

# 数据库系统基础

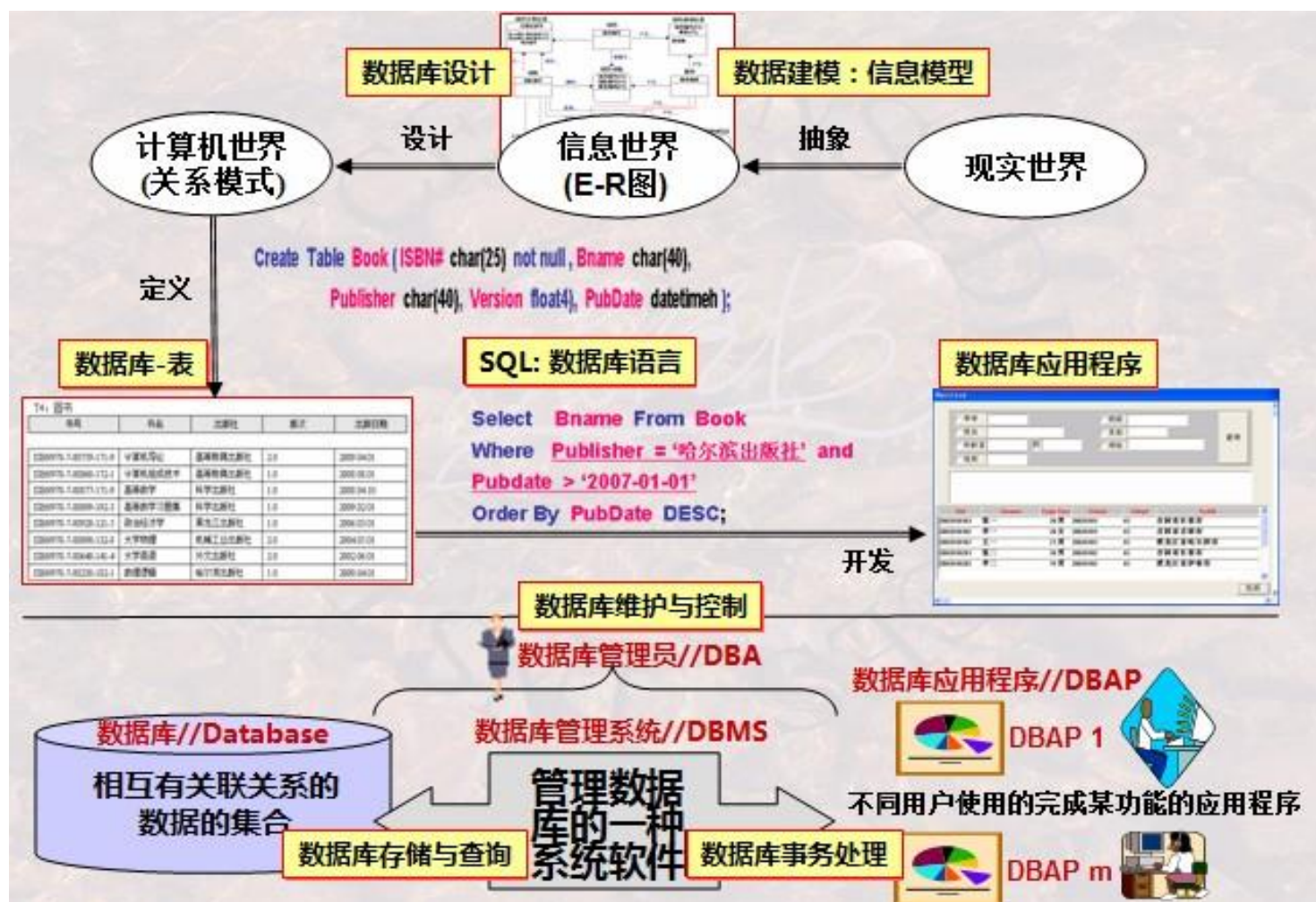
哈尔滨工业大学

# 第11章 数据建模之思想与方法

# 数据建模之思想与方法



## 数据建模与数据库设计在《数据库系统》中的位置



# 数据建模之思想与方法



- 理解数据建模与数据库设计

- 扎实地理解并掌握E-R模型：数据建模与数据库设计的重要工具

  - 能够正确绘制E-R图

- 能够用E-R模型准确理解现实世界并进行数据库设计

  - 理解现实世界并进行抽象的能力，理解并抽象得正确

- 能够分析数据库设计的正确性

# 数据建模之思想与方法



## 基本内容

1. 为什么要数据建模和数据库设计?
2. E-R模型--数据建模之基本思想
3. E-R模型--表达方法之Chen方法
4. E-R模型--表达方法之Crow' s foot方法
5. 数据建模之案例讲解
6. 数据库设计中的抽象

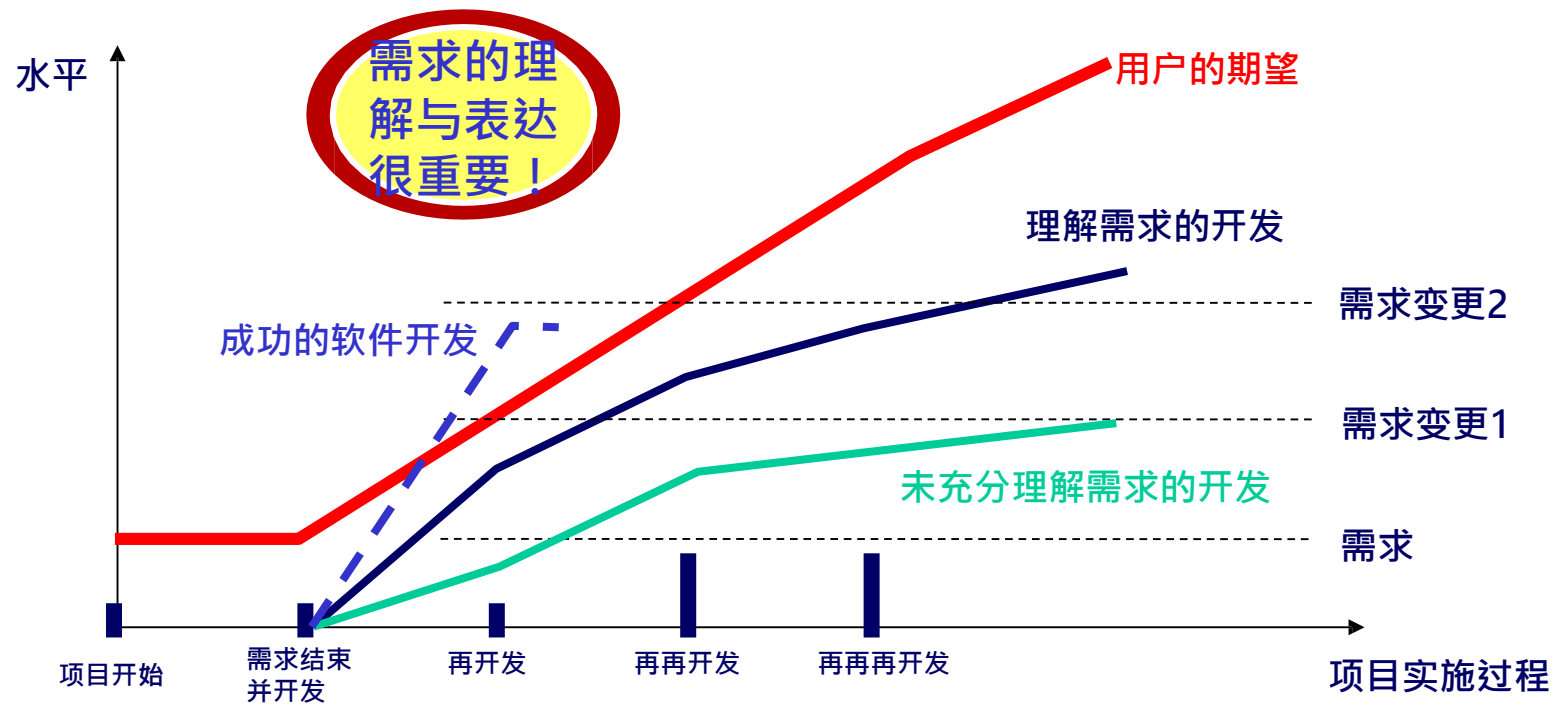
## 重点与难点

- 理解E-R模型进行数据建模的基本思想，熟练掌握E-R模型的Crow' s foot表达方法
- 熟练运用Crow's foot方法进行数据建模，即建模训练

# 为什么要数据建模和数据库设计?



研发出的软件为什么满足不了需求... ..



# 为什么要数据建模和数据库设计?



需求的理解，简单来讲，包括数据需求的理解和处理规则需求的理解... ..

数据需求的理解与表达是否更重要呢！

示例

需求：折扣政策管理

- 1) 不同客户有不同的折扣
- 2) 不同产品有不同的折扣
- 3) 不同客户购买不同产品有不同的折扣

怎样设计数据库呢？为什么要如此设计呢？每种设计有什么问题呢？

设计方案1

产品折扣表		
产品	折扣率	
客户折扣表		
客户	折扣率	
客户-产品折扣表		
客户	产品	折扣率

还需按类别管理呢

设计方案2

客户-产品折扣表			
序号	客户	产品	折扣率

还需按类别管理呢

设计方案3

客户-产品折扣表			
序号	客户类别	产品类别	折扣率

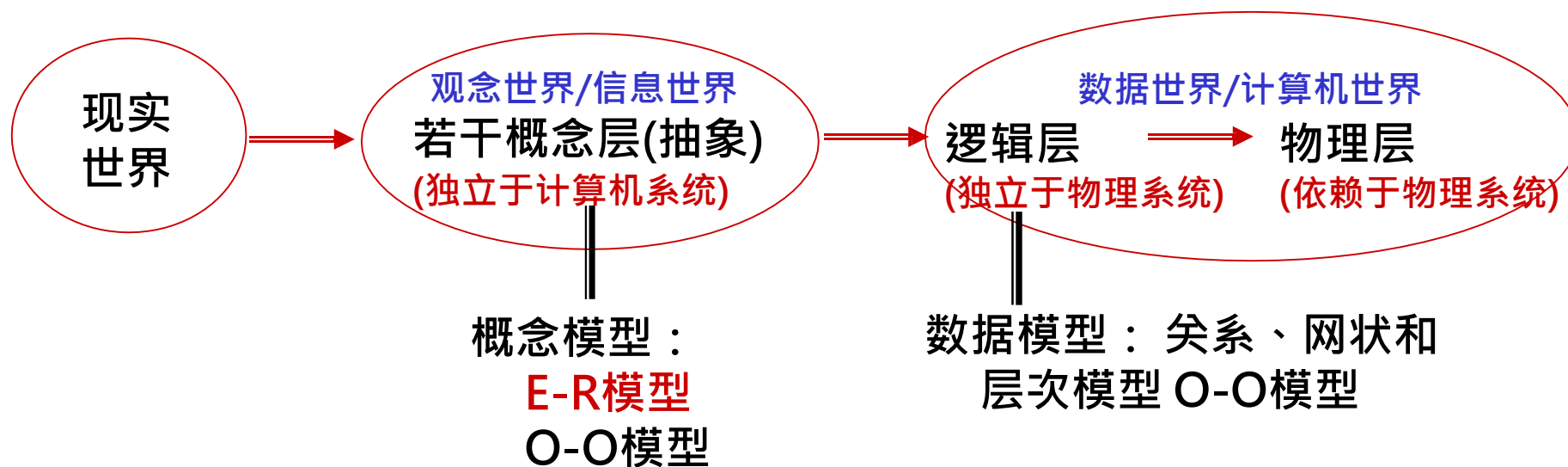
特殊的还需按个体管理呢

# 为什么要数据建模和数据库设计?



## 数据模型与概念模型

- 表达计算机世界的模型称**数据模型**；表达信息世界的模型称概念数据模型，简称**概念模型**，信息世界是对现实世界的理解与抽象





# 为什么要数据建模和数据库设计?



## 示例

便于人使用和理解

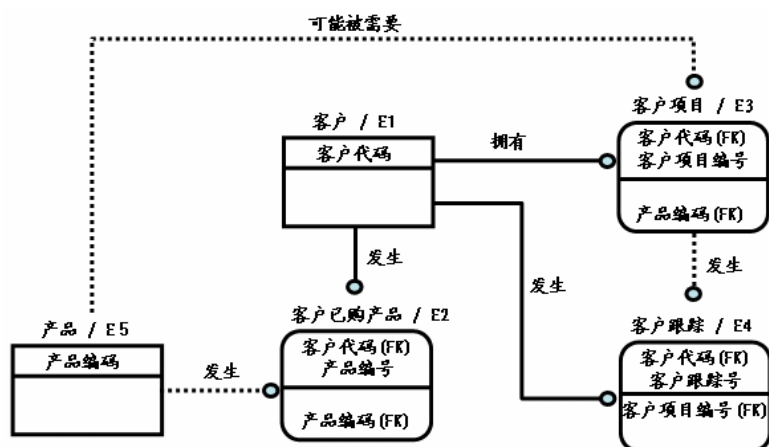
客户信息表					
客户姓名	客户地址	法人	电话	电传	联系人
客户已购买本厂产品					
客户现有项目	项目名称				负责人
	项目名称				负责人
跟踪	第一次				座谈人
座谈	第二次				座谈人
情况	第三次				座谈人

