



数据库原理

The Principle of Database

主讲：章优仕

12/13/2023





第二章 数据库设计和ER模型

- 1、数据库设计概述
- 2、ER模型基本概念
- 3、关系模型基本概念
- 4、ER模型到关系模型
- 5、ER模型实例分析
- 6、增强的ER模型





1、数据库设计概述

考核要求：到达“识记”层次

知识点：数据库设计有关的某些基本概念。

- 1.1 规划
- 1.2 需求分析
- 1.3 概念设计
- 1.4 逻辑设计
- 1.5 物理设计
- 1.6 实现
- 1.7 维护





1. 系统需求分析阶段

需求分析是整个数据库设计过程的基础，要搜集数据库全部顾客的信息内容和处理要求，并加以规格化和分析。

2. 概念构造设计阶段

概念设计是把顾客的信息要求统一到一种整体逻辑构造中，此构造能够体现顾客的要求，是一种独立于任何DBMS软件和硬件的概念模型。





3. 逻辑构造设计阶段

逻辑设计是将上一步所得到的概念模型转换为某个DBMS所支持的数据模型，并对其进行优化。

4. 物理设计阶段

物理设计是为逻辑数据模型建立一种完整的能实现的数据库构造，涉及存储构造和存取措施。





5. 数据库实施阶段

根据物理设计的成果把原始数据装入数据库，建立一种详细的数据库并编写和调试相应的应用程序。应用程序的开发目的是开发一种可依赖的有效的数据库存取程序，来满足顾客的处理要求。

6. 数据库运营与维护阶段

这一阶段主要是搜集和统计实际系统运营的数据，数据库运营的统计取来提升顾客要求的有效信息，用来评价数据库系统的性能，进一步调整和修改数据库。在运营中，必须保持数据库的完整性，并能有效地处理数据库故障和进行数据库恢复。在运营和维护阶段，可能要对数据库构造进行修改或扩充。





设计阶段	设计描述	
	数据	处理
需求分析	数据字典、全系统中数据项、数据流、数据存储的描述	数据流图和定表（鉴定树） 数据字典中处理过程的描述
概念构造设计	概念模型（E-R图） 数据字典	系统阐明书。涉及： (1) 新系统要求、方案和概图 (2) 反应新系统信息的数据流图
逻辑构造设计	某种数据模型 关系模型	系统构造图 非关系模型（模块构造图）
物理设计	存储安排 存取措施选择 存取途径建立	模块设计 IPO表
实施阶段	编写模式 装入数据 数据库试运营	程序编码 编译联结 测试
运营维护	性能测试，转储/恢复数据库 重组和重构	新旧系统转换、运营、维护（修正性、适应性、改善性维护）

数据库各个设计阶段的描述





- 数据库系统生存期
- 规划阶段：
 - 系统调查，可行性分析，拟定总目的
- 需求分析（内容）：
 - 业务流程图，系统关联图，数据流图，数据字典
- 概念设计（目的）：
 - 局部模型，全局模型，评审
- 逻辑设计（目的）：
 - 转换，外模型，接口，评价，修正





- 物理设计（定义）：
 - 设计统计构造，数据组合，存取措施，完整和安全，程序
- 数据库实现：
 - 数据库构造，数据装载，编调程序，试运营
- 数据库运营与维护：
 - 转储和恢复，完整性安全性控制，性能监督，数据库重组





