

企业信息化丛书

# 数据库设计与实现

徐兰芳 彭 冰 吴永英 编

冯玉才 主审

上海交通大学出版社

## 内 容 提 要

本书全面地介绍了数据库系统的基本概念、基本原理和基本技术。其中主要内容有：数据库系统引论、关系数据库、关系数据库标准语言 SQL、关系数据库设计理论、数据库设计、数据库安全与保护、基于 DM5 平台的数据库应用开发、数据库发展趋势与新技术、关系数据库设计实验等。本书在介绍基本原理的同时，利用国产数据库管理系统 DM5 平台，由浅入深、循序渐进地讲解了大量的设计实例，力求读者能抓住重点，尽快掌握数据库系统的基本原理和基本技术，并能将相应知识应用到设计和开发数据库及其应用系统中。

全书层次清楚、重点突出、图文并茂、实例丰富，可以作为高等学校研究生、本科及大专学生的教材，也可以作为各类工程技术人员培训以及自学的教材或参考书，或作为信息系统开发人员及系统维护人员的参考用书。

# 《企业信息化丛书》编委会



顾 问：覃 正 洪 帆

主 任：瞿彭志

副 主 任：刘 翔 徐兰芳

编 委：(按姓氏笔画为序)

刘 翔 江 胜 江文年 汤兵勇 吴永英

张志建 张建华 李明星 周培端 徐兰芳

殷贤亮 彭 冰 潘秋荣 瞿彭志

策 划：范荷英

# 序

随着社会进入信息时代，人们已经与信息这一重要资源息息相关。作为开发信息资源的主要技术之一的数据库技术，在信息的管理和开发领域发挥着越来越大的作用。近 30 年来，数据库系统理论和相关知识的研究不仅推动了数据库技术以及相关产业的发展，特别是关系数据库理论和技术的迅速发展，进一步促进了人工系统的规范化管理，使之产生了不可估量的社会效益。

本书作者 30 年来一直在华中科技大学教学和科研的第一线。在对多层次学生的数据库课程教学和亲身参与多项有关数据库系统的数据库设计和实现项目中，作者深切感受到学习数据库技术必须理论联系实际。而现今的数据库教学偏重于理论而缺乏实际能力的培养。因此，作者尝试在介绍数据库系统理论的同时，结合一个具体的数据库管理系统平台，由浅入深地讲解数据库设计和实现过程。这也是本书的特点和宗旨。

首先，注重理论联系实际，有的放矢。数据库技术近 30 年来发展很快，有关数据库的理论和技术的很多，涉及的面很广，在一本书里难以包括所有的内容。本书面对设计一个数据库及其应用系统所涉及的必要的理论知识，按照知识的连贯性对教材的内容进行编排，使得有待于了解数据库知识的各层次读者都可以在书中找到自己的切入点。例如，在第 1 章的最后一节详细介绍了 DM5 的安装和启动，让刚接触数据库基本概念的读者对数据库环境有一个感性认识。又如在讲解关系代数时，尽量结合检索实例，让读者感到的不仅仅是抽象的理论。读者若希望在网络环境下开发以数据库为基础的软件产品，可以在第 6、第 7 章中找到需要的内容。全书自始至终都结合一个实实在在的开发平台 DM5，使读者在接受理论知识的同时，可以在 DM5 平台上看到其操作结果。

其次，本书介绍的数据库平台 DM5，是我国自行研制的、商品化的、具有完全知识产权的数据库管理系统。DM5 除了具有国际流行的商用数据库管理系统的功能和性能外，还具有 B1 级安全级、能直接支持空间数据管理和多媒体数据管理全文检索、多字符集、跨平台互访等独特优势，可以与国外同类产品媲美。DM5 的问世，一举打破了长期以来国外数据库管理系统产品一统国内数据库市场的局面，在社会上引起强烈反响。至今已在国民经济许多领域，如军事、电力、邮电、公安、安全、水利、旅游、地质、房地产、商业、金融、医

