



第七章 数据库设计(续-1)



第七章 数据库设计

7.1 数据库设计概述

7.2 需求分析

7.3 概念结构设计

7.4 逻辑结构设计

7.5 数据库的物理设计

7.6 数据库实施

7.7 数据库运行与维护

7.8 小结



7.4 逻辑结构设计

□ 逻辑结构设计任务

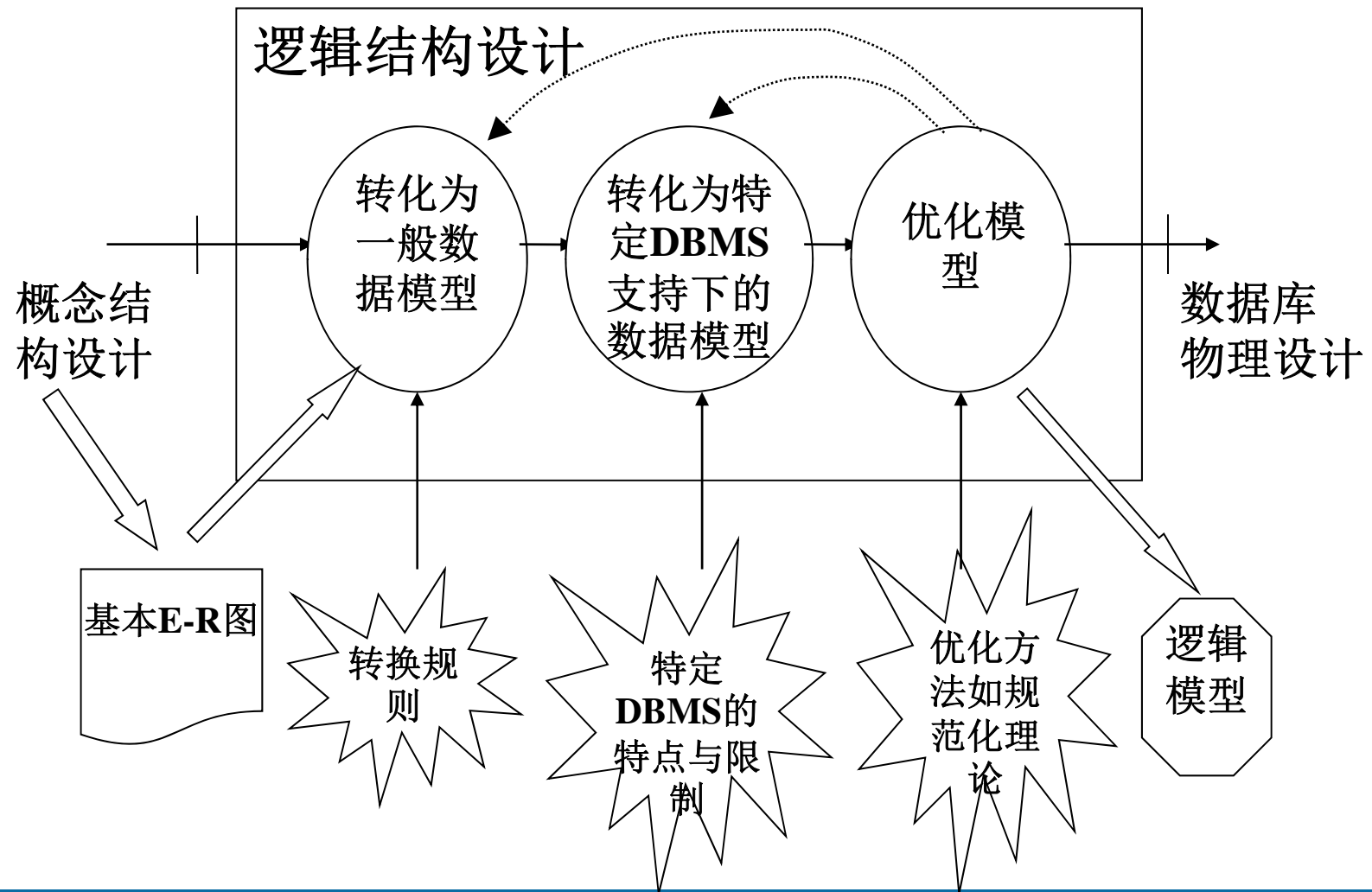
- ◆ 概念结构是各种数据模型的基础
- ◆ 为了能够用某一**DBMS**实现用户需求，还必须将概念结构进一步转化为相应的数据模型，这正是数据库逻辑结构设计所要完成的任务。