适合人类阅读习惯的信息表达  
===单据/报表

### 出库单

出库类别	出库日期	外协单位
领料单编号	领料部门	领料人
库房	库管员	备注
限额审核	部门制单	部门审核

序号	工程号	领料条码	领料名称	领料规格	申领数量	实发数量	单位	单价	金额



表征信息本质联系的概念表达  
===实体/联系

便于机器精确处理



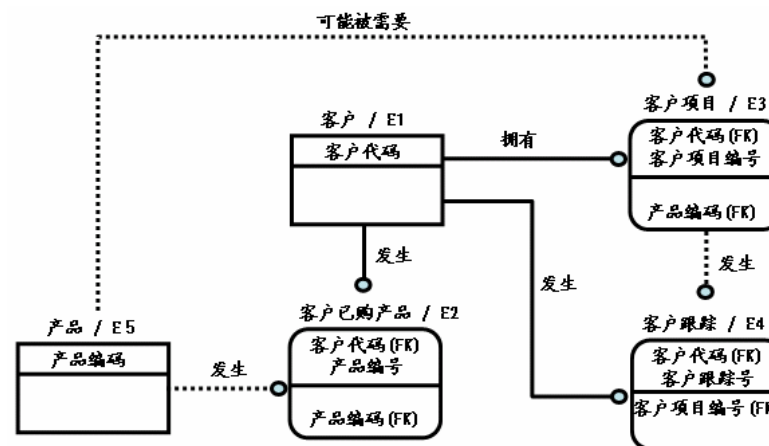
# 为什么要数据建模和数据库设计？



## 数据建模是抽象，抽象是理解-区分-命名-表达

- ❑ 现实世界需要理解：现实中的卡片、单据、表格、报表... ..
- ❑ 理解的标志是区分：表与表的区分，表内数据项的区分，数据项之间关系的区分，表之间关系的区分？
- ❑ 区分的标志是命名：命名表、命名数据项、命名表之间的联系
- ❑ 抽象的最终结果是正确的表达：用其他人能理解的表达方法来表达(E-R图/Crow's Foot/IDEF1X)

客户信息表					
客户姓名	客户地址	法人	电话	电传	联系人
客户已购买本厂产品					
客户现有项目	项目名称				负责人
	项目名称				负责人
跟踪	第一次				座谈人
座谈情况	第二次				座谈人
	第三次				座谈人





## E-R模型

- E-R模型：Entity-Relationship Model
- 1976年, P.P.S.Chen提出E-R模型，用E-R图来描述概念模型
- E-R模型的基本观点：
- 世界是由一组称作**实体**的基本对象和这些对象之间的**联系**构成的

建模思想--  
理解-区分-命名  
- 需要区分及命名哪些要素

表达方法-- 用  
不同的图形表  
达被区分的不同  
要素

先理解基本思想, 再学习表达方法

# E-R模型--数学建模之基本思想



□ E-R模型给出了一组基本概念，用这组概念可以刻画信息世界

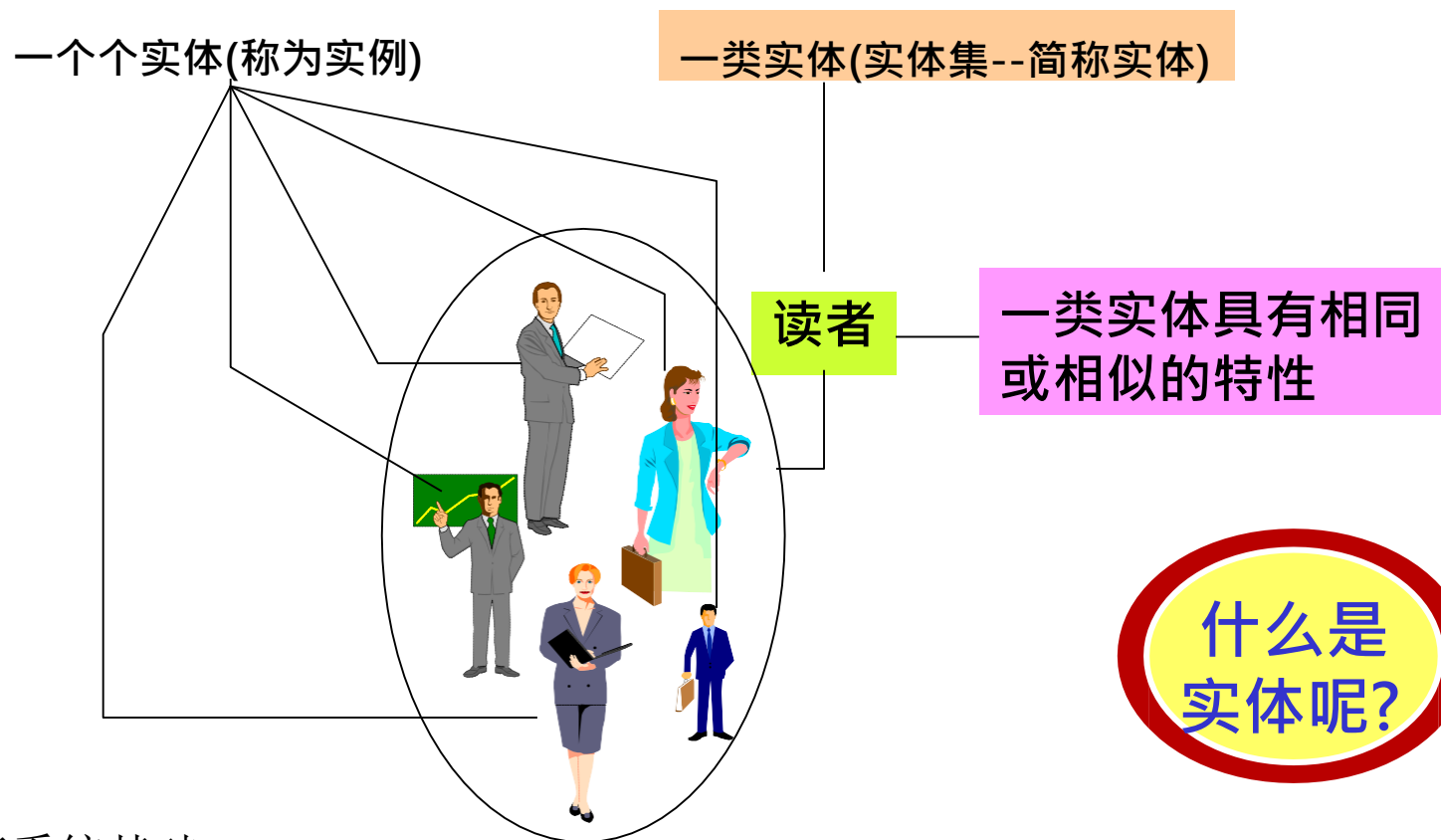
- 实体
- 属性
- 联系
- 关键字/码

抽象的概念(语义较少)	对应问题领域的概念(给抽象概念赋予一定语义的新的概念)
实体	学生(学号, 姓名, ...)
属性	课程(课程号, 课程名, ...)
联系	教师(教师编号, 教师名, ...)
关键字/码	任课(教师编号, 课程号, ...)
	... ..



## 实体与实例

- 实体：客观存在并可相互区分的事物
- 实体有类(实体，实体属性)和个体(实体的实例，实体的值)的概念



# E-R模型--数学建模之基本思想



□ 实体用属性来刻画

属性，实体所具有的某一方面特性

多方面的“  
型”与“值”



读者

借书证号	姓名	年龄	性别	家庭住址
1	张三	25	男	吉林
2	李四	22	女	黑龙江
3	王五	24	男	沈阳
4	杨六	23	女	黑龙江
5	李四	24	男	黑龙江

实体: 实体属性来刻画

实例: 实体的值来刻画

实体名称(属性名1, 属性名2,...)

实体名称(属性值1, 属性值2, ...)

读者(借书证号, 姓名, 年龄, 性别, 家庭住址)

读者(1, 张三, 25, 男, 吉林)  
读者(2, 李四, 22, 女, 黑龙江)  
读者(3, 王五, 24, 男, 沈阳)

# E-R模型--数学建模之基本思想



- 属性还有很多类型，注意区分：
  - 单一属性与复合属性,
  - **复合属性示例**：家庭住址：省份, 详细住址
  - **在关系模型中，复合属性一定要转化为单一属性(关系的第1范式)**
  - 单值属性和多值属性：每个实例的该属性值是一个还是多个
  - **多值属性示例**：电话号码，一个人可能有多个电话号码
  - **在关系模型中，多值属性一定要转化为单值属性(关系的第1范式)**
  - 可空值属性和非空值属性：每个实例的该属性值可以是或不能是空值
  - 导出属性
    - **由其他属性计算而得**
    - **例如由“出生年份”可以得出“年龄”**

# E-R模型--数学建模之基本思想

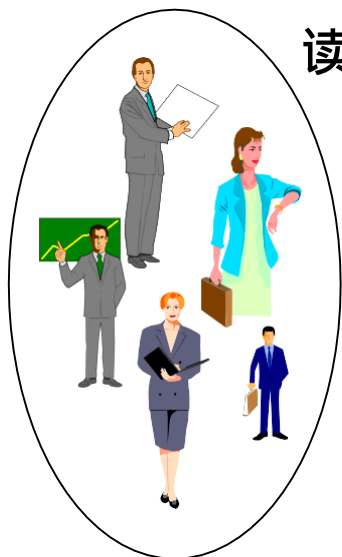


➤ 实体中每一实例如何区分？

**关键字/码**，实体中能够用其值唯一区分开  
每一实例的属性或属性组合

这是特殊并关键的  
属性，每一实体都  
必须给出关键字

读者( 借书证号, 姓名, 年龄, 性别, 家庭住址 )



读者

借书证号	姓名	年龄	性别	家庭住址
001	张三	25	男	吉林
002	李四	22	女	黑龙江
003	王五	24	男	沈阳
004	杨六	23	女	黑龙江
005	李四	24	男	黑龙江

借书证号是唯一的，其他属性有相重复的值  
所以借书证号是关键字, 其他不是... ..





**实体(Entity):** 一个“实体”表示一个现实和抽象事物的集合，这些事物必须具有相同的属性和特征。这个集合的一个元素就是该实体的一个实例。

- 一个个“用户”，一张张“车票”
- 实体被区分为**独立实体**和**从属实体**；
- 在扩展E-R图中，独立实体又称**强实体**，从属实体又称**弱实体**。



**弱实体**: 没有足够的属性构成主键

Course ( course\_id, course\_name,xxx); 主实体

(comp3010, DB, xxx);

(comm0426, Comm, xxx);

Section(sec\_id, semester, year) ;          弱实体/从属实体

( 001, Spring, 2019);    ->comp3010

( 002, Spring, 2019);    ->comp3010

( 001, Fall, 2019);        ->comp3010

( 002, Fall, 2019);        ->comp3010

( 001, Spring, 2019);    ->comm0426

- ER图，与强实体通过 course\_id做联系
- 模式转化：为上述Section模式(course\_id, sec\_id, semester, year)
- 弱实体模式： Section(course\_id, sec\_id, semester, year)