疗、交通等各行各业得到广泛应用。以具有自主知识产权的数据库管理系统 DM5 作为本书的讲解平台，是本书的又一大特点，也是作者多年的一个心愿。

希望作者的努力能得到同行专家、学者和广大师生、读者的支持和认可。

华中科技大学计算机科学与技术学院

冯玉才

# 致 谢

本书共分9章。第1, 2, 3, 4, 5, 9章由徐兰芳副教授编写, 第6, 7, 8章由彭冰博士编写。全书由冯玉才教授主审, 吴永英副教授认真地审阅了全书, 并提出了许多建议, 还亲自进行了修改。余院兰、胡怀飞、张大圣、邹艳、陈维和吕操等为本书的录稿、插图的制作和部分实例的调试做了大量的工作。在本书的写作和出版过程中, 得到了华中科技大学研究生院的关心和支持, 得到了华中科技大学计算机学院副院长卢炎生教授、洪帆教授和王元珍教授的关心, 得到了出版社编辑的大力协助, 在此一并表示衷心的感谢。

限于水平和时间, 书中不足和错误在所难免, 恳请广大读者批评指正。

读者联系: lfxu@hust.edu.cn

编者

2006年8月

# 目 录

1 概论.....	1
1.1 数据、信息、实体及其联系.....	1
1.1.1 实体及其联系.....	1
1.1.2 数据库发展简史.....	5
1.2 数据库.....	7
1.3 数据库系统和数据库管理系统.....	10
1.3.1 数据库系统的特点、组成和工作方式.....	10
1.3.2 数据库管理系统的目标、功能和组成.....	13
1.4 数据模型.....	15
1.4.1 数据模型和数据库.....	15
1.4.2 实体-联系模型.....	15
1.4.3 层次模型.....	19
1.4.4 网状模型(Network Model).....	21
1.4.5 关系模型.....	22
1.5 关系数据库管理系统实例.....	24
1.5.1 Oracle.....	24
1.5.2 SQL Server.....	25
1.5.3 DB2.....	27
1.5.4 Sybase.....	28
1.6 国产关系数据库管理系统 DM5.....	29
1.6.1 DM5 的特点.....	30
1.6.2 DM5 的优良特性.....	30
1.6.3 DM5 体系结构和组成.....	31
1.7 DM5 的安装和启动.....	33
1.7.1 硬件环境需求.....	33
1.7.2 软件环境需求.....	33
1.7.3 安装和启动.....	34

---

习题 1 .....	45
2 关系数据库系统 .....	46
2.1 关系模型的基本概念 .....	46
2.1.1 关系模型的特点 .....	46
2.1.2 基本定义 .....	47
2.2 关系完整性约束 .....	51
2.2.1 实体完整性 .....	51
2.2.2 参照完整性 .....	51
2.2.3 用户自定义的完整性 .....	52
2.3 关系代数 .....	53
2.3.1 关系代数的九种运算 .....	53
2.3.2 应用实例 .....	63
2.3.3 扩充的关系代数操作 .....	65
2.4 关系演算 .....	67
2.4.1 元组关系演算(Tuple Relational Calculus) .....	67
2.4.2 域关系演算 .....	72
2.4.3 元组演算表达式到域演算表达式的转换 .....	73
2.4.4 关系演算的安全性和等价性 .....	74
2.5 查询优化 .....	76
习题 2 .....	77
3 关系数据库标准语言 SQL .....	80
3.1 SQL 概述 .....	80
3.1.1 SQL 的产生和发展 .....	80
3.1.2 SQL 的组成和特点 .....	81
3.2 数据定义 .....	82
3.2.1 基本数据类型 .....	82
3.2.2 基本表的管理 .....	83
3.2.3 索引的管理 .....	86
3.2.4 视图的管理 .....	87
3.3 查询及优化 .....	88
3.3.1 简单查询 .....	91
3.3.2 组合查询 .....	100
3.3.3 计算查询 .....	105

---

3.3.4 查询优化 .....	107
3.4 数据更新 .....	109
3.4.1 插入数据 .....	109
3.4.2 修改数据 .....	109
3.4.2 删除数据 .....	110
3.5 数据控制 .....	112
3.5.1 权限的类型 .....	112
3.5.2 授权和收权 .....	112
3.6 嵌入式 SQL .....	114
3.6.1 SQL 语言的执行方式 .....	114
3.6.2 嵌入式 SQL 的使用 .....	115
3.6.3 动态的 SQL 语句 .....	126
习题 3 .....	126
4 关系数据库设计理论 .....	128
4.1 问题的提出 .....	128
4.2 函数依赖 .....	131
4.2.1 函数依赖和多值函数依赖 .....	131
4.2.2 函数依赖的逻辑蕴涵 .....	134
4.2.3 关键字 .....	134
4.3 函数依赖的公理系统 .....	135
4.3.1 Armstrong 公理系统 .....	136
4.3.2 公理系统的正确性 .....	136
4.3.3 公理系统的推论 .....	137
4.3.4 公理系统的完备性 .....	139
4.4 关系模式的分解 .....	147
4.4.1 分解原则 .....	147
4.4.2 无损连接性和依赖保持性 .....	147
4.4.3 无损连接性的分解 .....	149
4.4.4 无损连接性分解的检验算法 .....	150
4.4.5 简便算法 .....	152
4.4.6 算法中注意的问题 .....	153
4.4.7 检验的充要条件 .....	154
4.4.8 依赖保持性的分解 .....	156
4.5 关系模式的规范化 .....	156