7.4 逻辑结构设计

□ 逻辑结构设计的步骤

- ◆ 将概念结构转化为一般的关系、网状、层次模型
- ◆ 将转化来的关系、网状、层次模型向特定**DBMS**支持下的数据模型转换
- ◆ 对数据模型进行优化





7.4 逻辑结构设计

7.4.1 E-R图向关系模型的转换

7.4.2 向特定DBMS规定的模型进行转换

7.4.3 数据模型的优化

7.4.4 设计用户子模式



7.4.1 E-R图向关系模型的转换

□ 转换内容

□ 转换原则



E-R图向关系模型的转换（续）

□ 转换内容

- ◆ **E-R**图由实体、实体的属性和实体之间的联系三个要素组成
- ◆ 关系模型的逻辑结构是一组关系模式的集合
- ◆ 将**E-R**图转换为关系模型：将实体、实体的属性和实体之间的联系转化为关系模式。



E-R图向关系模型的转换（续）

□ 转换原则

1. 一个实体型转换为一个关系模式。

◆ **关系的属性**：实体型的属性

◆ **关系的码**：实体型的码

例，学生实体可以转换为如下关系模式：

学生（学号，姓名，出生日期，所在系，
 年级，平均成绩）

性别、宿舍、班级、档案材料、教师、课程、教室、教科书都分别转换为一个关系模式。



E-R图向关系模型的转换（续）

2. 一个m:n联系转换为一个关系模式。

- ◆ **关系的属性**: 与该联系相连的各实体的码以及联系本身的属性

- ◆ **关系的码**: 各实体码的组合

例，“选修”联系是一个m:n联系，可以将它转换为如下关系模式，其中学号与课程号为关系的组合码：

选修（学号，课程号，成绩）



E-R图向关系模型的转换（续）

3. 一个1:n联系可以转换为一个独立的关系模式，也可以与n端对应的关系模式合并。

◆ 1) 转换为一个独立的关系模式

- **关系的属性：** 与该联系相连的各实体的码以及联系本身的属性
- **关系的码：** n端实体的码



E-R图向关系模型的转换（续）

3. 一个1:n联系可以转换为一个独立的关系模式，也可以与n端对应的关系模式合并。

◆ 2) 与n端对应的关系模式合并

- **合并后关系的属性：** 在n端关系中加入1端关系的码和联系本身的属性
- **合并后关系的码：** 不变

◆ 可以减少系统中的关系个数，一般情况下更倾向于采用这种方法



E-R图向关系模型的转换（续）

例，“组成”联系为1:n联系。

将其转换为关系模式的两种方法：

1)使其成为一个独立的关系模式：

组成（学号，班级号）

2)将其学生关系模式合并：

学生（学号，姓名，出生日期，所在系，
年级，**班级号**，平均成绩）



E-R图向关系模型的转换（续）

4. 一个1:1联系可以转换为一个独立的关系模式，也可以与任意一端对应的关系模式合并。

◆ 1) 转换为一个独立的关系模式

- **关系的属性：** 与该联系相连的各实体的码以及联系本身的属性
- **关系的候选码：** 每个实体的码均是该关系的候选码



E-R图向关系模型的转换（续）

4. 一个1:1联系可以转换为一个独立的关系模式，也可以与任意一端对应的关系模式合并。

◆ 2) 与某一端对应的关系模式合并

- **合并后关系的属性：** 加入对应关系的码和联系本身的属性
- **合并后关系的码：** 不变



E-R图向关系模型的转换（续）

例，“管理”联系为1:1联系，可以有三种转换方法：

(1) 转换为一个独立的关系模式：

管理（职工号，班级号）

或

管理（职工号，班级号）

(2) “管理”联系与班级关系模式合并，则只需在班级关系中加入教师关系的码，即职工号：

班级：（班级号，学生人数，**职工号**）

(3) “管理”联系与教师关系模式合并，则只需在教师关系中加入班级关系的码，即班级号：

教师：（职工号，姓名，性别，职称，**班级号**，
是否为优秀班主任）



E-R图向关系模型的转换（续）

注意：

- ◆ 从理论上讲，1:1联系可以与任意一端对应的关系模式合并。
- ◆ 但在一些情况下，与不同的关系模式合并效率会大不一样。因此究竟应该与哪端的关系模式合并需要依应用的具体情况而定。
- ◆ 由于连接操作是最费时的操作，所以一般应以尽量减少连接操作为目标。

