复制表	429	19.5.4 继承隔离级别	449
18.7 常见问题	431	19.5.5 锁避免和隔离	449
18.7.1 加载的数据是 Excel 格式	431	19.6 小结	449
18.7.2 导出/加载的数据不是逗号/双引号分隔	432	19.7 实践与练习	450
18.7.3 文件中的列比要导入的表中的字段多	432		

第 1 篇


基础知识

- ▶▶ 第 1 章 初识 DB2
- ▶▶ 第 2 章 DB2 的体系结构与 DB2 工具
- ▶▶ 第 3 章 实例和服务器管理
- ▶▶ 第 4 章 DB2 数据库基础
- ▶▶ 第 5 章 管理表空间和表
- ▶▶ 第 6 章 表数据的基本操作
- ▶▶ 第 7 章 数据库对象

本篇通过对初识 DB2、DB2 的体系结构与 DB2 工具、实例和服务器管理、DB2 数据库基础、管理表空间和表、表数据的基本操作、数据库对象等内容的介绍，并结合大量的图示、举例、录像等使读者快速掌握 DB2 的基础内容，为以后 DB2 数据库管理的学习奠定坚实的基础。

第 1 章

初识 DB2

( 视频讲解：31 分钟)

本章主要介绍数据库的相关概念，包括 DB2 数据库的发展史概述、数据模型的概念、关系型数据库的 E-R 模型、关系型数据库的设计范式、关系数据库的设计原则、其他常见数据库以及 DB2 的安装与配置等内容。通过本章的学习，读者应该掌握数据库系统、数据模型以及数据库设计范式等概念，对比常见的关系数据库。

通过阅读本章，您可以：

- ▶▶ 了解 DB2 数据库的发展
- ▶▶ 掌握关系型数据库的基本理论
- ▶▶ 了解其他常见的关系数据库
- ▶▶ 掌握 DB2 数据库的安装与配置

1.1 DB2 的发展历史

对于每个最终站在领奖台上泪水盈面的奥运冠军来说，为此刻他（她）也许已经付出了 5 年甚至 10 年的艰苦努力。相对这些人类的冠军们，这个世界还有另外一种意义上的冠军，他们虽没有泪水，却依然在历史上留下了非凡的轨迹——DB2 就是这类冠军中的一员，这个数据库领域里当之无愧的冠军，它已用了足足 25 年来描绘自己的足迹。

DB2 LUW (IBM DB2 Database for Linux、UNIX and Windows, 在本书后续章节中统一简写为 DB2) 是 IBM 公司研制的一种关系型数据库系统。DB2 主要应用于大型应用系统，具有较好的可伸缩性。DB2 提供了高层次的数据利用性、完整性、安全性、可恢复性，以及小规模到大规模应用程序的执行能力，具有与平台无关的基本功能和 SQL 命令运行环境。

1968 年，在 IBM 360 计算机上研制成功了 IMS V1，这是第一个也是最著名的和最为典型的层次型数据库管理系统。

1970 年，IBM 公司的研究员 E.F.Codd 发表了业界第一篇关于关系型数据库理论的论文《A Relational Model of Data for Large Shared Data Banks》，首次提出了关系模型的概念，这篇论文是计算机科学史上最重要的论文之一，奠定了 Codd 博士“关系数据库之父”的地位。

1973 年，IBM 研究中心启动 System R 项目，为 DB2 的诞生打下了良好基础。System R 是 IBM 研究部门开发的一种产品，这种原型语言促进了技术的发展并将 DB2 带到了商业市场。

1981 年，IBM 的研究员 E.F.Codd 发明的关系型数据库模型获得了 ACM 图灵奖，这是计算机科学界的最高荣誉。Codd 博士是继查尔斯·巴赫曼 (Charles W. Bachman) 之后，又一位由于在数据库领域做出巨大贡献而获此殊荣的计算机科学家。

1983 年，IBM 推出的第一款面向大型企业的商业化关系数据库管理系统 (Database2 for MVS)，内部代号为“Eagle”，于是 DB2 正式诞生。

1986 年，IBM 发布了 System/38 V7，该系统首次配置查询优化器，能对应用的存取计划进行优化，而后 DB2 又完成了到 OS/2 V1.0 的扩展，这是 IBM 第一次把关系型数据库处理能力扩展到微机系统。也是 DB2 for OS/2、UNIX and Windows 的雏形。

