



华夏英才基金学术文库

马宗民 严丽 著

模糊信息XML与数据库 建模技术研究



科学出版社



华夏英才基金学术文库

模糊信息 XML 与数据库 建模技术研究

马宗民 严 丽 著

科学出版社

北 京

内 容 简 介

本书系统介绍了模糊数据建模的基本概念、原理、方法及应用技术,主要内容包括模糊概念数据模型、模糊 XML 模型、模糊数据库模型、模糊模型间映射、模糊数据库数据操作,最后介绍了一个模糊工程信息数据库建模的应用实例。

本书对信息领域从事非传统应用的专业人员具有重要的参考价值,可作为高等学校计算机、智能科学与技术、信息系统专业研究生和高年级本科生的教材,也可供计算机及相关专业科技工作者参考。

图书在版编目(CIP)数据

模糊信息 XML 与数据库建模技术研究/马宗民,严丽著. —北京:科学出版社, 2011

(华夏英才基金学术文库)

ISBN 978-7-03-029900-0

I. 模… II. ①马…②严… III. ①模糊集理论-应用-可扩充语言, XML-程序设计-研究 ②模糊集理论-应用-数据库技术-建立模型-研究 IV. ①TP312②TP311.13

中国版本图书馆 CIP 数据核字 (2010) 第 265240 号

责任编辑:姚庆爽 孙伯元/责任校对:林青梅

责任印制:赵 博/封面设计:陈 敬

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印刷

科学出版社发行 各地新华书店经销

*

2011 年 1 月第 一 版 开本: B5 (720×1000)

2011 年 1 月第一次印刷 印张: 20 3/4

印数: 1—3 000 字数: 401 000

定价: 65.00 元

(如有印装质量问题,我社负责调换)

前 言

数据库领域的一个重要研究课题是在数据模型中融入更多的语义。数据库模型从早期的层次和网状模型、经关系模型到面向对象和对象-关系模型的进化,已经见证了这种需求。在很多实际应用中广泛存在着信息的不精确性和不确定性,经典数据模型缺乏表示和处理这类信息的能力。因此,始于 20 世纪 80 年代早期, Zadeh 的模糊逻辑理论开始被用于各类数据模型的扩展。把模糊逻辑引入数据库建模技术中的目的是要增强传统数据模型的表达能力,实现不精确和不确定信息的表示和处理,为信息技术向知识化和智能化方向转化奠定基础。有关这方面的研究已经取得了较为丰硕的研究成果,而研究内容则主要集中在被广泛使用的关系数据库及其一些相关的数据模型上。

应当指出的是,随着计算机计算能力的提高,计算机应用领域在不断地拓宽,像 CAD/CAM、多媒体和地理信息系统等一些新出现的应用给数据库技术带来了巨大的机遇和挑战,这些应用需要表示和处理复杂对象及复杂的语义关系。以面向对象数据库为例,面向对象范例被证明是特别适合表示和处理复杂对象和语义关系的数据模型。经典关系数据库模型及其模糊扩展形式不能满足带不精确和不确定信息复杂对象的表示,因此,当前很多研究工作致力于模糊概念数据模型和模糊面向对象数据库模型的研究,以实现复杂对象和不精确及不确定数据的同时处理。除了模糊关系数据库,当前对模糊概念数据模型和模糊对象数据库的研究正受到越来越多的重视,并且随着网络和互联网技术的发展,数据库已被用于分布式环境下信息系统的构建,实现多模糊数据库的集成已变得十分必要。此外,数据库通常用于 XML 数据的存储和处理,这种情况导致了在 XML 中表示模糊信息,实现把模糊 XML 模型转化成模糊数据库模型的需求。有关上述这些研究问题,当前国内外的研究工作和研究成果还很零散、不全面。国内外仅有的少数介绍模糊数据库技术的学术著作,由于出版日期较早,内容主要集中在模糊关系数据库或少量的模糊概念数据模型上,缺乏最新的研究及开发实践方面的内容。

