



# 设计篇

## 第四章 数据库设计

主讲：王金宝





# Outline

- 数据库设计概述与需求分析
- 概念数据库设计
- 逻辑数据库设计
- 物理数据库设计
- 小结





## 4.1 数据库设计概述与需求分析





# 数据库设计概述

## • 数据库设计

- 对于一个给定的应用领域，设计优化的数据库逻辑和物理结构，使之满足用户的信息管理要求和数据操作要求，有效地支持各种应用系统的开发和运行
- 信息管理要求
  - 在数据库中应该存储和管理哪些数据对象。
- 数据操作要求
  - 对数据对象需要进行哪些操作，如查询、增、删、改、统计等操作。





# 数据库设计概述

- 数据库设计目标

- 为用户和各种应用系统提供一个信息基础设施和高效的运行环境

- 数据库数据的存取效率高

- 数据库存储空间的利用率高

- 数据库系统运行管理的效率高





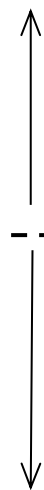
# 数据库设计概述

- 数据库设计特点
  - 结构（数据）设计
    - 概念结构设计、逻辑结构设计、物理结构设计
  - 行为（处理/查询）设计
    - 功能模型、事务设计、应用设计
  - 二者结合





# 数据库设计步骤



独立于数据库  
管理系统

与数据库  
管理系统  
相关

需求收集和分析

综合各个用户的应用需求(数据  
与处理)

设计概念结构

形成独立于机器特点, 独立于各个  
DBMS 产品的概念模式(如E-R图)

设计逻辑结构

基于具体DBMS支持的数据模型,  
将概念模式转换为数据库逻辑模式

数据模型优化

根据用户处理的要求、安全性的考  
虑, 在基本表的基础上再建立必要  
的视图(View), 形成数据的外模式

设计物理结构

根据DBMS特点和处理的需要,  
进行物理存储安排, 建立索引,  
形成数据库内模式

性能评价预测

物理实现

利用具体DBMS系统建立数据库,  
编写调试应用程序, 组织数据入  
库, 试运行

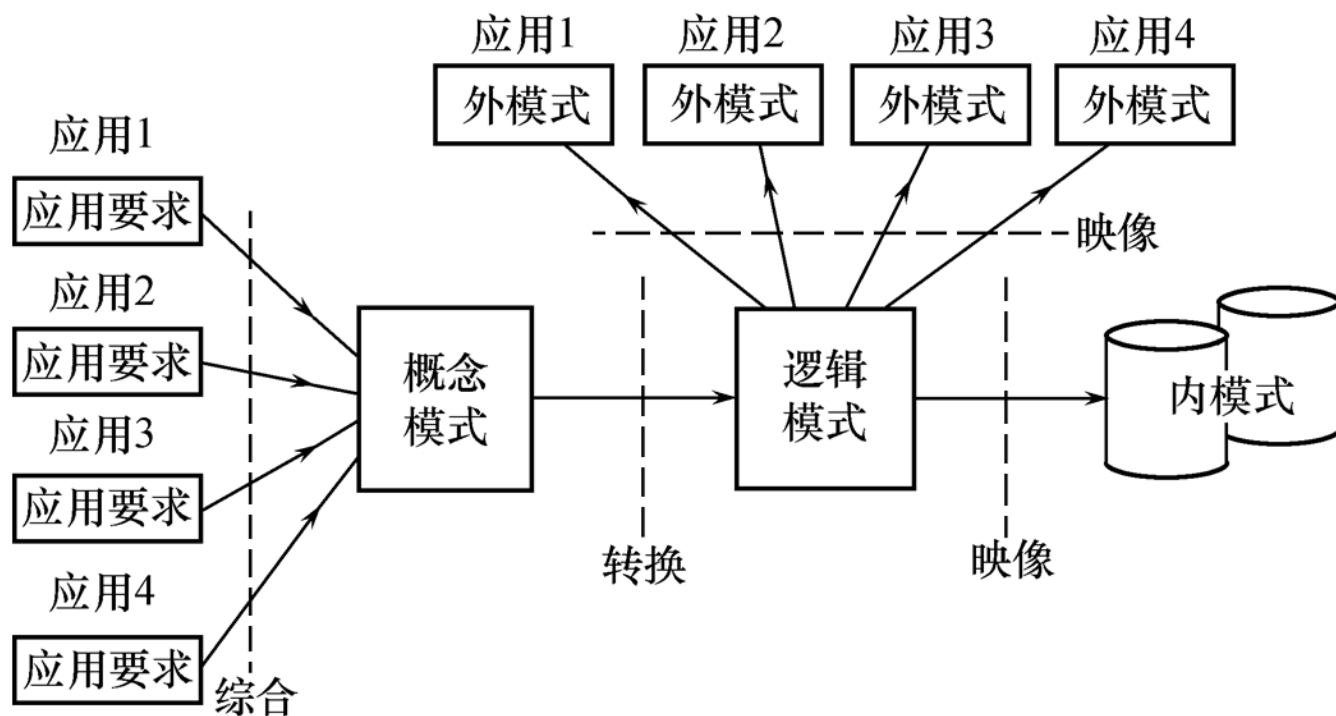
数据库运行维护

在数据库运行过程中, 不断对其  
评估、调整与修改





# 数据库设计过程中的各级模式



数据库的各级模式



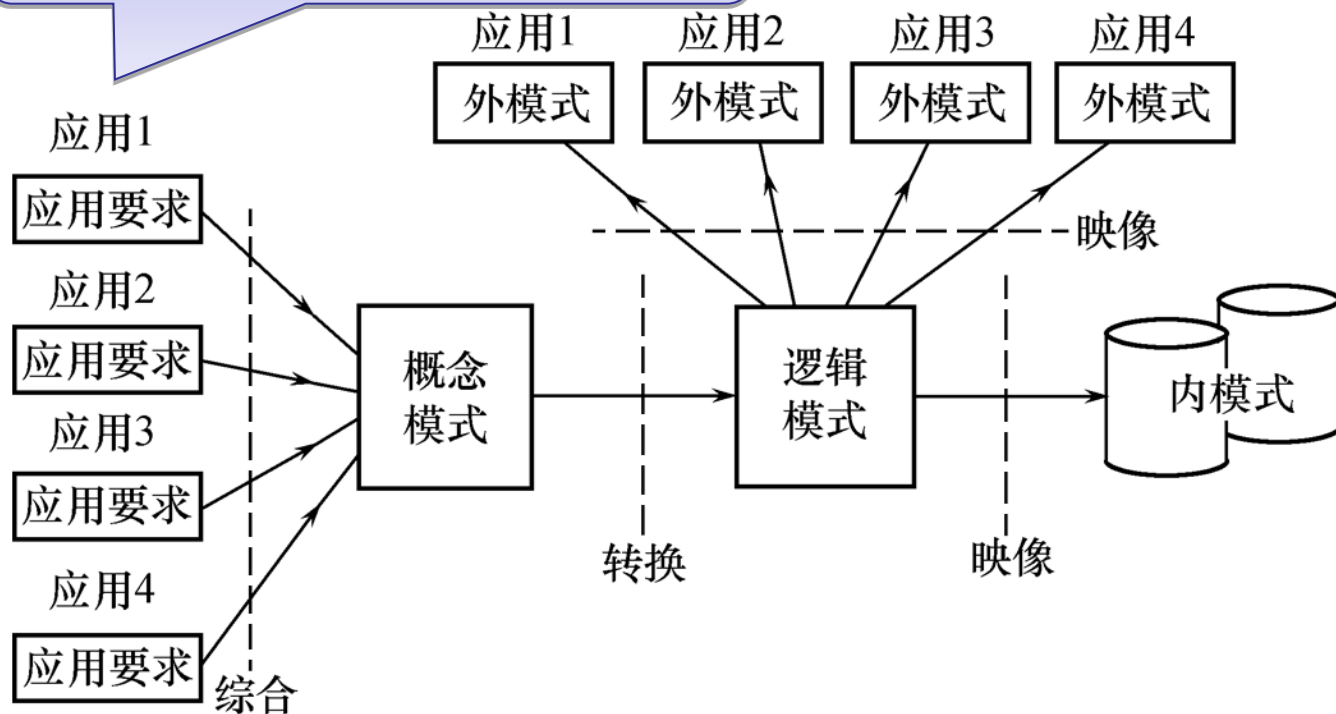




# 数据库设计过程中的各级模式

需求分析阶段：

综合各个用户的应用需求



数据库的各级模式

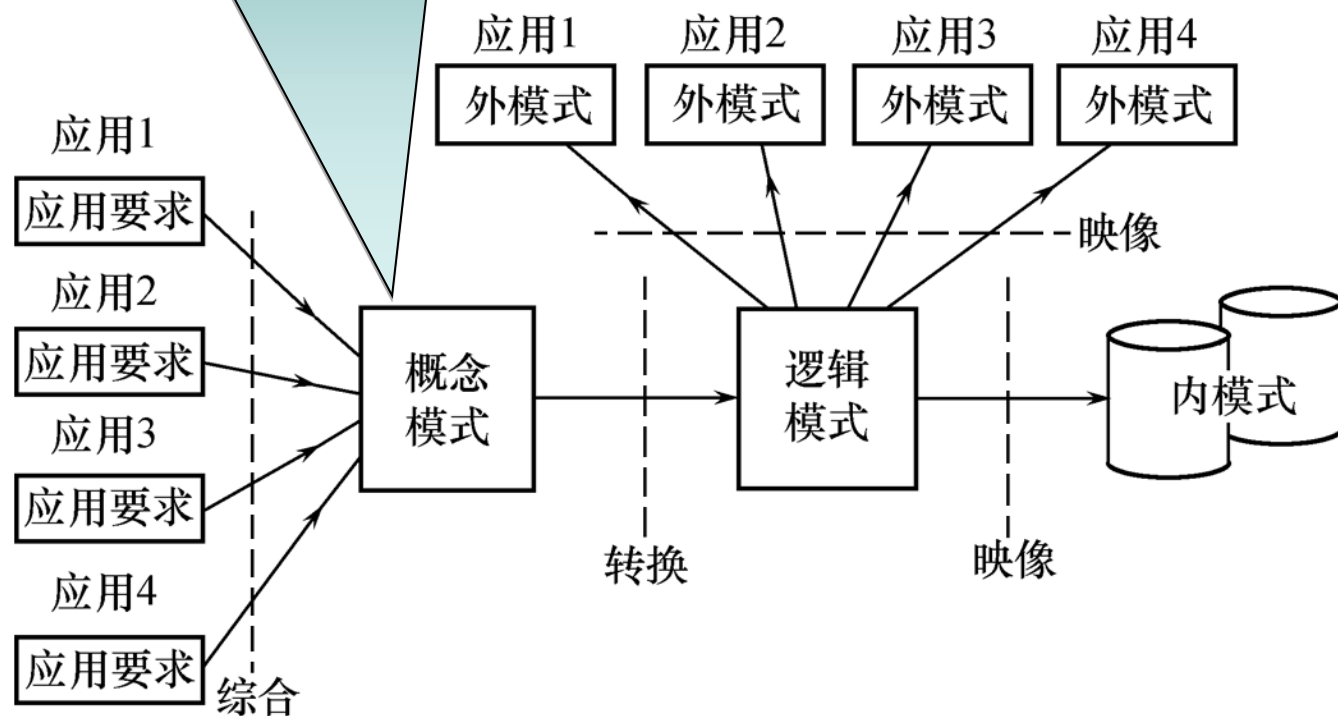




# 程中的各级模式

概念设计阶段：

形成独立于机器特点，独立于各个数据库管理系统产品的**概念模式**（E-R图）

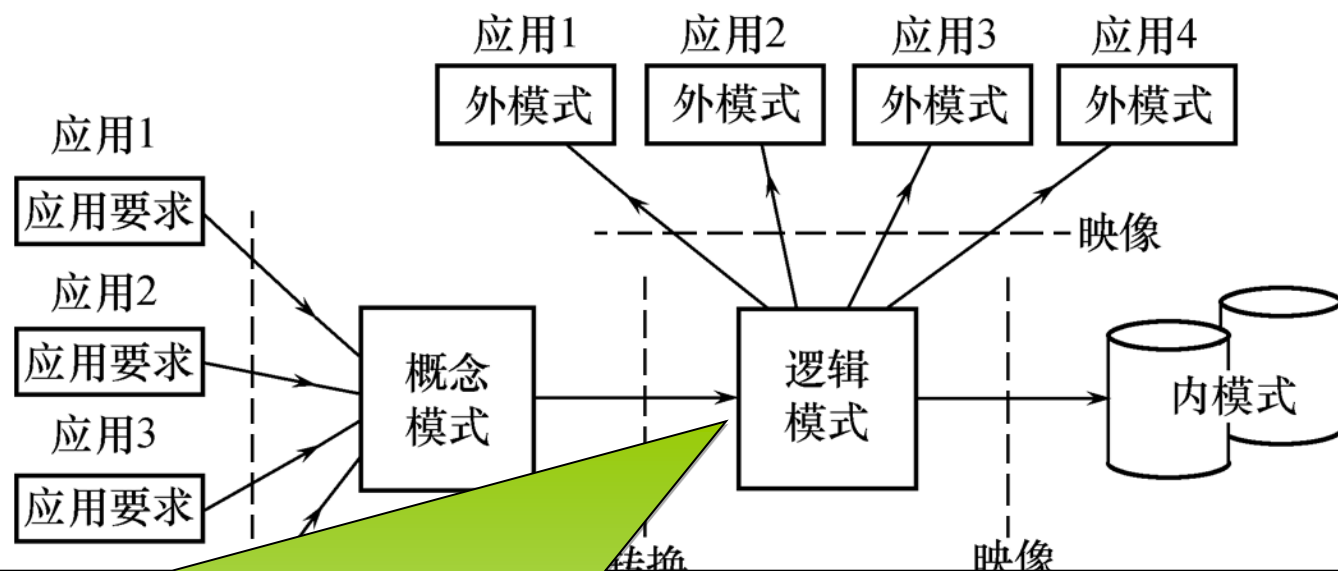


数据库的各级模式





# 数据库设计过程中的各级模式



## 逻辑设计阶段:

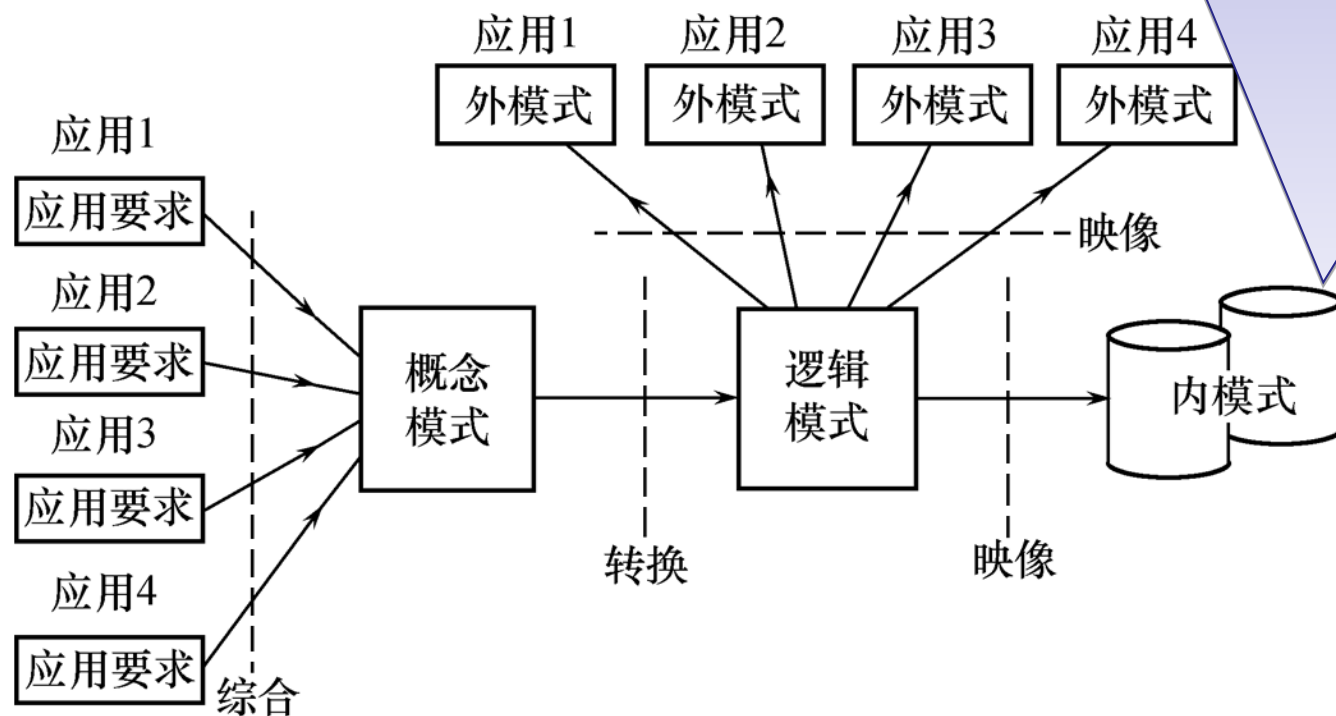
1. 首先将E-R图转换成具体的数据库产品支持的数据模型，如关系模型，形成数据库**逻辑模式**
2. 然后根据用户处理的要求、安全性的考虑，在基本表的基础上再建立必要的视图 (View)，形成数据的**外模式**



# 数据库设计过程中的各级模式

物理设计阶段：

根据数据库管理系统特点和处理的需要，进行物理存储安排，建立索引，形成数据库内模式



数据库的各级模式





# 需求分析

- 需求分析就是分析用户的要求
  - 是设计数据库的起点
  - 结果是否准确地反映了用户的实际要求，将直接影响到后面各个阶段的设计，并影响到设计结果是否合理和实用





# 需求分析

- 需求分析的任务

- 详细调查现实世界要处理的对象（组织、部门、企业等）
- 充分了解原系统（手工系统或计算机系统）工作概况
- 明确用户的各种需求
- 在此基础上确定新系统的功能
- 新系统必须充分考虑今后可能的扩充和改变





# 需求分析

- 需求分析的重点是

- “数据”和“处理”，获得用户对数据库要求

- 信息要求

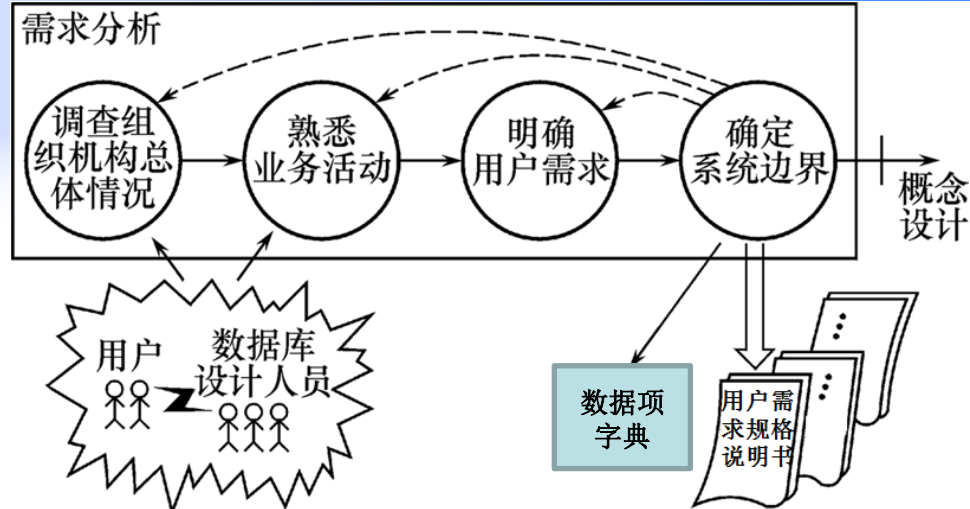
- 用户需要从数据库中获得信息的内容与性质
- 由用户的信息要求可以导出数据要求，即在数据库中需要存储哪些数据

- 处理要求

- 用户要完成什么处理功能
- 对处理的响应时间的要求
- 对处理方式的要求(批处理 / 联机处理)