1995 年，IBM 发布了 DB2 Common Server V2，这是第一个能够在多个平台上运行的“对象-关系型数据库” (ORDB) 产品，并能够对 Web 提供充分支持。DataJoiner for AIX 也诞生在这一年，该产品赋予了 DB2 对异构数据库的支持能力。

1996 年，IBM 发布 DB2 V2.1.2，这是第一个真正支持 Java 和 JDBC 的数据库产品。

1997 年，IBM 发布了可以支持 Web 的 DB2 for OS/390 V5，这是当时唯一能够支持 64000 个并发用户和百 TB 级别的数据库产品。

1999 年，IBM 为了对移动计算提供支持，发布了 DB2 UDB 卫星版和 DB2 Everywhere。

2000 年，IBM 发布了 DB2 XML Extender，成为在业界第一个为数据库提供内置 XML 支持的厂商。同年 IBM 将 Visual Warehouse 集成到 DB2 中，为 DB2 提供了内置的数据仓库管理功能。

2001 年，IBM 以 10 亿美金收购了 Informix 的数据库业务，这次收购扩大了 IBM 的分布式数据库业务。

2002 年, IBM 宣布计划收购 Rational Software Corp, 从而使得 IBM 软件能够支持从设计、开发、部署到管理和维护的完整过程。

DB2 的诞生不仅促进了与关系数据库概念相关的数学和科学的发展, 还创造性地开发出一种极具影响力的全新软件类型。今天 DB2 已经发展成为 IBM 信息管理 (IM) 软件组合的重要组成部分。在 IBM 信息按需应变策略和体系结构中, DB2 扮演数据基础服务平台的重要角色, 并且已经发展成同时支持传统关系数据和 XML 的混合型数据服务器, DB2 数据库产品及解决方案已经广泛应用在金融、电信、制造、零售等行业。

1.2 关系型数据库的基本理论

数据库技术是应对信息资源 (即大量数据) 的管理需求而产生的, 随着信息技术的不断发展, 尤其是人类迈入网络时代后, 社会信息资源在爆炸式地增长, 对数据管理技术也随之不断地提出更高的要求。数据库管理技术先后经历了人工管理、文件系统、数据库系统 3 个阶段。在数据库系统中, 数据模型主要有层次模型、网状模型和关系模型 3 种 (另外一种面向对象模型还处在探索研究中), 目前理论成熟、使用普及的模型就是关系模型——关系型数据库的理论基础。

1.2.1 数据模型的概念

数据模型是数据库系统的核心与基础, 是关于描述数据与数据之间的联系、数据的语义、数据一致性约束的概念性工具的集合。

数据模型通常是由数据结构、数据操作和完整性约束 3 部分组成的。

- ☑ 数据结构: 是对系统静态特征的描述。描述对象包括数据的类型、内容、性质和数据之间的相互关系。
- ☑ 数据操作: 是对系统动态特征的描述。是对数据库中各种对象实例的操作。
- ☑ 完整性约束: 是完整性规则的集合。它定义了给定数据模型中数据及其联系所具有的制约和依存规则。