本书的主体内容直接来源于笔者近年来在相关领域取得的系列研究成果,涉及模糊信息 XML 建模、模糊概念数据建模、模糊数据库建模以及模糊数据模型的映射与转换。很多成果已发表在包括 *Information Systems* 和 *Data & Knowledge Engineering* 在内的权威国际期刊上。本书较为系统和全面地介绍了模糊数据建模的基本概念、原理、方法及应用技术,主要内容包括模糊 ER/EER 模型、

模糊 UML 数据模型、模糊 XML 模型、模糊关系数据库、模糊面向对象数据库、模糊对象-关系数据库、模糊 ER 模型到模糊关系数据库映射、模糊 UML 数据模型到模糊面向对象数据库映射、模糊 EER 模型到模糊对象-关系数据库映射、模糊 ER/EER 和 UML 数据模型到模糊 XML 模型映射、模糊 XML 模型到模糊数据库映射, 本书最后介绍了一个模糊工程信息数据库建模的应用实例。关于模糊数据库建模, 本书除了介绍数据库模型及其相关的约束, 还给出了模糊数据库查询以及模糊关系数据库更新与集成等方面的内容。

本书是数据库与知识工程领域的学术专著, 编著的目的是通过系统介绍当前模糊信息数据建模技术在理论研究和应用方面的成果, 一方面为数据库研究人员提供国际前沿信息, 另一方面为信息领域从事非传统应用的专业人员提供技术帮助。本书可作为高等学校计算机、智能科学与技术、信息系统专业研究生和高年级本科生的教材, 也可作为计算机及相关专业科技工作者的参考书。

全书共分 12 章, 第一至四、八、九、十一、十二章由马宗民完成, 第五至七、十章由严丽完成。本书得到了笔者所在研究团队博士研究生和硕士研究生的大力协助。本书的研究工作得到国家自然科学基金(60873010 和 61073139)和教育部“新世纪优秀人才支持计划”(NCET-05-0288)的资助, 并部分得到教育部中央高校基本科研业务费专项(N090504005, N100604017 和 N090604012)的资助, 本书的出版得到了华夏英才基金资助, 在此一并表示感谢。书中参考了大量国内外模糊数据库研究领域的文献, 在此谨向文献的作者表示衷心的感谢。

由于数据库技术发展速度快、涉及面广, 同时由于时间仓促, 加之作者水平有限, 书中难免存在疏漏, 恳请读者批评指正。

作 者

2010 年 8 月于沈阳

目 录

前言

第一章 概念数据模型	1
1.1 实体-联系模型和扩展实体-联系模型	2
1.1.1 ER 模型	3
1.1.2 EER 模型	5
1.2 UML 类图模型	8
1.2.1 类	8
1.2.2 关系	9
1.3 XML 模型	10
1.3.1 XML 文档	11
1.3.2 XML DTD	11
1.3.3 XML Schema	13
1.3.4 XML 文档类型结构设计与 XML 数据库	14
第二章 逻辑数据库模型	18
2.1 关系数据库模型	18
2.1.1 属性与域	19
2.1.2 关系与元组	19
2.1.3 关键字	19
2.1.4 约束	19
2.1.5 关系代数	21
2.1.6 关系数据库设计	23
2.2 嵌套关系数据库模型	24
2.3 面向对象数据库模型	25
2.3.1 对象和标识符	26
2.3.2 类和实例	26
2.3.3 类结构层次和继承	26
2.4 对象-关系数据库模型	27
2.4.1 模型定义	27
2.4.2 模型的基本概念	28

第三章 模糊集和可能性分布理论	31
3.1 引言	31
3.2 信息的不精确和不确定性	33
3.2.1 语义与分类	33
3.2.2 表示方法	33
3.3 模糊集与可能性分布	35
3.3.1 形式化定义	35
3.3.2 模糊集相关基本概念	36
3.3.3 Zadeh 扩展原理	37
3.3.4 模糊集操作	38
3.4 小结	40
第四章 模糊 ER 模型和模糊 EER 模型	42
4.1 引言	42
4.2 模糊 ER 模型	43
4.2.1 实体、关系及属性中的三个级别模糊性	43
4.2.2 联系和约束	46
4.3 模糊 EER 模型	48
4.3.1 模糊属性	48
4.3.2 模糊实体和联系	49
4.3.3 模糊概化和特化	50
4.3.4 模糊范畴	51
4.3.5 模糊聚集	52
4.4 小结	53
第五章 模糊 UML 数据模型	55
5.1 引言	55
5.2 UML 数据模型与模糊数据建模	56
5.3 UML 中的模糊类	56
5.4 UML 中的模糊概化	59
5.5 UML 中的模糊聚集	62
5.6 UML 中的模糊关联	64
5.7 UML 中的模糊依赖	66
5.8 小结	68
第六章 模糊 XML 模型	69
6.1 引言	69
6.2 XML 文档中信息的模糊性	70

