

Chapter7 数据库设计

7.1 数据库设计概述

数据库设计

- 对于给定的软、硬件环境，针对现实问题，设计一个较优的数据模型，建立DB结构和DB应用系统。
- 对于一个给定的应用环境，**提供一个确定最优数据模型与处理模式的逻辑设计**，以及一个确定数据库存储结构与存取方法的物理设计，**建立起既能反映现实世界信息和信息联系**，满足用户数据要求和加工要求，又能被**某个数据库管理系统所接受**，同时能实现系统目标，并有效存取数据的数据库。

7.1.1 软件生存期

1. 软件生存期：指从软件的规划、研制、实现、投入运行 后的维护，直到它被新的软件所取代而停止使用的整个期间。
2. 六个阶段：规划阶段、需求分析阶段、设计阶段、程序编制阶段、调试阶段、运行维护阶段

7.1.2 数据库系统生存期

1. 数据库应用系统：以**数据库为基础的信息系统**。具有对信息的**采集、组织、加工、抽取和传播**等功能。
2. 数据库工程：数据库应用系统的开发。
3. 数据库系统生存期：数据库应用系统从**开始规划、设计、实现、维护到最后被新的系统取代而停止使用的整个期间**。

4. 七个阶段

1. 规划阶段

- 必要性和可行性分析。确定数据库系统的地位，以及数据库间联系。

2. 需求分析阶段

- 分析用户需求
- 信息要求（数据库中存储哪些数据）
- 处理要求（需要的处理功能，响应时间要求，批/联机处理）
- 安全性和完整性要求

3. 概念设计阶段

- 表达用户整体要求，且独立于DBMS和硬件结构。

4. 逻辑设计阶段

- **数据库逻辑结构设计**：ER图 => DBMS的DDL转换为逻辑数据库结构
- **应用程序设计**：DBMS的DML 进行结构式的程序设计。

5. 物理设计阶段

- **物理数据库结构的选择**
- **逻辑设计中程序模块说明的精确化**（开发）。
- 成果：一个完整的，能实现的数据库结构。

6. 实现

7. 运行维护

- 收集和记录系统实际运行的数据。

7.1.3 数据库设计的具体步骤

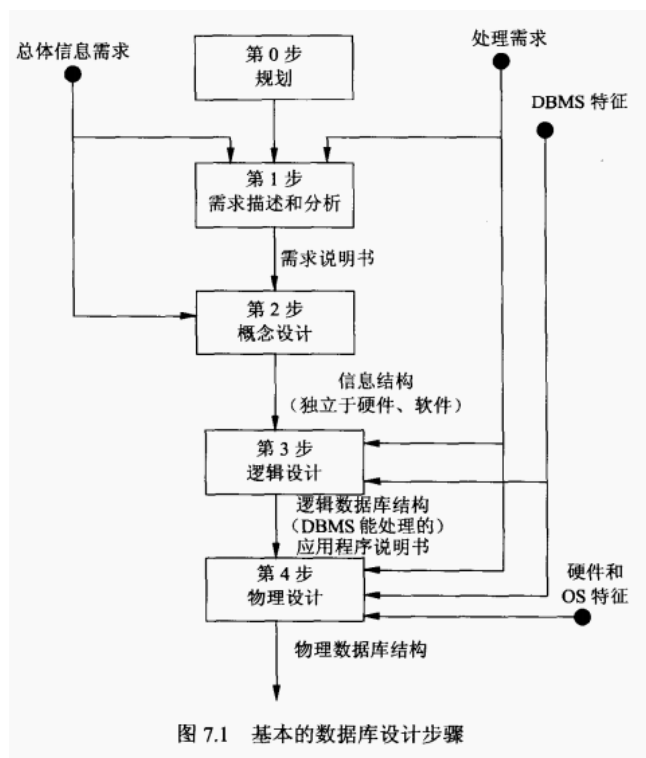


图 7.1 基本的数据库设计步骤

输入：

1. 总体信息需求：数据库的目标说明，数据元素的定义，数据在企业组织中的使用描述。
2. 处理需求：每个应用需要的数据项，数据量，应用执行的频率。
3. DBMS特征：DBMS说明+参数；DBMS模式+子模式，程序语法规则。
4. 硬件和OS特征：DBMS和OS访问方法特有的内容。