1.2.2 常见的数据模型

常用的数据库数据模型主要有网状模型、层次模型和关系模型, 下面分别进行介绍。

(1) 网状模型: 用有向图结构表示实体类型及实体间联系的数据模型称为网状模型。用网状模型编写应用程序极其复杂, 数据的独立性较差, 如图 1.1 所示。

(2) 层次模型: 用树形结构表示实体类型及实体间联系的数据模型称为层次模型, 它具有以下特点, 如图 1.2 所示。

- ☑ 每棵树有且仅有一个无双亲节点, 称为根。
- ☑ 树中除根外所有节点有且仅有一个双亲。

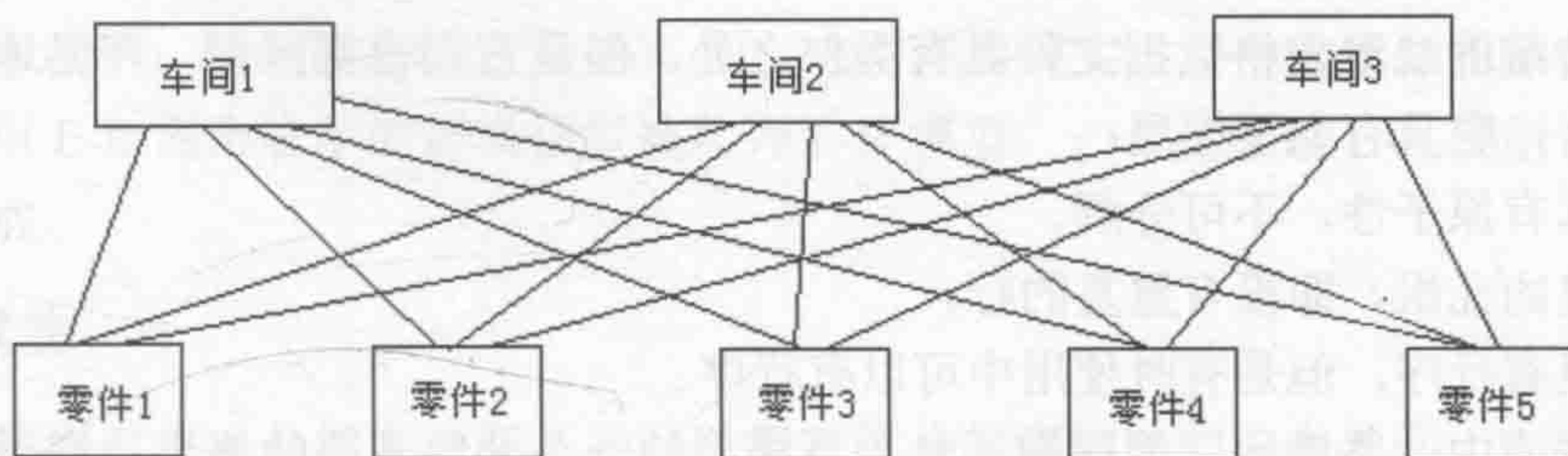


图 1.1 按网状模型组织的数据示例

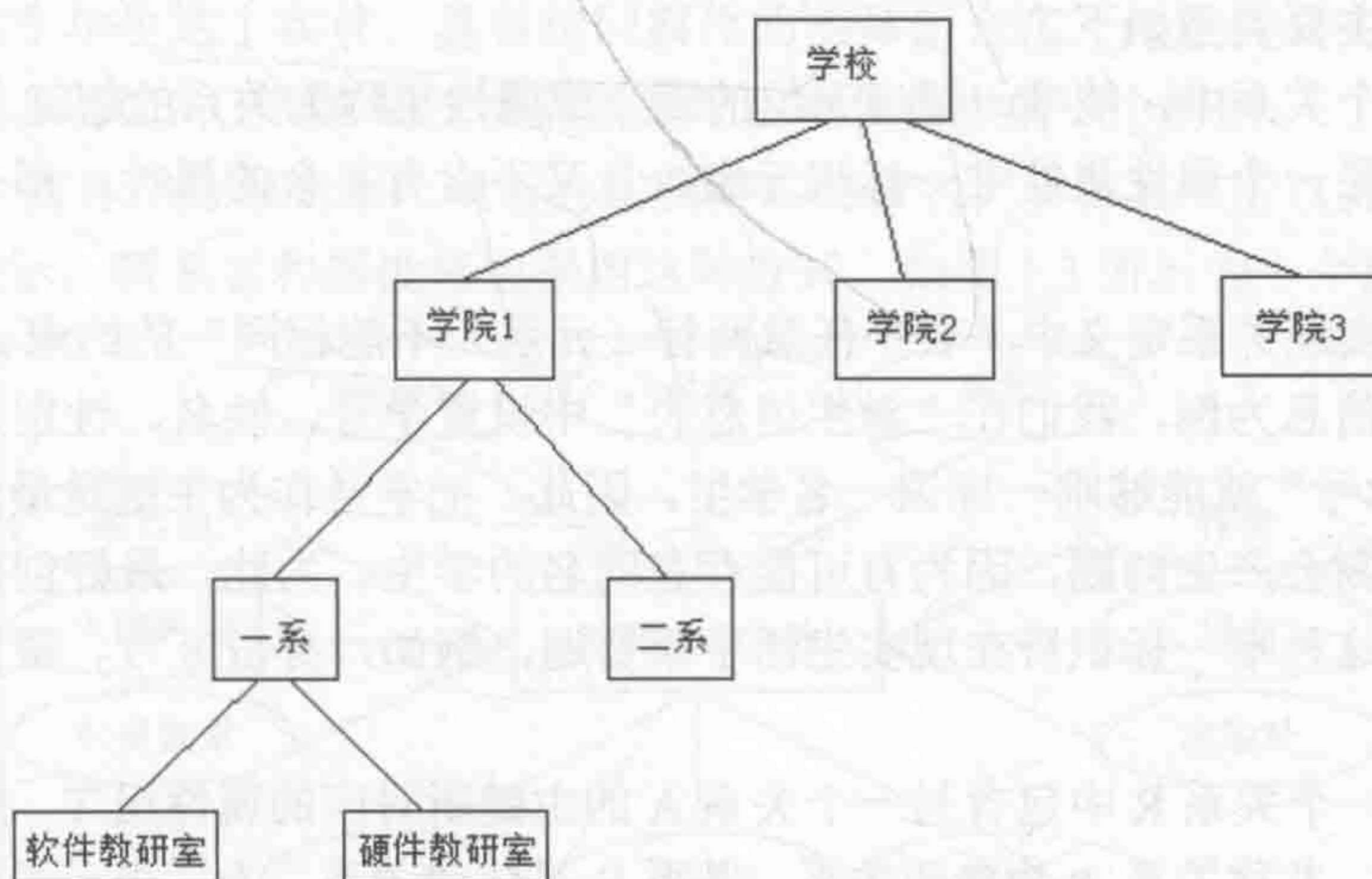


图 1.2 按层次模型组织的数据实例

(3) 关系模型：以二维表来描述数据。关系模型中，每个表有多个字段列和记录行，每个字段列有固定的属性（数字、字符、日期等）。关系模型数据结构简单、清晰、具有很高的数据独立性，因此是目前主流的数据库数据模型。

关系中的数据约束如下。

- 实体完整性约束：约束关系的主键中属性值不能为空值。
- 参照完整性约束：关系之间的基本约束。
- 用户定义的完整性约束：它反映了具体应用中数据的语义要求。

1.2.3 关系型数据库与数据库管理系统

关系型数据库是建立在关系模型基础上的数据库，借助于集合代数等数学概念和方法来处理数据库中的数据，现实世界中的各种实体以及实体之间的各种联系均可用关系模型来表示。