6.3 模糊 XML 的表示模型	73
6.3.1 XML 文档中模糊数据的表示	74
6.3.2 XML DTD 的修改	76
6.3.3 模糊 XML Schema 的定义	80
6.4 模糊 XML 代数操作	86
6.4.1 集合并运算	86
6.4.2 集合差运算	88
6.4.3 笛卡儿积	89
6.4.4 选择	89
6.4.5 投影	89
6.4.6 其他的代数操作	90
6.4.7 代数操作的性质及代数等价变换规则	92
6.4.8 XQuery 的转换	94
6.5 模糊 XML DTD 与模糊 UML 模型间转换	98
6.5.1 基于 UML 的模糊 XML DTD 概念设计	98
6.5.2 模糊 XML DTD 到 UML 转换	102
6.6 小结	106
第七章 模糊关系数据库	107
7.1 引言	107
7.2 模糊关系数据库模型	108
7.3 模糊数据语义度量及数据冗余	108
7.3.1 现有的方法	109
7.3.2 模糊数据之间的语义关系	110
7.3.3 语义度量评估	111
7.3.4 模糊数据冗余与消除	112
7.4 模糊数据完整性约束	114
7.4.1 模糊函数依赖	115
7.4.2 模糊多值依赖	117
7.4.3 模糊数据依赖的推理规则	118
7.5 模糊关系代数操作	120
7.5.1 模糊关系代数	121
7.5.2 模糊关系代数性质	127
7.6 SQL 柔性查询	129
7.6.1 模糊关系数据库查询	131
7.6.2 经典关系数据库的模糊查询	133

7.7 关系数据库更新	141
7.7.1 插入操作	142
7.7.2 删除操作	147
7.7.3 修改操作	148
7.8 多模糊关系数据库集成	151
7.8.1 多数据库集成中的冲突与解决	153
7.8.2 含兼容关键字多模糊数据库集成中的冲突与解决	156
7.8.3 含不兼容关键字多模糊数据库实体标识	166
7.9 小结	171
第八章 模糊面向对象数据库	173
8.1 引言	173
8.2 模糊对象和模糊类	174
8.2.1 模糊对象	174
8.2.2 模糊类	174
8.2.3 模糊对象-类关系	175
8.3 模糊子类-超类关系	180
8.4 模糊对象的比较	182
8.4.1 同一个类中两个对象的等价	183
8.4.2 两个不同类对象的等价	184
8.5 柔性约束	185
8.5.1 约束及其分类	186
8.5.2 柔性约束比较	187
8.5.3 柔性约束上的操作	190
8.6 模糊面向对象数据库模型	194
8.7 模糊面向对象数据库的查询与操作	195
8.7.1 模糊乘操作	196
8.7.2 模糊交操作	196
8.7.3 模糊并操作	196
8.7.4 模糊查询操作	197
8.8 小结	198
第九章 模糊嵌套关系数据库和模糊对象-关系数据库	200
9.1 引言	200
9.2 模糊嵌套关系数据库	201
9.2.1 模糊嵌套关系模型	201
9.2.2 代数操作	202