---

4.5.1	什么是关系模式的规范化.....	156
4.5.2	范式.....	157
4.6	求解候选关键字.....	165
4.6.1	求解候选关键字的充分条件.....	166
4.6.2	求解候选关键字的方法.....	167
习题 4	.....	175
5	数据库设计.....	177
5.1	数据库设计概述.....	177
5.1.1	数据库设计的内容和特点.....	177
5.1.2	数据库设计方法.....	179
5.1.3	数据库设计步骤.....	180
5.1.4	数据库设计工具.....	182
5.2	需求分析.....	184
5.2.1	需求分析的内容.....	184
5.2.2	方法和步骤.....	185
5.3	概念结构设计.....	192
5.3.1	概念设计方法.....	193
5.3.2	视图综合设计方法.....	194
5.4	逻辑结构设计.....	197
5.4.1	E-R 模型到关系模式的转换.....	198
5.4.2	逻辑模式的调整.....	201
5.5	物理结构设计.....	203
5.5.1	簇集设计.....	204
5.5.2	索引的选择.....	206
5.5.3	分区设计.....	207
5.6	数据库应用系统设计.....	207
5.6.1	总体设计.....	208
5.6.2	详细设计.....	209
5.7	数据库的实施与维护.....	209
5.7.1	数据库的定义.....	210
5.7.2	数据的装入.....	210
5.7.3	数据库的试运行.....	211
5.7.4	数据库的运行和维护.....	212
习题 5	.....	213

---

6 数据库安全与保护 .....	215
6.1 事务处理 .....	215
6.1.1 事务 .....	215
6.1.2 事务的执行 .....	216
6.1.3 事务的可串行化 .....	218
6.1.4 事务恢复 .....	221
6.1.5 SQL 中事务定义 .....	223
6.2 并发控制 .....	224
6.2.1 并发控制概述 .....	224
6.2.2 封锁 .....	228
6.2.3 活锁和死锁 .....	231
6.2.4 两段封锁协议 .....	234
6.2.5 基于时间标志的并发控制 .....	236
6.3 数据库的安全性 .....	238
6.3.1 数据库安全性概述 .....	238
6.3.2 用户标识与鉴别 .....	240
6.3.3 存取控制 .....	241
6.3.4 数据加密 .....	245
6.4 数据库的完整性 .....	246
6.4.1 完整性的概念 .....	246
6.4.2 完整性约束条件 .....	247
6.4.3 完整性控制 .....	250
6.5 数据库恢复和备份 .....	253
6.5.1 数据库故障分类 .....	253
6.5.2 数据库恢复原理 .....	255
6.5.3 数据库恢复策略 .....	258
6.5.4 数据库镜像 .....	260
习题 6 .....	261
7 基于 DM5 平台的数据库应用开发 .....	263
7.1 DM5 应用开发基础 .....	263
7.1.1 数据类型 .....	263
7.1.2 表达式 .....	267
7.1.3 函数 .....	271