- 安全性与完整性要求





## • 需求分析的步骤

### — 应用领域的调查分析

- 组织结构、业务活动
- 目标：协助用户明确要求
- 确定新系统的边界

### — 定义数据库支持的信息与应用

### — 定义数据库操作任务

### — 定义数据项字典

- 关于数据库中数据的描述，即元数据

- 数据项、数据结构、数据流、数据存储、处理过程

### — 预测应用领域的未来改变

#### 常用调查方法：

跟班作业  
开调查会  
专人介绍  
询问  
调查问卷  
查询记录





# 需求分析

## • 数据项字典的组成

### — 数据项

完整性约束条件

函数依赖

- 不可再分的数据单位
- 数据项描述：数据项名、数据项含义说明、别名、数据类型、长度、取值范围、取值含义、与其他数据项的逻辑关系、数据项之间的联系

### — 数据结构

- 反映数据之间的组合关系，由数据项和其他数据结构组成

数据结构描述：数据结构名、含义说明、组成





# 需求分析

## • 数据项字典的组成（续）

### — 数据流

- 数据结构在系统内传输的路径
- 数据流描述：数据流名、说明、数据流来源、数据流去向、组成（包括哪些数据结构）、平均流量、高峰期流量

### — 数据存储

- 数据结构停留或保存的地方，数据流的来源和去向之一。可以是手工文档、凭单或计算机文档

数据存储描述：数据存储名、说明、编号、输入的数据流、输出的数据流、组成、数据量、存取频度、存取方式





# 需求分析

## • 数据项字典的组成（续）

### — 处理过程

- 具体处理逻辑一般用判定表或判定树来描述，数据项字典中只描述其说明性信息
- 处理过程描述：处理过程名、说明、输入（数据流）、输出（数据流）、处理（简要说明，包括功能和要求）





## 4.2 概念数据库设计





# Outline

- 概述
- 实体联系模型
- 概念设计的方法与策略
- 视图综合设计方法
- 事务的设计





## 4.2.1 概述

- 概念数据库设计的任务包括两方面

- 概念数据库模式设计

- 以需求分析阶段所识别的数据项和应用领域的未来改变信息为基础，使用高级数据模型建立概念数据库模式

- 事务设计

- 考察需求分析阶段提出的数据库操作任务，形成数据库事务的高级说明





## • 概念数据库模式设计的目标

- 准确描述应用领域的信息模式，支持用户的各种应用
- 既易于转换为逻辑数据库模式，又容易为用户理解

概念模型：用于将需求分析阶段所得到的应用需求抽象为信息世界的结构

概念模型的特点：

- 能真实、充分地反映现实世界，包括事物和事物之间的联系
- 易于理解，可用于和不熟悉计算机的用户交换意见
- 易于改变，应用环境和应用要求改变时容易对概念模型进行修改
- 易于向关系、网状、层次模型转换

概念模型是各种数据模型的基础，更加独立于DBMS和机器（更高级）



- 概念数据库模式独立于任何数据库管理系统，不能直接用于数据库的实现。
- 用于概念数据库设计的高级数据模型：**实体联系模型**







# Outline

- 概述
- **实体联系模型**
- 概念设计的方法与策略
- 视图综合设计方法
- 事务的设计





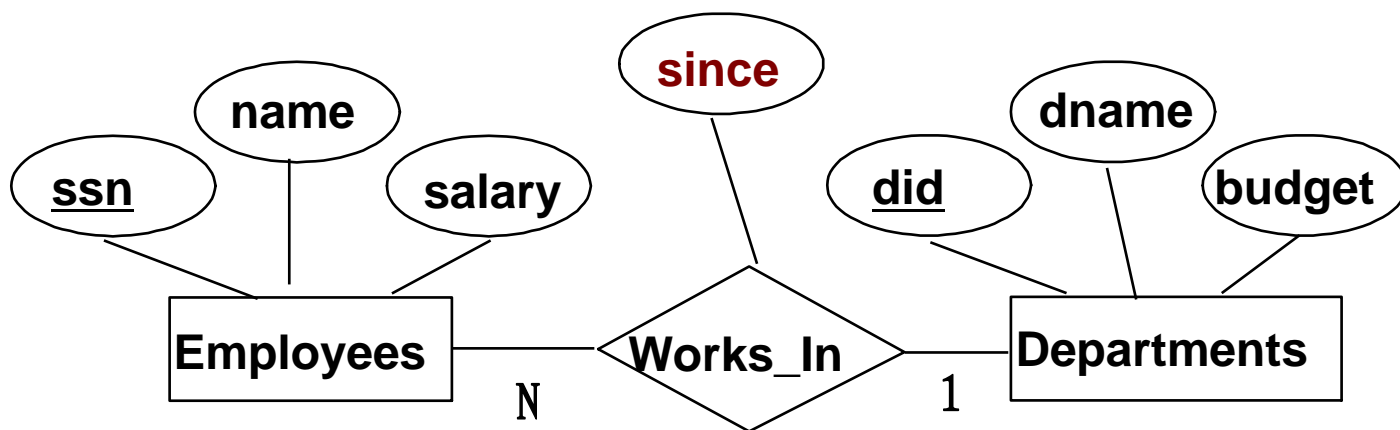
- 实体联系模型

- Entity-Relationship Model, 简称ER模型

- 表示成“实体-联系”图

- 三个主要元素

- 实体
- 属性
- 联系





## 4.2.2 实体联系模型

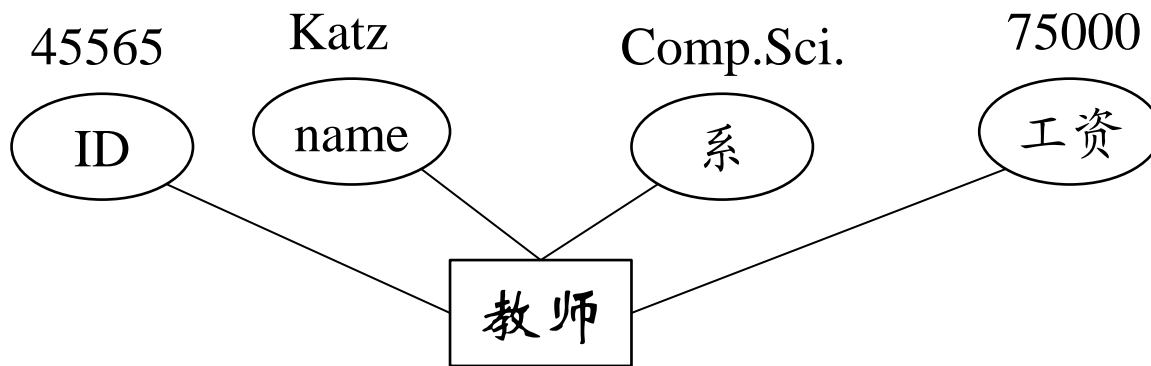
### • 基本概念

- **实体**是ER模型的基本对象。实体是现实世界中各种事物的抽象。
  - 实体可以是物理存在的事物，如人、汽车；
  - 也可以是抽象的概念，如学校、课程。
- 每个实体都有一组特征或性质，称为实体的**属性**。实体属性的一组特定值确定了一个特定的实体。**实体的属性值**是数据库中存储的主要数据。
  - 例如，学生实体具有名字、年龄、性别等属性





- 基本概念(续)
  - 实体例子





- 基本概念(续)

- 实体集

- 是相同类型(即具有相同性质或属性)的实体集合

- 例如, 某个大学所有教师的集合可被定义为实体集*Teacher*.

- 实体集不必互不相交





ID	name	dept_name	salary
76766	Crick	Biology	72000
45565	Katz	Comp. Sci.	75000
10101	Srinivasan	Comp. Sci.	65000
83821	Brandt	Comp. Sci.	92000
98345	Kim	Elec. Eng.	80000
12121	Wu	Finance	90000
76543	Singh	Finance	80000
32343	El Said	History	60000
58583	Califieri	History	62000
15151	Mozart	Music	40000
33456	Gold	Physics	87000
22222	Einstein	Physics	95000

实体集: *Teacher, student*

学号 Sno	姓名 Sname	性别 Ssex	年龄 Sage	所在系 Sdept
200215121	李勇	男	20	CS
200215122	刘晨	女	19	CS
200215123	王敏	女	18	MA





# 基本概念(属性)

- 基本概念(续)

- 属性(Attribute)

- 实体的特征或性质
    - 是实体集映射到域的函数

- 属性可包括

- 单值(single-valued)、多值(multivalued)属性
    - 简单(simple)、复合(composite)属性
    - 派生(derived)属性 (导出属性)





# 基本概念(属性)

## • 基本概念(续)

### - 实体属性

- 多数实体属性都是**单值属性**，即对于同一个实体只能取一个值。
  - 例如，同一个人只能具有一个年龄，所以人的年龄属性是一个单值属性。
- 但是，在某些情况下，实体的一些属性可能取多个值。这样的属性称为**多值属性**。
  - 例如：职务、联系方式等





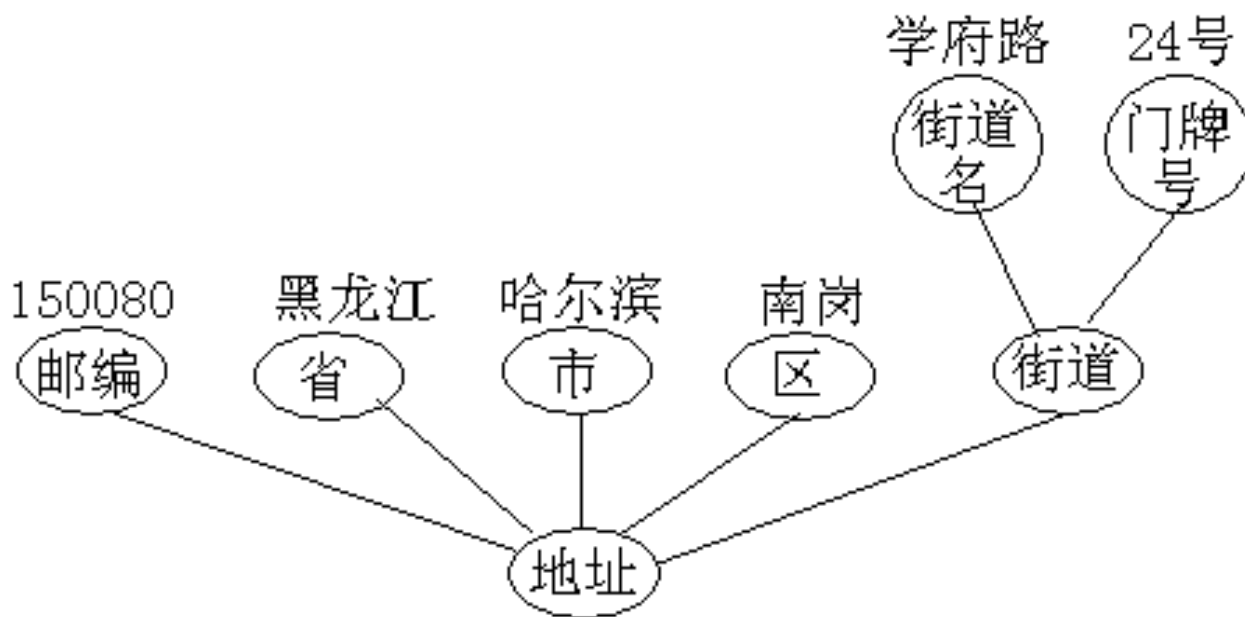


# 基本概念(属性)

- 基本概念(续)

- 实体属性

- 可以划分为多个具有独立意义的子属性。我们称这类属性为**复合属性**。复合属性具有层次结构





# 基本概念(属性)

- 基本概念(续)

- 实体属性

- 导出属性

- 导出属性的值
      - 从当前日期和生日属性的值可以确定年龄属性的值，即年龄属性的值可以由其他属性导出。我们称这样的属性为**导出属性**。
      - 导出属性的值不仅可以从另外的属性导出，也可以从有关的实体导出。例如，一个公司实体的雇员数属性的值可以通过累计该公司所有雇员数得到。

- 空值

- 在某些情况下，实体的有些属性可能没有适当值可设置。这些属性通常被设置一个称为**空值**的特殊值。
      - 例如，一个未获得任何学位的人的学位属性只能被设置为空值





# 基本概念(属性)

- 基本概念(续)

- 码：在ER模型中每个实体集具有一个由一个或多个属性组成的**码**，用来区别不同的实体。

- 例如，由于不同学生不能具有相同的学号，学生实体集的学号属性是键。
    - **超码**
    - **候选码**
    - **主码**





- 基本概念(续)

- 联系

- 不同实体集之间可能具有某种关联，我们称这种关联为实体间的联系

- 例如，学生实体集和系实体集；学生实体集与课程实体集

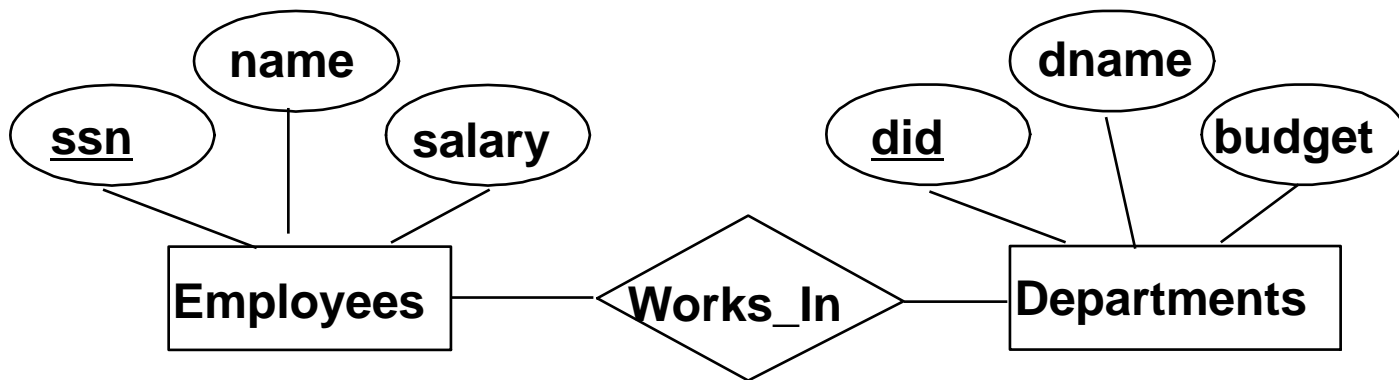




## • 基本概念(续)

### — 联系集

#### • 同类联系的集合



称一个联系集所关联的实体集的数量为这个联系集的阶。阶为 $n$ 的联系集称为 $n$ 元联系集

如果 $E_1, \dots, E_n$ 是实体集，那么联系集 $R$ 可以表示为

$$R = \{(e_1, \dots, e_n) | e_1 \in E_1, \dots, e_n \in E_n\}$$

实体集 $E_1, \dots, E_n$ 参与联系集 $R$





## • 基本概念(续)

— 实体之间的联系既可以使用联系集定义，也可以通过实体属性来表示。

• 例如，可以用属性表示实体教研室和系之间的所属联系。

— 用教研室实体的属性“所属系”的值来表示这个实体所属的系实体。

— 如何在系实体里表示教研室和系之间的所属关系？

— 单值属性？

— 多值属性？

— 复合属性？





## • 基本概念(续)

### — 实体之间的联系 (3类映射基数)

#### 映射基数

一个实体通过一个联系能关联的实体的个数

#### • 一对一

— 一个系只有一个系主任，一个大学只有一个校长

#### • 一对多 (或多对一)

— 一个系有很多学生

#### • 多对多

— 一个学生可选修多门课程，每门课程可被多个学生选修

#### • 参与约束

- 如果实体集E中的每个实体都参与到联系集R中的至少一个联系中，实体集E在联系集R中的参与是全部的

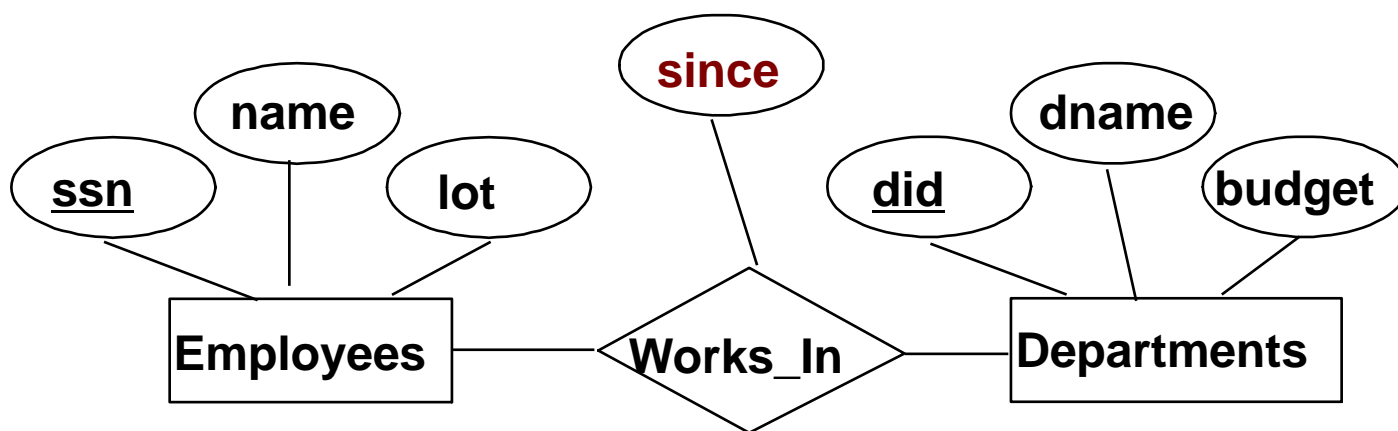
- 如果E中只有部分实体参与到R的联系中，实体集E到联系集R的参与是部分的





- 基本概念(续)

— 联系集的属性：描述性属性







## • 基本概念(续)

### — 弱实体

- 现实世界中存在这样的一些实体集，它们没有足够的属性形成自己的**主码**。为了区分各个实体，它们必须与其它实体集相关联。这样的实体集称为弱实体集合

— 弱实体集中的不同实体的属性可能完全相同

- 与弱实体集相关联的实体集：**识别实体集(主实体集)**。
- 主实体集与它的弱实体集之间的联系称为**识别联系**。
- 主实体与弱实体的关系都为**一对多**的关系，弱实体为**多方**。

- 一个主实体对应的多个弱实体之间可以相互区别



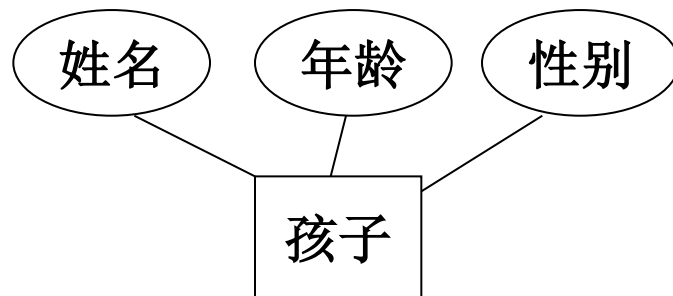


## • 基本概念(续)

### — 弱实体

#### • 例：父亲实体集与孩子实体集

- 不同父亲的孩子可以具有相同的姓名、年龄和性别
- 同一个父亲的孩子一定具有不同的名字
- 显然，孩子实体集是弱实体集





## • 基本概念(续)

### — 弱实体

- 弱实体集必须具有一个或多个属性，使得这些属性可以与主实体集的主码相结合，形成相应弱实体集的主码
- 这样的弱实体属性称为弱实体集的部分码
  - 给定一个弱实体集，可以使用它的主实体集的码和它的部分码识别不同的弱实体。
  - 例如，上例中的孩子名是弱实体集孩子的部分码



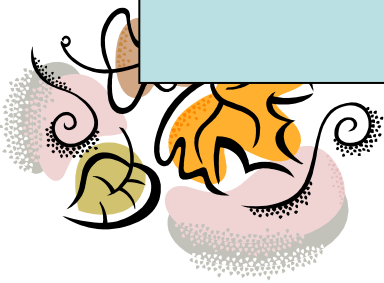
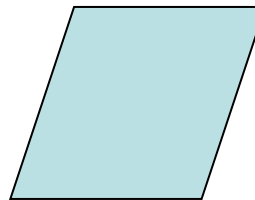
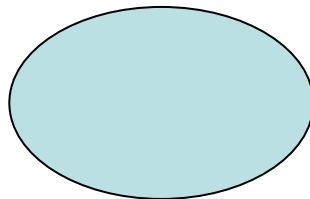