在关系数据模型中，关系可以看成由行和列交叉组成的二维表格，表中一行称为一个元组，可以用来标识实体集（实体的集合）中的一个实体。表中的列称为属性，给每一列起一个名称即为属性名，表中的属性名不能相同。列的取值范围称为域，同列具有相同的域，不同的列也有相同的域。表中任意两行（元组）不能相同。能唯一标识表中不同行的属性或属性组（即多个属性的组合）称为主键或复合主键。

尽管关系与传统的二维表格数据文件具有类似之处，但是它们也有区别，严格地说，关系是一种规范化的二维表格，它具有如下性质：

- ☑ 属性值具有原子性，不可分解。
- ☑ 没有重复的元组，即没有重复的行。
- ☑ 理论上没有行序，但是有时使用中可以有行序。

在关系型数据库中，关键码（简称键）是关系模型的一个非常重要的概念，它通常是行（元组）的一个或几个列（属性）。如果键是由一个列组成，则称之为唯一键；若是由多个列（属性）组成，则称之为复合键，键的主要类型如下。

- ☑ 超键：在一个关系中，能唯一标识元组的属性或属性集称为关系的超键。
- ☑ 候选键：如果一个属性集能唯一标识元组，且又不含有多余的属性，那么这个属性集称为关系的候选键。
- ☑ 用主键可以实现关系定义中“表中任意两行（元组）不能相同”的约束。

这里以管理学生信息为例，我们在“学生信息表”中设置学号、姓名、性别、年龄、院系、班级等列。在该表中，“学号”就能够唯一标识一名学生，因此，把学号作为主键是最佳的选择，而如果把“姓名”列作为主键则会产生问题，因为有可能存在同名的学生。为此，最好创建一个单独的键将其明确地指定为主键，这种唯一标识符在现实生活中很普遍，例如，身份证号、银行卡号、手机号、发票号等。

- ☑ 外键：如果一个关系 R 中包含另一个关系 A 的主键所对应的属性组 T，则称此属性组 T 为关系 R 的外键，并称关系 A 为参照关系，关系 R 是依赖关系。为了表示关联，可以将一个关系的主键作为属性放入另外一个关系中，第二个关系中的那些属性就称为外键。