例如，如果经常要查询某个班级的班主任姓名，则将管理联系与教师关系合并更好些。



E-R图向关系模型的转换（续）

5. 三个或三个以上实体间的一个多元联系转换为一个关系模式。

◆ **关系的属性**：与该多元联系相连的各实体的码以及联系本身的属性

◆ **关系的码**：各实体码的组合

例，“讲授”联系是一个三元联系，可以将它转换为如下关系模式，其中课程号、职工号和书号为关系的组合码：

讲授（课程号，职工号，书号）



E-R图向关系模型的转换（续）

6. 同一实体集的实体间的联系，即**自联系**，也可按上述1:1、1:n和m:n三种情况分别处理。

例，如果教师实体集内部存在领导与被领导的1:n自联系，我们可以将该联系与教师实体合并，这时主码职工号将多次出现，但作用不同，可用不同的属性名加以区分：

教师：{ 职工号，姓名，性别，职称，**系主任** }



E-R图向关系模型的转换（续）

7. 具有相同码的关系模式可合并。

- ◆ **目的**：减少系统中的关系个数。
- ◆ **合并方法**：将其中一个关系模式的全部属性加入到另一个关系模式中，然后去掉其中的同义属性（可能同名也可能不同名），并适当调整属性的次序。



E-R图向关系模型的转换（续）

例，“拥有”关系模式：

拥有（学号，性别）

与学生关系模式：

学生（学号，姓名，出生日期，所在系，年级，
班级号，平均成绩）

都以学号为码，可以将它们合并为一个关系模式：

学生（学号，姓名，性别，出生日期，所在系，
年级，班级号，平均成绩）

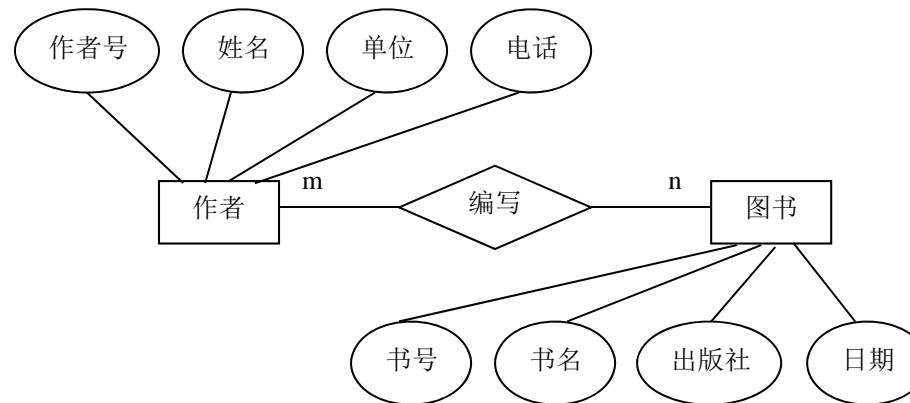


应用

- 例：
- 在著书工作中，一位作者可以编写多本图书，一本书可以由多位作者编写。
设作者的属性有：作者号，姓名，单位，电话；书的属性有：书号，书名，出版社，日期。试完成如下问题：
- （1）设计该系统的E-R图；
- （2）将该E-R图转换为关系模型结构；
- （3）指出转换结果中每一个关系模式的候选码。



□ (1) E-R图



□ (2) 将该E-R图转换为关系模型结构如下：

□ 作者（作者号，姓名，单位，电话）

□ 图书（书号，书名，出版社，日期）

□ 编写（作者号，书号）

□ (3) 以上关系模式中带下划线的为该关系模式的候选码



7.4 逻辑结构设计

7.4.1 E-R图向关系模型的转换

7.4.2 向特定DBMS规定的模型进行转换

7.4.3 数据模型的优化

7.4.4 设计用户子模式



7.4.2 向特定DBMS规定的模型进行转换

- 一般的数据模型还需要向特定**DBMS**规定的模型进行转换。
- 转换的主要依据是所选用的**DBMS**的功能及限制，没有通用规则。
- 对于关系模型来说，这种转换通常都比较简单。



7.4 逻辑结构设计

7.4.1 E-R图向关系模型的转换

7.4.2 向特定DBMS规定的模型进行转换

7.4.3 数据模型的优化

7.4.4 设计用户子模式



7.4.3 数据模型的优化

- 数据库逻辑设计的结果**不是唯一的**。
- 得到初步数据模型后，还应该适当地修改、调整数据模型的结构，以进一步提高数据库应用系统的性能，这就是数据模型的优化。
- 关系数据模型的优化通常以**规范化理论**为指导。