## • 实体联系图

— 是表示ER模型的图形工具，简称**ER图**

— ER图用来表示实体集和实体联系集

- 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。





# ER 图中 基本 符号

1. 实体集:

2. 弱实体集:

3. 联系集:

4. 识别联系集:

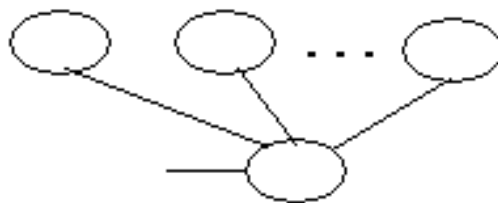
5. 属性:

6. 主属性:

7. 多值属性:

8. 导出属性:

8. 复合属性:



10. B具有全域实体关联约束: 参与约束(B到R的参与是全部的)

11. 1:1联系集:

12. 1:N联系集:

13. M:N联系集:

映射基数

14. 实体集E上的结构约束:

基数约束

E中一个实体最少与D中0个实体存在联系, 最多和D中无穷多个实体存在联系

基数约束

D中一个实体最少与E中min个实体存在联系, 最多和E中max个实体存在联系

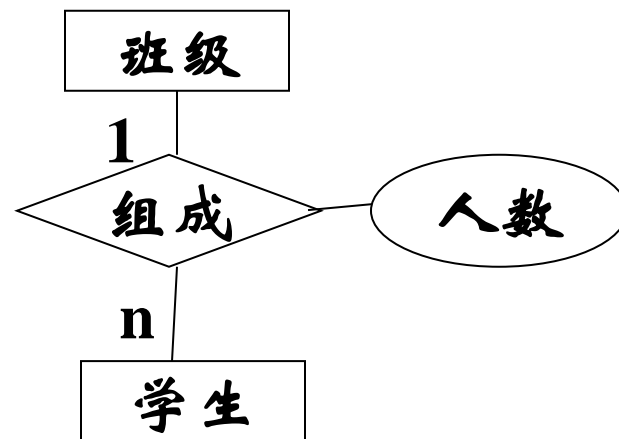
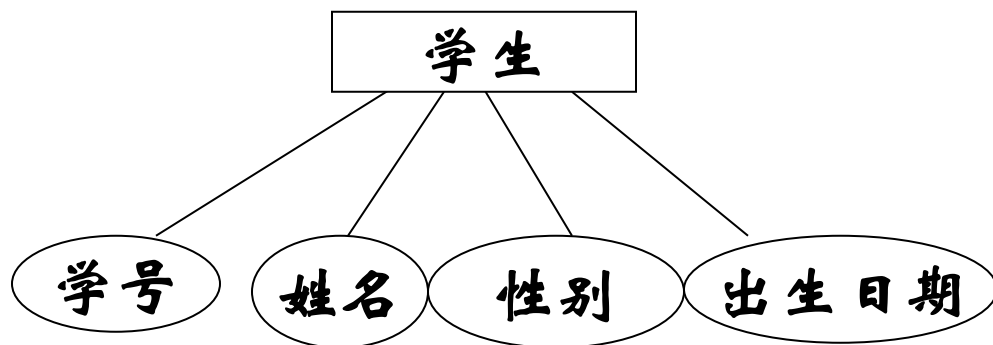




实体名

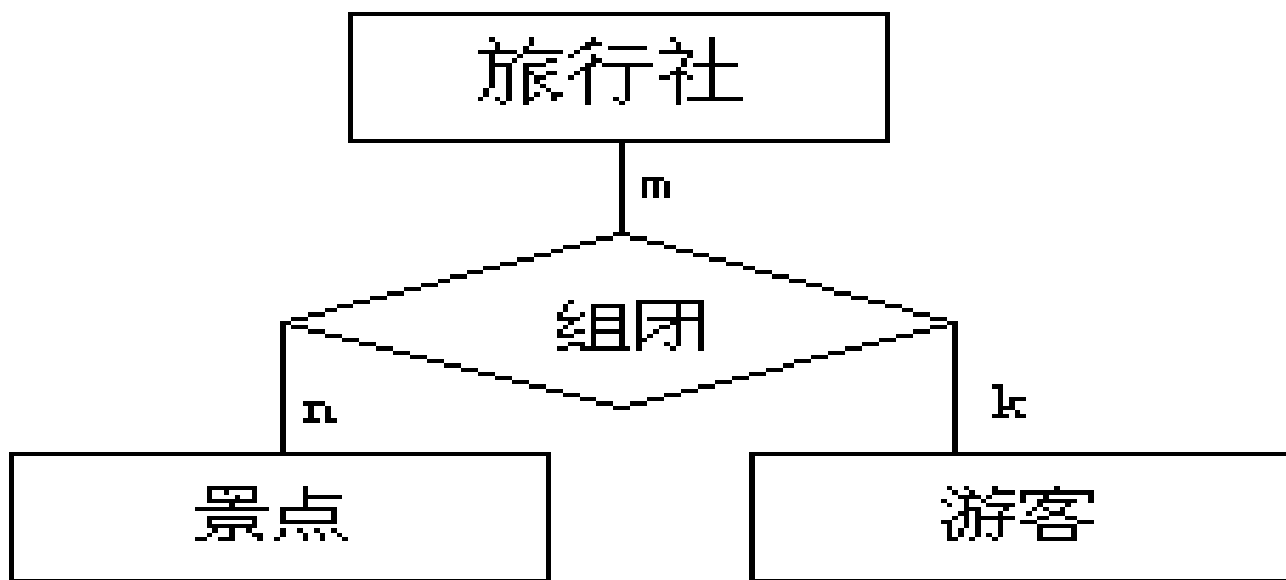
属性名

联系名





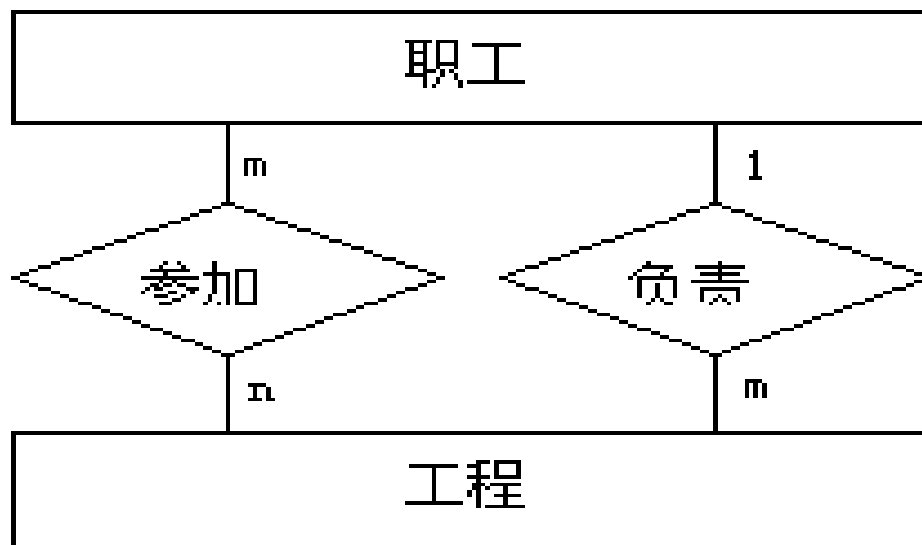
- 两个不同实体集间的联系
- 两个以上实体间的多元联系





## • 两个不同实体集间的多种联系

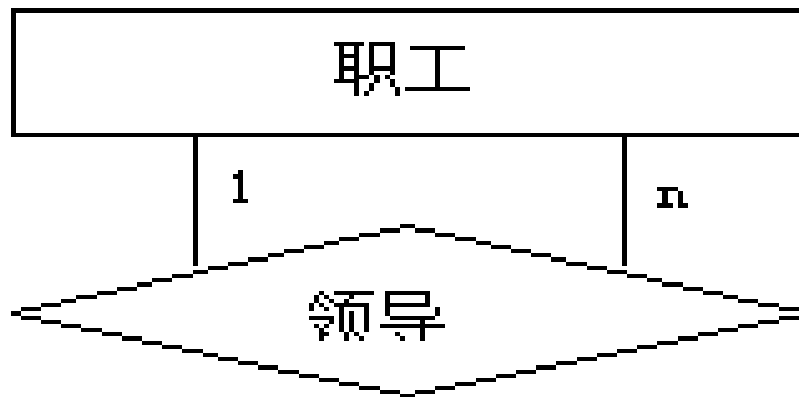
- 职工与工程间，一个职工可以参加多个工程，一个工程可以有多个职工参加，同时一个工程由一个职工负责，一个职工可以负责多个工程







- 同一实体内部个体间的二元联系





## • ER模型设计实例

根据财经学院数据库应用需求，运用E-R模型的基本方法设计财经学院教学数据库的概念模型。

1、根据需求调查，初步确定所关心的数据对象：

学院、系、教师、班级、学生、课程、成绩

2、根据业务规则，设计初步E-R模型

(1) 学院有多个系，每个系只能属于一个学院。

(2) 每个系有多个班级，而每个班级只能属于一个系。

(3) 每个系聘任多名教师，而每个教师又只能属于一个系。

(4) 每个班级有多个学生，每个学生只能属于一个班级。

(5) 每个教师讲授多个课程，每个课程只由一个教师讲授。

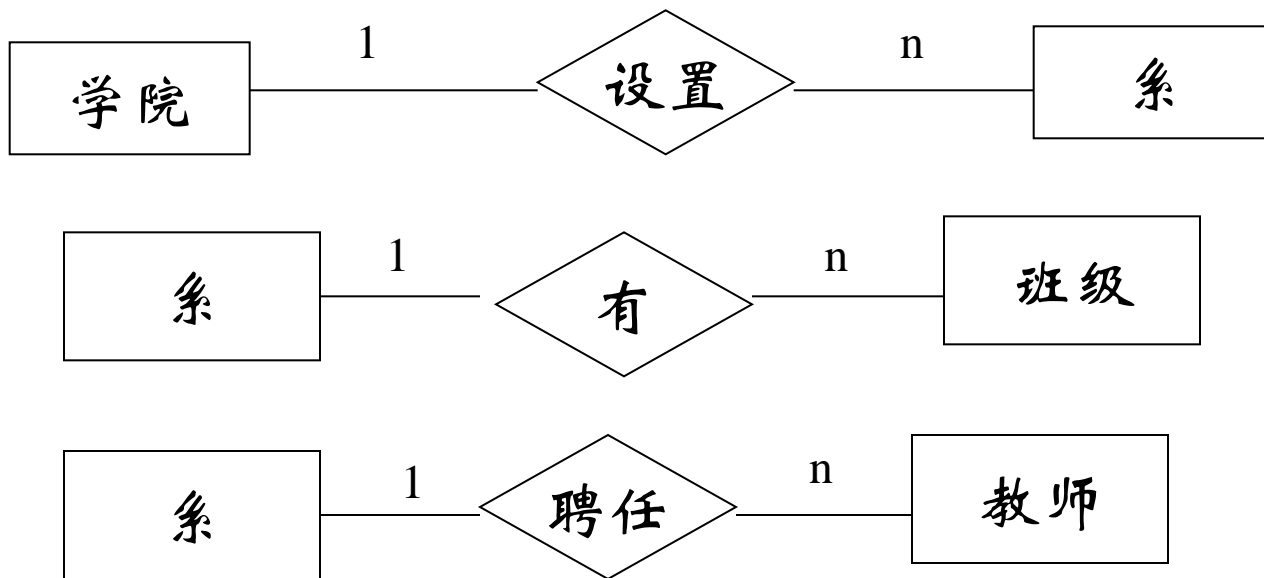
(6) 每个学生选修多个课程，每个课程由多个学生选修。





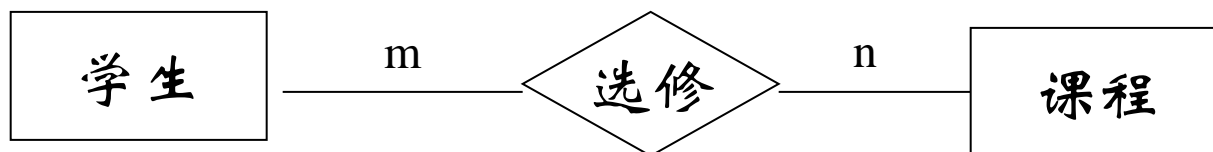
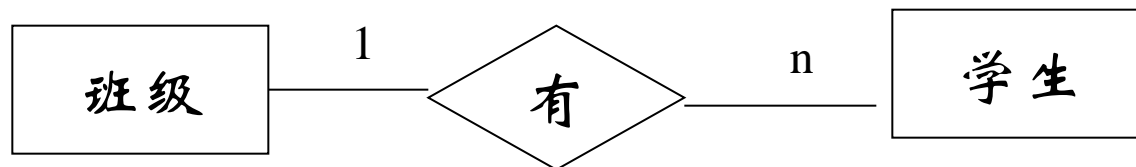
- (1) 学院有多个系，每个系只能属于一个学院。
- (2) 每个系有多个班级，而每个班级只能属于一个系。
- (3) 每个系聘任多名教师，而每个教师又只能属于一个系

归纳上述6项可定义6个实体:学院、系、班级、教师、学生、课程。



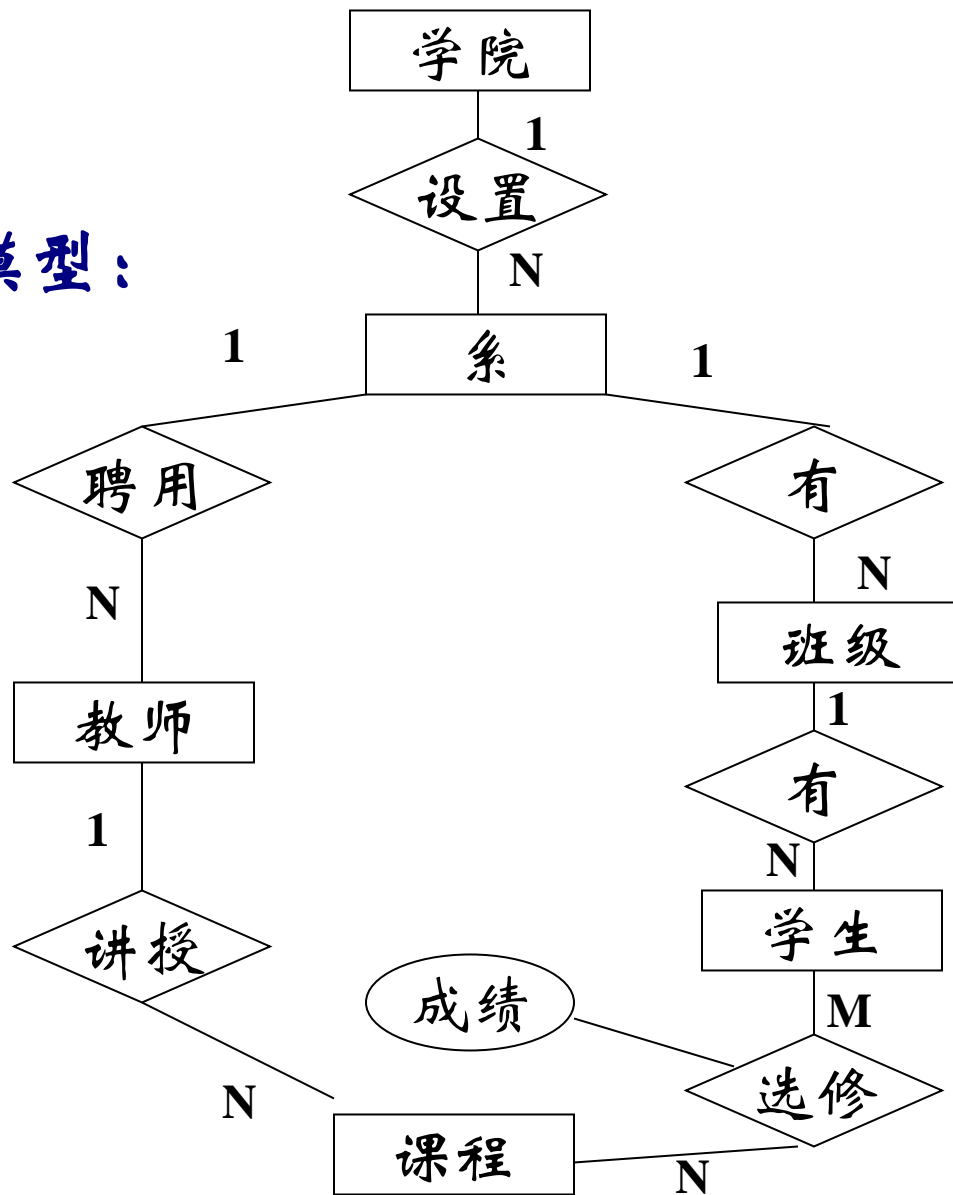


- (4) 每个班级有多个学生，每个学生只能属于一个班级。  
(5) 每个教师讲授多个课程，每个课程只由一个教师讲授。  
(6) 每个学生选修多个课程，每个课程由多个学生选修。





## 整体E-R模型：



# • 实例

- 用E-R图来表示某个工厂物资管理的概念模型
- 物资管理涉及的实体有
  - 仓库 属性有仓库号、面积、电话号码
  - 零件 属性有零件号、名称、规格、单价、描述
  - 供应商 属性有供应商号、姓名、地址、电话号码、账号
  - 项目 属性有项目号、预算、开工日期
  - 职工 属性有职工号、姓名、年龄、职称

这些实体之间的联系如下：

(1) 一个仓库可以存放多种零件，一种零件可以存放在多个仓库中，因此仓库和零件具有多对多的联系。用库存量来表示某种零件在某个仓库中的数量。

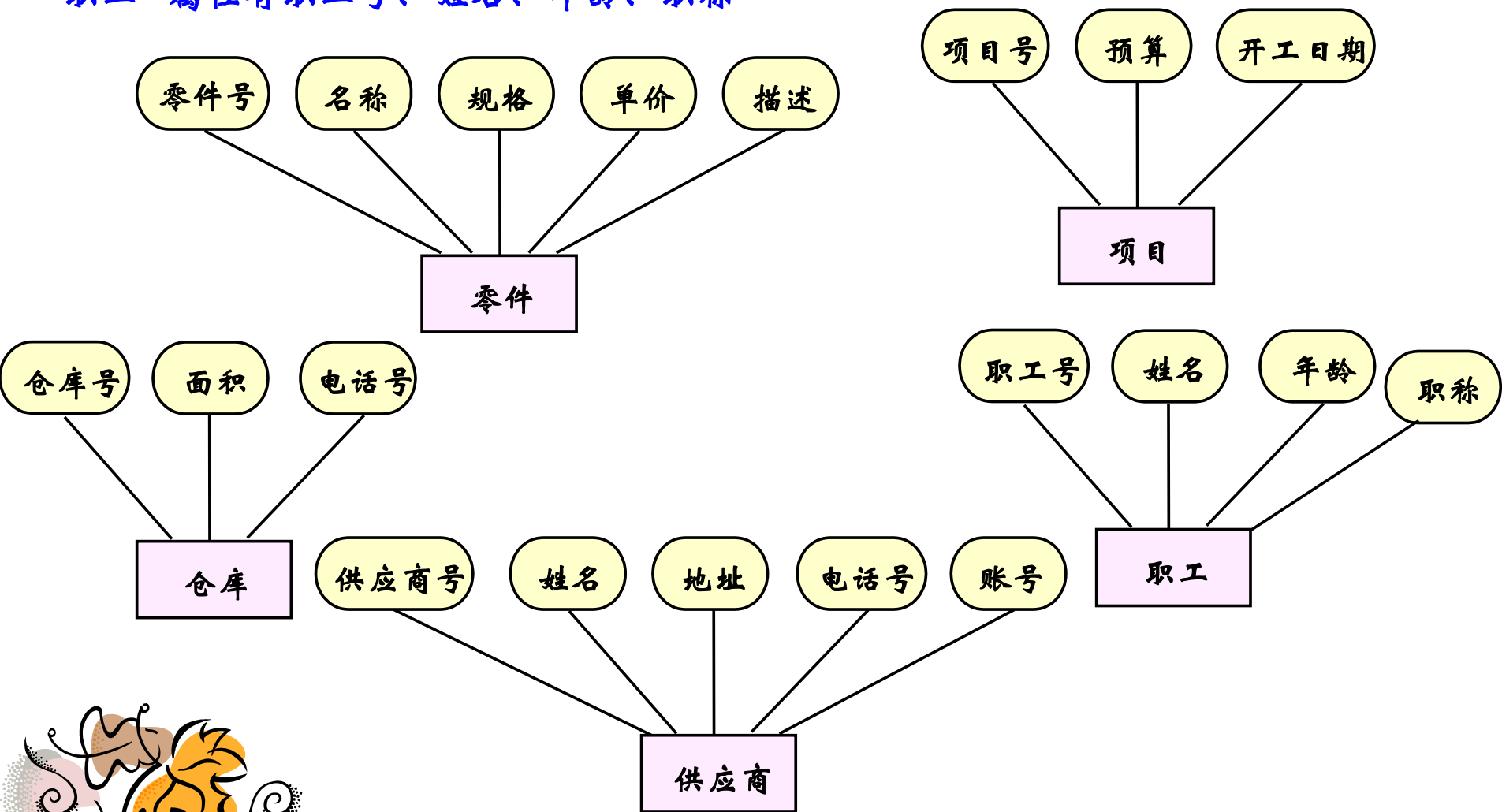
(2) 一个仓库有多个职工当仓库保管员，一个职工只能在一个仓库工作，因此仓库和职工之间是一对多的联系