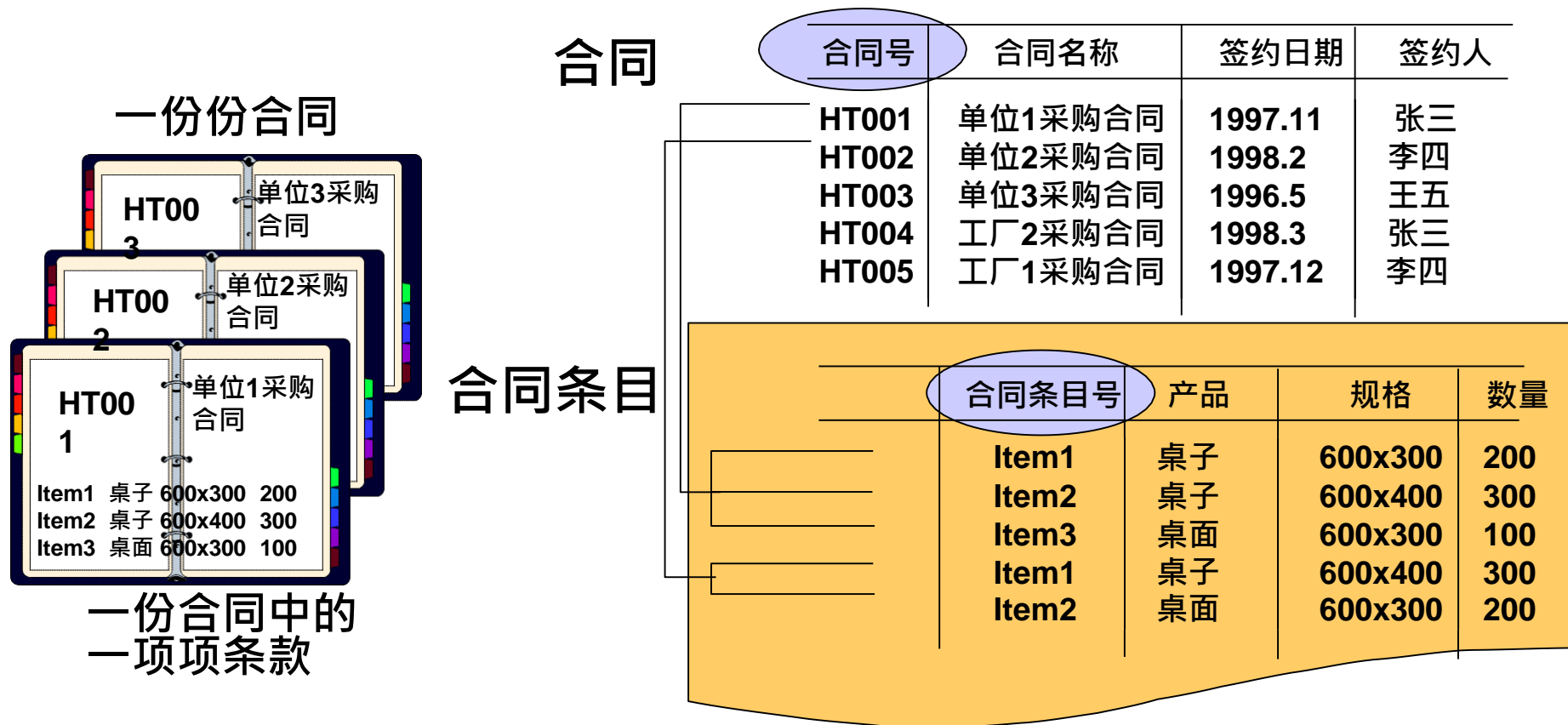
**独立实体**：一个实体的实例都被唯一的标识而不决定于它与其他实体的联系

合同号	合同名称	签约日期	签约人
HT001	单位1采购合同	1997.11	张三
HT002	单位2采购合同	1998.2	李四
HT003	单位3采购合同	1996.5	王五
HT004	工厂2采购合同	1998.3	张三
HT005	工厂1采购合同	1997.12	李四

独立实体的关键字属性是自身拥有的属性



**从属实体**：一个实体的实例的唯一标识需要依赖于该实体与其他实体的联系



# E-R模型--数学建模之基本思想



➤ 实体之间是有联系的

**联系**，指一个实体的实例和其他实体实例之间所可能发生的联系

➤ 如，哪位《读者》借阅了哪本《图书》，哪本《图书》放在哪个《书架》上

读者(借书证号,姓名,性别,家庭住址)

借书证号	姓名	性别	家庭住址
001	张三	男	吉林
002	李四	女	黑龙江
003	王五	男	沈阳

图书(书号,书名,作者,出版商)

书号	书名	作者	出版商
B01	数据库	张力	高教
B02	高数	李明	电子
B03	计算思维	王英	高教

借阅(借书证号,书号)

借书证号	书号
001	B01
002	B02
003	B01

联系是要表达的要素。无联系的实体是没有意义的？

有什么样的联系要区分呢？

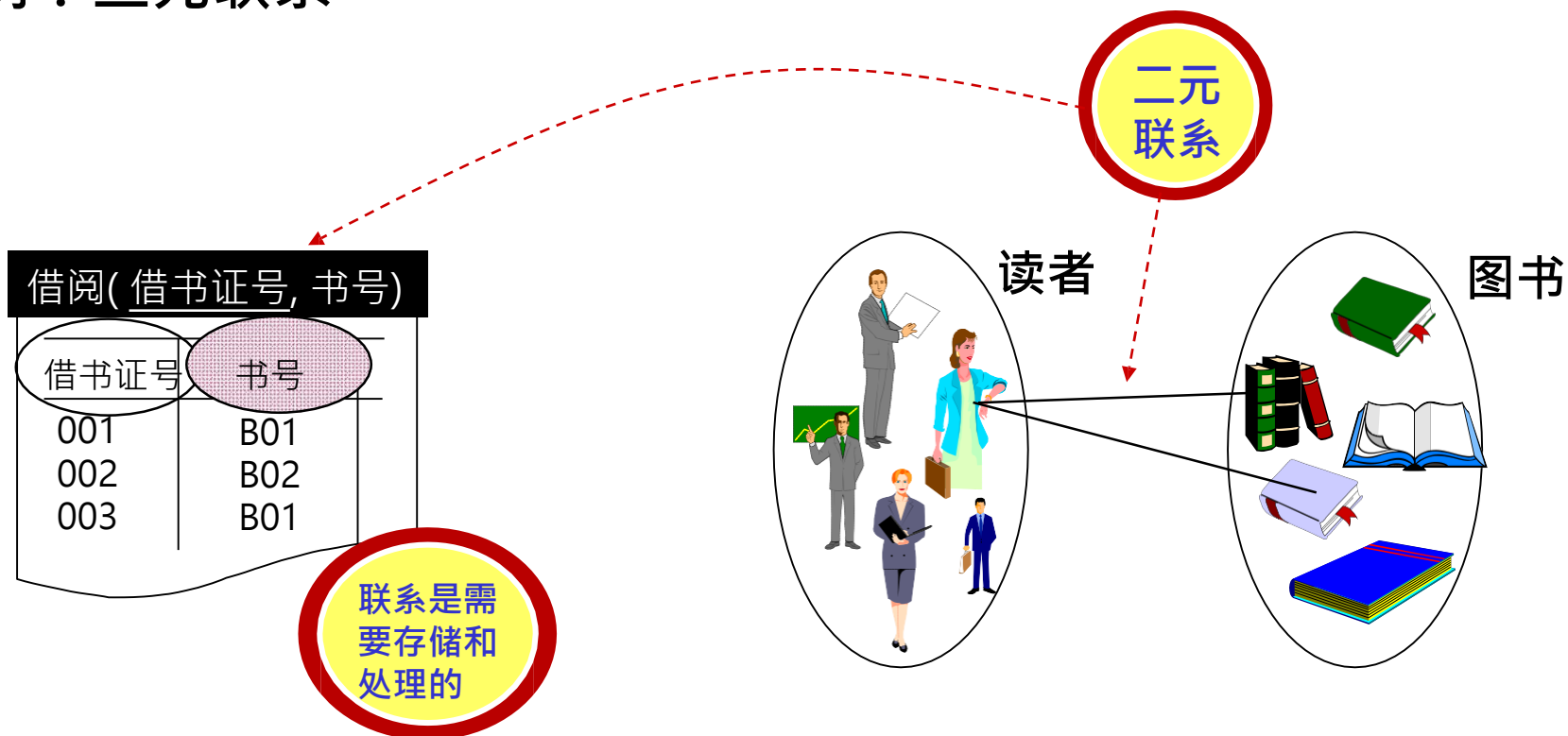
# E-R模型--数学建模之基本思想



参与发生联系的实体的数目，称为联系的度或元。

✓ 联系有一元联系、二元联系和多元联系

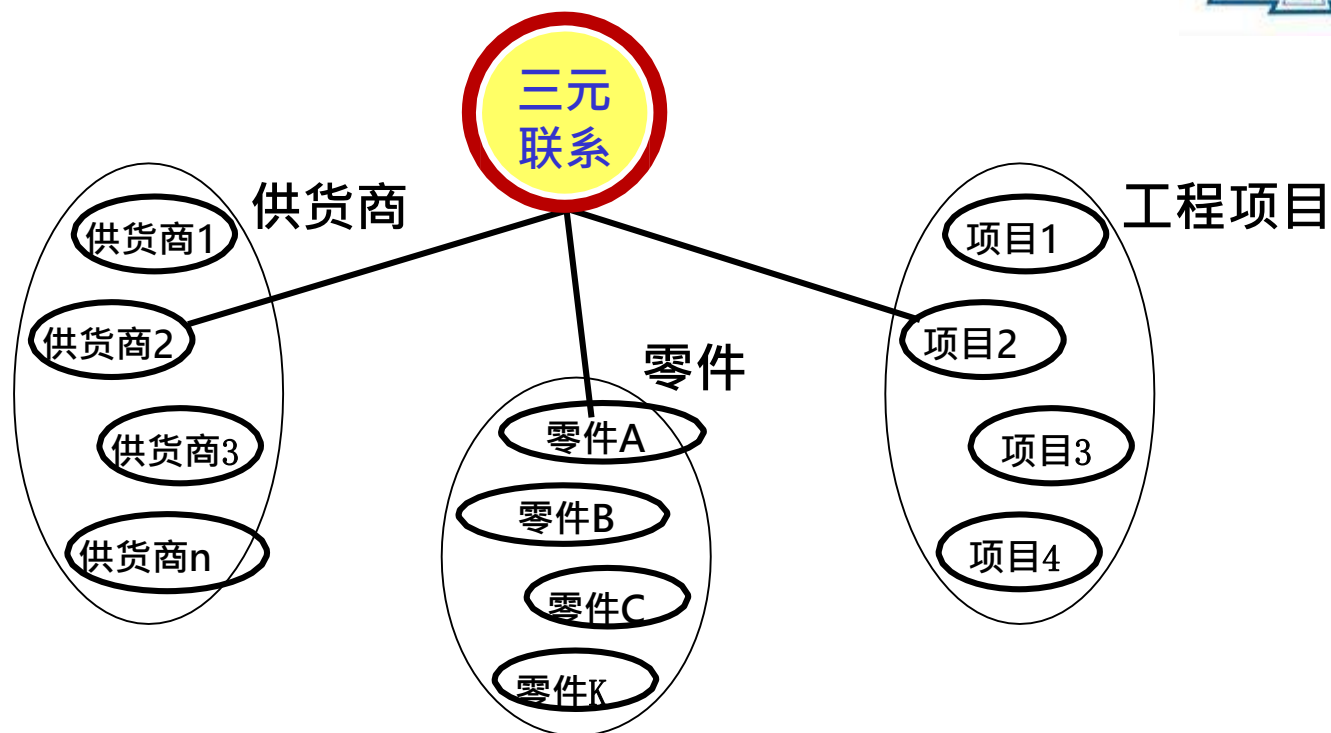
示例：二元联系



# E-R模型--数学建模之基本思想



## 示例：三元联系



实体是相对稳定的，但联系是多样化的

已发生供货(供货商号，零件号，工程项目号，数量，日期)

供货商号	零件号	项目号	数量	日期
供货商1	零件A	项目1	100	20160304
供货商2	零件B	项目2	200	20160208
供货商3	两件C	项目3	100	20160405

允许发生供货(供货商号，零件号，工程项目号)

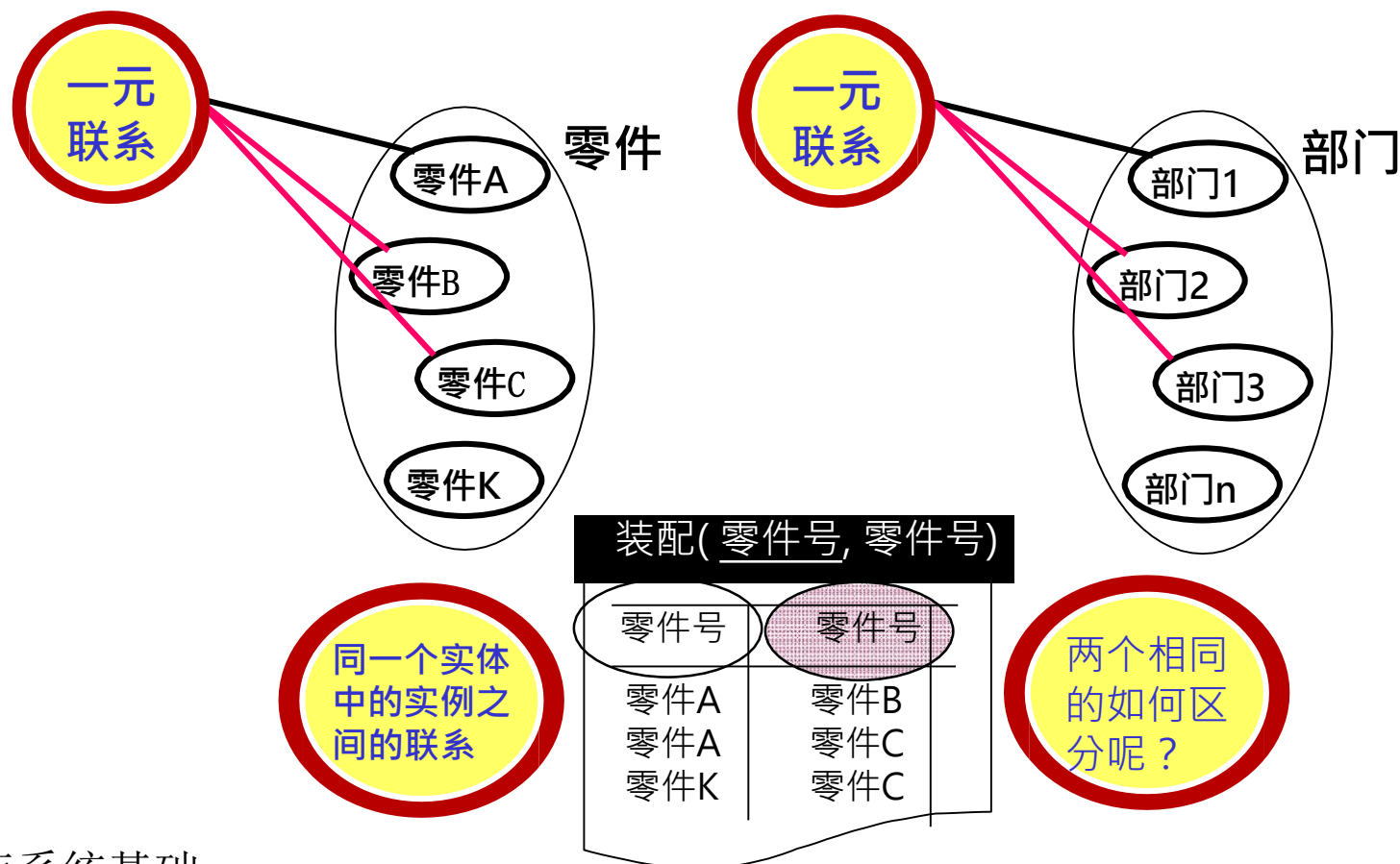
供货商号	零件号	项目号
供货商1	零件A	项目1
供货商2	零件B	项目2
供货商3	两件C	项目3