这里以商品销售为例，在填写一张商品销售单时，可以将商品销售信息分为两大类：第一类是单据的主体信息（销售主表），例如，销售单号、销售金额、销售日期、收款人；第二类是单据的明细信息（销售明细表），例如，商品序号、商品名称、商品数量等。在数据库的“销售主表”中通常以“销售单号”作为主键；在“销售明细表”中，为了标识被销售出去的商品隶属于哪张单据，需要对每一条商品销售记录标明“单据编号”。在这种情况下，销售明细表中的“销售单号”就被称为外键，因为“销售单号”是其所在表以外（主体表）的一个主键。

当出现外键时，主键与外键的列名称可以是不同的。但必须要求它们的值集相同，即“销售明细表”中出现的“销售单号”一定要和主体表中的值匹配。

对于上面提到的“二维表格”中存储的数据信息，通常以物理文件的形式存储在磁盘上，这种物理文件称之为“数据文件”，用户会使用一种数据库软件实现与磁盘上的数据文件进行交互，这种数据库软件就被称之为数据库管理系统（英文缩写为 DBMS）。DBMS 是建立在操作系统基础上的，它可以实现对数据库文件进行统一管理和控制。用户对数据库提出的访问请求都是由 DBMS 来处理的。另外，DBMS 还提供了多种用于管理数据的实用工具。

1.2.4 关系型数据库的 E-R 模型

在设计关系型数据库时，首先需要为它建立逻辑模型。关系型数据库的逻辑模型可以通过实体和

关系组成的图形来表示，这种图形称之为 E-R 图，它实现将现实世界中的实体和实体之间的联系转换为逻辑模型。使用 E-R 图形表示的逻辑模型被称为 E-R 模型，一个标准的 E-R 模型主要由实体、属性和联系 3 部分组成。

1. 实体和属性

实体是一个数据对象，是指客观存在并可以相互区分的事物，如一个教师、一个学生、一个雇员等。每个实体由一组属性来表示，如一个具体的学生拥有学号、姓名、性别和班级等属性，其中学号可以唯一标识具体某个学生这个实体。具有相同属性的实体组合在一起就构成实体集，而实体则是实体集中的某一个特例，例如，王同学这个实体就是学生实体集中的一个特例。

在 E-R 模型中，实体用矩形表示，矩形内注明实体的命名。实体名常用以大写字母开头的有具体意义的英文名词来表示，联系名和属性名也采用这种方式。如图 1.3 所示为一个图书档案的 E-R 图。