(3) 职工之间具有领导-被领导关系。即仓库主任领导若干保管员，因此职工实体集中具有一对多的联系。

(4) 供应商、项目和零件三者之间具有多对多的联系。即一个供应商可以供给若干项目多种零件，每个项目可以使用不同供应商供应的零件，每种零件可由不同供应商供给。



- 仓库 属性有仓库号、面积、电话号码
- 零件 属性有零件号、名称、规格、单价、描述
- 供应商 属性有供应商号、姓名、地址、电话号码、账号
- 项目 属性有项目号、预算、开工日期
- 职工 属性有职工号、姓名、年龄、职称

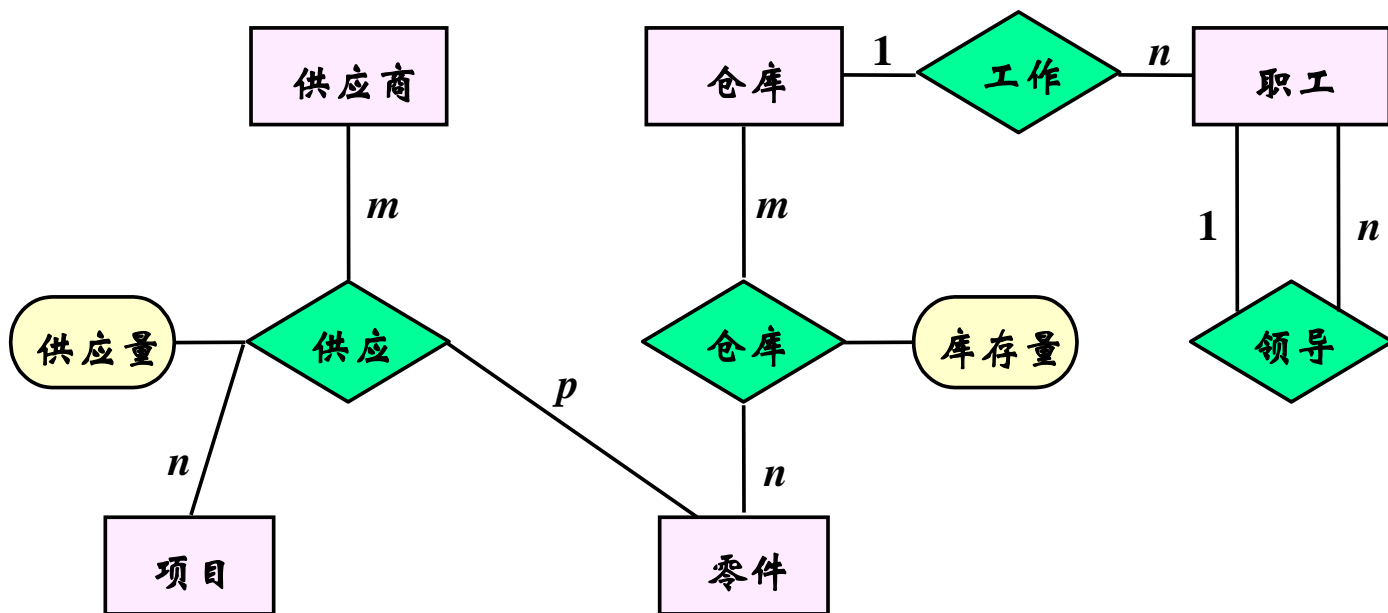


(a) 实体及其属性图

## 实体之间的联系:

- (1) 一个仓库可以存放多种零件，一种零件可以存放在多个仓库中，因此仓库和零件具有多对多的联系。用库存量来表示某种零件在某个仓库中的数量。
- (2) 一个仓库有多个职工当仓库保管员，一个职工只能在一个仓库工作，因此仓库和职工之间是一对多的联系

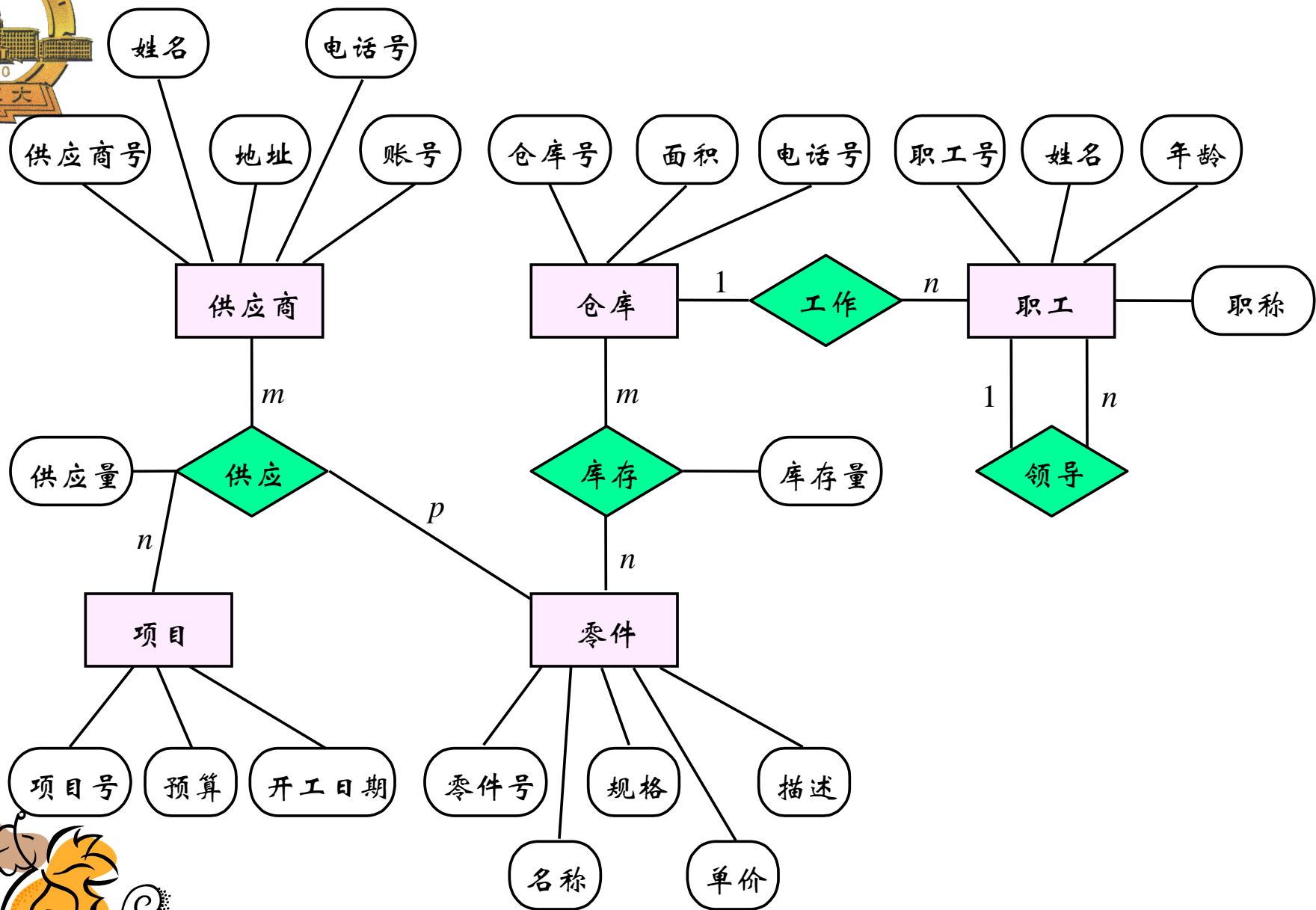
- (3) 职工之间具有领导-被领导关系。即仓库主任领导若干保管员，因此职工实体集中具有一对多的联系。
- (4) 供应商、项目和零件三者之间具有多对多的联系。即一个供应商可以供给若干项目多种零件，每个项目可以使用不同供应商供应的零件，每种零件可由不同供应商供给。



(b) 实体及其联系图







(c) 完整的实体联系图



# • E-R模型的扩展及其E-R图表示

## – ISA联系

- 一个实体集是另一个实体集的子类型

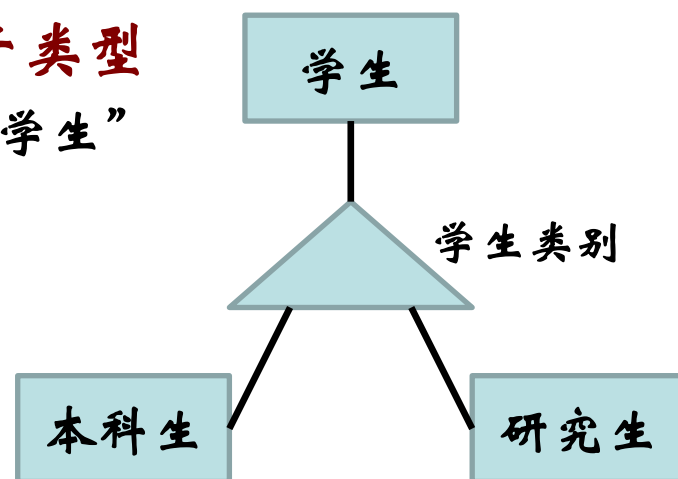
- “本科生”、“研究生”之于“学生”
- ISA联系用三角形表示
- 特化（父到子）、概化（子到父）

- 子类继承父类的所有属性

- 学号、系...

- 子类可以有自己的属性

- 研究生：导师姓名、...



- 分类属性

- 根据分类属性的值把父实体集的实体分派到子实体集中
  - » 学生类别=‘本科生’、学生类别=‘研究生’





# • E-R模型的扩展及其E-R图表示

## – ISA联系

### • 不相交约束、可重叠约束

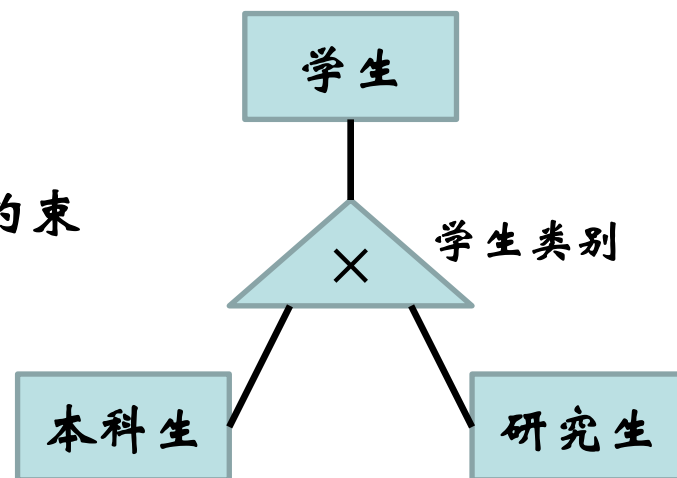
- 不相交约束描述父实体集中的一个实体不能同时属于多个子实体集,即父实体集中的一个实体最多属于一个子实体集

» 不相交约束用三角形和其中的一个×表示

- 可重叠约束允许父实体集中的一个实体能同时属于多个子实体集

» 三角形内没有×

不相交约束



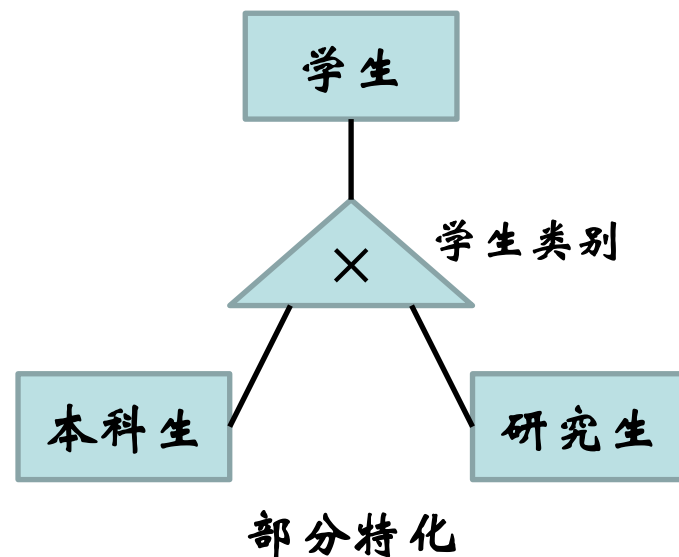
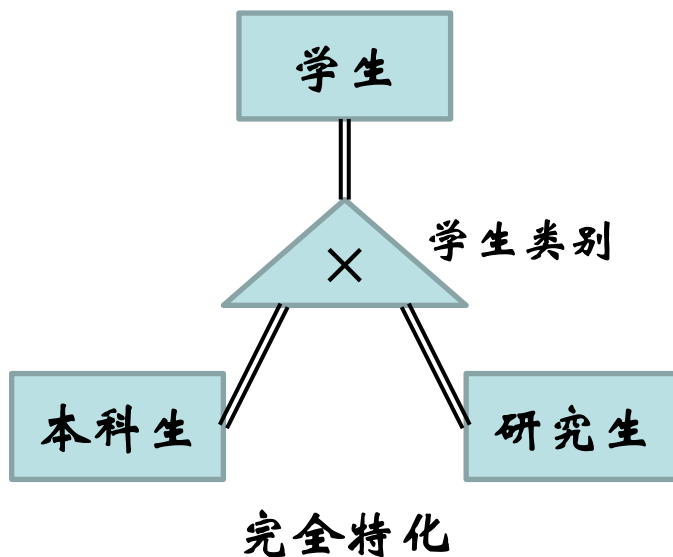


# • E-R模型的扩展及其E-R图表示

## – ISA联系

### • 完备性约束

- 完全特化：父实体集中的一个实体必须是属于某一个子实体集中的实体
- 部分特化：父实体集中的存在实体不属于任何子实体集





# • E-R模型的扩展及其E-R图表示

## — 基数约束

- 对一对一、一对多和多对多的细化
- 参与联系的每个实体集用基数约束说明其中任何一个实体可以在联系中出现的最小次数和最大次数
  - 约束用一对min、max表示,  $0 \leq \min \leq \max$
  - 0..1、1..3、1..\*(\*代表无穷大)
  - min=1的约束为强制参与约束, 对应实体集在相应联系集的参与是全部的
  - min=0的约束为非强制参与约束, 对应实体集在相应联系集的参与是部分的

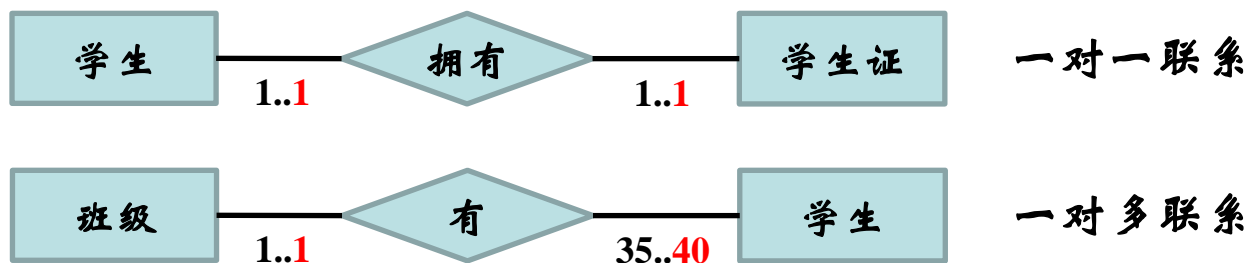
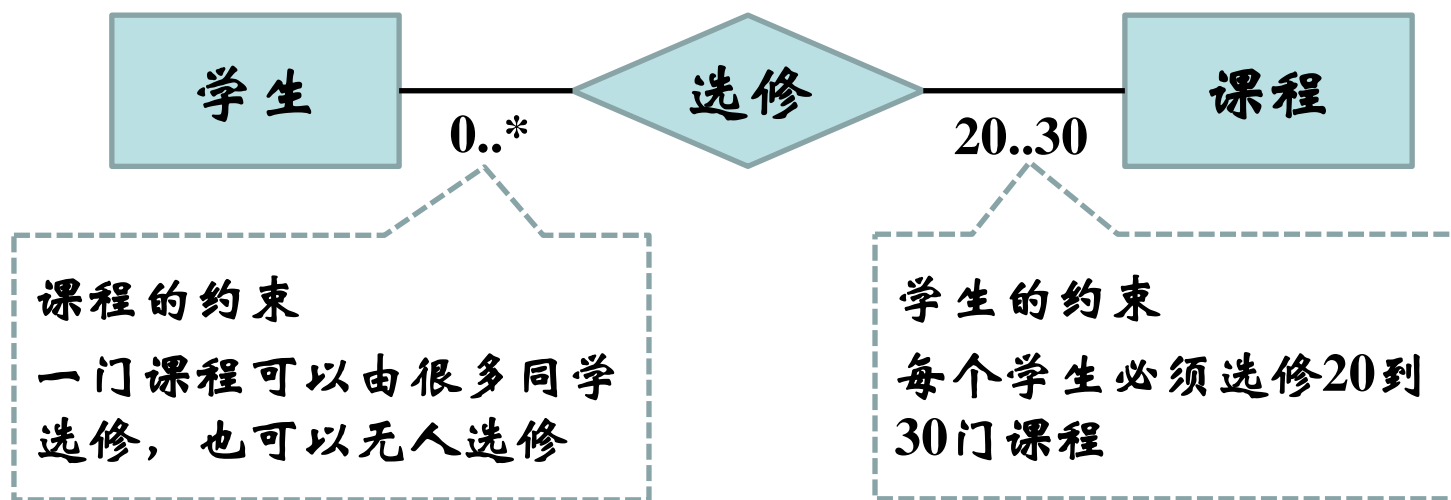




# • E-R模型的扩展及其E-R图表示

## — 基数约束

### • 远端标注基数约束





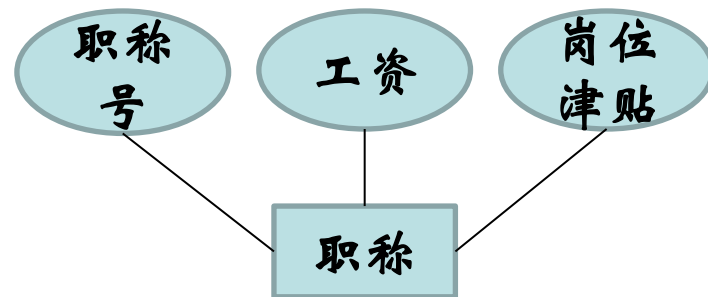
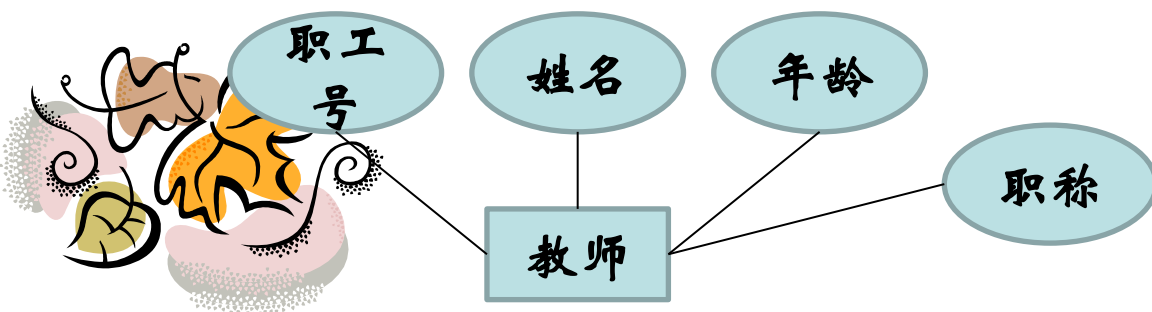
## • E-R模型设计中的一些讨论