2、ER模型基本概念

考核要求：到达“简朴应用”层次
知识点：ER图的设计

- 2.1 基本元素
- 2.2 属性
- 2.3 联络
- 2.4 操作
- 2.5 ER模型与概念设计





2.1、ER模型基本元素

- 实体：
 - 实体，实体集，实体类型
- 联络：
 - 联络，联络集，联络类型
- 属性：
 - 实体标识符，属性域





矩形: 表示实体集。

椭圆: 表示属性。

菱形: 表示联系集。

线段: 将属性连接到实体集或将实体集连接到联系集。

双椭圆: 表示多值属性。

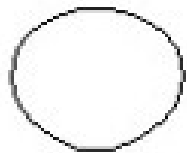
虚椭圆: 表示派生属性。

双线: 表示一个实体全部参与到联系集中。

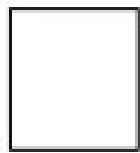
双边框的矩形: 表示弱实体集。

双边框的菱形: 表示弱实体集对应的标识性联系。

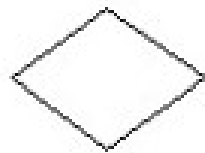




椭圆



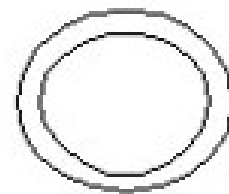
矩形



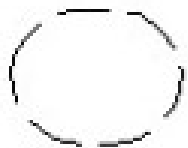
菱形



线段



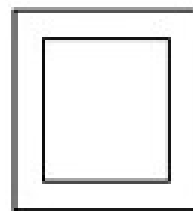
双椭圆



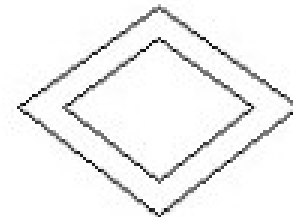
虚椭圆



双线



双边框的矩形



双边框的菱形

ER图中使用的多种符号



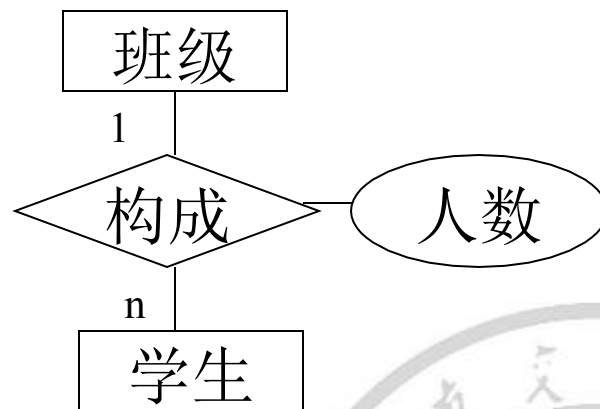
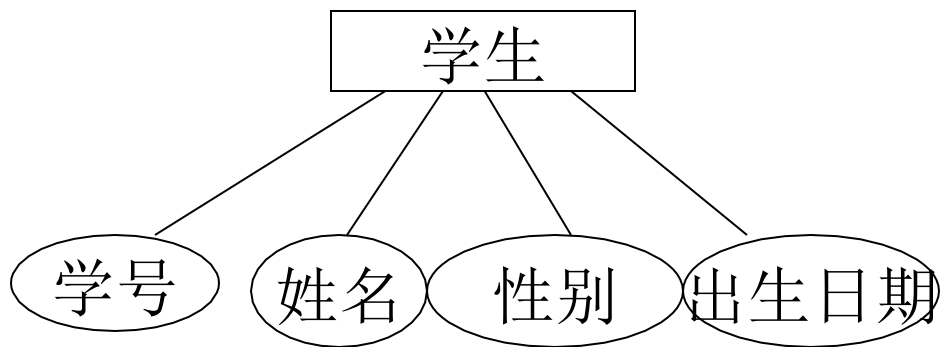


实体名

属性名

联络名

——





2.2、属性

- 简朴属性和复合属性
- 单值属性和多值属性
- 多值属性的分解，弱实体
- 派生属性和存储属性
- 空值属性





根据属性取值的不同种类，可将属性划分类型如下：

- (1) **简单属性**：指它们不能再划分为更小的部分。例如，课程名是简单属性。
- (2) **复合属性**：指它们可以再划分为更小的部分。例如，出生日期可被设计成包括出生年、月、日的成分属性。
- (3) **单值属性**：指所定义的属性对一个特定实体都只有单独的一个值。例如，学号属性只对应一个学号号码。
- (4) **多值属性**：指对某个特定实体而言，一个属性可能对应于一组值。例如，学生的社会关系这个属性。
- (5) **NULL属性**：当实体在某个属性上没有值或属性值未知时使用NULL值。例如，某个学生无亲属，那么该学生的社会关系属性值是NULL，表示“无意义”。
- (6) **派生属性**：这类属性的值可以从别的相关属性或实体派生出来。例如，学生的年龄可以通过其出生日期计算出来。





2.3、联络

- ▶ 元数：一元，二元，三元……
- ▶ 类型的约束：基数约束，参加约束

映射的基数（或基数比例）

是指通过一个联系集能同另一实体相联系的实体数目。

映射的基数比约束

实体间联系的约束。

实体的参与度

实体参与联系的最小和最大次数。

其一般形式可表示为(min, max)，式中， $0 \leq \min \leq \max$ ，且 $\max \geq 1$ 。

分部分参与和全参与两种方式。





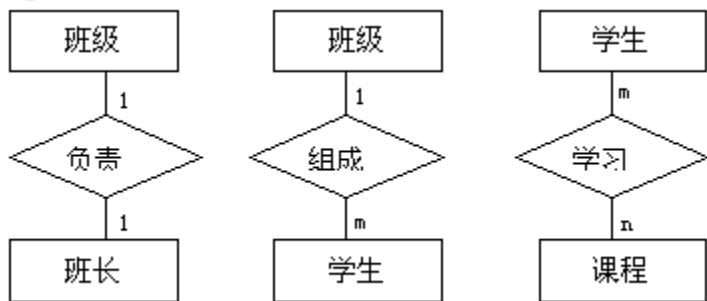
- 因为联络本身也是一种实体型，所以联络也能够有属性。假如一种联络具有属性，则这些联络也要用无向边与该联络连接起来。例如，学生选修的课程有相应的成绩。这里的“成绩”既不是学生的属性，也不是课程的属性，只能是学生选修课程的联络的属性。
- 设计局部E-R模型的关键就是对划分实体和属性。实体和属性之间在形式上并无能够明显辨别的界线，一般是按照现实世界中事物的自然划分来定义实体和属性，将现实世界中的事物进行数据抽象，得到实体和属性