---

7.2	DM5 数据库应用开发环境.....	276
7.2.1	数据库系统体系结构概述.....	277
7.2.2	数据访问接口.....	280
7.2.3	DM5 体系结构.....	284
7.3	DM5 工具及应用开发工具支持.....	290
7.3.1	管理工具 JManager.....	290
7.3.2	交互式工具 JISQL.....	293
7.3.3	逻辑备份与恢复工具 JImpExp.....	294
7.3.4	数据迁移工具 JDTS.....	295
7.3.5	控制台工具 JConsole.....	297
7.3.6	性能监控工具 JMonitor.....	298
7.3.7	命令行工具.....	298
7.4	数据库应用系统设计案例.....	299
7.4.1	案例说明.....	299
7.4.2	数据库创建.....	300
7.4.3	DM ODBC 编程简述.....	302
7.4.4	参考源代码.....	304
	习题 7.....	311
8	数据库发展趋势与新技术.....	313
8.1	新型数据库系统.....	313
8.1.1	面向对象数据库系统.....	313
8.1.2	分布式数据库系统.....	317
8.1.3	并行数据库系统.....	320
8.1.4	Web 数据库系统.....	324
8.2	数据挖掘和数据仓库技术.....	328
8.2.1	数据挖掘技术.....	328
8.2.2	数据仓库技术.....	332
8.3	联机分析处理技术.....	336
8.3.1	OLAP 概述.....	336
8.3.2	OLAP 的结构与数据组织.....	338
8.3.3	OLAP 与 OLTP 的比较.....	340
	习题 8.....	341
9	关系数据库设计试验.....	342

---

9.1 引言 .....	342
9.2 课题 .....	342
9.3 设计关系数据库的步骤 .....	342
9.4 开发概念数据库 .....	343
9.5 修正数据需求 .....	345
9.6 使设计适应于数据库管理系统的使用 .....	347
9.7 数据库窗口 .....	348
9.8 实验内容 .....	351
实验一 设计并实现一个关系数据库 .....	351
实验二 建立、测试、维护数据库 .....	351
实验三 开发数据库应用系统 .....	351
实验四 设计并实现具有游标查询和错误检测等功能的应用程序 .....	352
A 附录 .....	353
A.1 SQL 语法描述说明 .....	353
A.2 SQL 命令列表 .....	353
A.3 DM4 保留字 .....	355
A.4 DM5 技术支持 .....	356
参考文献 .....	358

# 1 概 论

数据库技术是计算机科学的主要分支之一，近 30 年来有了长足的发展。随着社会进入信息化时代，人们已经与信息这一重要资源的认识、应用和管理息息相关。作为开发信息资源的主要技术——数据库技术，在这一领域中发挥着越来越大的作用。因此，人们普遍要求了解、学习和掌握有关数据库技术的知识。

本章首先介绍涉及数据库技术的基本知识，如数据、数据库、数据库系统、数据库管理系统，接着介绍当前的几种商用关系数据库管理系统，最后具体介绍国产数据库管理系统 DM5 的特性、功能、安装和启动等内容。

## 1.1 数据、信息、实体及其联系

人类生存的现实世界，是一个物质的世界，也是一个信息的世界。物质是信息的基础，信息是物质的抽象反映。一方面，物质的存在伴随着信息的存在。例如，学校有许多学生，那么学生就是信息源，学生的姓名、性别、年龄等就是反映学生的信息。又如，商店有许多商品，必定有关于这些商品的各种特性的信息：商品的名称、价格、颜色等。另一方面，物质的变化会引起信息的改变。例如某一名学生离校、商店新增一种商品等，这时信息必然随着物质的增、减而发生变化，所以信息反映了事物的物理状态，是消息、新闻、情报和知识的总和。

大千世界物质丰富多彩，反映他(它)们本来面目的信息的形式随之也灵活多样，可以用文字，也可以用声音、图形表示。表示信息或者说载荷信息的物理符号称为数据。数据是一切数字、文字、声音、图形、图像等的有意义的组合。信息依赖数据而存在，数据生动具体地表现出信息。

### 1.1.1 实体及其联系

现实世界中的各类物质经过人们认识、选择、分类、命名后形成有意义的信息，并用数据表达出来。对这些数据进行管理，也即间接地管理了现实世界。这些数据也需要存储，存放数据的仓库称为数据库。数据的形成范围是相对的，可以是一所学校、一个企业、一个商店，或者多所学校、多个相关企业等。将