### — 实体集和联系集的定义方式并不精确

- 定义一组实体和它们的相互联系可能有多种不同方式
- 个人选择 / 基本原则

### — 实体集 or 属性

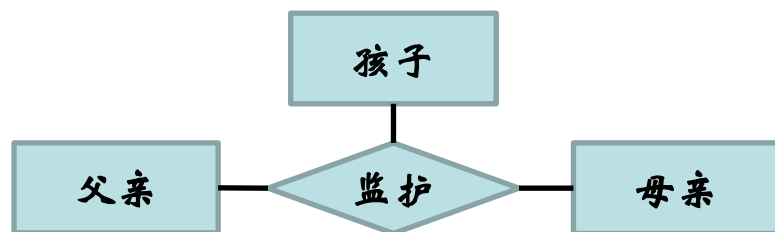
- 为了简化E-R模型，现实世界中的事物能用作属性对待的尽量用作属性
  - 属性不要有需要描述的性质
  - 属性不能与其他实体有联系





# • E-R模型设计中的一些讨论

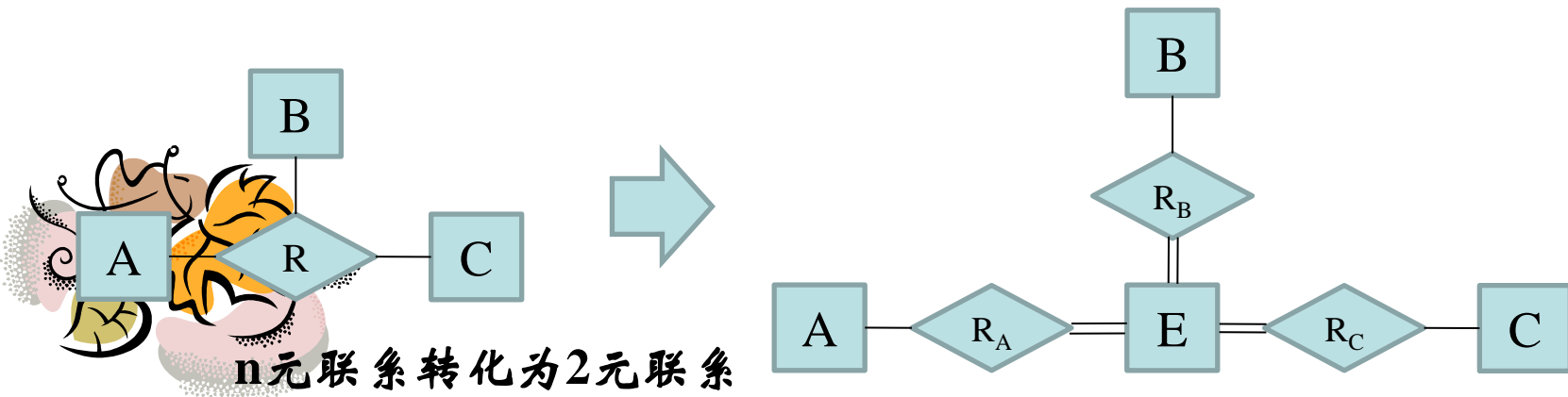
## — 二元联系 or n元联系



若父母信息只有其一，  
无法使用该联系



独立的二元联系表达父母与孩子的联系更为自然



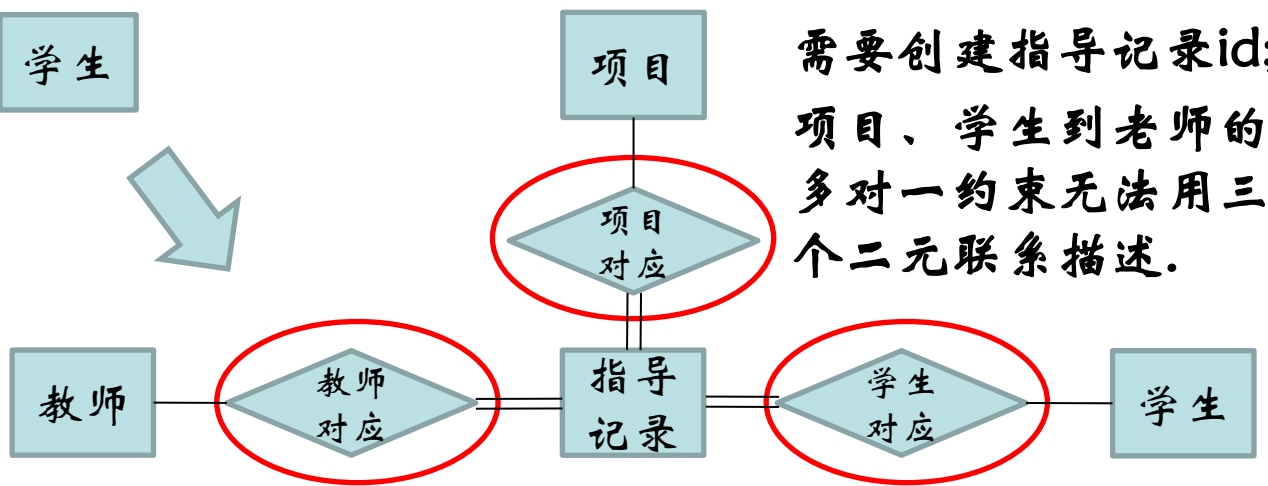
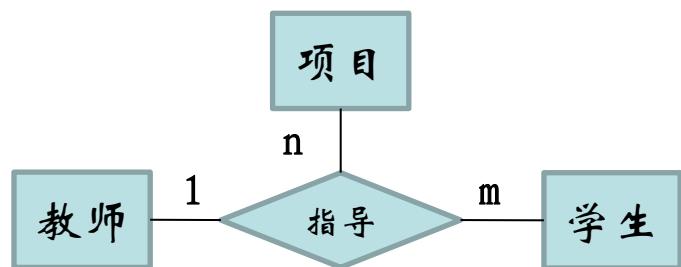




# • E-R模型设计中的一些讨论

## — 二元联系 or n元联系

- 根据语义，选择更自然的方式

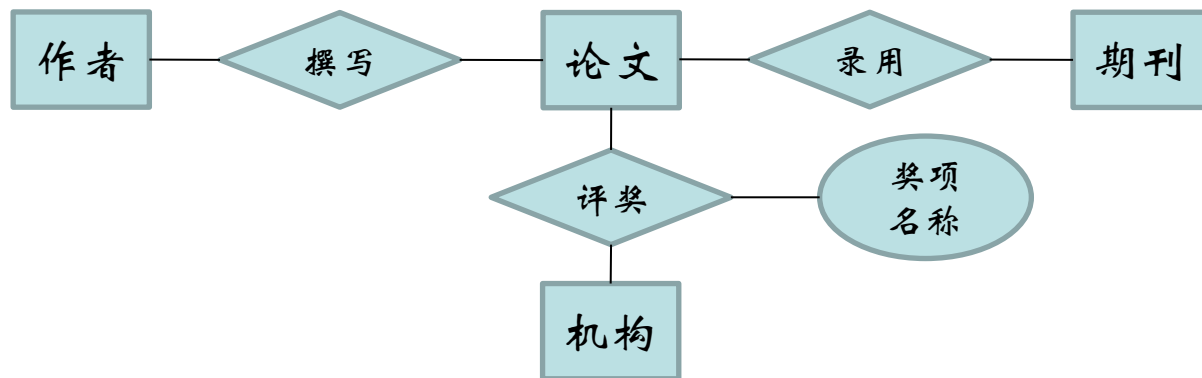
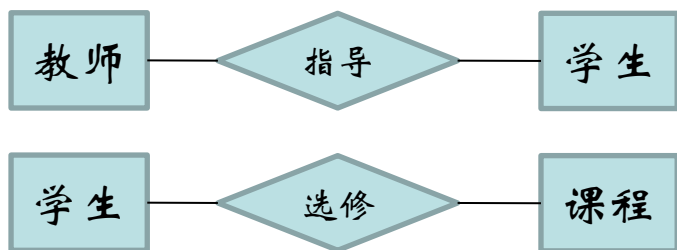