9.3 模糊对象-关系数据库	206
9.3.1 模糊数据类型	206
9.3.2 模糊属性值	207
9.3.3 模糊对象-关系数据库模型实例	208
9.4 小结	209
第十章 模糊数据库概念设计	211
10.1 引言	211
10.2 基于 UML 数据模型的模糊关系数据库概念设计	211
10.2.1 模糊类的转换	212
10.2.2 模糊聚集的转换	214
10.2.3 模糊关联的转换	215
10.3 基于 UML 数据模型的模糊面向对象数据库概念设计	216
10.3.1 模糊类的转换	217
10.3.2 模糊聚集的转换	219
10.3.3 模糊关联的转换	219
10.3.4 模糊依赖的转换	221
10.4 模糊 EER 模型到模糊嵌套关系数据库的映射	222
10.4.1 模糊实体的转换	223
10.4.2 模糊联系的转换	224
10.4.3 模糊概化的转换	225
10.4.4 模糊特化的转换	227
10.4.5 模糊范畴的转换	227
10.4.6 模糊聚集的转换	228
10.5 模糊 XML 模型与模糊关系数据库间的转换	229
10.5.1 模糊 XML DTD 到模糊关系数据库的映射	230
10.5.2 模糊关系数据库到模糊 XML DTD 的映射	233
10.6 小结	239
第十一章 模糊工程数据建模语言 EXPRESS/STEP	240
11.1 引言	240
11.2 基本元素的模糊性	240
11.2.1 保留字	241
11.2.2 值符	241
11.3 EXPRESS 的模糊数据类型建模	244
11.3.1 伪类型	245
11.3.2 简单数据类型	246

11.3.3 汇集数据类型	249
11.3.4 枚举类型	251
11.3.5 选择类型	252
11.4 模糊声明	253
11.4.1 常量	253
11.4.2 类型	254
11.4.3 实体	254
11.4.4 算法	259
11.4.5 规则	260
11.4.6 模式	260
11.5 带有模糊信息的表达式	261
11.5.1 算术操作符	261
11.5.2 关系操作符	263
11.5.3 逻辑操作符	267
11.5.4 串和二进制操作符	268
11.5.5 聚集操作符	268
11.6 接口与可执行语句的模糊扩展	271
11.6.1 接口说明	271
11.6.2 可执行语句	271
11.7 模糊 EXPRESS-G 模型	274
11.7.1 实体级模型中的模糊性	274
11.7.2 模式级模型的模糊性	277
11.8 模糊 EXPRESS 的适用性	277
11.9 模糊 EXPRESS 模型的概念设计	279
11.10 小结	281
第十二章 模糊 EXPRESS 数据模型的面向对象数据库实现	283
12.1 引言	283
12.2 STEP 及其实现	284
12.2.1 STEP 的结构与信息模型	284
12.2.2 STEP 实现方法和层次	288
12.2.3 STEP 数据库实现过程	290
12.3 模糊 EXPRESS 模型到模糊面向对象数据库的转换	291
12.3.1 面向对象数据库对 EXPRESS 数据模型的支持	292
12.3.2 形式化转换	293
12.4 操作模糊 EXPRESS 模型的功能需求	295

12.4.1 数据查询需求	296
12.4.2 数据更新需求	297
12.4.3 结构查询要求	298
12.4.4 数据验证需求	299
12.5 模糊面向对象数据库中的 SDAI 实现	300
12.5.1 数据查询	300
12.5.2 数据创建	301
12.5.3 数据删除	301
12.5.4 实体结构查询	302
12.6 小结	302
参考文献	303

第一章 概念数据模型

为了实现对现实世界的抽象和表示，需要使用数据模型（data model）。根据数据抽象的级别，可以把数据模型分成四种形式，即概念数据模型、逻辑数据模型、外部数据模型和内部数据模型。数据库是当今信息领域使用最为广泛的数据模型，从数据库的角度来看，概念数据模型用来表达用户需求观点下的数据库全局逻辑结构；逻辑数据模型用来表达计算机实现观点下的数据库全局逻辑结构；外部数据模型用来表达用户使用观点下的数据库局部逻辑结构；内部数据模型用来表达数据库的物理结构。概念数据模型也被称为语义数据模型，逻辑数据模型也被称为逻辑数据库模型，这里把概念数据模型和逻辑数据模型统称为数据库模型。在这四类数据模型中，概念数据模型和逻辑数据模型与面向应用数据库的设计最为相关，并且概念数据模型的抽象级别最高，逻辑数据模型的抽象级别较低，数据库的设计通常从概念数据模型入手，之后再吧概念数据模型映射为逻辑数据模型。我们把与概念数据模型、逻辑数据模型，及这两种数据模型内部及其之间的转换等相关的处理统称为数据库建模（database modeling）。