输出：说明书

1. **完整的数据库结构**：逻辑结构+物理结构。
2. 基于数据库结构和处理需求的**应用程序的设计原则**。

7.2 规划

任务：建立数据库的必要性及可行性分析，确定数据库系统在组织中和信息系统中的地位，以及各个数据库之间的联系。

输出：**可行性分析报告 + 数据库系统规划纲要**

后者包括：信息范围、信息来源、人力资源、设备资源、软件及支持工具资源、开发成本估算、开发进度计划、现行系统向新系统过渡计划等。

7.3 需求分析

7.3.1 需求描述与分析

1. 对系统的整个应用情况做**全面的详细的调查**，**确定企业组织的目标**
2. 收集**支持系统总的设计目标的基础数据**和**对这些数据的要求**，**确定用户的需求**
3. 并把这些要求写成用户和数据库设计者都能接受的文档。

7.3.2 需求分析阶段的输入和输出

1. 输入：总体信息需求（数据本质和概念上联系），处理需求（数据处理）
2. 输出：需求说明书（数据流图 + 数据字典）

7.3.3 需求分析的步骤

步骤

1. 分析用户活动，产生用户活动图（即**用户的业务流程图**）；
2. 确定系统范围，产生系统范围图（即确定**人机界面**）；
3. 分析用户活动所涉及的数据，产生数据流图（**数据的流向及加工**）；
4. 分析系统数据，产生数据字典。

数据流图：→ 数据流；○ 加工 / 处理；= 文件；□ 外部实体

数据字典：

- 定义：对系统中数据的详尽描述，提供了对数据库数据描述的集中管理。
- 内容：数据项，数据结构，数据流，数据存储，加工过程
- 功能：存储和检索各种数据描述，为DBA提供有关的报告。
- 在需求分析阶段建立，并在数据库设计过程中不断改进、充实和完善。

7.4 概念设计

目标：**产生**反映企业组织信息需求的数据库概念结构（**概念模式**）。

概念模式独立于数据库逻辑结构，也独立于支持数据库的DBMS。

7.4.1 概念设计的必要性

1. 各阶段的任务相对单一化，设计复杂程度大大降低，便于组织管理。
2. 不受特定的DBMS的限制，也独立于存储安排和效率方面的考虑，比逻辑模式更稳定。
3. 不含具体DBMS附加技术细节，用户理解容易，可以更准确反应用户的信息需求。

7.4.2 概念模型

1. ~是对现实世界的抽象和概括。
2. ~应简洁，明晰，独立于机器，容易理解。
3. ~应易于变动。
4. ~应很容易想关系，层次，网状等各种数据模型转换。

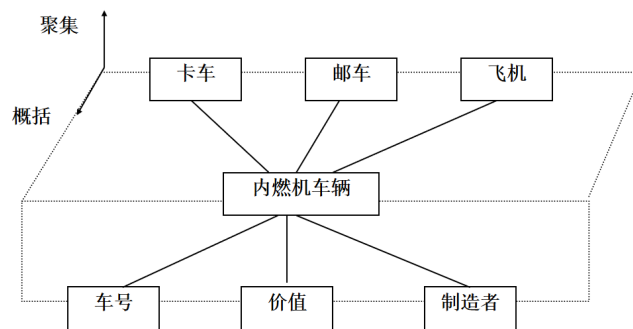
7.4.3 概念设计的主要步骤

1. 进行数据抽象，设计局部概念模式；（“聚集” + “概括”）
2. 将局部概念模式综合成全局概念模式；（解决冲突：消除冗余，统一命名）
3. 评审。（用户评审 + DBA和应用开发人员评审）

7.4.4 数据抽象

1. 抽象：对实际的人、物、事或概念的人为处理，抽取人们关心的共同特性，忽略非本质的细节，将这些特性用各种概念精确地加以描述，这些概念组成了某种模型。
2. 两种形式：
 - 抽象对象：系统状态的抽象

- 抽象运算：系统转换的抽象
- 3. 聚集：笛卡尔积，形成对象之间的一个联系对象。
 - 聚集层次表示：“是.....的一部分” (is part of) 的关系。
- 4. 概括：从一类对象形成一个对象。
 - 概括层次表示：“是.....一种”(is a)的关系。
- 5. 数据抽象层次
 - 每个对象既可以是聚集对象，也可以是概括对象。
 - 反复进行数据抽象，形成层次关系。