数据模型的优化（续）

□ 优化数据模型的方法

1. 确定数据依赖

- ◆ 按需求分析阶段所得到的语义，分别写出每个关系模式内部各属性之间的数据依赖以及不同关系模式属性之间数据依赖。



数据模型的优化（续）

例，课程关系模式内部存在下列数据依赖：

课程号 \rightarrow 课程名

课程号 \rightarrow 学分

课程号 \rightarrow 教室号

选修关系模式中存在下列数据依赖：

(学号, 课程号) \rightarrow 成绩



数据模型的优化（续）

学生关系模式中存在下列数据依赖：

学号→姓名

学号→性别

学号→出生日期

学号→所在系

学号→年级

学号→班级号

学号→平均成绩

学号→档案号



数据模型的优化（续）

学生关系模式的学号与选修关系模式的学号之间存在数据依赖：

学生.学号 \rightarrow 选修.学号



数据模型的优化（续）

2. 对于各个关系模式之间的数据依赖进行极小化处理，消除冗余的联系。



数据模型的优化（续）

3. 按照数据依赖的理论对关系模式逐一进行分析，考查是否存在部分函数依赖、传递函数依赖、多值依赖等，确定各关系模式分别属于第几范式。



数据模型的优化（续）

4. 按照需求分析阶段得到的各种应用对数据处理的要求，分析对于这样的应用环境这些模式是否合适，确定是否要对它们进行合并或分解。



数据模型的优化（续）

- ◆ 并不是规范化程度越高的关系就越优。
 - 当一个应用的查询中经常涉及到两个或多个关系模式的属性时，系统必须经常地进行连接运算，而连接运算的代价是相当高的，可以说关系模型低效的主要原因就是做连接运算引起的，因此在这种情况下，第二范式甚至第一范式也许是最好的。



数据模型的优化（续）

非BCNF的关系模式虽然从理论上分析会存在不同程度的更新异常，但如果在实际应用中对此关系模式只是查询，并不执行更新操作，则就不会产生实际影响。

对于一个具体应用来说，到底规范化进行到什么程度，需要权衡响应时间和潜在问题两者的利弊才能决定。一般说来，第三范式就足够了。



数据模型的优化（续）

例：在关系模式

学生成绩单(学号,英语,数学,语文,平均成绩) 中存在下列
函数依赖：

学号→英语

学号→数学

学号→语文

学号→平均成绩

(英语, 数学, 语文)→平均成绩



数据模型的优化（续）

显然有：

学号 \rightarrow (英语, 数学, 语文)

因此该关系模式中存在传递函数依赖，是2NF关系。

虽然平均成绩可以由其他属性推算出来，但如果应用中需要经常查询学生的平均成绩，为提高效率，我们仍然可保留该冗余数据，对关系模式不再做进一步分解。



数据模型的优化（续）

5. 按照需求分析阶段得到的各种应用对数据处理的要求，对关系模式进行必要的分解或合并，以提高数据操作的效率和存储空间的利用率

◆ 常用分解方法

- 水平分解
- 垂直分解



数据模型的优化（续）

◆ 水平分解

- 什么是水平分解

- 把(基本)关系的元组分为若干子集合，定义每个子集合为一个子关系，以提高系统的效率。



数据模型的优化（续）

- 水平分解的适用范围

- 1. 满足“80/20原则”的应用

- ✓ 80/20原则：一个大关系中，经常被使用的数据只是关系的一部分，约20%
 - ✓ 把经常使用的数据分解出来，形成一个子关系，可以减少查询的数据量。



数据模型的优化（续）

- 水平分解的适用范围

- 2. 并发事务经常存取不相交的数据

- ✓ 如果关系 R 上具有 n 个事务，而且多数事务存取的数据不相交，则 R 可分解为少于或等于 n 个子关系，使每个事务存取的数据对应一个关系。



数据模型的优化（续）

◆ 垂直分解

- 什么是垂直分解

- 把关系模式 R 的属性分解为若干子集合，形成若干子关系模式。

- 垂直分解的原则

- 经常在一起使用的属性从 R 中分解出来形成一个子关系模式。



数据模型的优化（续）

- 垂直分解的优点
 - 可以提高某些事务的效率
- 垂直分解的缺点
 - 可能使另一些事务不得不执行连接操作，从而降低了效率。



数据模型的优化（续）

- 垂直分解的适用范围
 - 取决于分解后 R 上的所有事务的总效率是否得到了提高。
- 进行垂直分解的方法
 - 简单情况：直观分解
 - 复杂情况：用第六章中的模式分解算法
 - 垂直分解必须不损失关系模式的语义(保持无损连接性和保持函数依赖)。