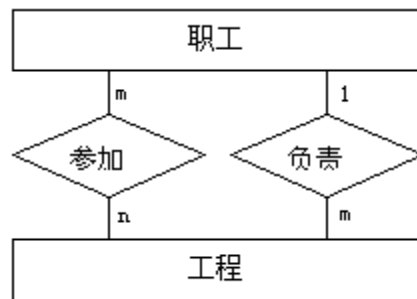
联络的类型

两个不同型实体间的联络



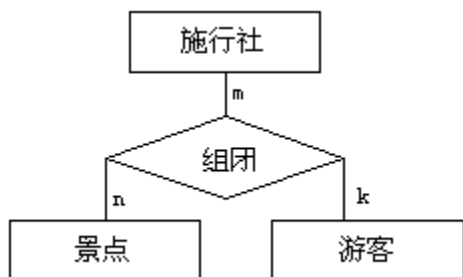
两个不同型实体间的多种联络

职员与工程间，一种职员能够参加多种工程，一种工程能够有多种职员参加，同步一种工程由一种职员负责，一种职员能够负责多种工程。

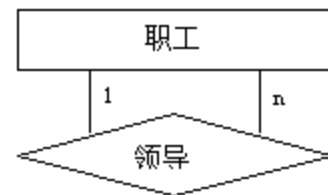


两个以上实体间的多元联络

旅行社和景点及游客三个实体间存在三元联络



同一实体内部个体间的一元联络





2.4、ER模型的操作

- ▶ 水平分裂是把关系的元组分为若干子集合，定义每个子集合为一种子关系。
- ▶ 例如，有学生关系（学号，姓名，类别……），其中类别涉及大专生、本科生和硕士。假如多数查询一次只涉及其中的一类学生，就应该把整个学生关系水平分割为大专生、本科生和硕士三个关系。
- ▶ 垂直分裂是把关系模式的属性分解为若干子集合，形成若干子关系模式。
- ▶ 例如，有教师关系（教师号，姓名，性别，年龄，职称，工资，岗位津贴，住址， ），假如经常查询的仅是前六项，而后三项极少使用，则能够将教师



2.5、ER模型与概念设计

- 局部ER模型
 - 局部构造范围，定义实体，定义联络，分配属性
- 全局ER模型
 - 拟定公共实体，合并局模型，消除冲突
- 优化
 - 合并实体类型，消除冗余属性，消除冗余联络





➤ 局部模型设计

- 实体和属性是相对而言的，往往要根据实际情况进行必要的调整
- 同一数据项，可能因为环境和要求的不同，有时作为属性，有时则作为实体，此时必须根据实际情况而定。一般情况下，凡能作为属性看待的，应尽量作为属性，以简化E-R图的处理。





- ✓①实体具有描述信息，而属性没有。即属性必须是不可分的数据项，不能再由另某些属性构成。
- ✓②属性不能与其他实体具有联络，联络只能发生在实体之间。
- ✓例如：学生是一种实体，学号、姓名、性别、年龄、系别等是学生实体的属性，系别只体现学生属于哪个系，不涉及系的详细情况，换句话说，没有需要进一步描述的特征，即是不可分的数据项，则根据原则①能够作为学生实体的属性。但假如考虑一种系的系主任、学生人数、教师人数、办公地点等，则系别应看作一种实体。
- ✓又如，“职称”为教师实体的属性，但在涉及住房分配时，因为分房与职称有关，即职称与住房实体之间有联络，则根据原则②，职称应作为一种实体。