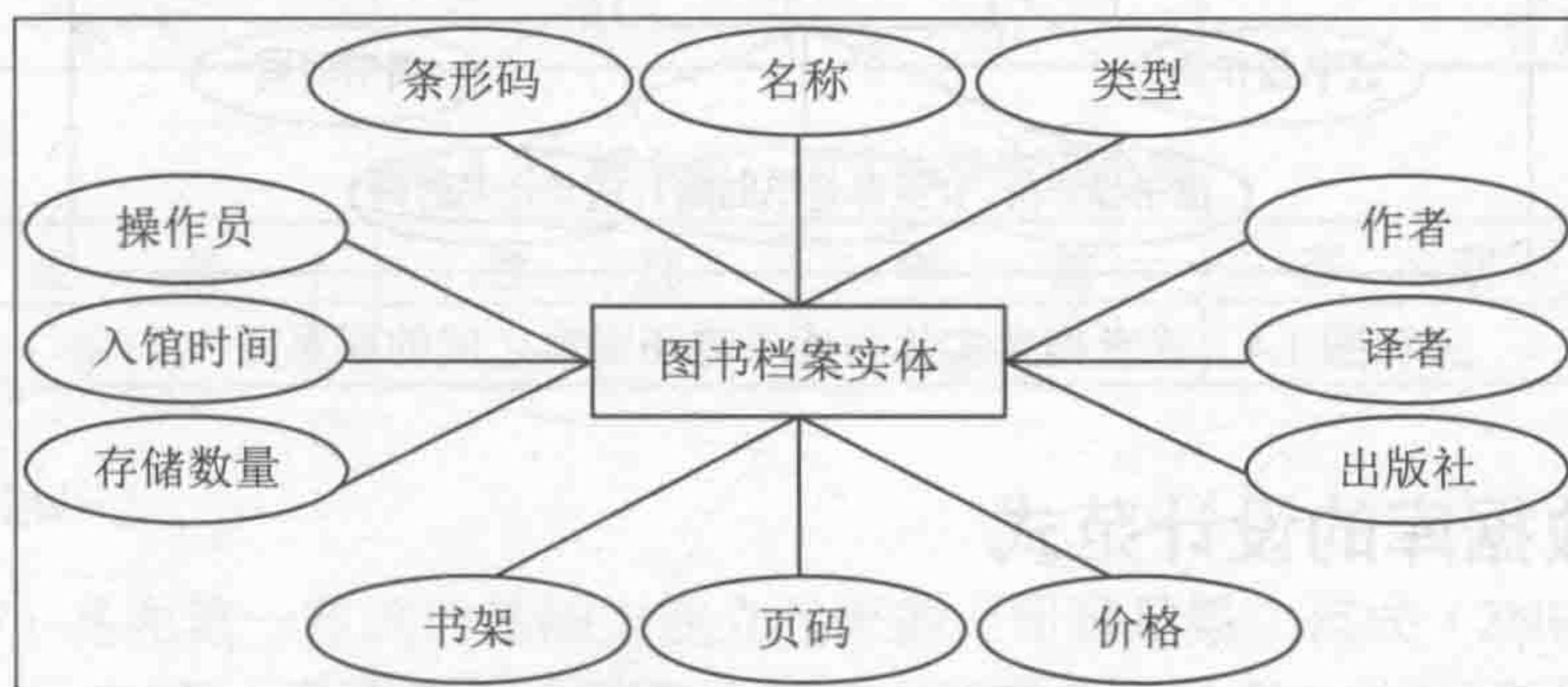


图 1.3 图书档案实体 E-R 图

2. 联系

在实际应用中，实体之间是存在联系的，这种联系必须在逻辑模型中表现出来。在 E-R 模型中，联系用菱形表示，菱形框内写明“联系名”，并用“连接线”将有关实体连接起来，同时在“连接线”的旁边标注上联系的类型，两个实体之间的联系类型可以分为以下 3 类。

- ☑ 一对一：若对于实体集 A 中的每一个实体，在实体集 B 中最多有一个实体与之相关，反之亦然，则称实体集 A 与实体集 B 具有一对一的联系，可标记联系为 1:1。
- ☑ 一对多：若对于实体集 A 中的每一个实体，在实体集 B 中有多个实体与之相关；反之，对于实体集 B 中的每一个实体，实体集 A 中最多有一个实体与之相关，则称实体集 A 与实体集 B 具有一对多的联系，可标记联系为 1:n。
- ☑ 多对多：若对于实体集 A 中的每一个实体，在实体集 B 中有多个实体与之相关；反之，对于实体集 B 中的每一个实体，实体集 A 中也有多个实体与之相关，则称实体集 A 与实体集 B 具有多对多的联系，可标记联系为 m:n。

例如，一个读者可以有多个图书借还记录，而一个借还记录只能隶属于一个读者，这样“读者档案实体”与“读者借还实体”之间就存在一对多的联系（即 1:n），那么这两个实体之间的联系如图 1.4 所示。