# • E-R模型设计中的一些讨论

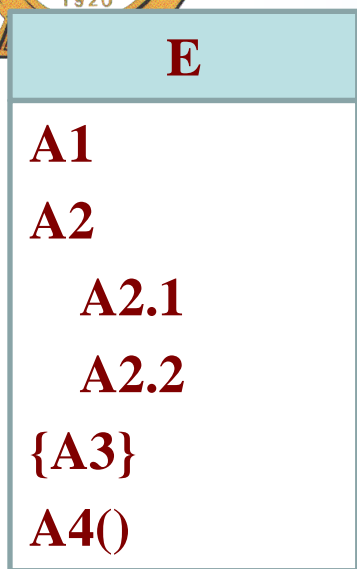
## — 实体集 or 联系集

- 当单纯地描述实体间的行为时，采用联系集

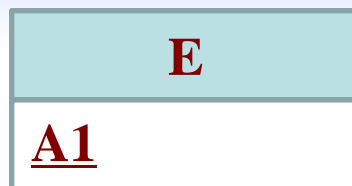




# E-R图其他表达方式示例



实体集E, 其中包括:  
简单属性A1,  
复合属性A2,  
多值属性A3,  
派生属性A4



强实体集E的主码A1



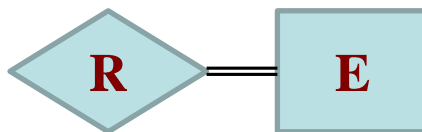
弱实体集E的部分码A1



联系集R



弱实体集的标识性联系集R



实体集E全部参与联系集R



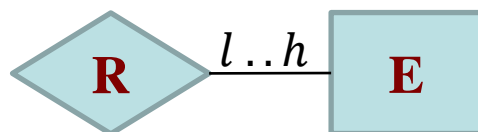
1对1联系



1对多联系



多对多联系

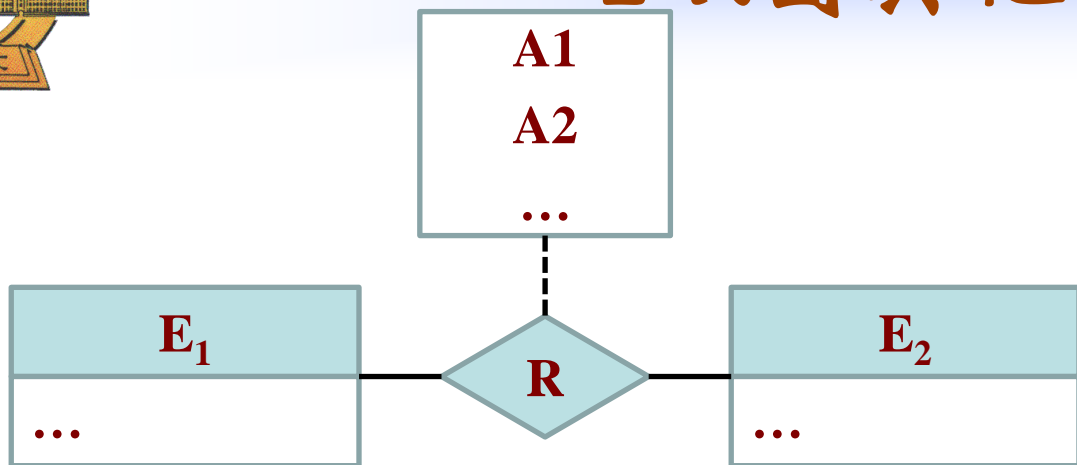


基数约束

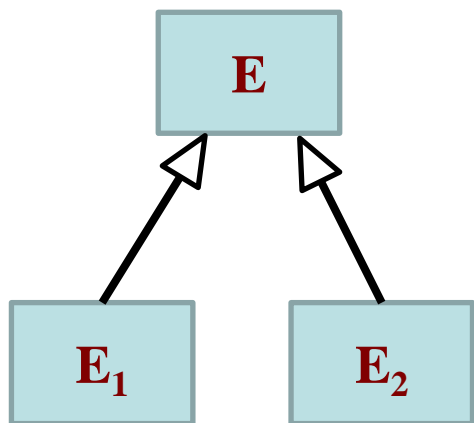




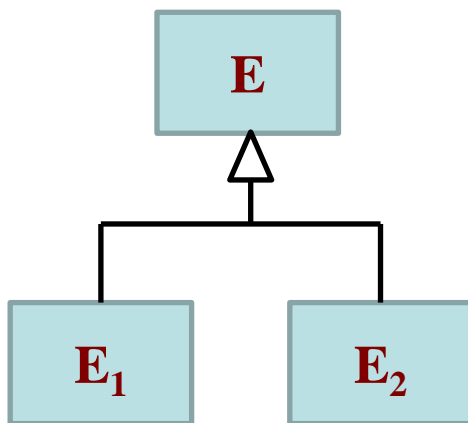
# E-R图其他表达方式示例



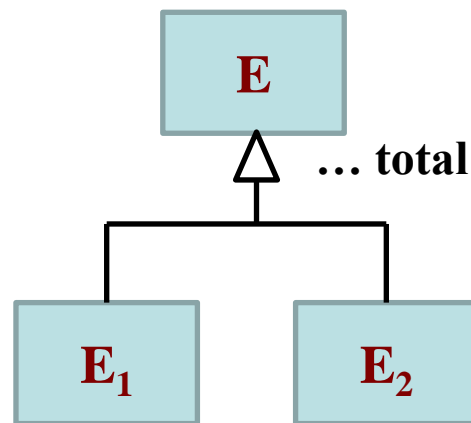
联系  $R$  上的描述性  
属性  $A_1, A_2, \dots$



ISA: 特化或概化



不相交特化

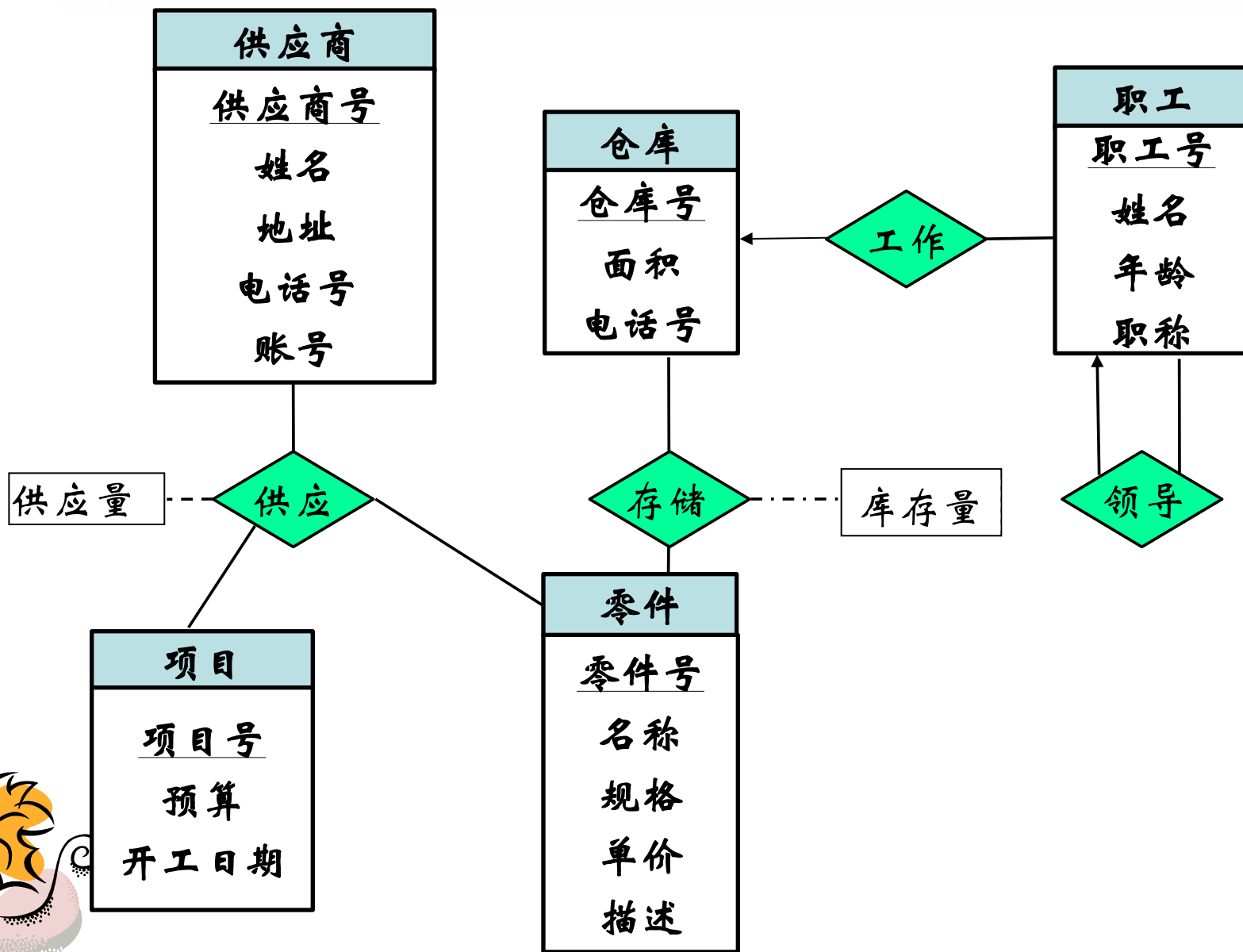


全部、不相交特化





# E-R图其他表达方式示例





# Outline

- 概述
- 实体联系模型
- **概念数据库设计的基本步骤**
- 视图综合设计方法
- 事务的设计





# 概念数据库设计的基本步骤

- 确定实体集、实体集的属性、主码
- 确定实体集之间的联系及其相关的约束
- 形成概念数据库的E-R图





# 概念数据库设计方法

- 集中式设计方法

- 合并在需求分析阶段得到的各种应用需求
- 在上述基础上设计一个概念数据库模式，满足所有应用的需求

- 视图综合设计法

- 不要求应用需求合并

- 视图设计阶段

- 根据每个应用的需求，独立地为每个用户和应用设计一个概念数据库模式。每个应用的概念数据库模式称为一个视图

- 视图合并阶段

- 把所有视图有机合并成一个统一的概念数据库模式，支持所有应用







# 概念数据库设计策略

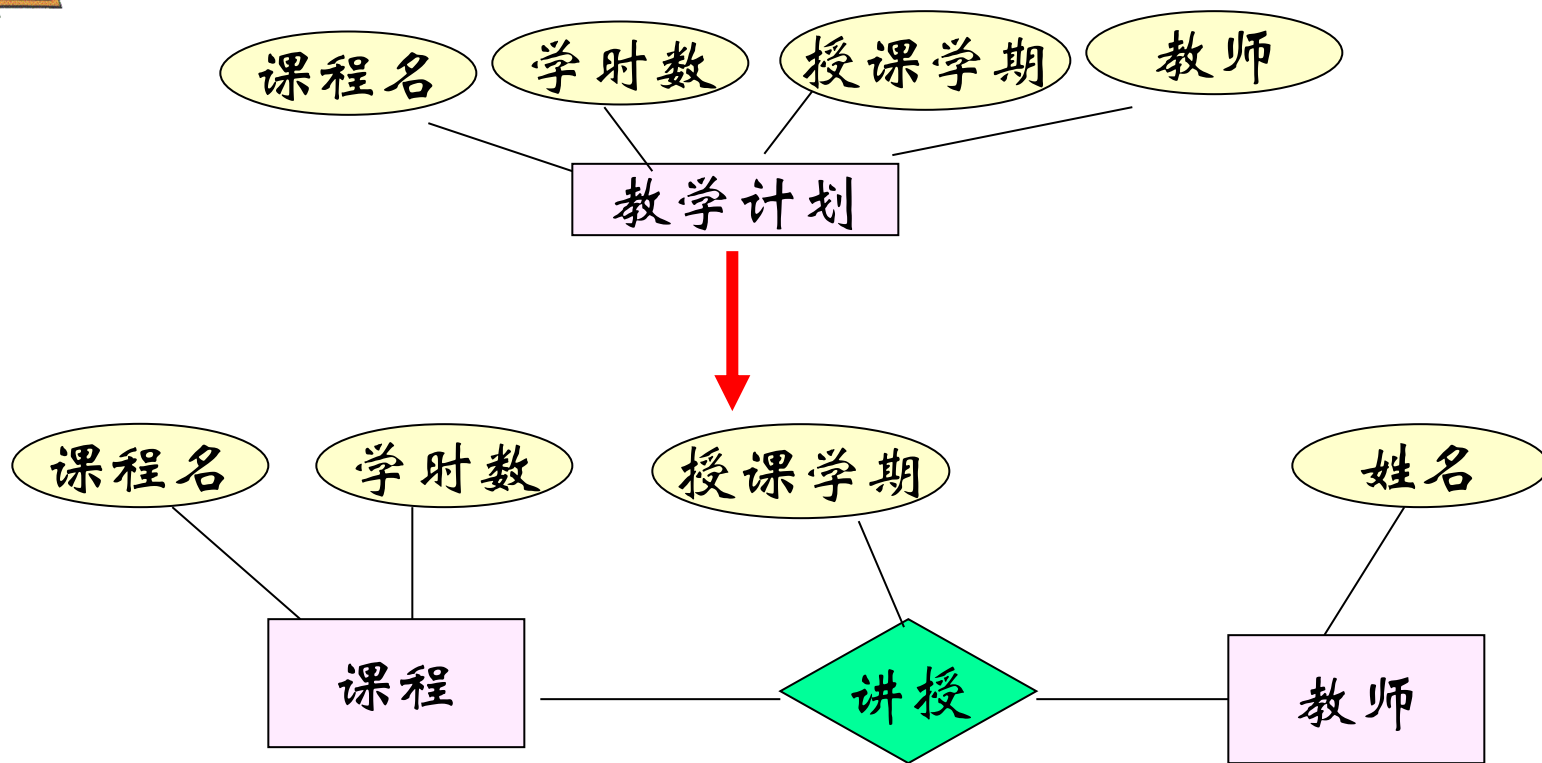
- 自顶向下的策略

- 首先定义全局概念结构框架，然后逐步细化





## 一个实体集细分为两个实体集和一个联系集的实例





- 自底向上的策略

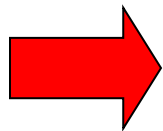
- 首先定义各局部应用的概念结构，然后将它们集成起来，得到全局概念结构





教师

学生



导师

教师

辅导员

学生

增加两个联系





# Outline

- 概述
- 实体联系模型
- 概念数据库设计的基本步骤
- **视图综合设计方法**
- 事务的设计





# 视图综合设计方法

- 局部概念模式设计

- 抽取各局部应用涉及的数据，标定各局部应用中的实体、实体的属性、标识实体的码，确定实体之间的联系及其类型 (1:1, 1:n, m:n)
- 形成局部E-R图

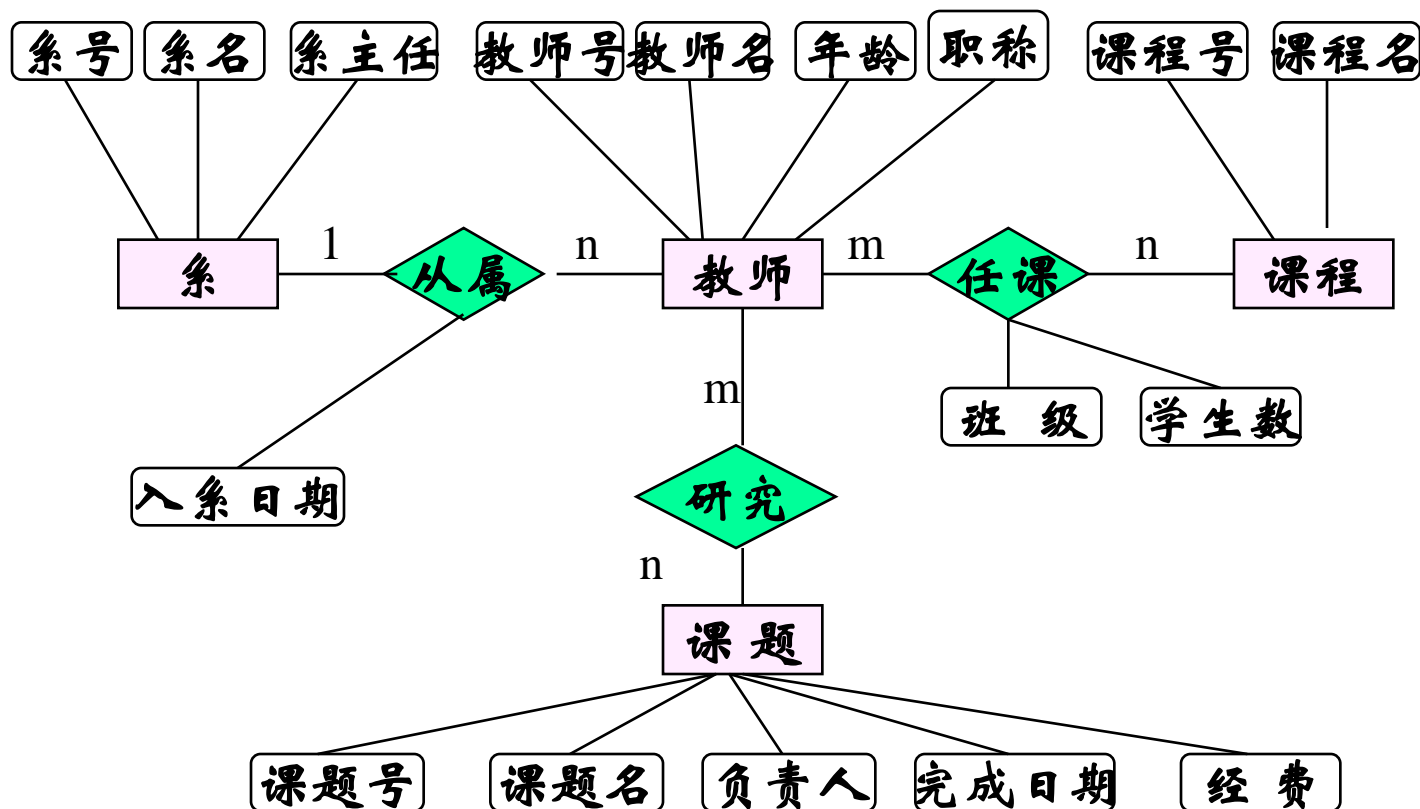




# 局部E-R设计实例：学校教学管理

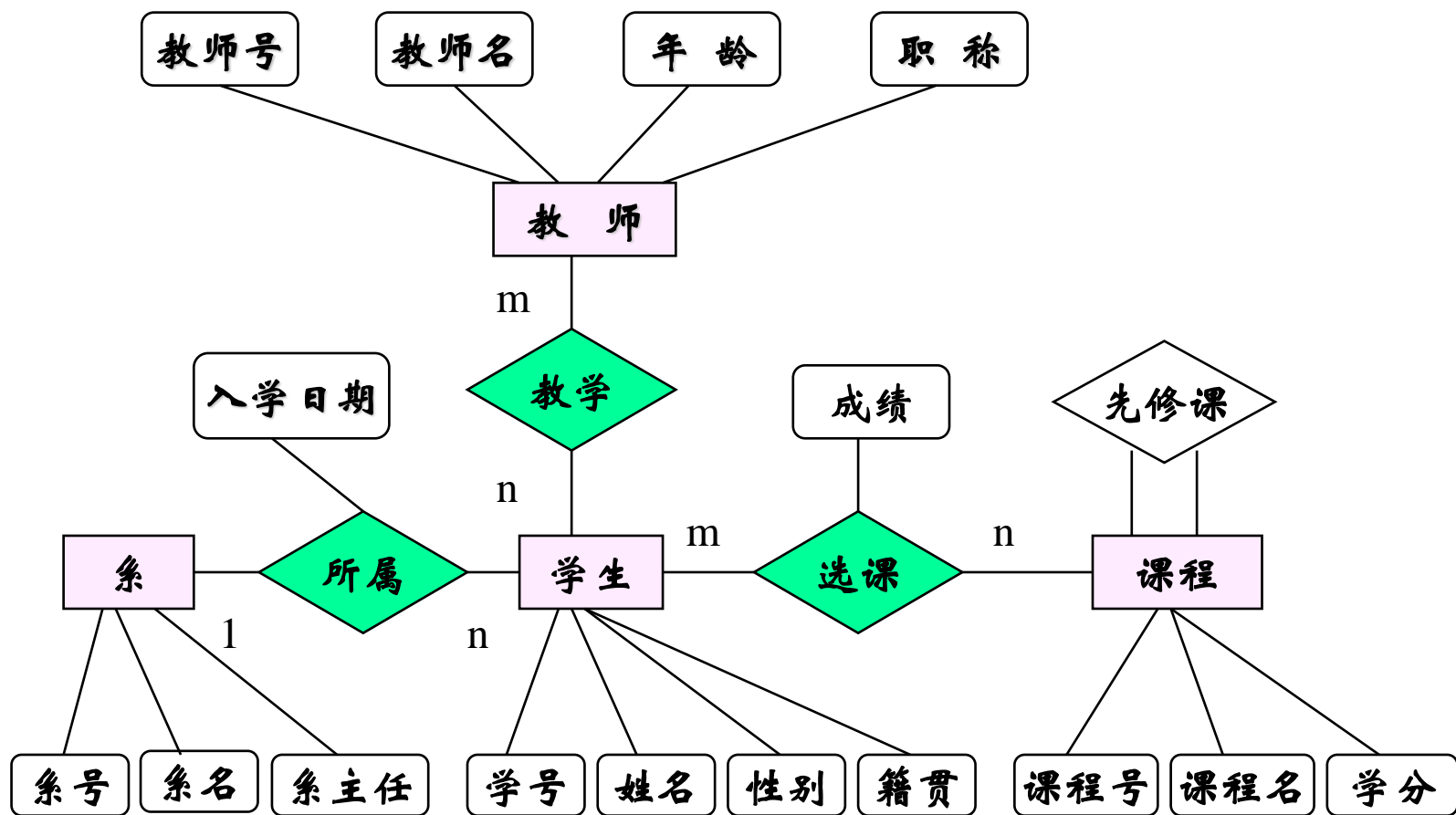
假定它涉及到师资、教务等部门。

师资部门  
局部  
E-R  
图





# 教务部门局部E-R图







## • 全局概念模式的合成

- 各个局部概念模式即局部E-R图建立好后，还需要对它们进行合并，集成为一个整体的数据概念结构即总E-R图
- 集成局部E-R图的步骤
  - 1、识别局部概念模式间的冲突
  - 2、修改局部模式
  - 3、局部模式合并
  - 4、消除不必要的冗余





# 识别局部概念模式间的冲突

- 各局部E-R图存在冲突
  - 各个局部应用所面向的问题不同
  - 由不同的设计人员进行设计
  - 各个局部E-R图之间必定会存在许多不一致的地方
  - 合并分E-R图的主要工作与关键所在：合理消除各局部E-R图的冲突