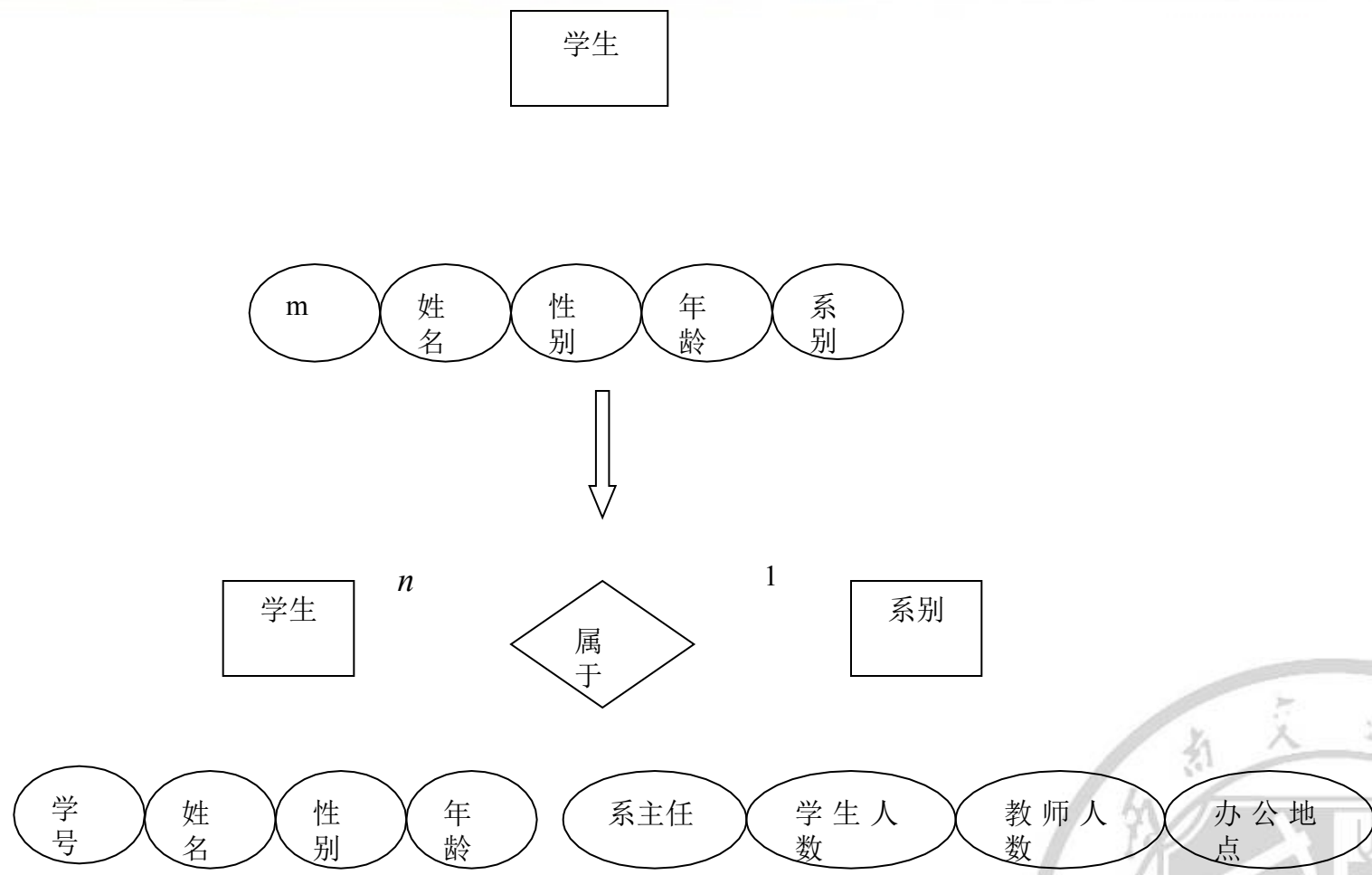


图1 系别作为一种属性或实体



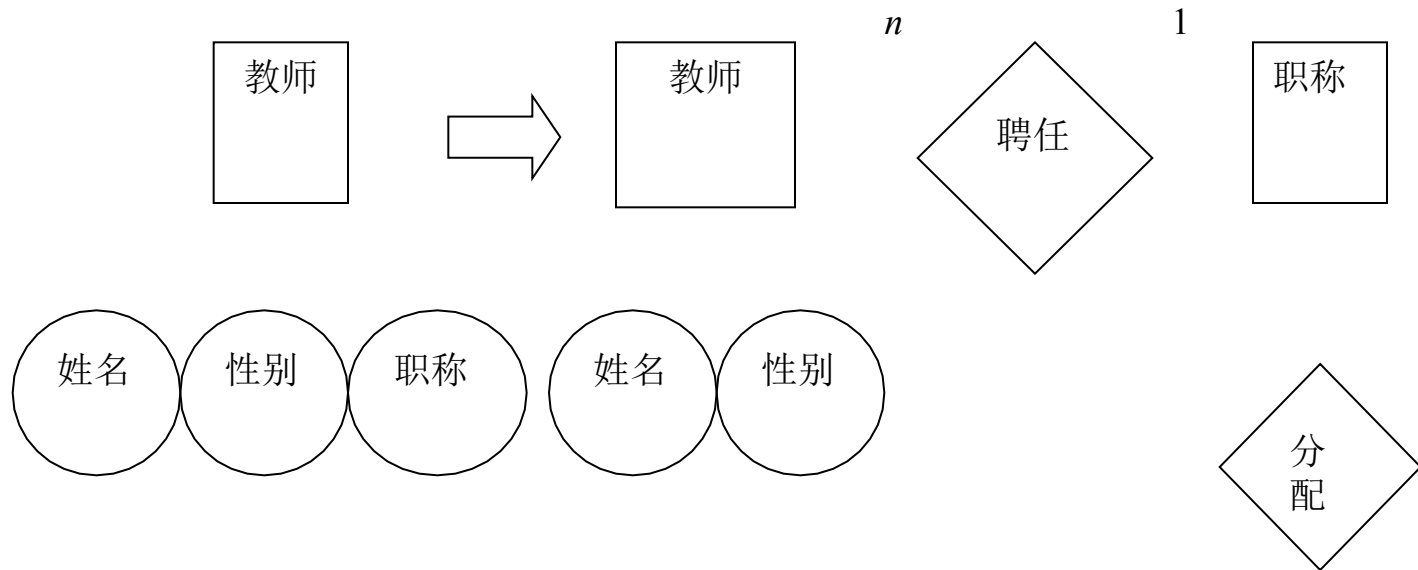


图2 职称作为一种属性或实体





下面举例阐明局部ER模型设计。

在简朴的教务管理系统中，有如下语义约束。

- ①一种学生可选修多门课程，一门课程可为多种学生选修，所以学生和课程是多对多的联络；
- ②一种教师可讲授多门课程，一门课程可为多种教师讲授，所以教师和课程也是多对多的联络；
- ③一种系可有多种教师，一种教师只能属于一种系，所以系和教师是一对多的联络，一样系和学生也是一对多的联络。