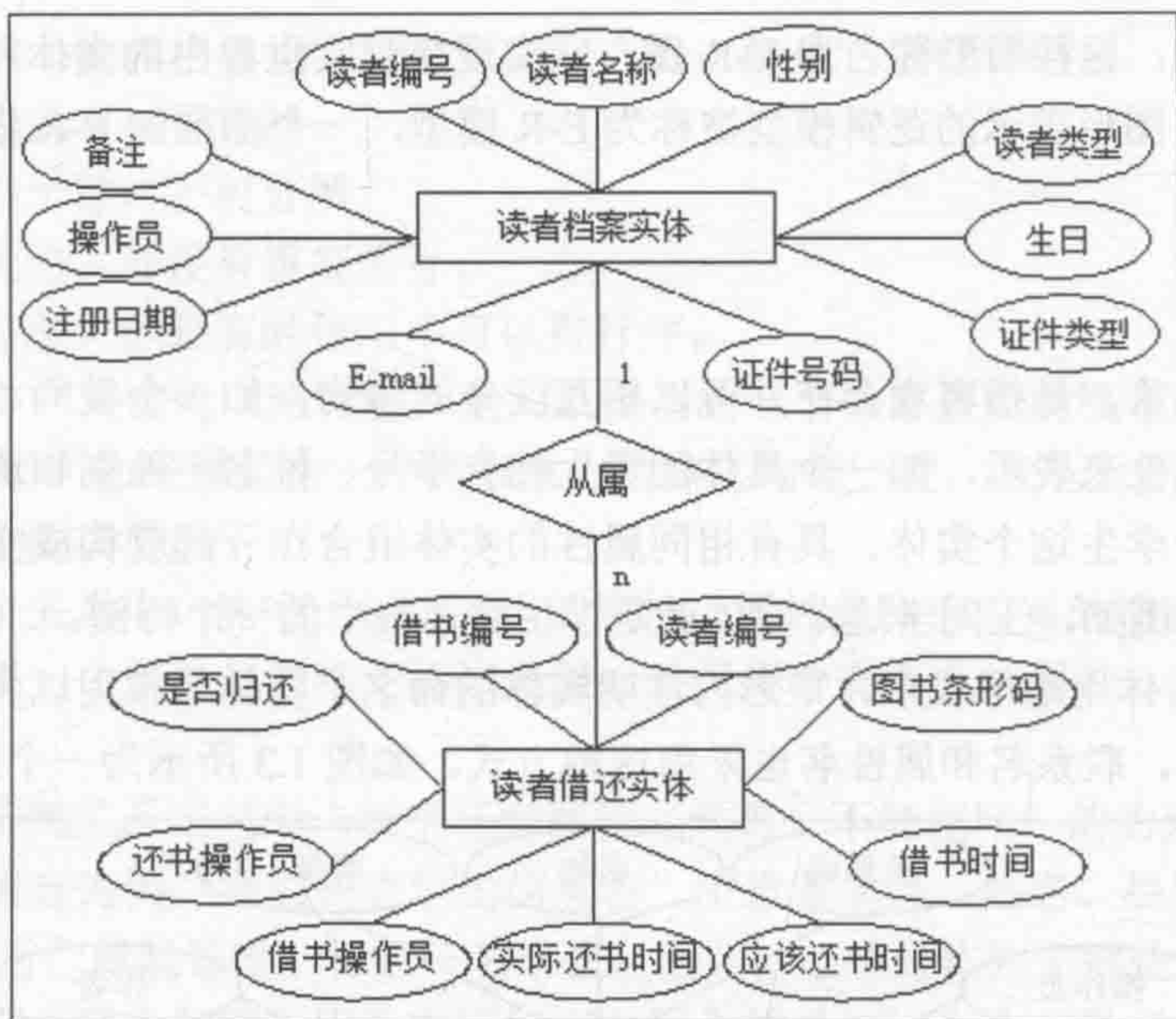


图 1.4 读者档案实体与读者借还实体之间的联系图

1.2.5 关系型数据库的设计范式

在数据库中，数据之间存在着密切的联系。关系型数据库由相互联系的一组关系所组成，每个关系包括关系模式和关系值两个方面。关系模式是对关系的抽象定义，给出关系的具体结构；关系的值是关系的具体内容，反映关系在某一时刻的状态。一个关系包含许多元组（记录行），每个元组都是符合关系模式结构的一个具体值，并且都分属于相应的属性。在关系数据库中的每个关系都需要进行规范化，使之达到一定的规范化程度，从而提高数据的结构化、共享性、一致性和可操作性。

规范化是把数据库组织成在保持存储数据完整性的同时最小化冗余数据的结构的过程。规范化的数据库必须符合关系模型的范式规则。范式可以防止在使用数据库时出现不一致的数据，并防止数据丢失。关系模型的范式有第一范式（1NF）、第二范式（2NF）、第三范式（3NF）、第四范式（4NF）、第五范式（5NF）、第六范式（6NF）和 BCNF 范式等多种。通常数据库只要满足前 3 个范式就足够用了，下面举例介绍前 3 个范式。

1. 第一范式（1NF）

第一范式是第二和第三范式的基础，是最基本的范式。第一范式包括下列指导原则：

- ☑ 数据组的每个属性只可以包含一个值。
- ☑ 关系中的每个数组必须包含相同数量的值。
- ☑ 关系中的每个数组一定不能相同。

在任何一个关系数据库中，第一范式是对关系模式的基本要求，不满足第一范式（1NF）的数据库就不是关系型数据库。

如果数据表中的每一个列都是不可再分割的基本数据项——即同一列中不能有多个值，那么就称