现实世界中的物体经抽象整理后送到计算机进行管理和控制要经过现实、信息、数据等三个阶段，如图 1.1 所示。

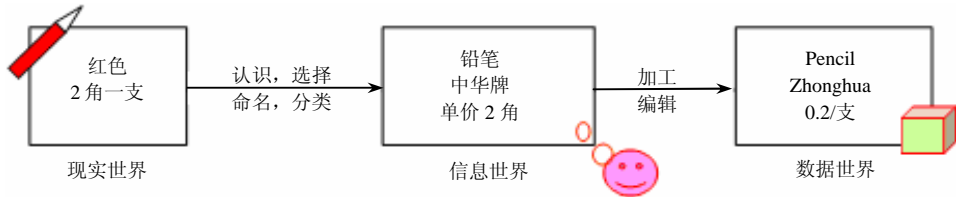


图 1.1 数据处理的三个阶段

为了达到数据的标准处理目标，每一阶段都必须对不同的处理对象规范化。相应于现实、信息、数据三个阶段的规范处理对象如图 1.2 所示。

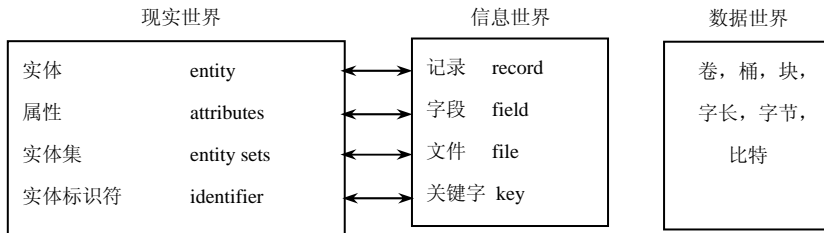


图 1.2 现实、信息、数据三个阶段的规范处理对象

客观存在并可相互区别的事物称为实体(entity)，实体可以是具体的，如一辆汽车、一位学生、一个城市，也可以是抽象的，如社会主义、资本主义等。实体具有的某一特性称为属性(attributes)。例如，学生的属性有：姓名、性别、生日等。书的属性可以有：书名、单价、作者名等。性质相同的同类实体的集合称为实体集(entity sets)，如一个班的全体同学、一个工厂生产的所有产品。实体集中的每一个具体实体用实体标识符来区分。例如，用学号作为学生的区分属性，产品用产品编号来区分。图 1.3 表示了华中科技大学计算机学院 2005 级学生实体集。

实体	型	学号	姓名	性别	生日	型	实体集
		05101	张一	男	1981		
	值	05106	王飞	男	1982	值:	
		05344	李力	男	1980		
		05655	余兰	女	1983		
属性: 型、值					文件		
实体标识符							

图 1.3 实体集、实体和属性

每一个实体由实体型和值组成。例如，学生实体的型是学号、姓名、性别、

生日, 由其决定的值是 05101、张一、男、1981。实体集的类型也是实体的型。实体集的值是相应于该实体集型的全体值, 在信息世界中也是一个数据文件。属性也有型和值之分, 属性学号为型, 其值有 05101、05106 等。属性的取值范围也是该属性的定义域。实体标识符的型和值与属性相似, 不同的是, 实体标识符可以是一个属性, 也可以是几个属性。其值是唯一区别于其他实体的实体标识符值。

现实世界中的实体在信息世界中被称为一个记录, 同理, 属性被称为字段, 实体集被称为文件, 实体标识符被称为关键字。

现实世界中的各种事物互相之间都有着千丝万缕的牵连, 存在于现实中的实体也不是孤立的, 经归纳和规范, 一类是一个实体内部的诸属性之间存在多种联系, 如实体(学号, 姓名, 性别, …), 一个学号仅对应一个姓名, 一个性别对应多个学号等。另一类是实体集之间, 例如学生实体集和图书实体集, 学生实体集中的一名学生可以借图书实体集中的多本图书, 即对应多个个体。

实体间的上述两类联系连接规律经抽象后分为 1 对 1、1 对多、多对多等几种类型。

**【定义 1.1】** 1 对 1 联系。

设实体集  $A$ 、 $B$ , 若  $A$  中的每一个实体  $a_i(i=0, 1, 2, \dots, n)$  至多与  $B$  中的某一个实体  $b_k(k=0, 1, 2, \dots, m)$  有联系, 如果  $k=0$ , 表示  $a_i$  与  $B$  中无任何联系,  $B$  中的每一个实体  $b_k(1 \leq k \leq m)$  至多与  $A$  中的某一个实体  $a_i$  有联系, 同理, 此时  $i$  可以为 0, 称  $A$  与  $B$  之间具有 1 对 1 的联系, 如图 1.4(a)所示。