根据上述约定，能够得到如图3所示的学生选课局部ER图和如图4所示的教师任课局部ER图。



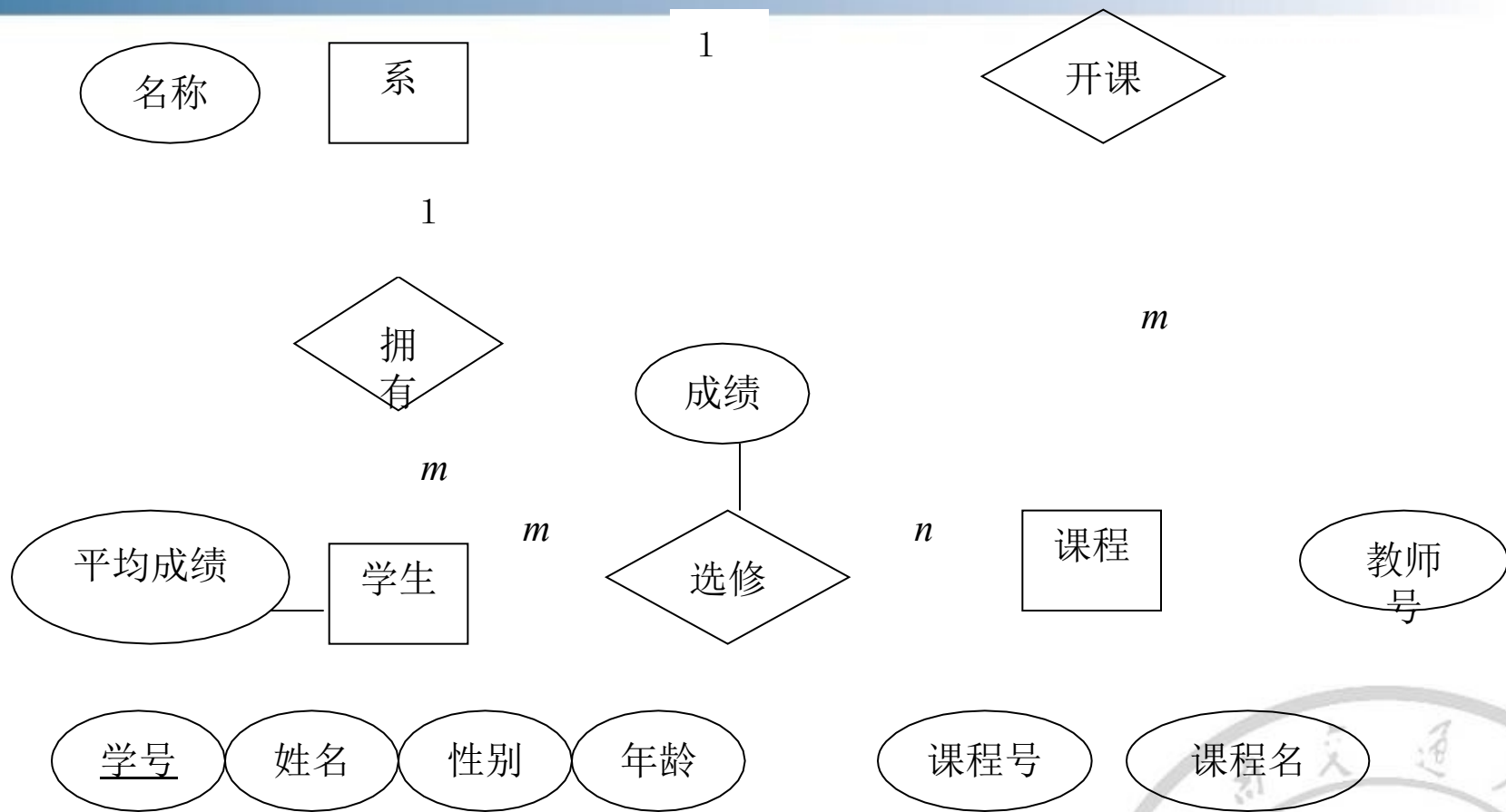


图3 学生选课局部ER图

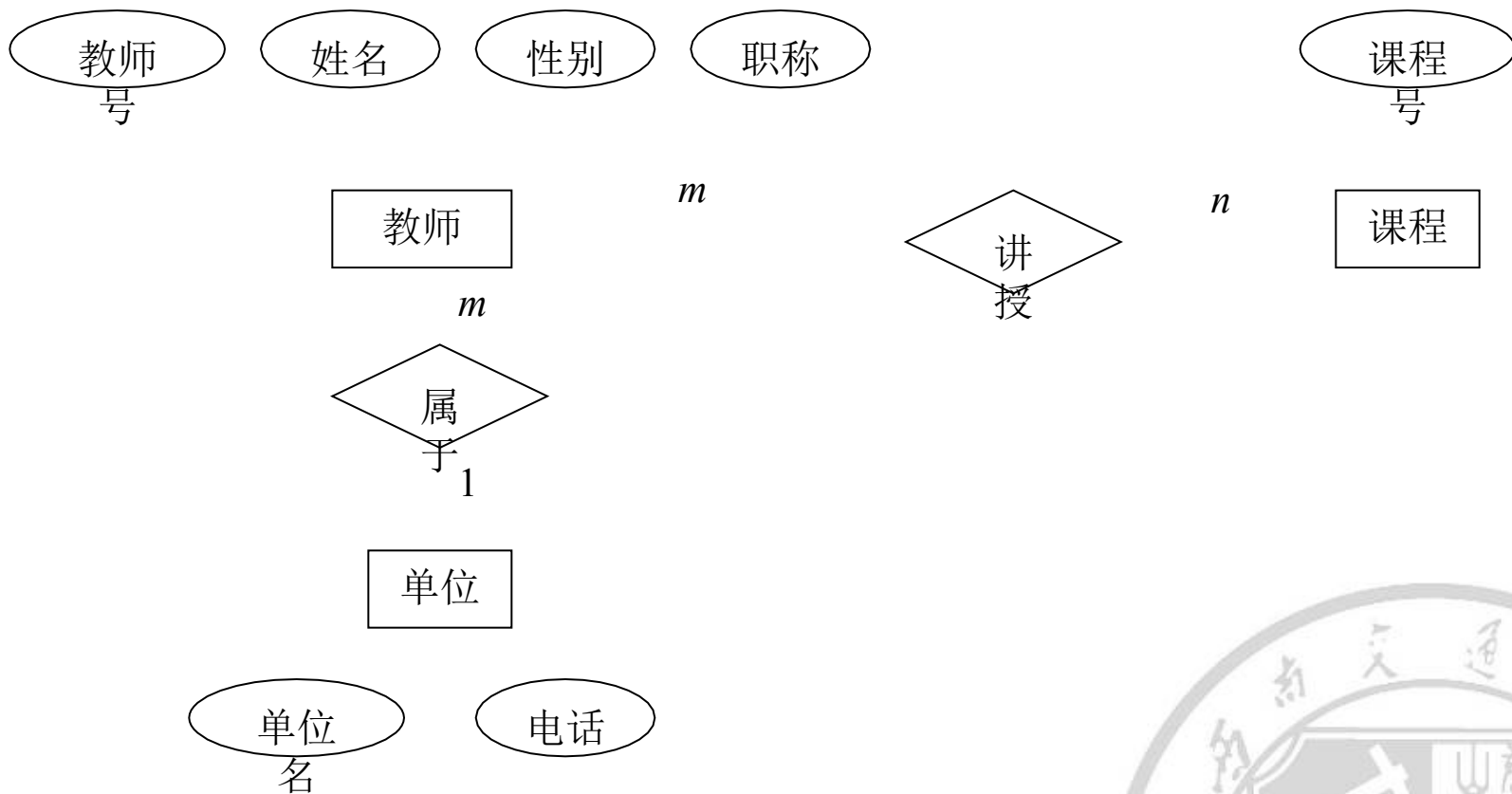


图4 教师任课局部ER图





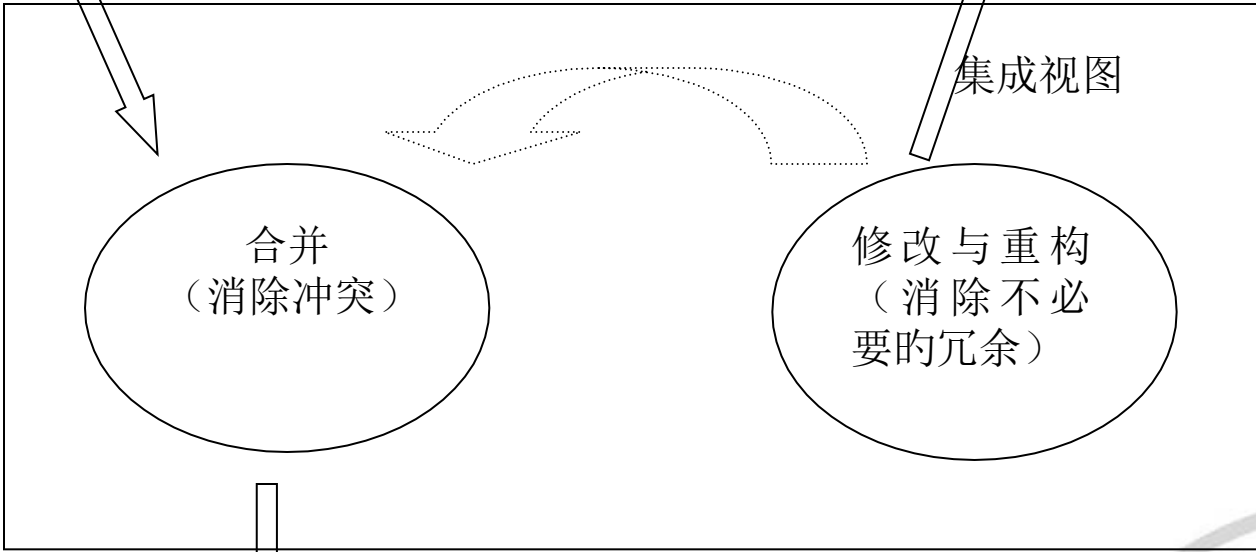
➤全局模型设计

- 局部E-R模型设计完毕之后，下一步就是集成各局部E-R模型，形成全局ER模型，即视图的集成。视图集成的措施有两种：
 - ①多元集成法，一次性将多种局部E-R图合并为一种全局E-R图
 - ②二元集成法，首先集成两个主要的局部视图，后来用累加的措施逐渐将一种新的视图集成进来
- 视图集成均提成两个环节，
 - ①合并，消除各局部E-R图之间的冲突，生成初步E-R图。
 - ②优化，消除不必要的冗余，生成基本E-R图。



局部 E - R 图

基本 E - R 图



初步 E - R 图

分析

规范化理论

视图集成