# E-R模型--数学建模之基本思想



## 示例：一元联系

- ✓ 零件A 由 零件B 和 零件C 装配构成
- ✓ 部门1 下设两个子部门 部门2 和 部门3

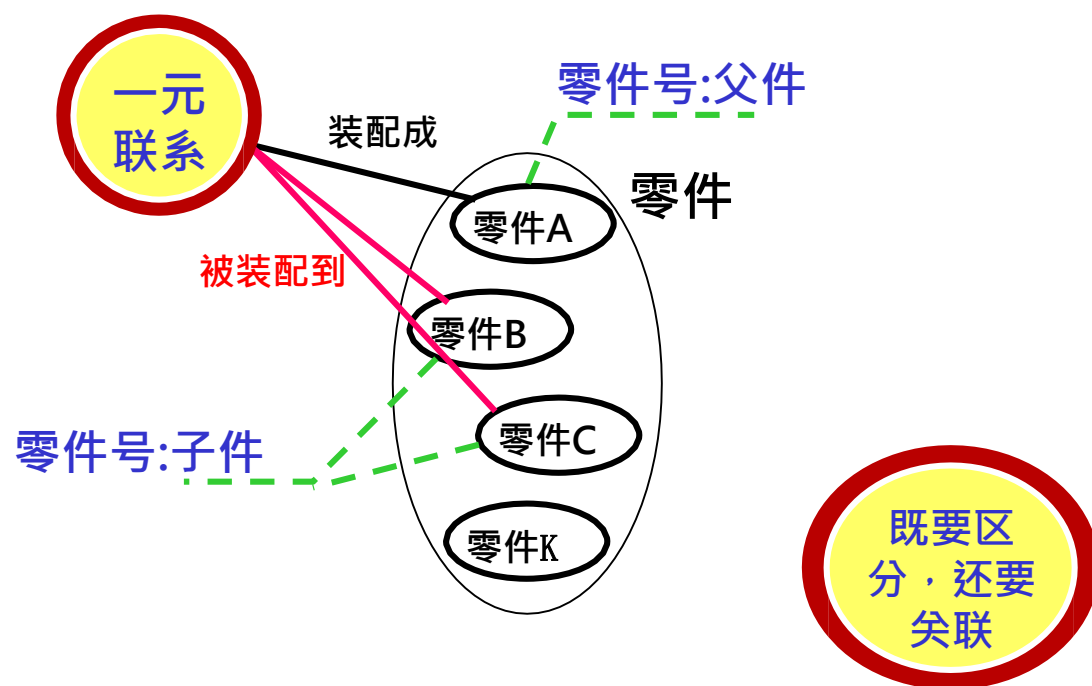






## 角色(作用)

- ✓ 实体在联系中的作用称为实体的角色
- ✓ 当同一实体的不同实例参与一个联系时，为区别各实例参与联系的方式，需要显式指明其角色(role)



零件(零件号, 零件名,...)

零件号	零件名
零件A	紧固件
零件B	螺栓
零件C	螺母

装配(父件号, 子件号)

父件号	子件号
零件A	零件B
零件A	零件C
零件K	零件C

# E-R模型--数学建模之基本思想



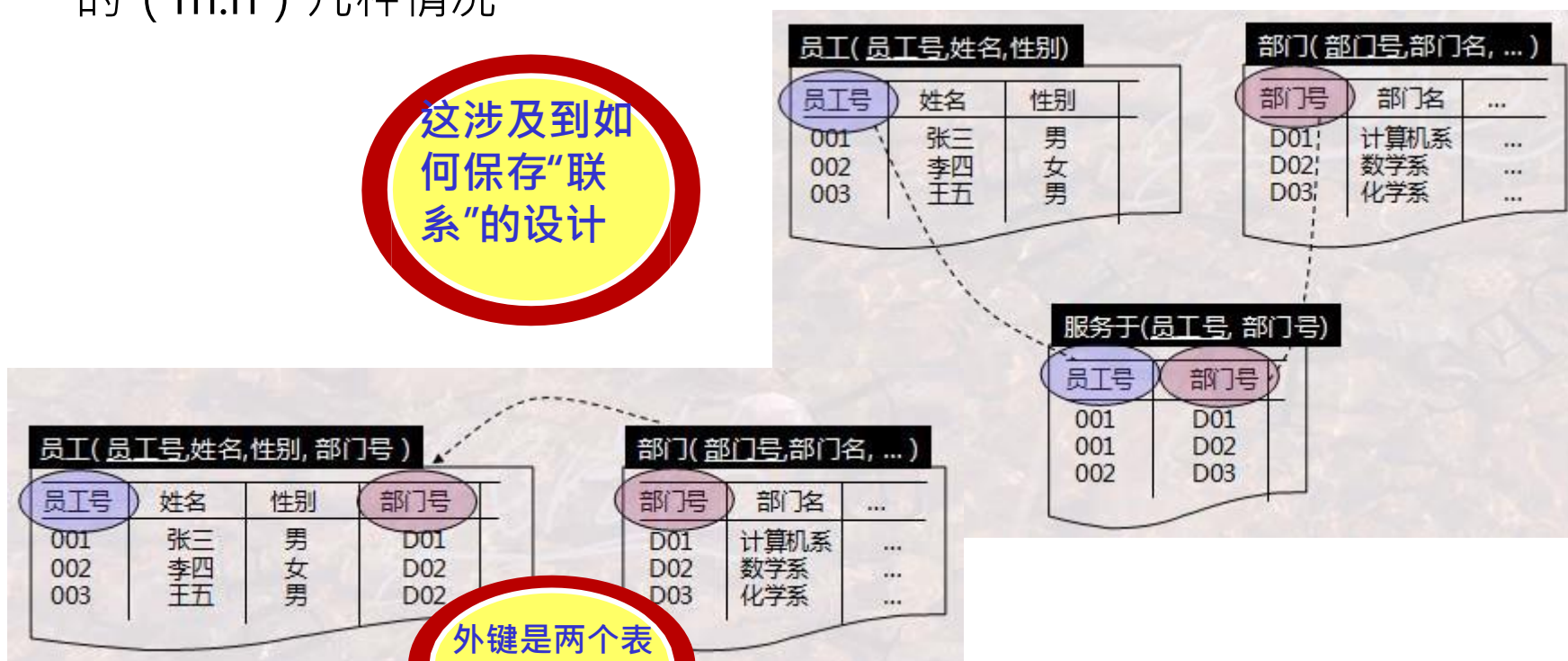
- 实体之间的联系有很多种类
- 二元联系：一对一、一对多和多对多联系
- **一对一联系(1:1)**：实体A的实例只能和实体B的一个实例发生联系，反之，实体B的实例也只能和实体A的一个实例发生联系。  
一个“经理”只管理一个“商店”，一个“商店”只能有一个“经理”
- **一对多联系(1:m和m:1)**：实体A的实例能和实体B的多个实例发生联系，反之，实体B的实例只能和实体A的一个实例发生联系。  
一个“画家”可以绘制多幅“作品”，一幅“作品”只能由一个“画家”来完成
- **多对多联系(m:n)**：实体A的实例可以和实体B的多个实例发生联系，反之，实体B的实例也可以和实体A的多个实例发生联系。  
一位同学可以选学多门课程，一门课程可由多个人来选学

# E-R模型--数学建模之基本思想



- 联系的**基数** ( Cardinalities ) : 实体实例之间的联系的数量，即一个实体的实例通过一个联系能与另一实体中相关联的实例的数目
- 常见的映射基数如上，有一对一的 ( 1:1 ) ，一对多的 ( 1:m ) ，多对多的 ( m:n ) 几种情况

这涉及到如何保存“联系”的设计



外键是两个表之间“联系”的纽带

# E-R模型--数学建模之基本思想



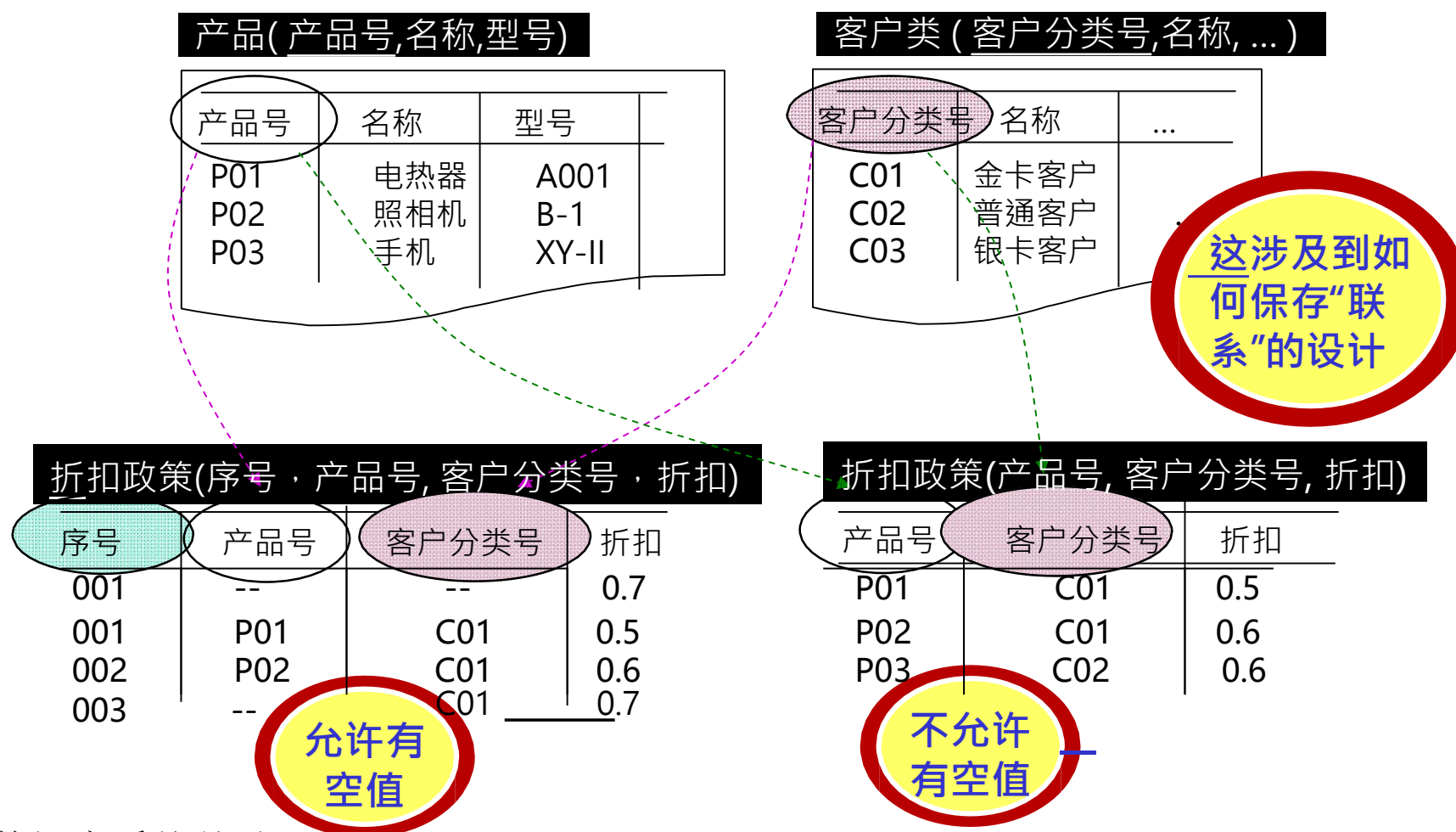
- 进一步，联系的基数还要区分是0个、1个、不定数目的多个还是固定数目的多个(即，对每个实体的实例而言是否必须存在)
- 通常以实体参与联系的最小基数和最大基数来标记(MinCard..MaxCard)  
“书架”参与“存放图书”联系的基数为(0..m)，而“图书”参与此联系的基数为(1..1)  
一个“书架”可以存放0或多本“图书”，一本“图书”只能存放在1个“书架”
- ✓ **完全参与联系**，即该端实例至少有一个参与到联系中, 最小基数为1 (1..m)；
- ✓ **部分参与联系**，即该端实例可以不参与联系，最小基数为0 (0..m)