## • 冲突的种类

- 命名冲突：异名同义，同名异义

- 值域冲突

  - 属性值的类型、取值范围、计量单位等不同

- 模式结构冲突

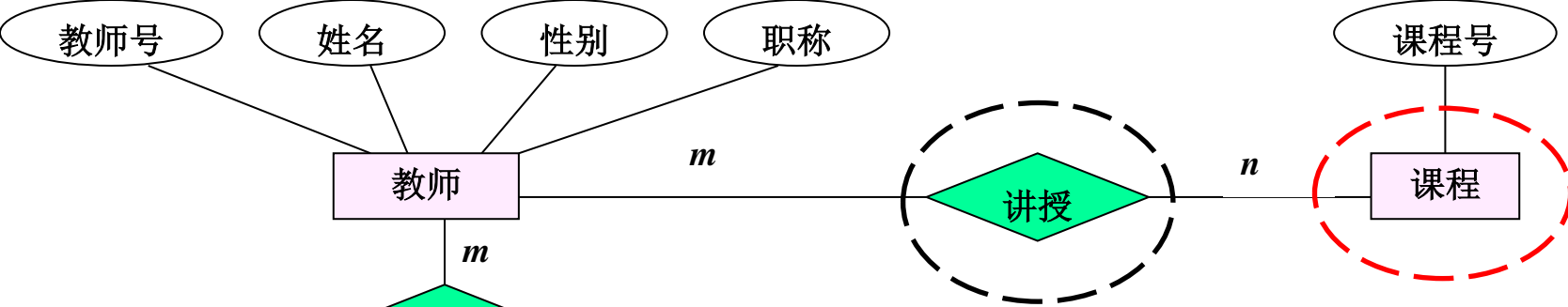
  - 相同概念在不同的局部模式中使用不同的概念模式表示

    - 例如，“系”在一个局部模式中表示为一个实体，而在另一个模式中则被表示成一个复合属性

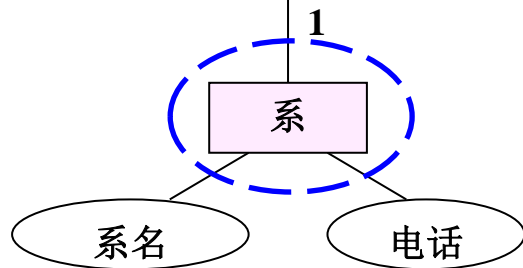
    - 同一实体在不同子系统的E-R图中包含不同的属性

    - 约束冲突

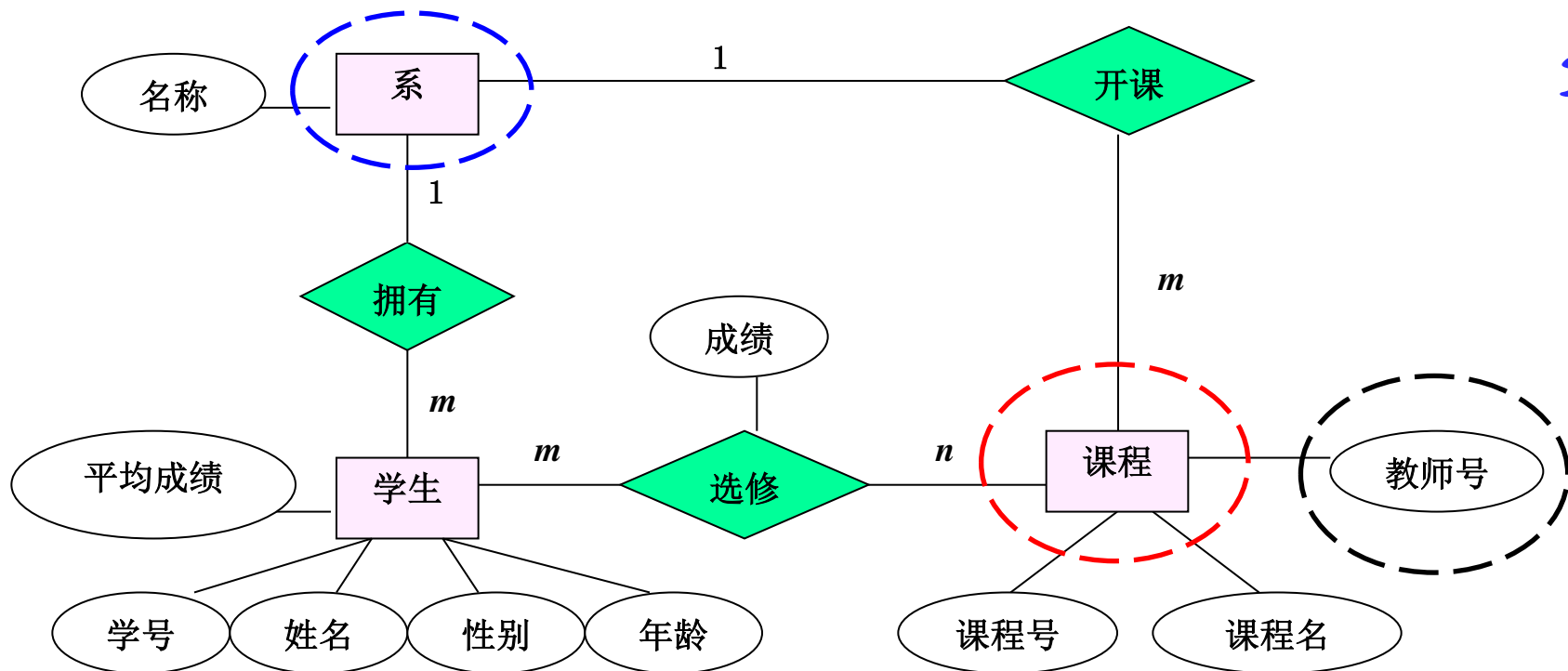




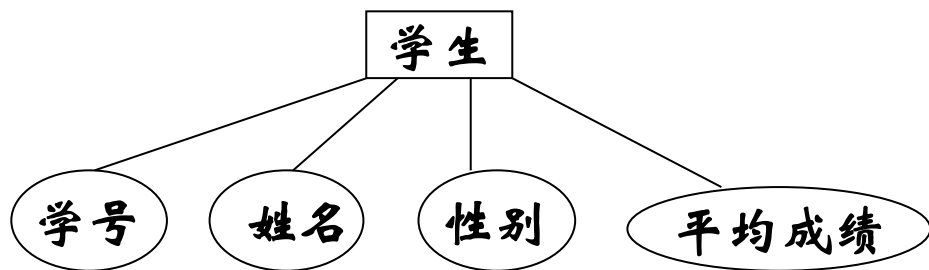
图a 教师任课局部 E-R 图



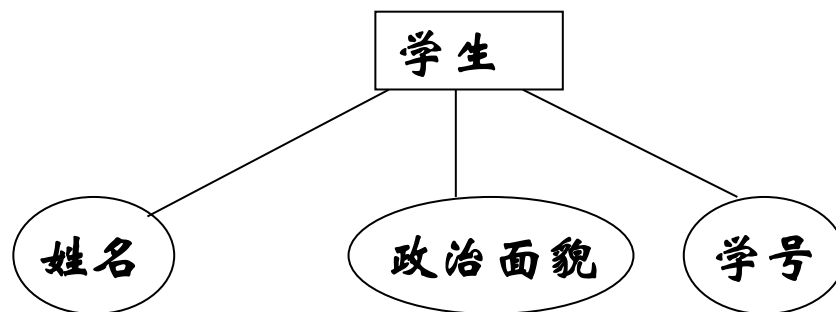
图b 学生选课局部 E-R 图



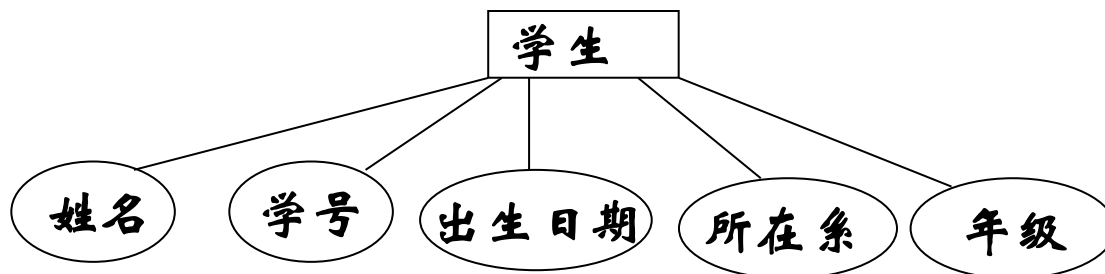
结  
构  
冲  
突



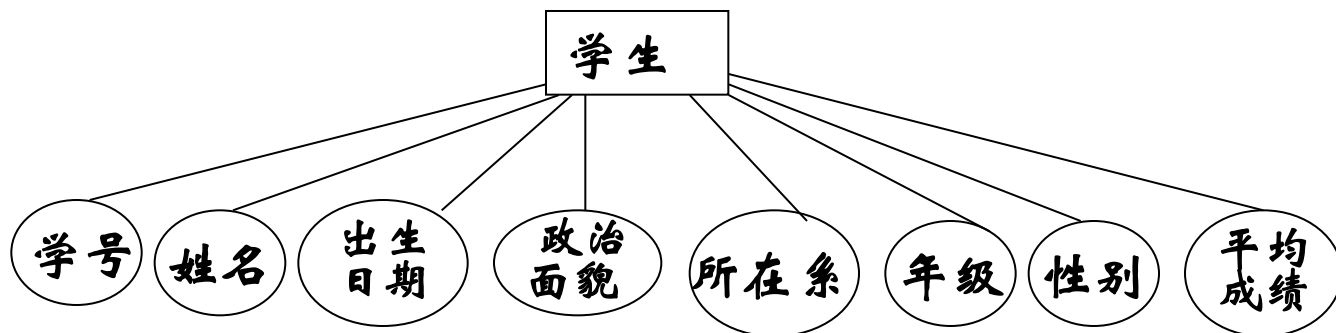
(a)在局部应用A中



(c)在局部应用C中



(b)在局部应用B中



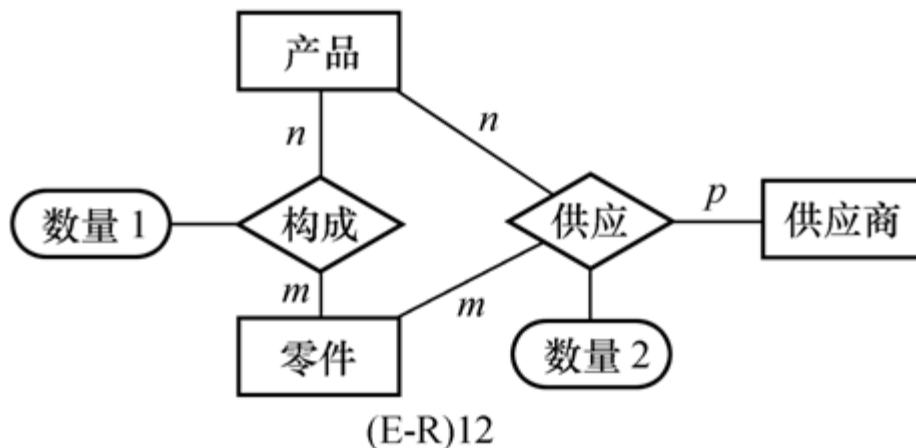
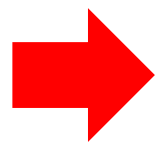
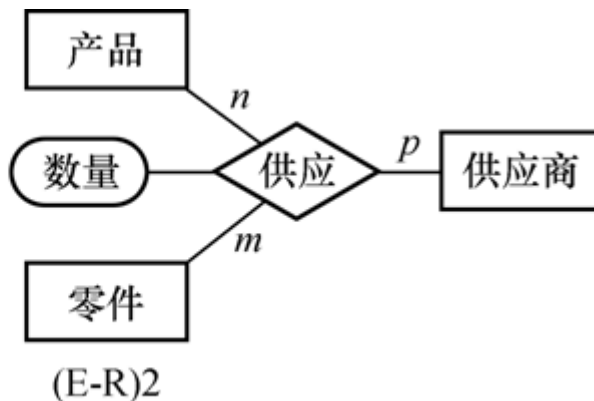
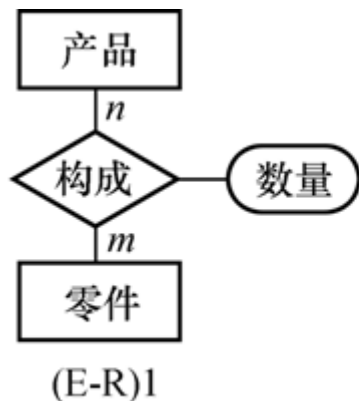
(d)合并后

结构冲突





## 结构冲突



构成:

零件与产品之间存在多对多的联系

供应:

产品、零件与供应商三者之间还存在多对多的联系

这两个联系互相不能包含, 在合并两个分E-R图时就应把它们综合起来





- 不必要的冗余

- 冗余的数据

- 可由其他基本数据导出的数据

- 冗余的联系

- 可由其他的联系导出的联系

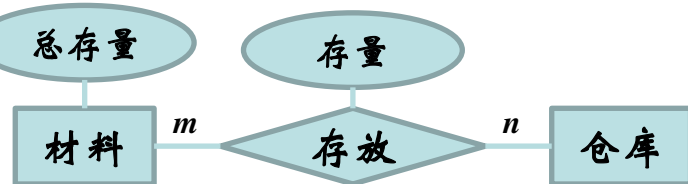
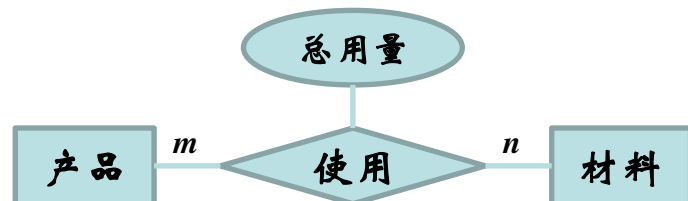
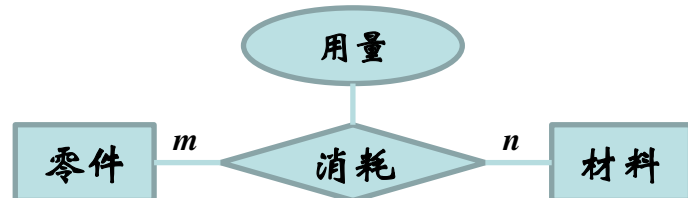
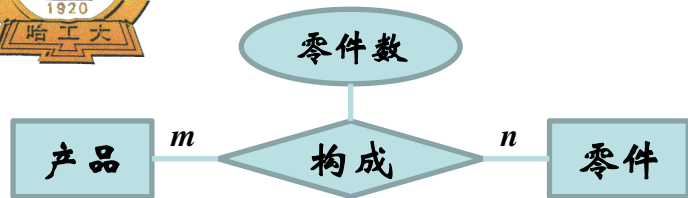
- 冗余的潜在危害

- 破坏数据库的完整性
    - 给数据库维护增加负担

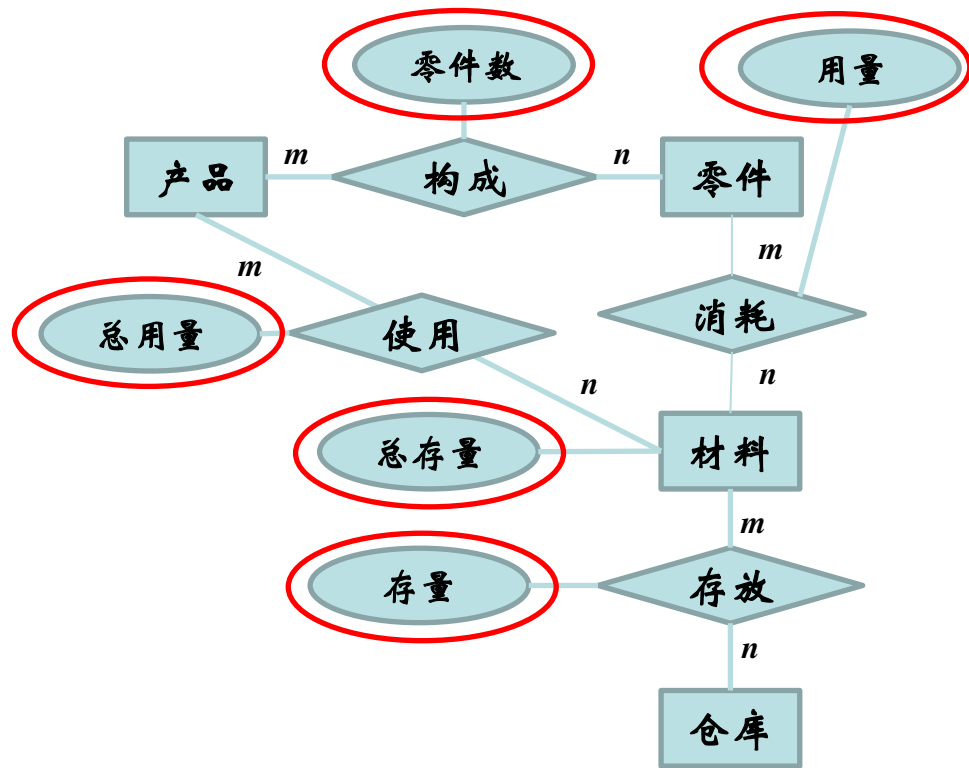




## 不必要的冗余——数据



合并



冗余的数据（参考数据项字典）：

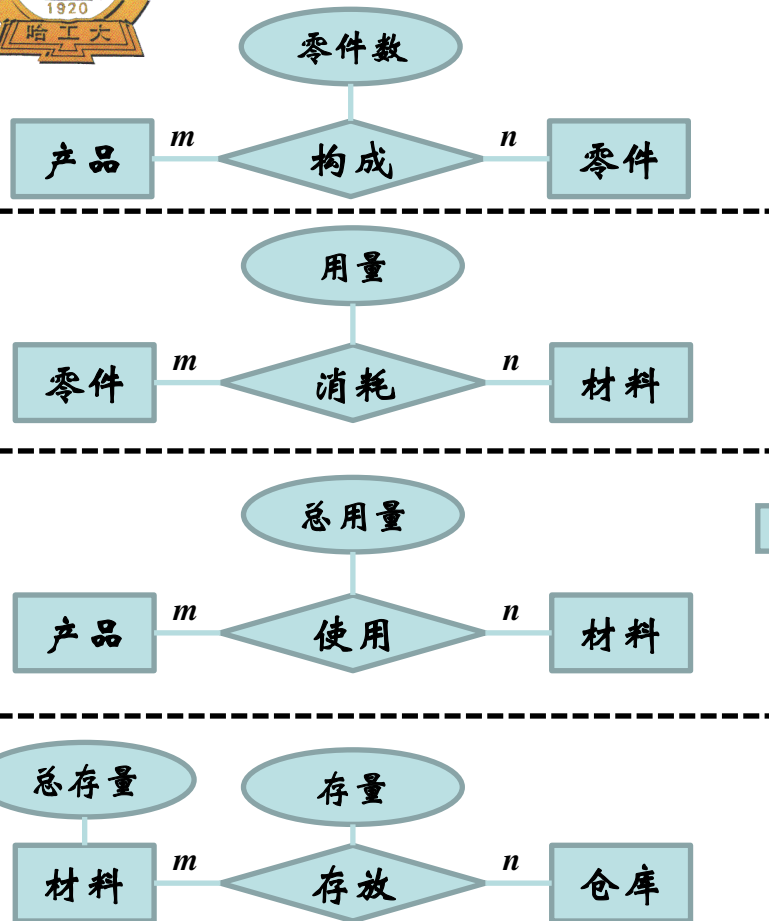
“总用量”可以通过“零件数”和“用量”导出；

“总存量”可以通过“存量”导出

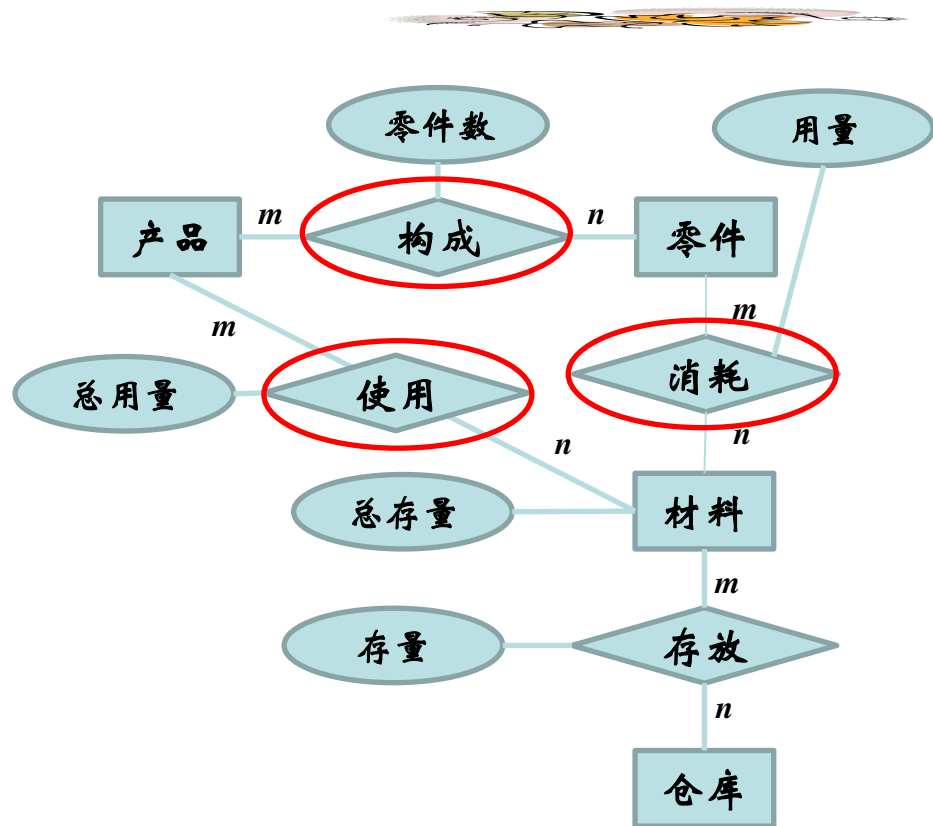




## 不必要的冗余——联系



合并



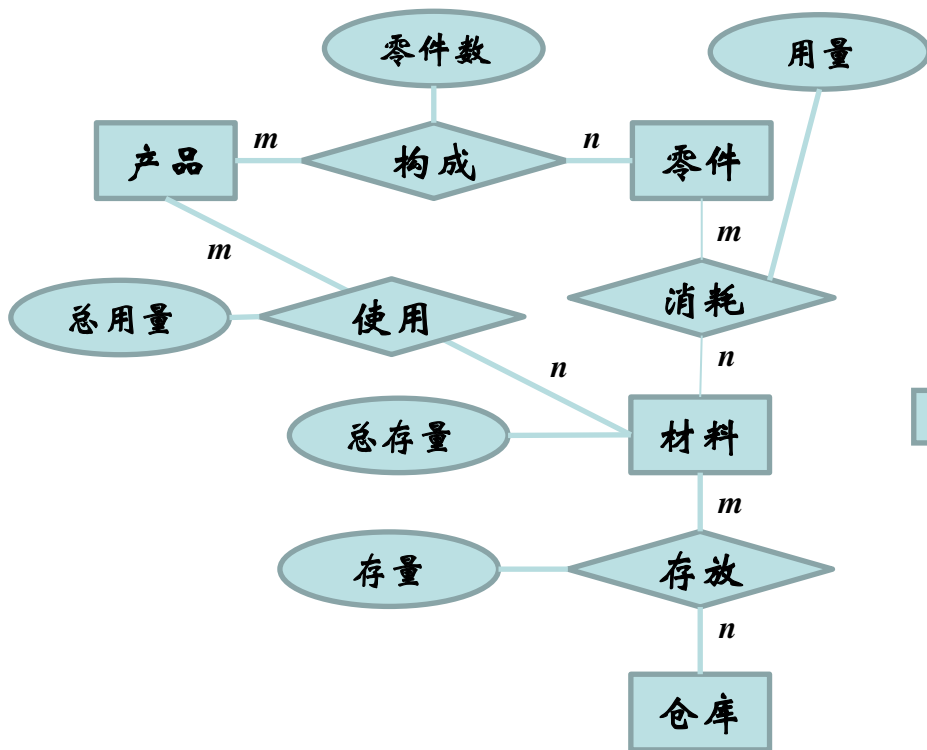
冗余的联系（参考数据项字典）：

“使用”可以通过“构成”和“消耗”导出

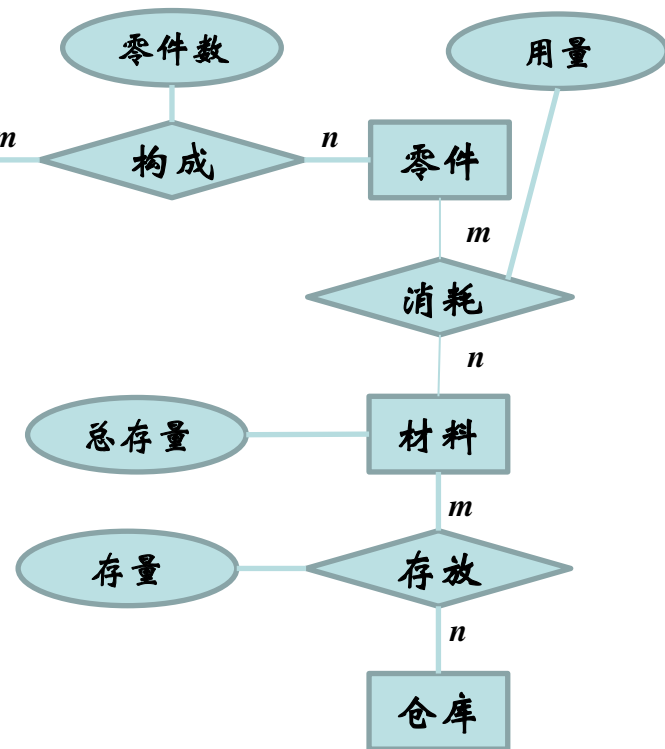




## 不必要的冗余



消除冗余



合并之后的E-R图

可以人为保留某些冗余:

数据操作的需求中包括“频繁地查询各种材料的总存量”，每次查询都从“存量”中导出“总存量”效率较低



# 图书馆、教材科、宿管科的应用需求



图书馆应用：

学生：学号、姓名

图书：书号、书名、出版社、出版时间

阅览室：楼号、房间号

学生可以借阅多本图书，每次借阅记录借阅时间、归还时间；  
多本图书存放在一个阅览室。

教材科应用：

图书：书名、出版社、出版时间

课程：课程号、课程名

每门课程设置一本图书为参考教材，每本书作为一门课程的参考教材。

宿管科应用：

学生：学号、姓名、性别

公寓：公寓号、房间数、最大人数、现有人数

房间：房间号、最大人数、现有人数

每个学生住在一个房间，每个房间有多位同学；

每个房间在一个公寓中，每个公寓有多个房间。





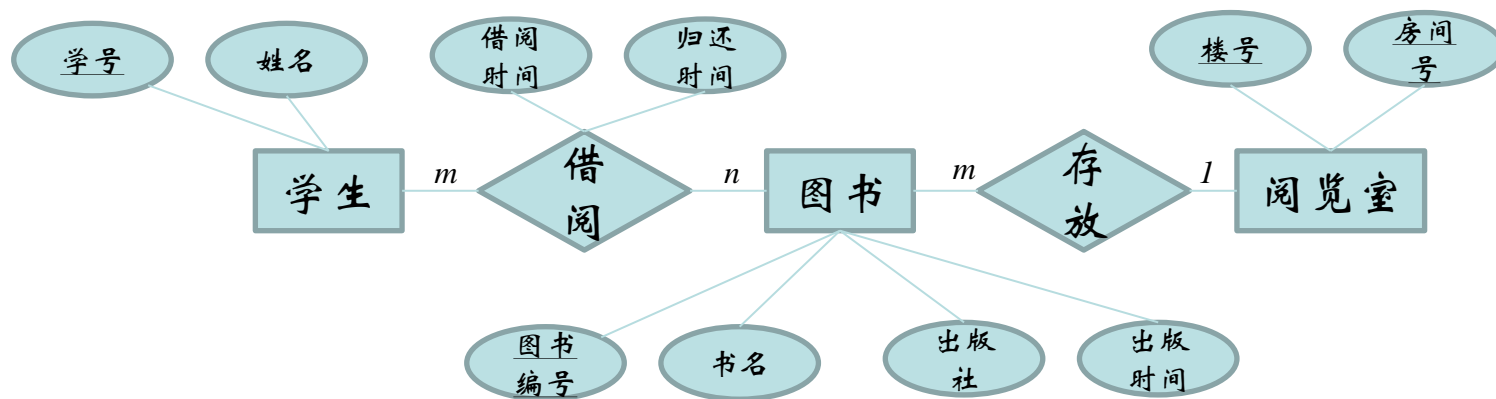
图书馆应用：

学生：学号、姓名

图书：书号、书名、出版社、出版时间

阅览室：楼号、房间号

学生可以借阅多本图书，每次借阅记录借阅时间、归还时间；  
多本图书存放在一个阅览室。



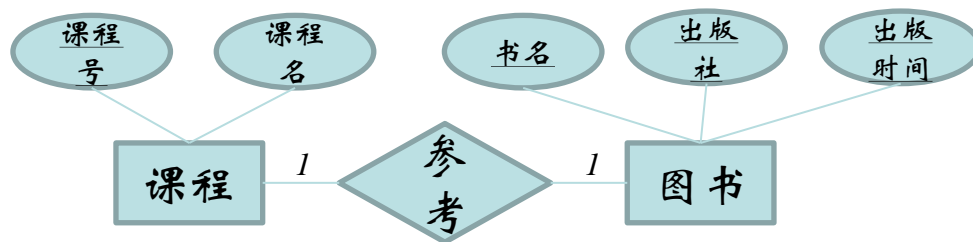


教材科应用:

图书: 书名、出版社、出版时间

课程: 课程号、课程名

每门课程设置一本图书为参考教材, 每本书作为一门课程的参考教材。





宿管科应用:

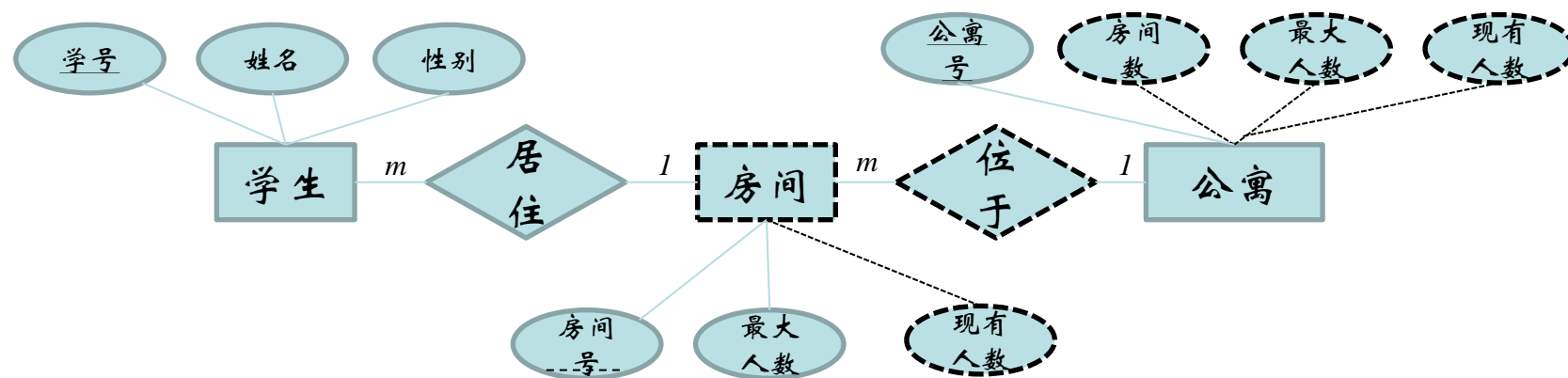
学生: 学号、姓名、性别

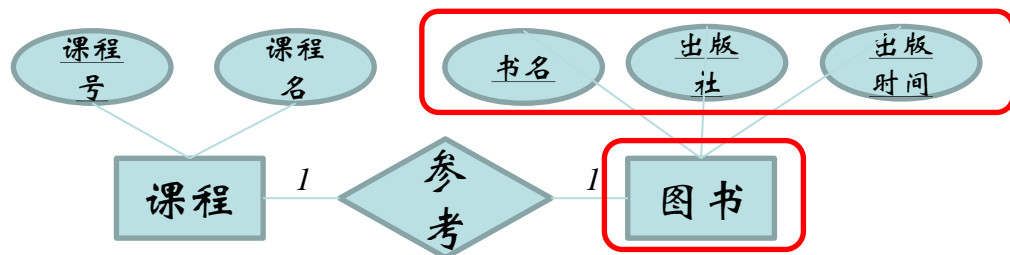
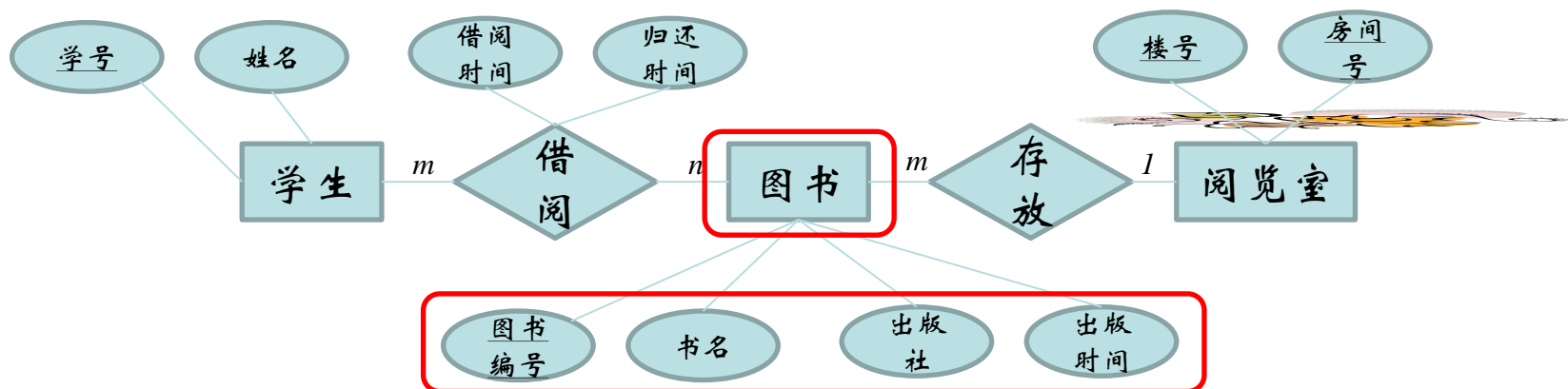
公寓: 公寓号、房间数、最大人数、现有人数

房间: 房间号、最大人数、现有人数

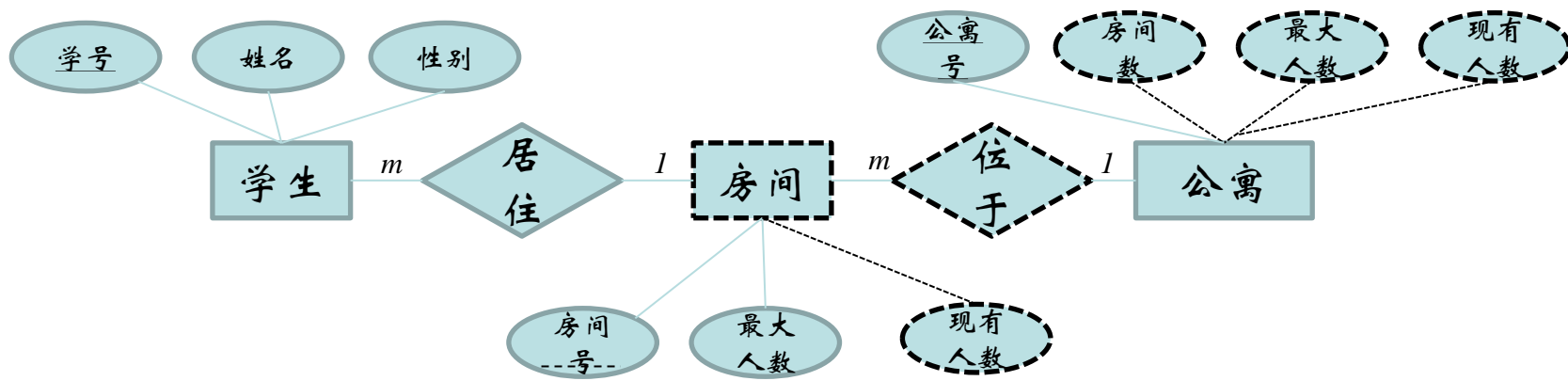
每个学生住在一个房间, 每个房间有多位同学;

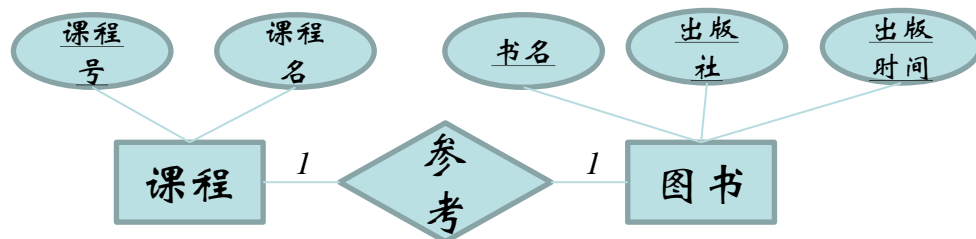
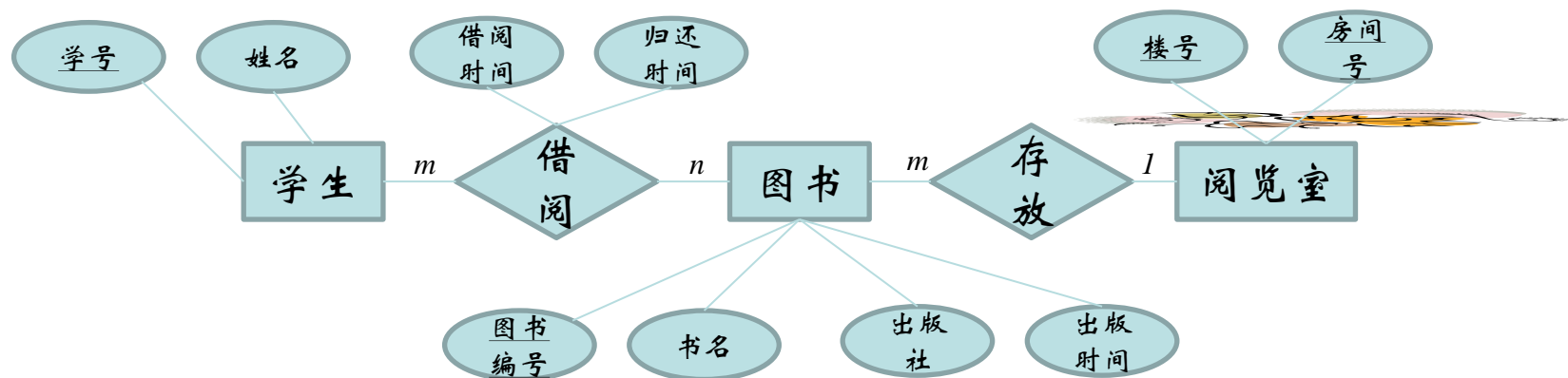
每个房间在一个公寓中, 每个公寓有多个房间。



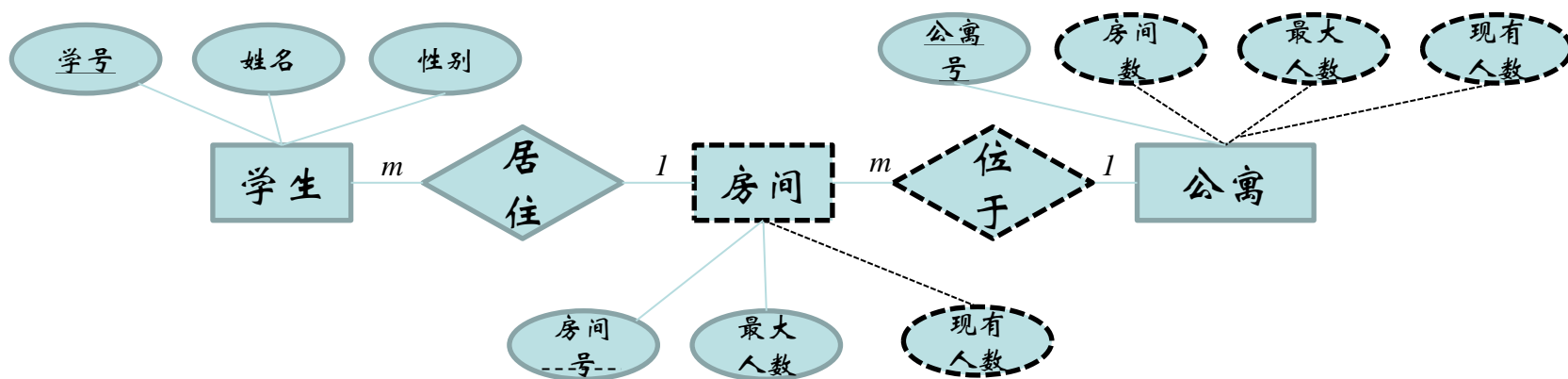


结构冲突：图书实体  
处理方式1：合并属性，成为单一实体集——图书；  
处理方式2：作为两个实体集，图书和书目，建立二者之间的联系。

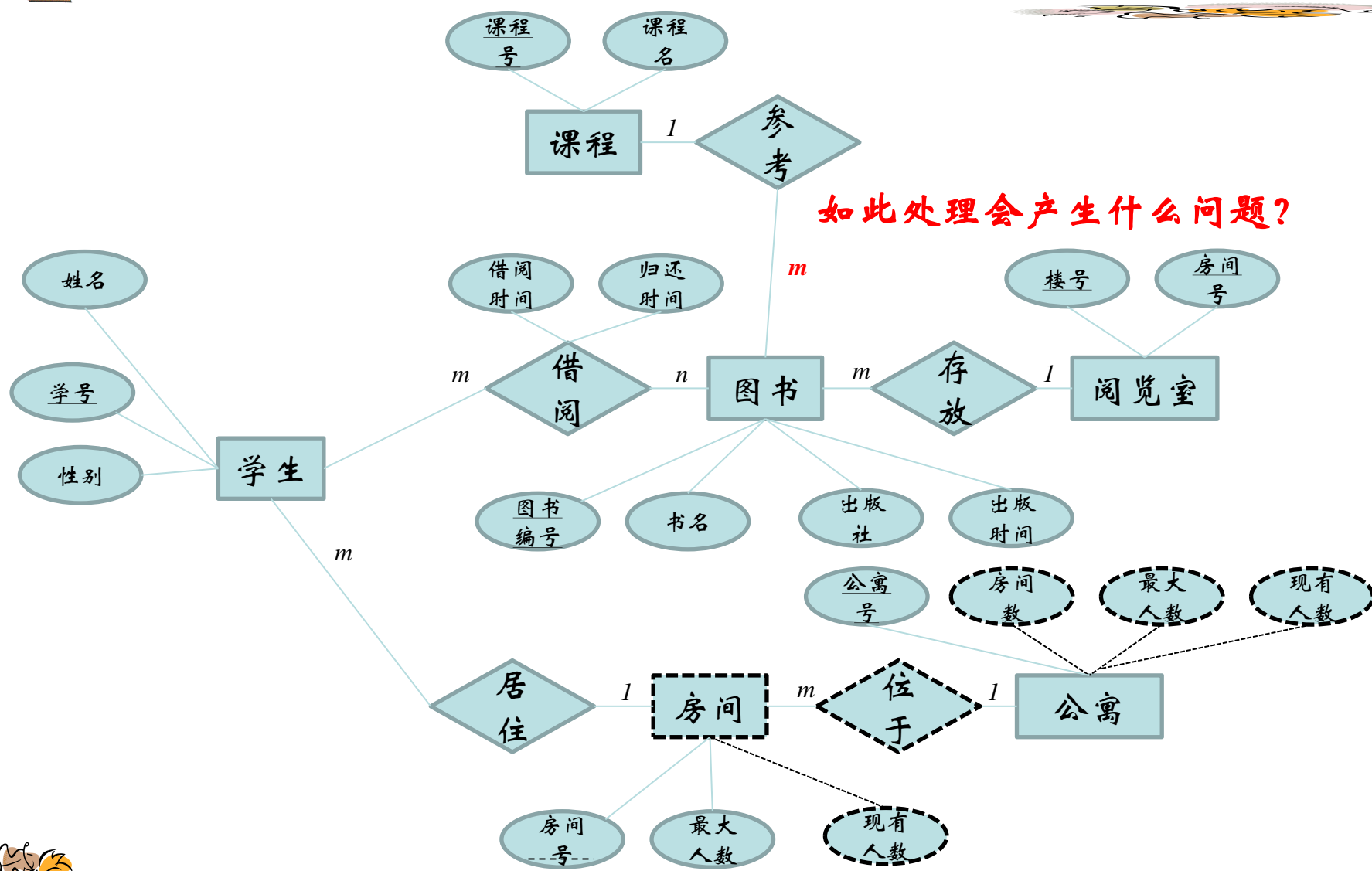


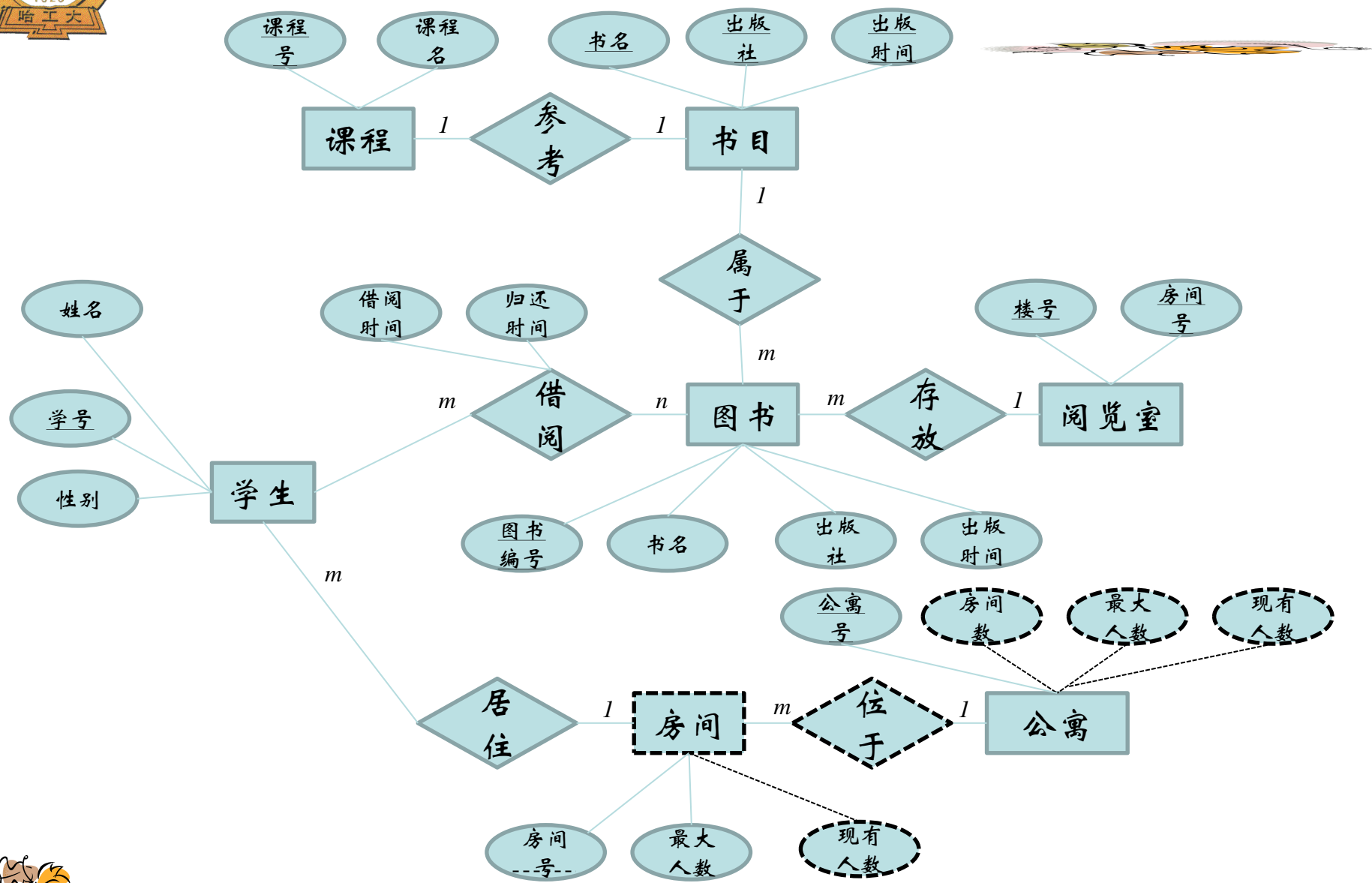


数据冗余：房间和公寓实体集的属性——现有人数  
处理方式1：去除冗余；  
处理方式2：保留冗余。











# Outline

- 概述
- 实体联系模型
- 概念数据库设计的基本步骤
- 视图综合设计方法
- 事务的设计





# 事务的设计

- 事务

- 一个或多个数据操作构成的集合，这组操作满足原子性。

- 例如，银行从账户A到账户B的一次资金转账操作

- 事务设计任务：定义事务功能

- 说明事务的输入、输出