传统上采用的概念数据模型主要是实体-联系（entity-relationship，ER）模型，当前 UML 类图也广泛用做概念数据模型。逻辑数据库模型方面，早期采用的是层次数据库模型（hierarchical database model）和网状数据库模型（network database model），之后是广泛使用的关系数据库模型，此外还有面向对象数据库模型（object-oriented database model）和对象-关系数据库模型（object-relational database model）。近年来，随着 Web 技术的发展和普遍使用，出现了大量 Web 数据以及 Web 数据管理的新内容。作为 Web 时代数据管理的两个重要技术方法，XML 和数据库模型之间存在着密切的关系，实现两者的集成已成为必要，并成为 Web 领域和数据库领域重要的研究内容。一方面，包括关系数据库、面向对象数据库和对象-关系数据库在内的逻辑数据库模型可用于 XML 数据的存储，并进而实现基于数据库的 XML 数据查询与更新处理；另一方面，逻辑数据库模型可通过 XML 的形式在 Web 上实现发布与共享，XML 和数据库之间可实现相互映射与转换。此外，由于 XML 缺乏充分表示现实世界数据以及它们之间复杂内在语义关系的能力，因此可以使用概念数据模型首先开发出概念数据模式，之后再吧这样的概念模式转换成 XML。

从数据库建模的角度来看，传统数据库建模考虑的是概念数据模型、逻辑数据库模型以及模型间的映射转化。当 XML 这种新的数据管理形式出现以后，在

Web 环境下, 数据库建模的覆盖范围和研究内容已发生了变化。体现上述 XML、逻辑数据库模型以及概念数据模型三者之间关系的数据库建模形式如图 1-1 所示。

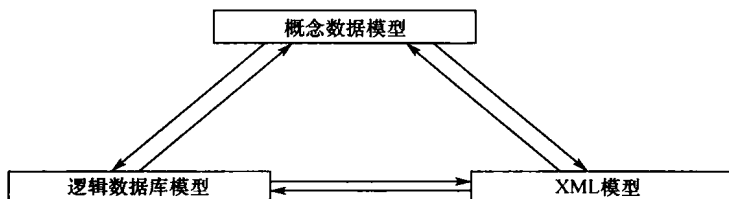


图 1-1 Web 框架下的数据库建模

逻辑数据库模型将在下面第二章中进行讨论, 本章将主要介绍概念数据模型。我们将简单地介绍实体-关系数据模型、增强（扩展）的实体-联系模型 EER 以及统一建模语言（unified modeling language, UML）的类图模型。此外, 随着基于 Web 应用的普及, 出现了 Web 上数据交换和数据共享的需求, 而可扩展标记语言（eXtensible Markup Language, XML）为 Web 上数据交换提供了 Web 友善且易于理解的语法, 并对 Web 上数据的定义和共享产生影响（Seligman et al., 2001）。因此, 本章还将简要介绍 XML 的相关概念。

1.1 实体-联系模型和扩展实体-联系模型

1976 年, Chen 教授在 *ACM Transactions on Database Systems* 杂志上发表了其创新性文章 “The entity-relationship model: Toward a unified view of data”（Chen, 1976）, 从而创建了实体-联系模型。ER 模型提出的初衷是为了给数据库提供一个统一的数据模型, 但是在数据库语义建模和概念设计方面取得了巨大成功。ER 模型的提出为数据库设计和信息系统分析, 提供了一个十分有效的工具。

需要指出的是, 含有实体、联系和属性的 ER 模型虽然能够抽象和表示很多现实世界的状态, 并且已经在数据库设计及信息系统的分析中得到了广泛的应用, 但是在一些复杂对象建模和复杂关系表达方面, ER 模型却存在着语义表达能力不足的限制。为了更准确更全面地表示现实世界应用, 很多研究者把一些新的概念引入到 ER 模型中（Gegolla et al., 1991; Elmasri et al., 1985; Dos Santos et al., 1979; Scheuermann et al., 1979）, 以提高其表达能力和可用性, 从而形成了增强（扩展）的实体-联系（enhanced ER or extended ER, EER）模型。