为什么需要  
区分呢？

# E-R模型--数学建模之基本思想

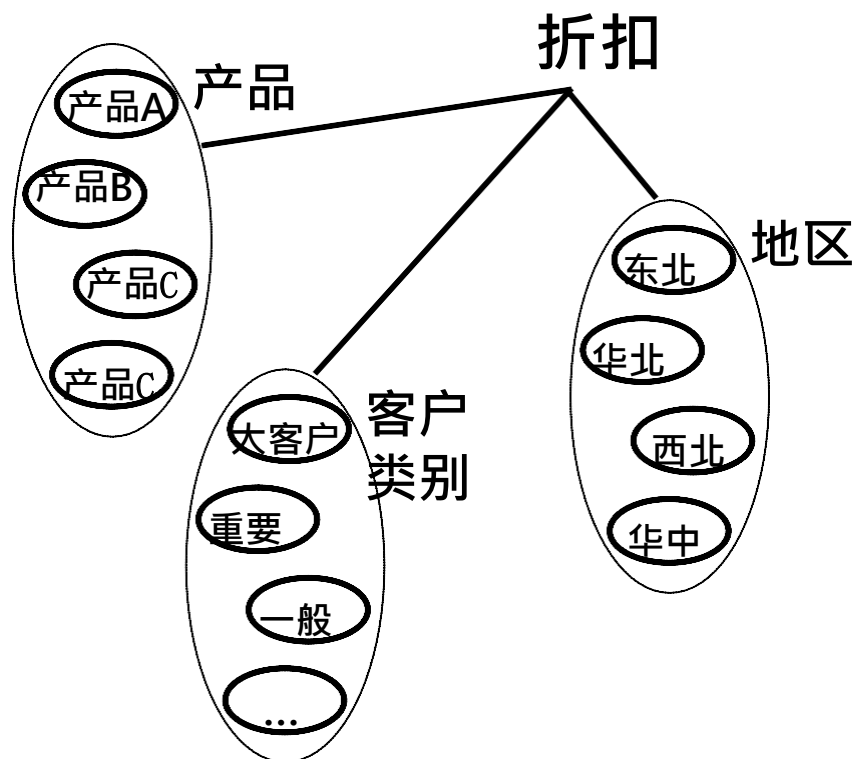


## 示例：完全参与联系和部分参与联系





## 示例：如何解读此三元联系



折扣政策表

产品	客户类别	地区	时间周期	...	折扣
产品 A	大客户	东北	8-10		0.95
产品 A	----	----	----		0.90
----	大客户	----	----		0.90
产品 A	----	东北	----		0.93
产品 A	大客户	东北			
产品 B	重要客户	华北			
产品 C	一般客户	西北			
	所有客户				

涉及到空  
值含义的  
解读

# E-R模型—表达方法之Chen方法



## E-R模型的几种图示化表达方法

❖ Chen 方法

❖ Crow's Foot方法（讲）

IDEF1X方法(工程化方法)（不讲）





## Chen方法

实体：**矩形框**

属性：**椭圆**

多值属性：**双线椭圆**

导出属性：**虚线椭圆**

关键字/码：**下划线**

连接实体和属性：**直线**

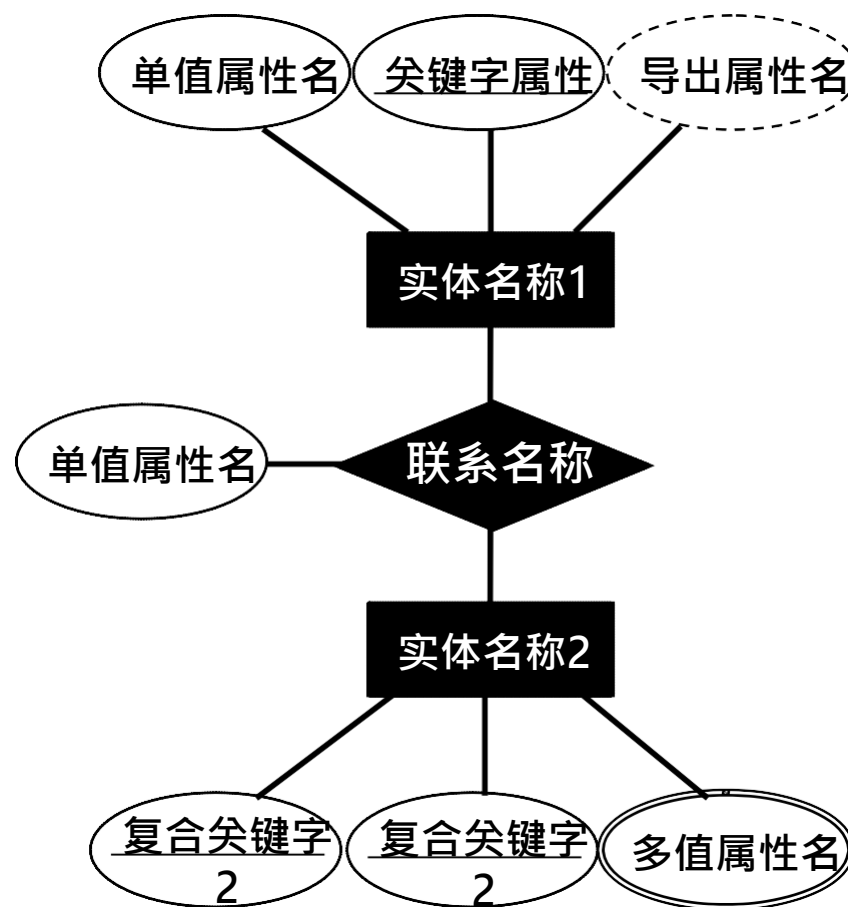
联系：**菱形框**

连接实体与联系：**直线**

连接联系和属性：**直线**

复合关键字：**标有相同数字**

多组关键字：**标有不同数字**

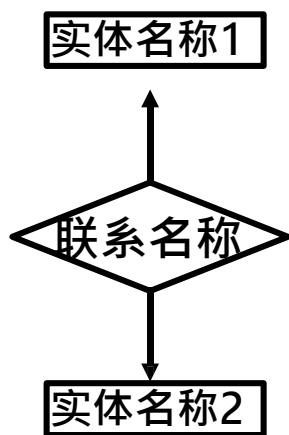




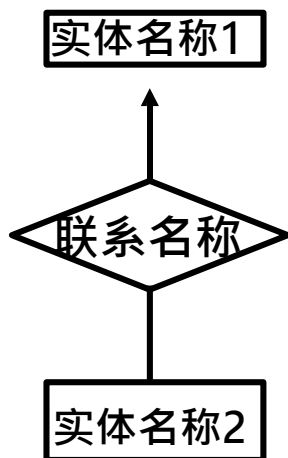
# E-R模型—表达方法之Chen方法



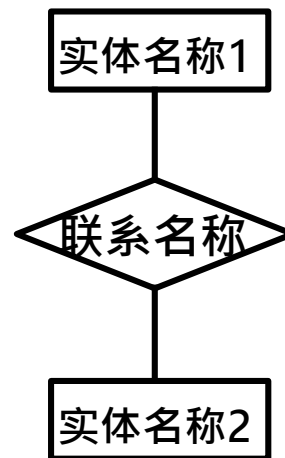
- ✓ 1:1联系：箭头直线，由联系指向实体
- ✓ 1:m联系：指向1端为箭头直线，指向多端为无箭头直线
- ✓ m:n联系：无箭头直线
- ✓ 完全参与联系：双直线
- ✓ 部分参与联系：单直线



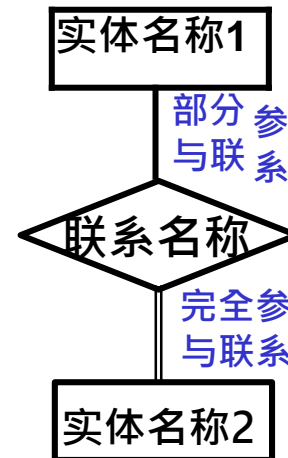
1:1联系



1:m联系



m:n联系

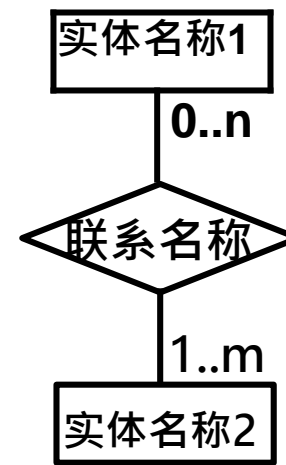
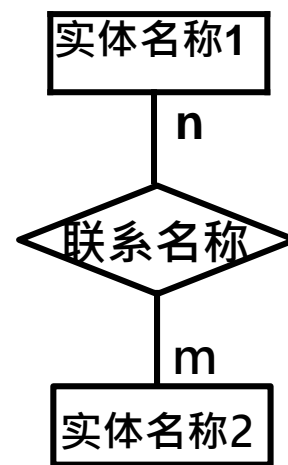
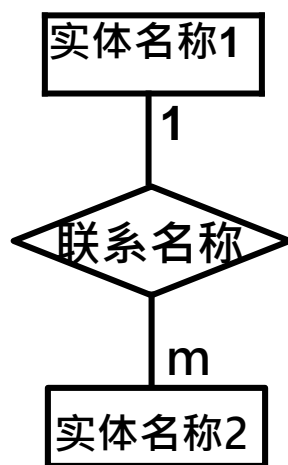
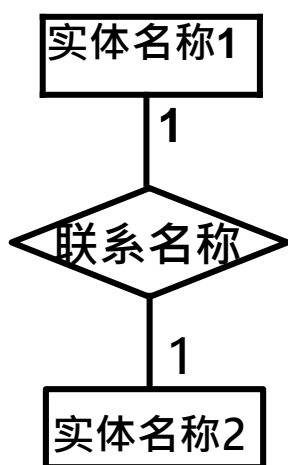


# E-R模型—表达方法之Chen方法



1:1, 1:m, m:n的联系也可以如下区分：

- 1端实体-直线旁标1
- 多端实体-直线旁标m或n 完全/部分参与联系也可以标注最小基数..最大基数进行区分，最小基数0 的为部分参与联系，最小基数1的为完全参与联系
- 直线旁标1..1, 0..1, 1..m, 0..m：