- 全局模型不但要支持全部的局部ER模型，而且必须合理地体现一种完整、一致的数据库概念构造。
- 因为各个局部应用不同，一般由不同的设计人员进行局部E-R图设计，所以，各局部E-R图不可预防地会有许多不一致的地方，我们称之为冲突。
- 合并局部E-R图时并不能简朴地将各个ER图画到一起，而必须消除各个局部ER图中的不一致，使合并后的全局概念构造不但支持全部的局部E-R模型，而且必须是一种能为全系统中全部顾客共同了解和接受的完整的概念模型。
- 合并局部ER图的关键就是合理消除各局部E-R图中的冲突。





➤ 属性冲突

- 属性冲突又分为属性值域冲突和属性的取值单位冲突。
- a. 属性值域冲突，即属性值的类型、取值范围或取值集合不同。例如学号，有些部门将其定义为数值型，而有些部门将其定义为字符型。又如年龄，有的可能用出生年月体现，有的则用整数体现。
- b. 属性的取值单位冲突。例如零件的重量，有的以公斤为单位，有的以斤为单位，有的则以克为单位。
- 属性冲突属于顾客业务上的约定，必须与顾客协商后处理。





命名冲突

- 命名不一致可能发生在实体名、属性名或联络名之间，其中属性的命名冲突更为常见。