1.1.1 ER 模型

由 Chen (1976) 提出的 ER 模型通过使用实体、联系和属性三个概念来描述现实世界的语义。由 ER 模型描述的数据模型一般用 ER 图进行表示。

1. 实体

实体 (entity) 是一个可辨识、易于理解的具体事物或者抽象概念, 具有相同特征的实体的集合称为实体集 (entity set)。一个命名的实体集可视为一个实体类型 (entity type) 的描述, 实体集中的每个实体都是该实体类型的一个实例。举例来说, “轿车” 可看做是一个实体集, 对轿车特征的描述则属于实体类型, 而一个实际型号的轿车, 比如 “Honda Civic DX”, 则是一个轿车实体。

通常, 在不致混淆的情况下, 实体、实体集和实体类型被简称为实体。因此, 对于实体所代表的实际含义, 应根据上下文进行理解。

2. 属性和关键字

实体的特征称为实体的属性 (attribute)。实体的每个属性都有一个取值范围, 称做值集 (value set), 值集实际上与关系数据库中的属性域相同。从属性值集的元素形式上来看, 实体中的属性不仅可以是简单属性, 此时值集中的元素为一个值, 还可以是复杂属性, 此时值集中的元素为多个值, 这样的复杂属性被称为复合属性。来看一个例子, 对于一个人的 “名字” 和 “通信地址” 两个属性, 它们就分别是简单属性和复杂属性。此外, 从属性取值的数量上来看, 属性可以是单值属性, 也可以是多值属性。例如, 一个人的属性 “年龄” 和 “电子邮件地址” 分别是单值属性和多值属性。

对于实体包含的若干个属性来说, 属性对实体识别所发挥的作用有所不同。把可以区分出一个实体的最小属性或属性集称做该实体的关键字, 一个实体可能有多个关键字, 而把其中的一个关键字指定为主关键字, 简称为主键。

3. 联系

现实世界中的实体不是孤立的, 实体间存在着一定的联系 (relationship)。设 $e_1, e_2, \dots, e_n$ 为 n 个实体, 它们之间的联系表示为 $r(e_1, e_2, \dots, e_n)$ 。其中, r 是联系名。如果 $n=2$, 则上述联系是二元联系, 如果 $n>2$, 则上述联系是多元联系。相同类型的联系所组成的集合称为联系集 (relationship set), 一个联系集可以看做是实体集之间的联系。形式上, 设 $E_1, E_2, \dots, E_n$ 为 n 个实体集, 它们之间的联系集表示为 $R(E_1, E_2, \dots, E_n)$ 。联系集是实体联系的类型描述, 具体实体之间的一个联系是相应联系集的一个实例, 并且相同的实体集

合可以在一个联系集中出现多次。一个命名的联系集可看做是一个联系类型 (relationship type) 的描述。与实体的情况类似, 在不引起混淆的情况下, 联系、联系集和联系类型被简称为联系, 对联系所代表实际含义的理解也需要依据上下文。

根据参与联系实体集的个数, 可以把联系区分为二元联系和多元联系 (例如三元联系, 四元联系等)。同一个实体集内实体之间的联系被称为一元联系。在 ER 模型中, 根据参与一个联系中实体的数目, 一个二元联系可以是一对一、一对多或多对多的联系, 分别表示一个实体集中的一个实体只与另一个实体集中的一个实体联系, 一个实体集中的一个实体与另一个实体集中的多个实体联系, 以及一个实体集中的多个实体与另一个实体集中的多个实体联系。这种联系分类同样适用于 n 元联系。上述实体间联系所体现出的约束被称为基数比约束 (cardinality ratio constraint)。在 ER 模型中, 除了基数比约束之外, 联系中还存在着一种重要的语义约束, 称为参与约束 (participation constraint), 它规定了实体参与联系的方式。参与度 (participation degree) 的概念用于表达联系中相关实体集里参与联系实体的最大和最小数目, 形式化地记为 “min, max”。其中, $\max \geq \min \geq 0$ 且 $\max \geq 1$ 。当 $\min = 0$ 时, 联系中实体参与的方式称为部分参与 (partial participation), 否则, 称为完全参与 (total participation)。ER 模型中的基数比约束和参与约束有时也看做结构化约束。

在现实世界中, 存在一种被称为标识联系 (identifying relationship) 的特殊联系, 用于表示实体间的拥有联系 (ownership)。标识联系有以下两个特征:

(1) 被另一个实体所拥有的实体依赖于其拥有者实体, 这样的实体不能够单独存在, 必须完全地参与联系。

(2) 被另一个实体所拥有的实体不能是其自身的实体键。

由于被另一个实体所拥有的实体具有这样的特征, 因此被称为弱实体 (weak entity)。弱实体可以看做是一个实体, 也可以看做是其拥有者实体的一个复杂属性。

与实体类似, ER 模型中的联系包含属性, 称为联系属性。

4. ER 图

ER 模型中实体、属性和联系可以用图形化符号进行表示, 由此构成了 ER 图 (ER diagram) 数据模型, 简称为 ER 图。

在 ER 图中, 一个实体集用一个矩形来表示, 一个弱实体集用一个双线的矩形表示, 一个联系用一个菱形表示, 一个标识联系用一个双线的菱形表示。实体的名字和联系的名字要求写在相应的图形框内。矩形和菱形用连线连接起来, 并且可以标记出基数比。如果连线是单线, 则表明相应的实体是部分参与; 如果连

线是双线，则表明相应的实体是全部参与。参与度可以在需要的时候给出。

ER 图中的属性用圆圈或椭圆表示，属性名字写在其中，如果属性是多值属性，则用双线圆圈或椭圆来表示。在 ER 图中，属性符号需要用线同相应的实体集或联系集连接起来。如果一个属性是实体的关键字，或者是实体关键字的一部分，则需要在 ER 图中把这样的属性标识出来，采用的方法是在属性名下面加上下划线，或者在连接属性的边上加上一个与之垂直的短线。如果一个属性是复杂的属性，则在 ER 图中使用一个树结构进行表示。

图 1-2 给出了 ER 图中使用的符号。

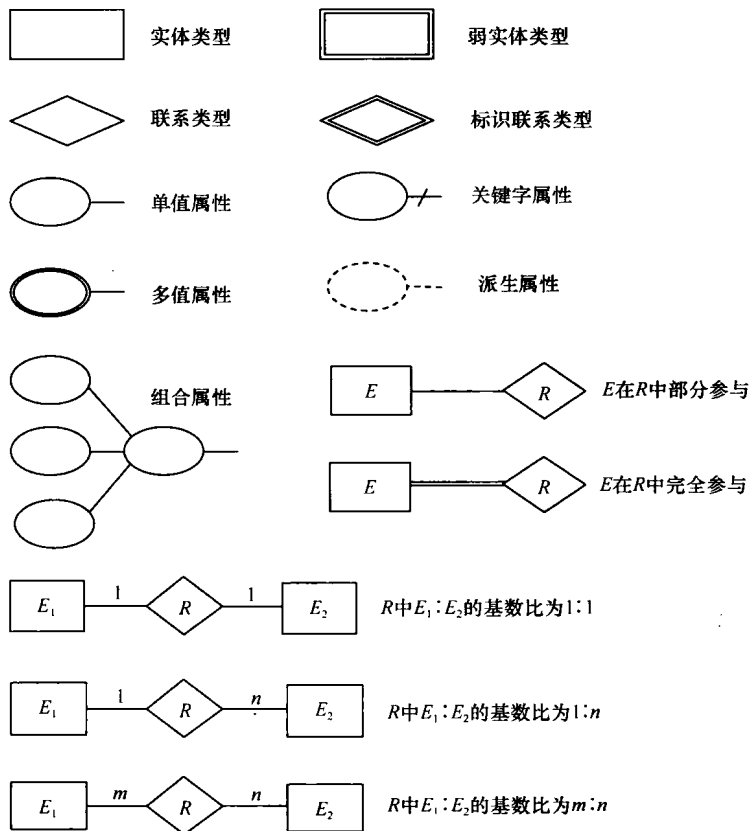


图 1-2 ER 图的记法

1.1.2 EER 模型

基于实体、联系和属性的 ER 模型被称做基本 ER 模型。为了能够建模 CAD/CAM、CASE、GIS 等应用中的复杂语义和复杂关系，一些新的概念已经引入到了 ER 模型中，于是形成了扩展实体-联系（或叫增强实体-联系）模型，