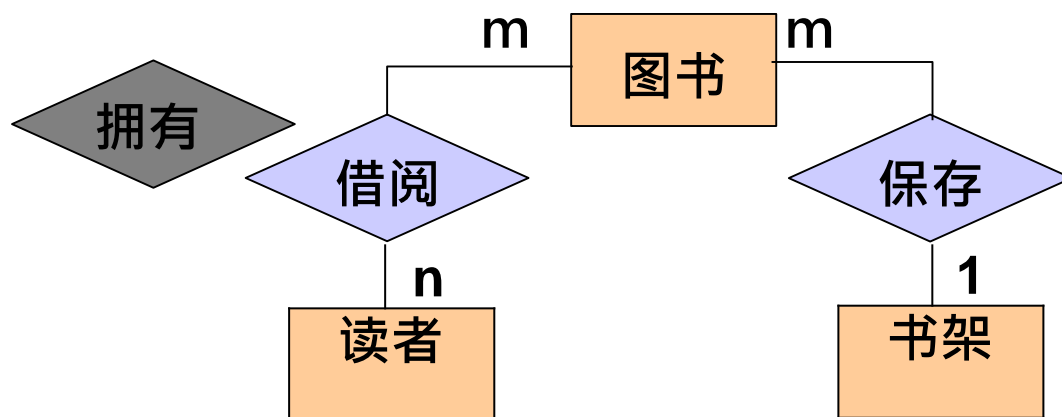


# E-R模型—表达方法之Chen方法



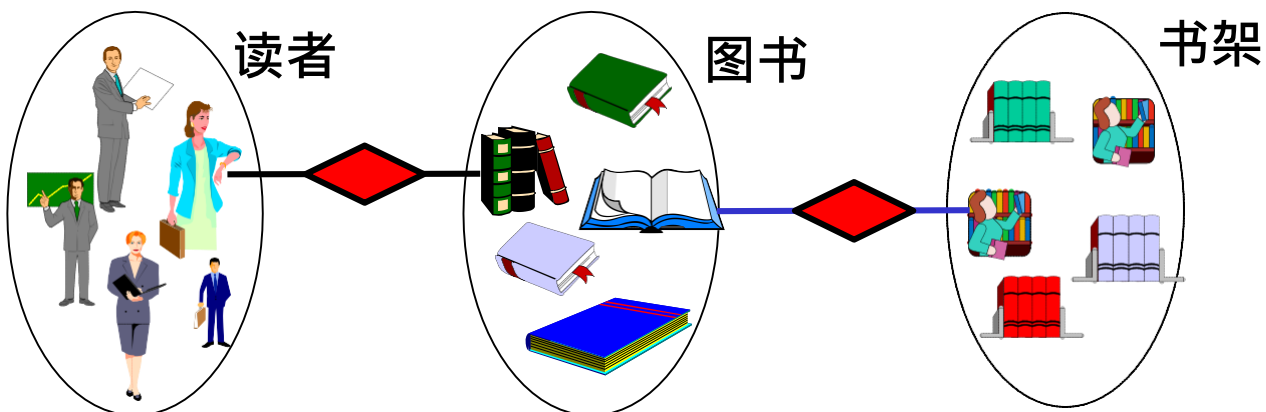
## Chen方法表达示例—基本的实体-联系表达

示例：图书管理的E-R Diagram



强调：联系  
也需要命名  
和表达

强调：实体之间  
可能有多种含义  
的联系—你关注  
的是哪一个？

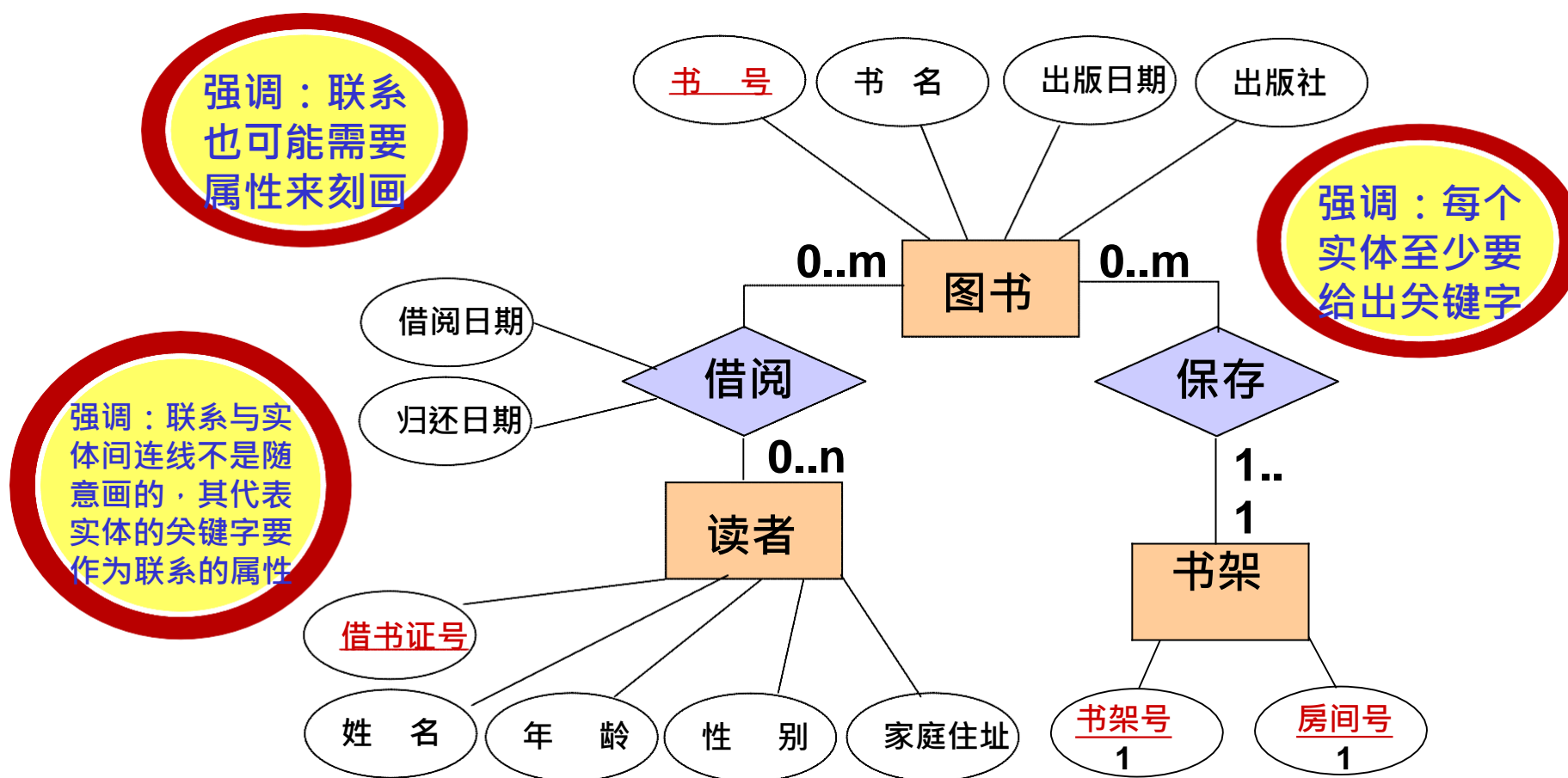


# E-R模型—表达方法之Chen方法



## Chen方法示例—完整的实体-属性-联系的表达

示例：图书管理的E-R Diagram

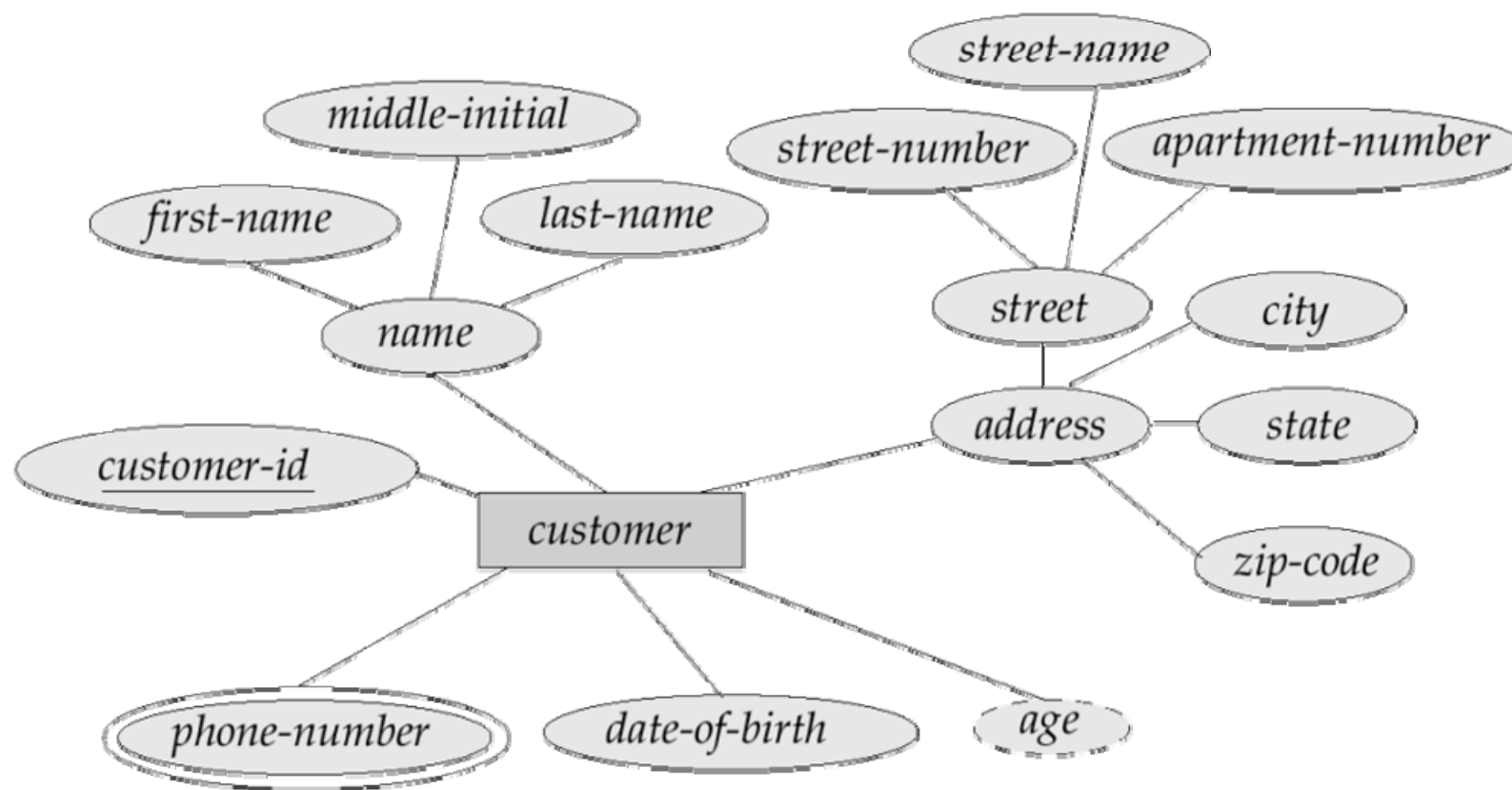


# E-R模型—表达方法之Chen方法



## Chen方法示例—带组合、多值和导出属性的表达

示例：客户实体的E-R Diagram

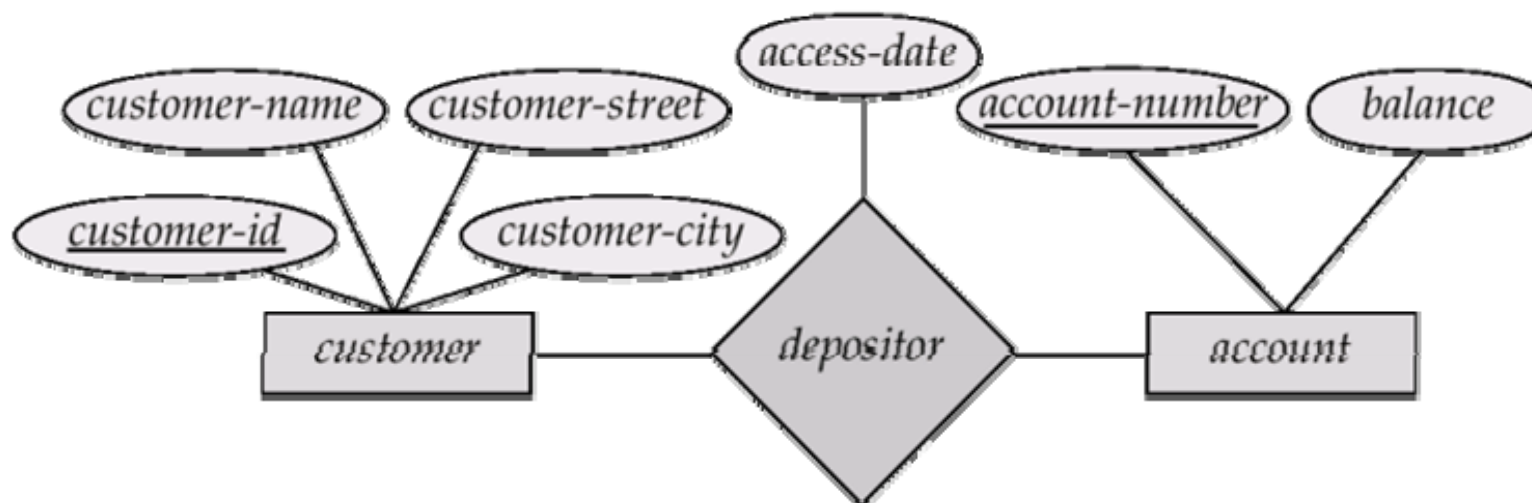


# E-R模型—表达方法之Chen方法



## Chen方法示例—带组合、多值和导出属性的表达

示例：账户管理的E-R Diagram



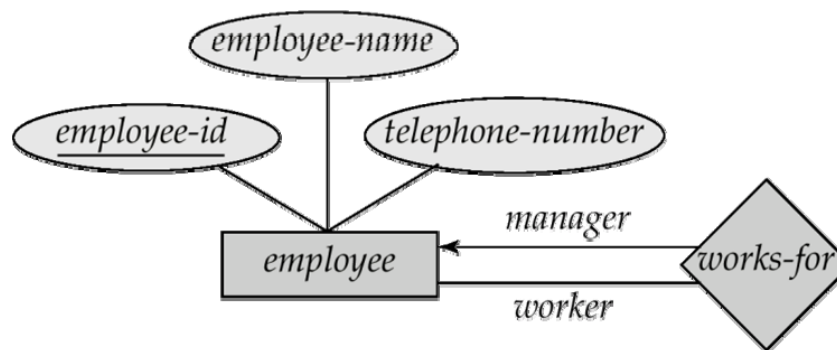
# E-R模型—表达方法之Chen方法



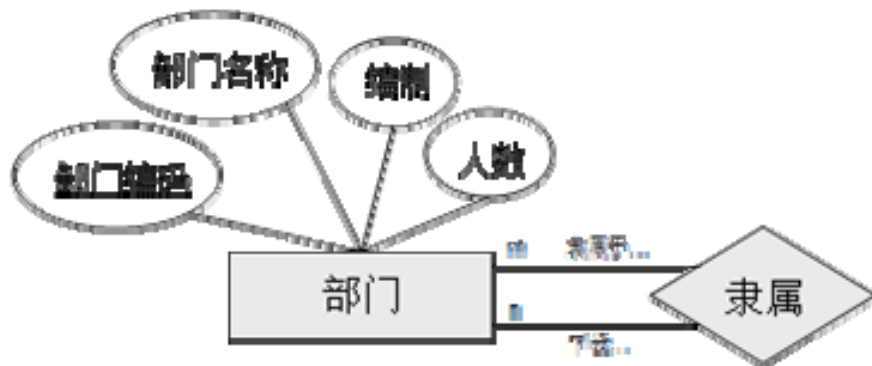
## Chen方法示例—联系的角色表达

直线上标记有文字：联系的角色

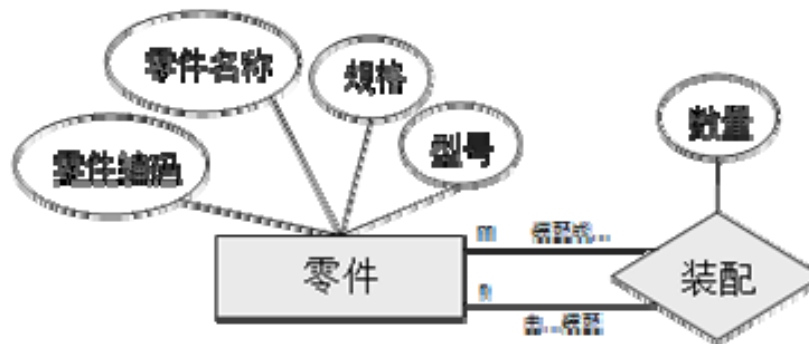
示例：雇员关系的E-R Diagram



示例：组织之间关系的E-R Diagram



示例：产品构成关系的E-R Diagram

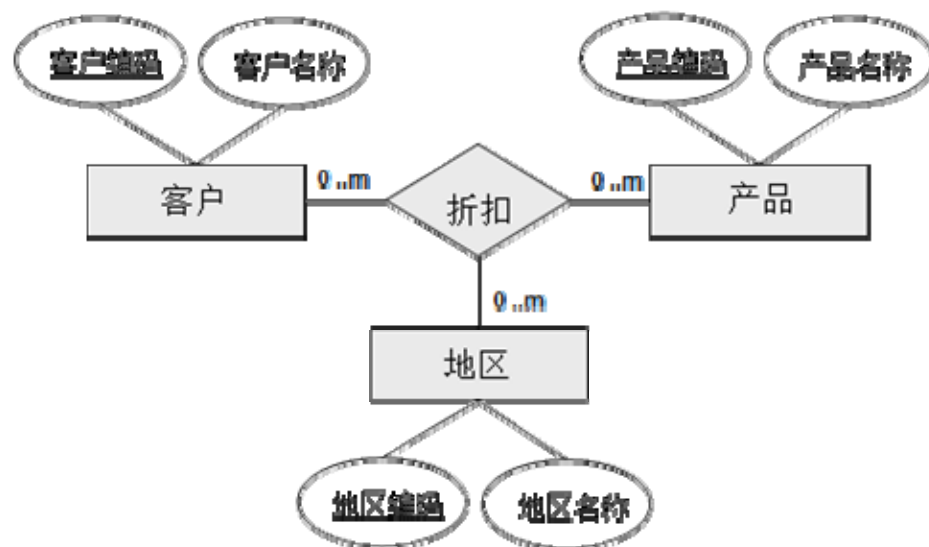


# E-R模型—表达方法之Chen方法



## Chen方法示例—参与联系基数的表达

示例：客户-产品折扣关系的E-R Diagram