- 一般体现为同名异义或异名同义（实体、属性、联络名）。
 - a. 同名异义，即同一名字的对象在不同的部门中具有不同的意义。例如，“单位”在某些部门体现为人员所在的部门，而在某些部门可能体现物品的重量、长度等属性。
 - b. 异名同义，即同一意义的对象在不同的部门中具有不同的名称。例如，对于“房间”这个名称，在教务管理部门中相应着为教室，而在后勤管理部门相应为学生宿舍。
- 命名冲突的处理措施同属性冲突，需要与各部门协商、讨论后加以处理。





➤ 构造冲突

- a. 同一对象在不同应用中有不同的抽象，可能为实体，也可能为属性。例如，教师的职称在某一局部应用中被看成实体，而在另一局部应用中被看成属性。
- 此类冲突在处理时，就是使同一对象在不同应用中具有相同的抽象，或把实体转换为属性，或把属性转换为实体。
- b. 同一实体在不同应用中属性构成不同，可能是属性个数或属性顺序不同。
- 处理措施是，合并后实体的属性构成为各局部E-R图中的同名实体属性的并集，然后再合适调整属性的顺序。
- c. 同一联络在不同应用中呈现不同的类型。
- 这种情况应该根据应用的语义对实体联络的类型进行综合或调整。