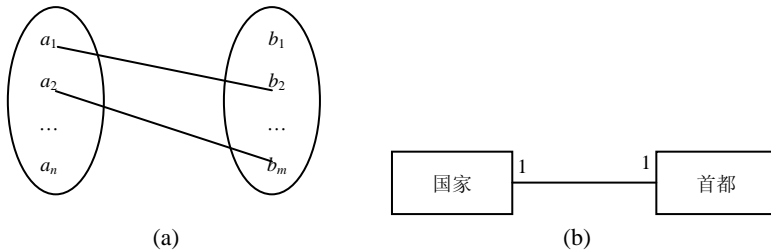


图 1.4 1 对 1 联系

例如, 实体集  $A$ 、 $B$  分别为国家集、首都集。对国家集中的每一个实体, 某国家仅拥有一座城市为首都, 反之, 首都集中每一座城市仅属于一个国家, 所以集合国家、首都之间是 1 对 1 联系, 如图 1.4(b)所示。又如, 在一个实体集内, 如实体集学生具有属性学号、姓名、性别、身份证号等, 属性学号仅对应一个身份证号, 反之亦然, 所以属性学号与身份证号之间是 1 对 1 联系。

**【定义 1.2】** 1 对多联系。

设实体集  $A$ 、 $B$ , 若  $A$  中的每一个实体  $a_i(i=0, 1, 2, \dots, n)$  至多与  $B$  中

的任意个实体  $b_k(k=0, 1, 2, \dots, m)$  有联系,  $B$  中的每一个实体  $b_k(1 \leq k \leq m)$  至多与  $A$  中的某一个实体  $a_i$  有联系, 称  $A$  与  $B$  之间具有 1 对多的联系。如图 1.5(a) 所示。

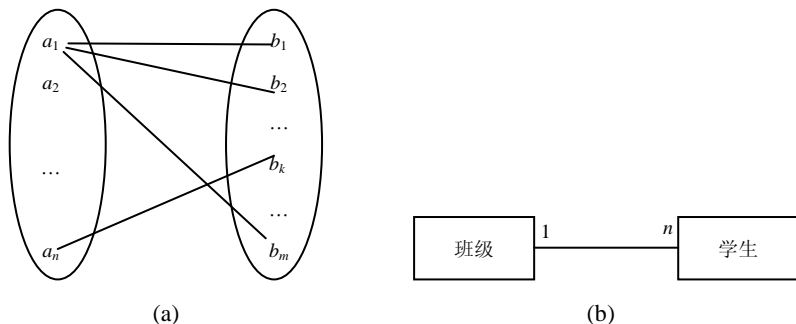


图 1.5 1 对多联系

例如, 若实体集分别为班级、学生, 班级集中的每一个班有多名学生, 学生集中的每一名学生仅属于一个班, 所以集合班级与学生之间是 1 对多联系, 如图 1.5(b) 所示。又如, 在一个实体集内, 如上例的实体集学生(学号, 姓名, 性别, 身份证号, ...)属性性别对应多个学号, 反之, 每一个学号仅对应一个性别。所以属性学号与性别之间是 1 对多联系。

如果交换具有 1 对多联系的两个实体集的位置, 1 对多联系类型也可以用来定义多对 1 联系。

**【定义 1.3】** 多对多联系。

设实体集  $A$ 、 $B$ , 若  $A$  中的每一个实体  $a_i(i=0, 1, 2, \dots, n)$  与  $B$  中的任意个实体  $b_k(k=0, 1, 2, \dots, m)$  有联系,  $B$  中的每一个实体  $b_k(1 \leq k \leq m)$  与  $A$  中的任意个实体有联系, 称  $A$ 、 $B$  之间具有多对多的联系。如图 1.6(a) 所示。

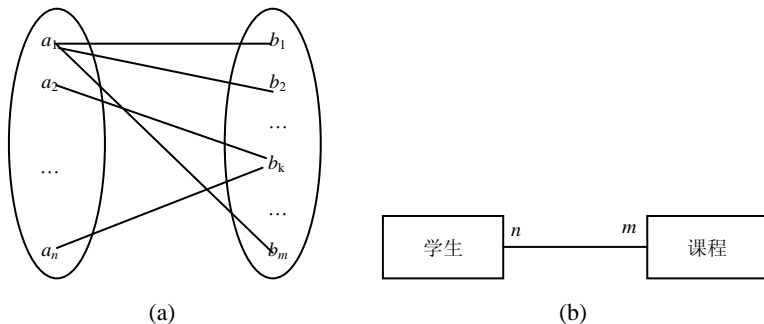


图 1.6 多对多联系

例如, 若实体集分别为学生、课程, 对学生集中的每一个实体学生可以选