# E-R模型—建模案例讲解（Chen方法）



## 示例：仓储管理的E-R Diagram

- 需求理解：
- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库---哪个零件存在哪个仓库中
- 管理职工---哪个职工管理哪个仓库

这些需求相互之间是有关联的

# E-R模型—建模案例讲解（Chen方法）



## Step1 理解需求，寻找实体

- E-R图建模首先要找出问题领域的实体，即找出有哪些类/实体
- 能够用一个个、一件件、一串串等重叠量词形容的，而不是一个、一件...

供应商

仓库

职工

项目

零件

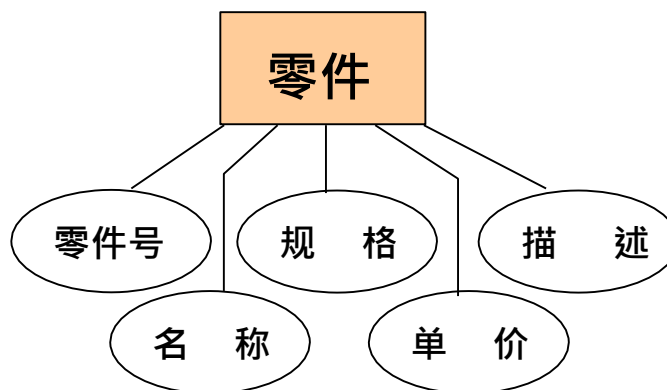
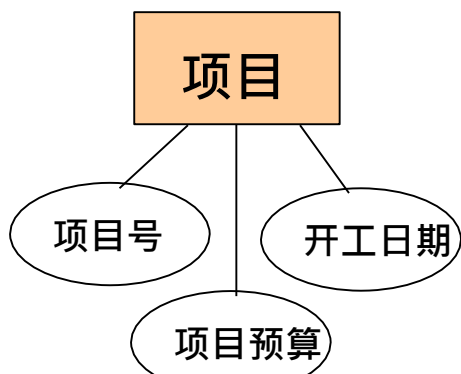
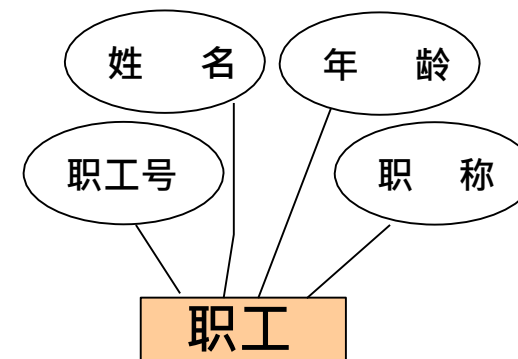
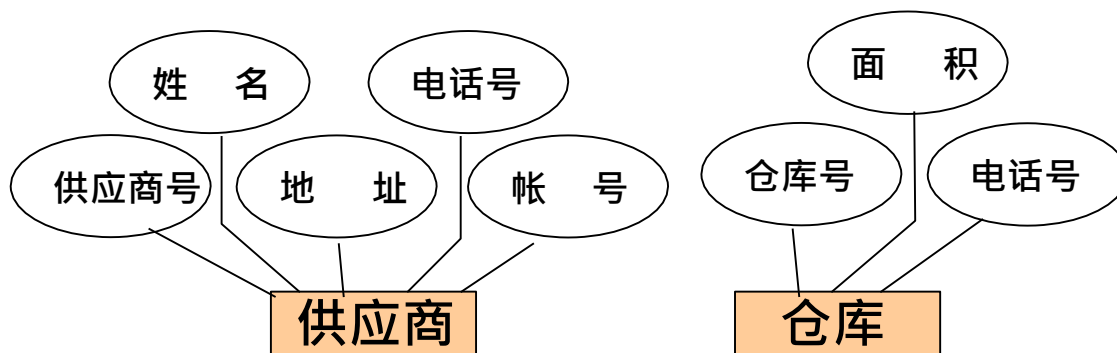
要覆盖需求涉及到的可独立管理的每一类事物

- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库---哪个零件存在哪个仓库中
- 管理职工---哪个职工管理哪个仓库

# E-R模型—建模案例讲解（Chen方法）



## Step2用属性刻画每一个实体



至少要给出重要的属性

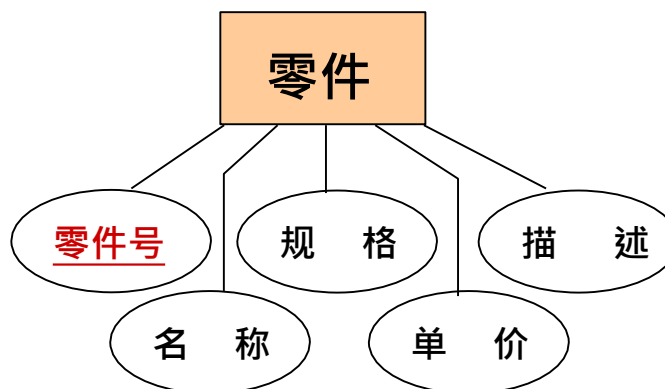
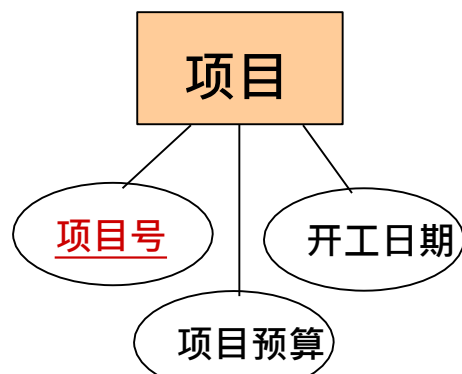
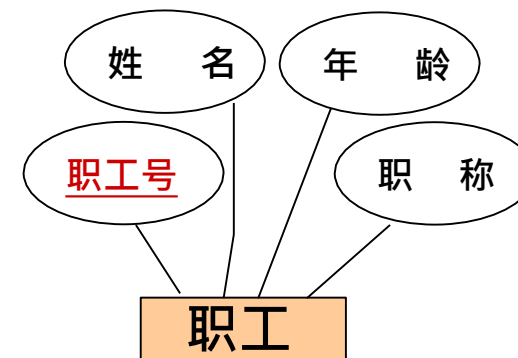
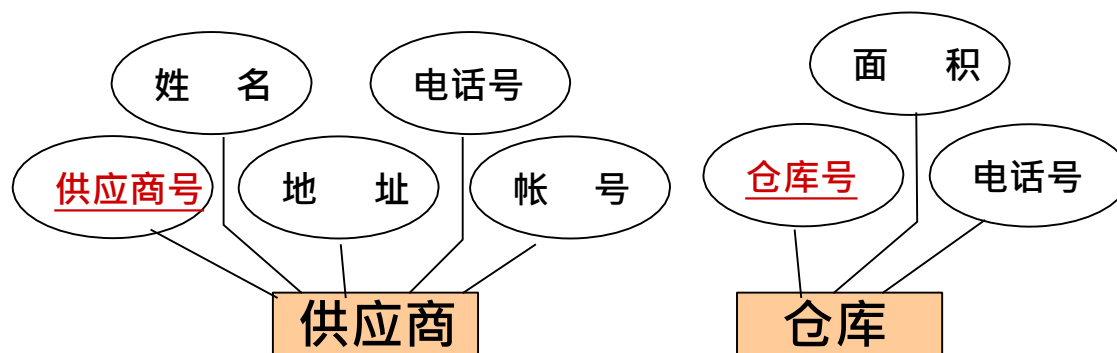
零件中为什么没有供应商和仓库呢？

- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库---哪个零件存在哪个仓库中
- 管理职工---哪个职工管理哪个仓库