7.4.5 ER模型的操作

ER模型的操作：在利用ER模型进行数据库概念设计的过程中，常常对ER图进行的种种变换。（**实体类型、联系类型和属性的分裂、合并和增删……**）

1. 实体类型的分裂

1. 垂直分割：把一个实体类型的属性分成若干组，然后按组形成若干实体类型。**键必须在所有实体中体现。**

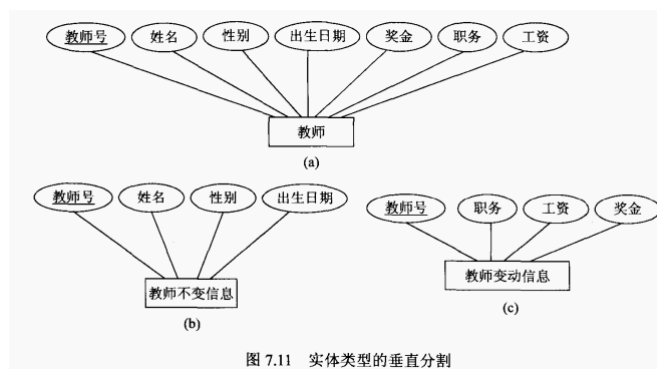


图 7.11 实体类型的垂直分割

- 水平分割：分裂成互补相交的子类。 e.g. 教师 => 男教师 + 女教师

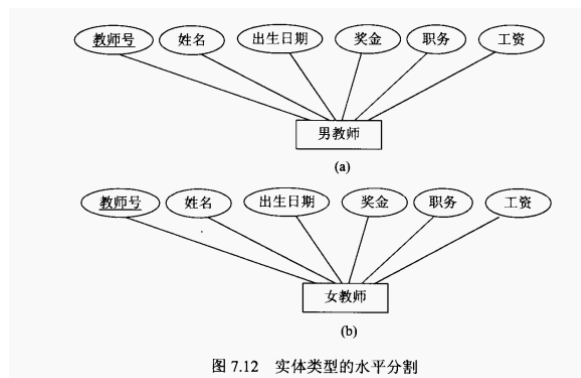


图 7.12 实体类型的水平分割

2. 实体类型的合并

1. 水平合并，垂直合并。
2. 是否会产生新联系视情况而定。

3. 联系类型的分裂

1. 一个联系类型可分裂成几个新联系类型。新联系类型可能和原联系类型不同。

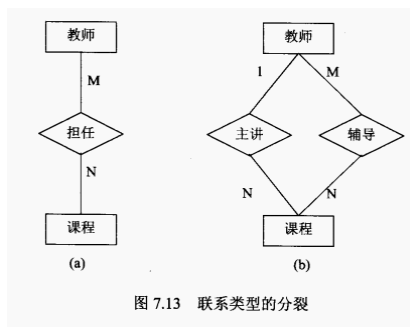


图 7.13 联系类型的分裂

4. 联系类型的合并

1. 合并的联系类型必须是定义在相同的实体类型组合中，否则，不合法。

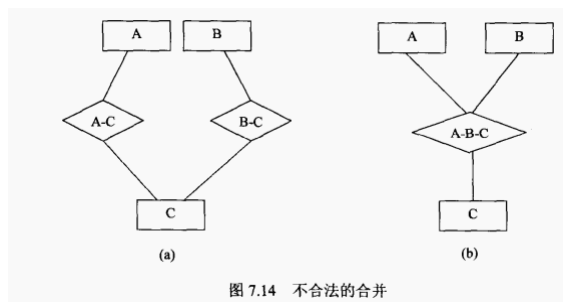


图 7.14 不合法的合并

7.4.6 采用ER方法的数据概念设计

第一步：设计局部ER模式：

1. 确定局部结构范围划分；
2. 实体定义；
3. 联系定义。

第二步：设计全局ER模式：

1. 确定公共实体类型；（同名实体类型；相同主键的实体）
2. 局部ER模式的合并；
3. 消除冲突（属性冲突，结构冲突，命名冲突）

第三步：全局ER模式的优化原则：

1. 相关实体类型的合并：1:1的实体合并，部分相同键的实体
2. 冗余属性的消除；
3. 冗余联系的消除。

7.5 逻辑设计

目的：把概念设计阶段设计好的全局ER模式转换成与选用的具体机器上的DBMS所支持的数据模型相符合的逻辑结构（包括数据库模式和外模式）。

7.5.1 逻辑设计环境

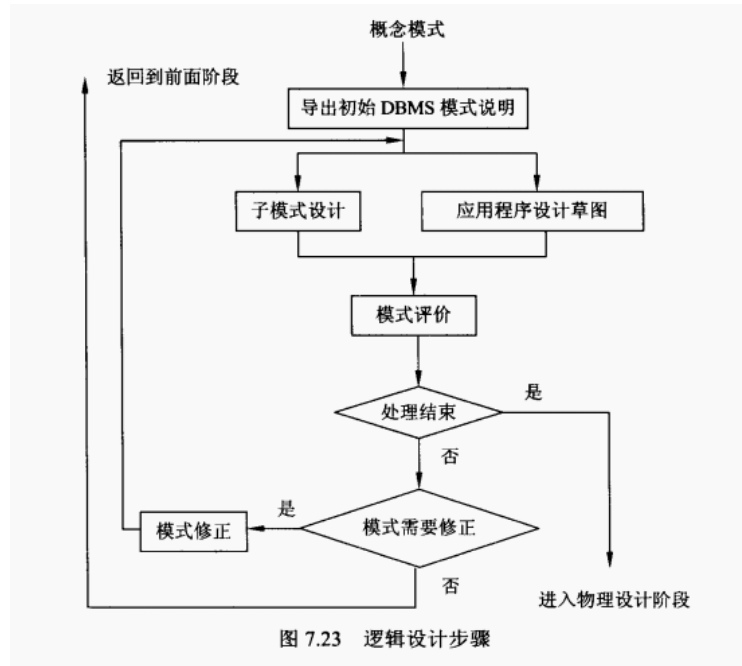
1. 输入

1. 独立于DBMS的概念模式：局部，全局概念模式（概念设计阶段）