- ▶ 下面以教务管理系统中的两个局部ER图为例，来阐明怎样消除各局部ER图之间的冲突，进行局部ER模型的合并，从而生成初步E-R图。
- ▶ 首先，这两个局部ER图中存在着命名冲突，学生选课局部ER图中的实体“系”与教师任课局部ER图中的实体“单位”，都是指“系”，即所谓的异名同义，合并后统一改为“系”，这么属性“名称”和“单位”即可统一为“系名”。
- ▶ 另一方面，还存在着构造冲突，实体“系”和实体“课程”在两个不同应用中的属性构成不同，合并后这两个实体的属性构成为原来局部E-R图中的同名实体属性的并集。处理上述冲突后，合并两个局部E-R图，生成如图所示的初步的全局ER图。



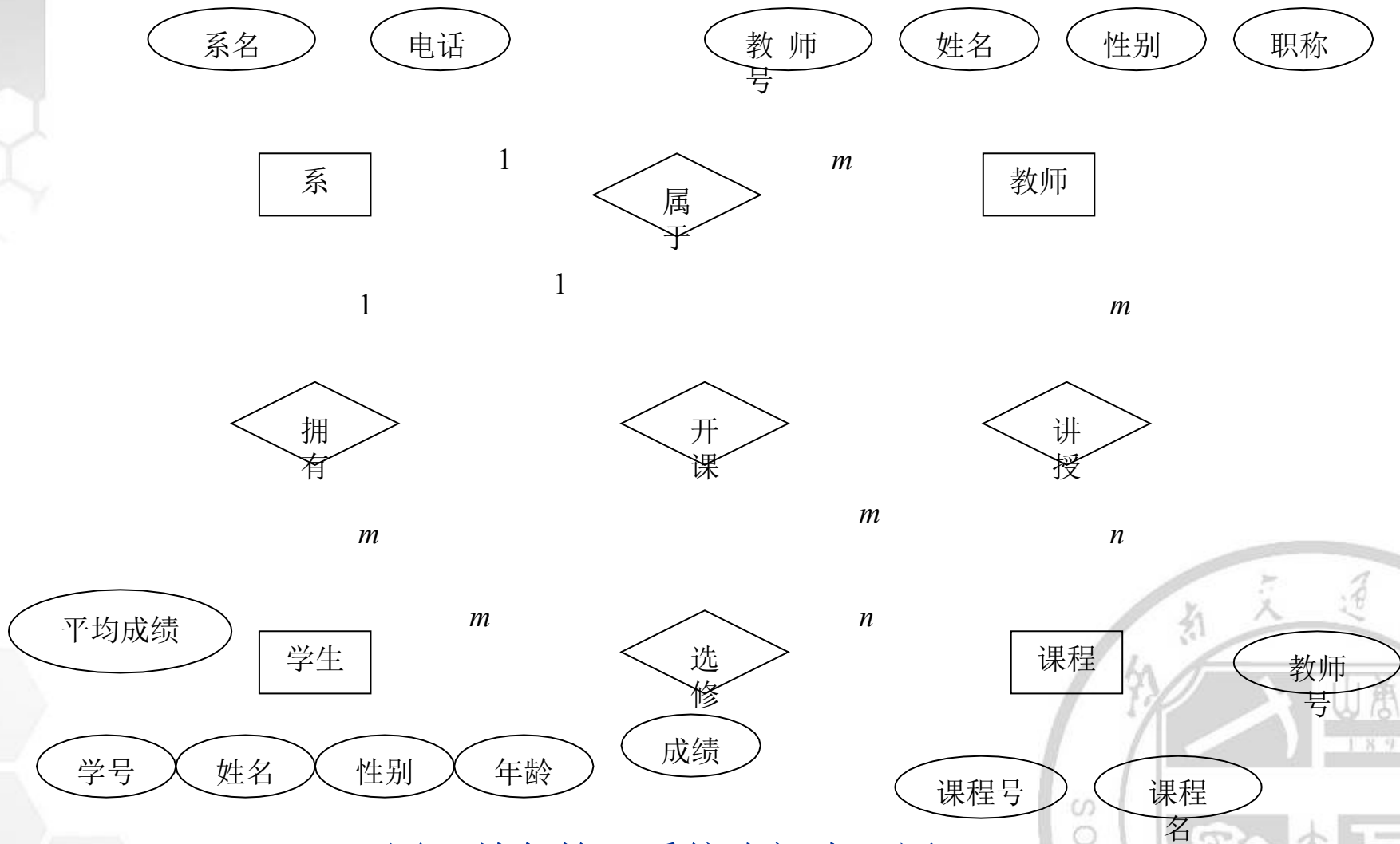


图5 教务管理系统的初步ER图

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/385331303310011303>