7.4 逻辑结构设计

7.4.1 E-R图向关系模型的转换

7.4.2 向特定DBMS规定的模型进行转换

7.4.3 数据模型的优化

7.4.4 设计用户子模式



7.4.4 设计用户模式

- 定义数据库模式主要是从系统的**时间效率、空间效率、易维护**等角度出发。
- 定义用户模式时应该更注重考虑用户的习惯与方便。包括三个方面：



设计用户子模式（续）

(1) 使用更符合用户习惯的别名

- ◆ 合并各分**E-R**图曾做了消除命名冲突的工作，以使数据库系统中同一关系和属性具有唯一的名字。这在设计数据库整体结构时是非常必要的。
- ◆ 但对于某些局部应用，由于改用了不符合用户习惯的属性名，可能会使他们感到不方便。



设计用户子模式（续）

(1) 使用更符合用户习惯的别名(续)

- ◆ 因此在设计用户的子模式时可以重新定义某些属性名，使其与用户习惯一致。
- ◆ 当然，为了应用的规范化，我们也不应该一味地迁就用户。

例：负责学籍管理的用户习惯于称教师模式的职工号为教师编号。因此可以定义视图，在视图中医工号重定义为教师编号。



设计用户子模式（续）

(2) 针对不同级别的用户定义不同的外模式，以满足系统对安全性的要求。



设计用户子模式（续）

例：

教师关系模式中包括职工号、姓名、性别、出生日期、婚姻状况、学历、学位、政治面貌、职称、职务、工资、工龄、教学效果等属性。

学籍管理应用只能查询教师的职工号、姓名、性别、职称数据；

课程管理应用只能查询教师的职工号、姓名、性别、学历、学位、职称、教学效果数据；

教师管理应用则可以查询教师的全部数据。



设计用户子模式（续）

定义两个外模式：

教师_学籍管理(职工号，姓名，性别，职称)

教师_课程管理(工号，姓名，性别，学历，
学位，职称，教学效果)

授权学籍管理应用只能访问教师_学籍管理视图

授权课程管理应用只能访问教师_课程管理视图

授权教师管理应用能访问教师表

这样就可以防止用户非法访问本来不允许他们查询的数据，
保证了系统的安全性。



设计用户子模式（续）

(3) 简化用户对系统的使用

- ◆ 如果某些局部应用中经常要使用某些很复杂的查询，为了方便用户，可以将这些复杂查询定义为视图。



逻辑结构设计小结

□ 任务

- ◆ 将概念结构转化为具体的数据模型

□ 逻辑结构设计的步骤

- ◆ 将概念结构转化为一般的关系、网状、层次模型
- ◆ 将转化来的关系、网状、层次模型向特定**DBMS**支持下的数据模型转换
- ◆ 对数据模型进行优化
- ◆ 设计用户子模式



逻辑结构设计小结

□ E-R图向关系模型的转换内容

- ◆ 将**E-R**图转换为关系模型：将实体、实体的属性和实体之间的联系转化为关系模式。



逻辑结构设计小结

□ E-R图向关系模型的转换原则

1. 一个**实体型**转换为一个关系模式。
2. 一个**m:n联系**转换为一个关系模式。
3. 一个**1:n联系**可以转换为一个独立的关系模式，也可以与n端对应的关系模式合并。
4. 一个**1:1联系**可以转换为一个独立的关系模式，也可以与任意一端对应的关系模式合并。



逻辑结构设计小结

□ E-R图向关系模型的转换原则

5. 三个或三个以上实体间的一个**多元联系**转换为一个关系模式。
6. 同一实体集的实体间的联系，即**自联系**，也可按上述1:1、1:n和m:n三种情况分别处理。
7. 具有**相同码**的关系模式可合并。



逻辑结构设计小结

□ 优化数据模型的方法

1. 确定数据依赖
2. 对于各个关系模式之间的数据依赖进行极小化处理，消除冗余的联系。
3. 确定各关系模式分别属于第几范式。
4. 分析对于应用环境这些模式是否合适，确定是否要对它们进行合并或分解。
5. 对关系模式进行必要的分解或合并。



逻辑结构设计小结

□ 设计用户子模式

1. 使用更符合用户习惯的别名
2. 针对不同级别的用户定义不同的外模式，以满足系统对安全性的要求。
3. 简化用户对系统的使用