2. 处理需求：业务活动分析结果（需求分析阶段）
3. 约束条件
4. DBMS特性

2. 输出

1. DBMS可处理的模式：说明
2. 子模式
3. 应用程序设计指南
4. 物理设计指南：完全文档化的模式和子模式。

7.5.2 逻辑设计的步骤

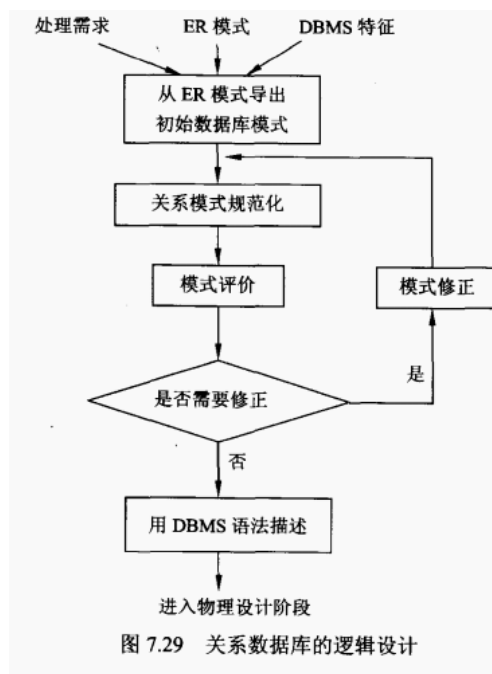


7.5.3 ER模型象关系模型的转换

ER模型转换为关系模型的一般规则：

1. 实体类型的转换：1个实体1张表。
2. 联系类型的转换
 1. 1:1 任意一端加入另一端的主键，作为外键。
 2. 1: N N端加入1端主键，作为外键。
 3. 弱实体 1: N, N端为弱实体 N端中加入1端主键，作为外键。且N端主键= 1端主键 + N端外键
 4. M: N 新增一张表，主键为两端的外键构成。
3. 超类和子类的转换规则：子类 = 父类主键 + 新增属性

7.5.4 关系数据库的逻辑设计



1. 导出初始关系模式
2. 运用模式设计理论，对初始关系模式进行规范化处理
 1. 确定规范级别，3NF / BCNF
 2. 实施规范化处理
3. 模式评价：功能 + 性能
4. 模式修正

7.6 物理设计

1. 物理设计：对于给定的基本数据模型选取一个最适合应用环境的物理结构。
2. 数据库的物理结构：数据库的存储记录格式，存储记录安排，存取方法。
3. 物理设计
 1. **存储记录结构设计**：记录的组成、数据项的类型、长度，以及逻辑记录到存储记录的映射。
 2. **确定数据存放位置**：把经常同时被访问的数据组合在一起。
 3. **存取方法的设计**：主存取路径，辅存取路径。
 4. **完整性和安全性考虑**
 5. **程序设计**

7.7 数据库的实现

1. 数据库的实现阶段：根据逻辑设计和物理设计的结果，在计算机系统上建立起实际数据库结构、装入数据、测试和试运行的过程。
2. 实现阶段：
 1. 建立实际数据库结构
 2. 装入试验数据对应用程序进行调试
 3. 装入实际数据，进入试运行状态

7.8 数据库的运行和维护

1. 维护数据库的安全性与完整性
2. 监测并改善数据库运行性能
3. 根据用户要求对数据库现有功能进行扩充

4. 及时改正运行中发现的系统错误