# E-R模型—建模案例讲解（Chen方法）



## Step3确定每一个实体的关键字/码



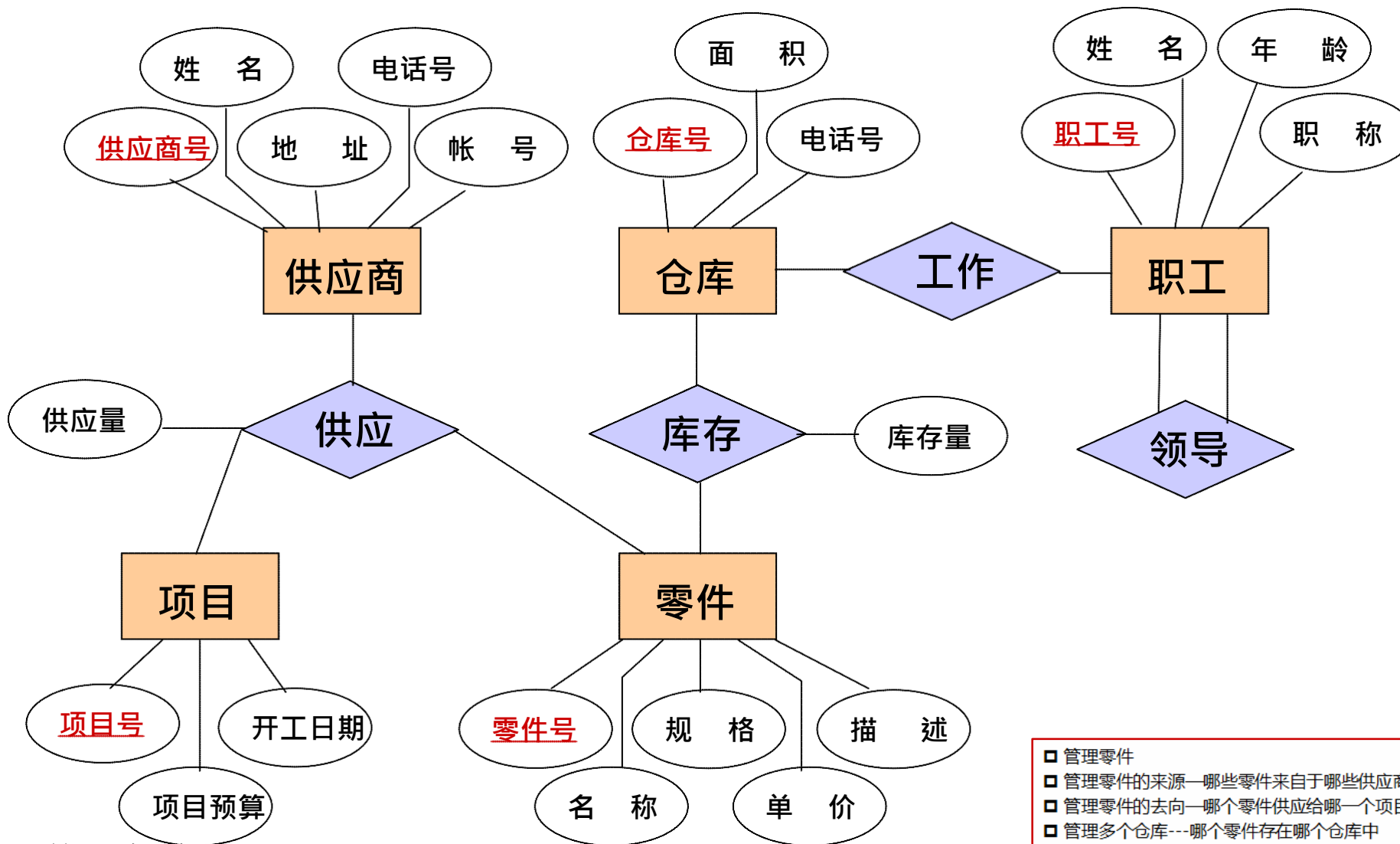
关键字属性  
是必须要标  
记清晰的

- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库---哪个零件存在哪个仓库中
- 管理职工---哪个职工管理哪个仓库

# E-R模型—建模案例讲解（Chen方法）



## Step4数据建模的重点是分析实体之间的联系

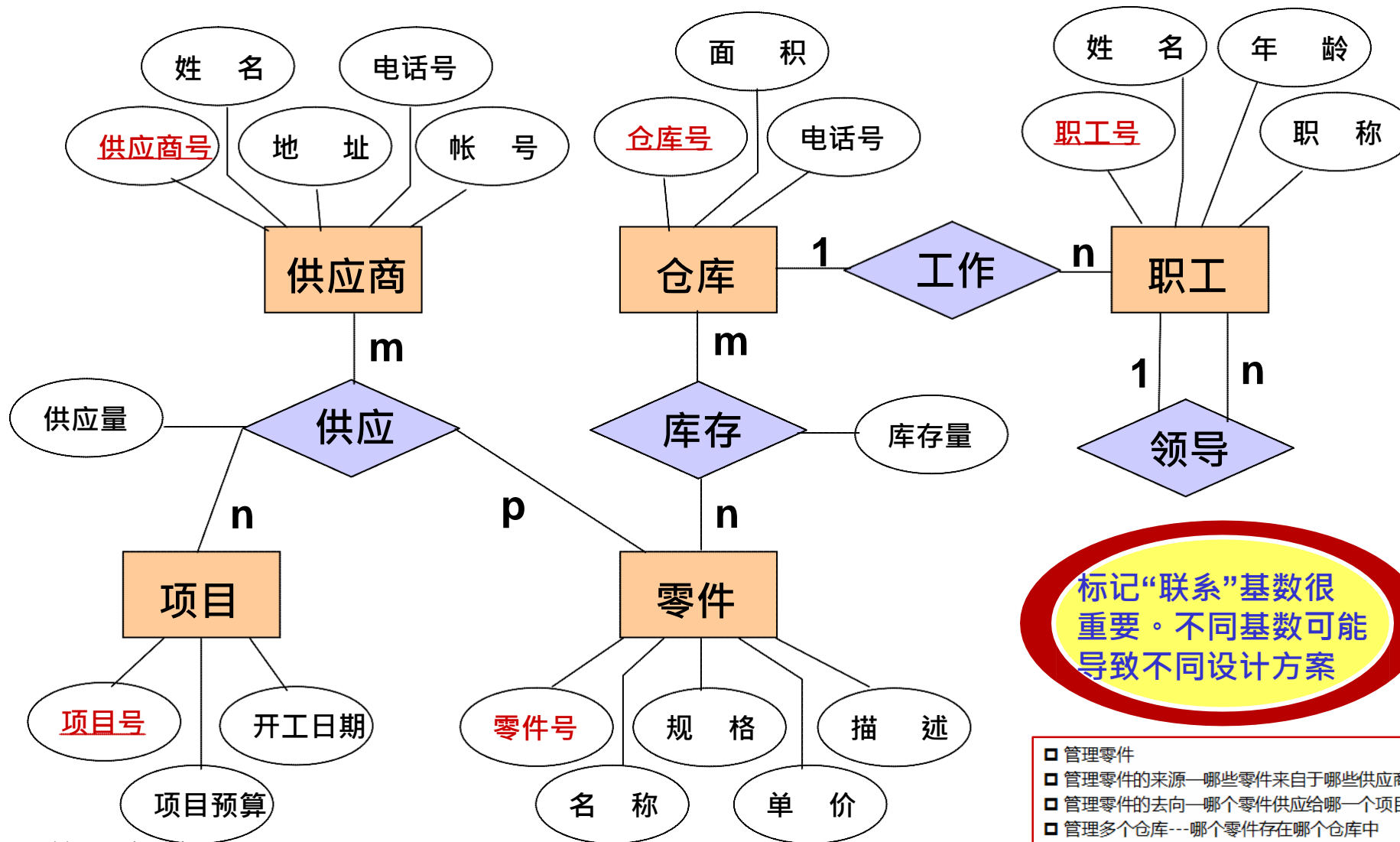


- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库—哪个零件存在哪个仓库中
- 管理职工—哪个职工管理哪个仓库

# E-R模型—建模案例讲解（Chen方法）



## Step4数据建模的重点是分析实体之间的联系



标记“联系”基数很重要。不同基数可能导致不同设计方案

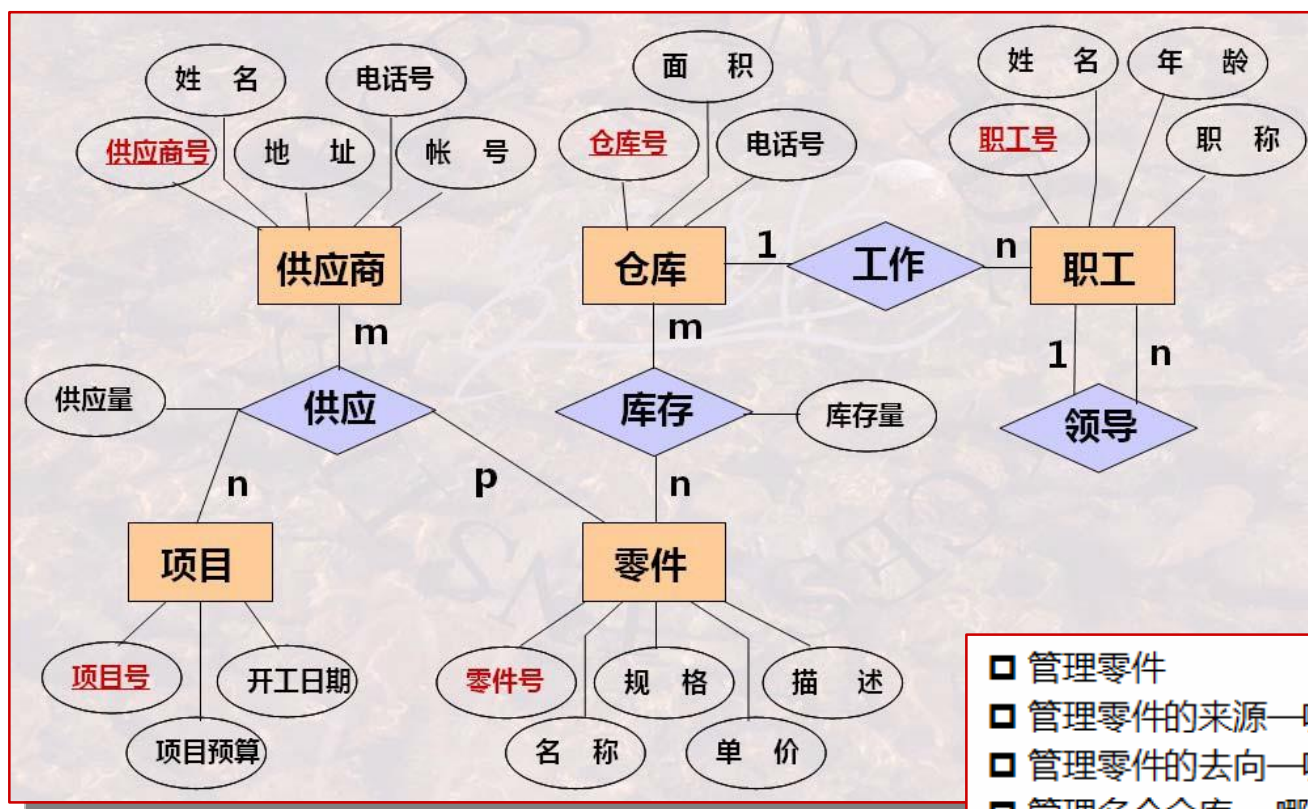
- 管理零件
- 管理零件的来源—哪些零件来自于哪些供应商
- 管理零件的去向—哪个零件供应给哪一个项目使用
- 管理多个仓库—哪个零件存在哪个仓库中
- 管理职工—哪个职工管理哪个仓库



# E-R模型—建模案例讲解（Chen方法）



## Step5检查是否覆盖了需求



E-R图是对需求的  
理解与表达--  
要覆盖需求

E-R图的绘制要  
符合规范--  
Chen方法

- ❑ 管理零件
- ❑ 管理零件的来源—哪些零件来自于哪些供应商
- ❑ 管理零件的去向—哪个零件供应给哪一个项目使用
- ❑ 管理多个仓库---哪个零件存在哪个仓库中
- ❑ 管理职工---哪个职工管理哪个仓库

# E-R模型—表达方法之Crow's Foot方法



## E-R模型的几种图示化表达方法

❖ Chen 方法

❖ **Crow's Foot方法**



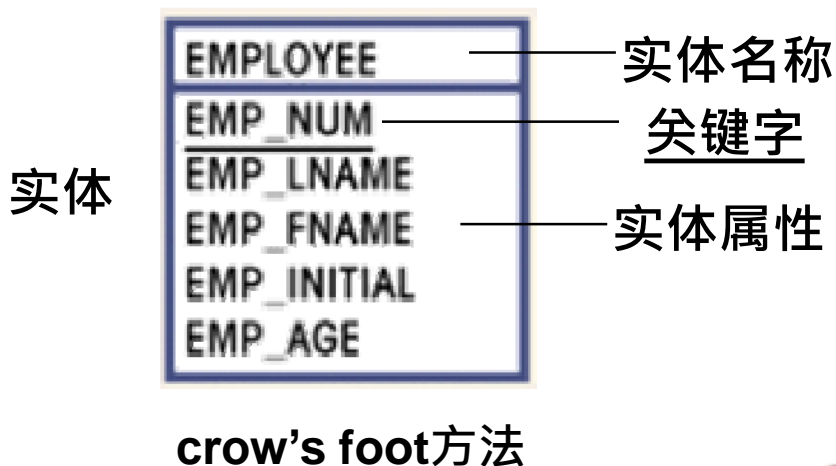
# E-R模型—表达方法之Crow's Foot方法



## Crow's foot方法

- 实体：矩形框，实体的名称写在横线上面
- 属性：实体框横线的下面
- 关键字：属性下加下划线

Crow's foot的优点

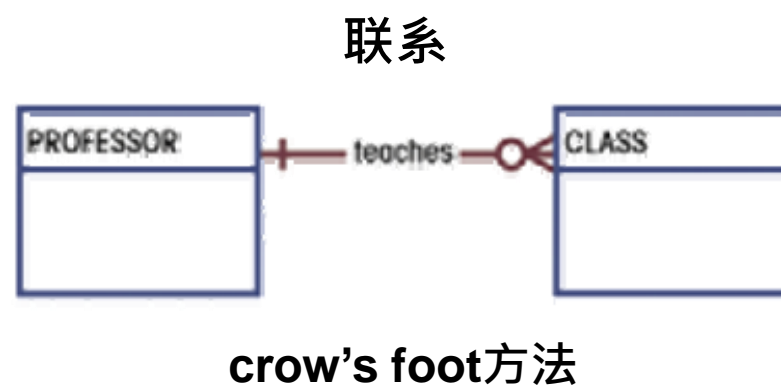
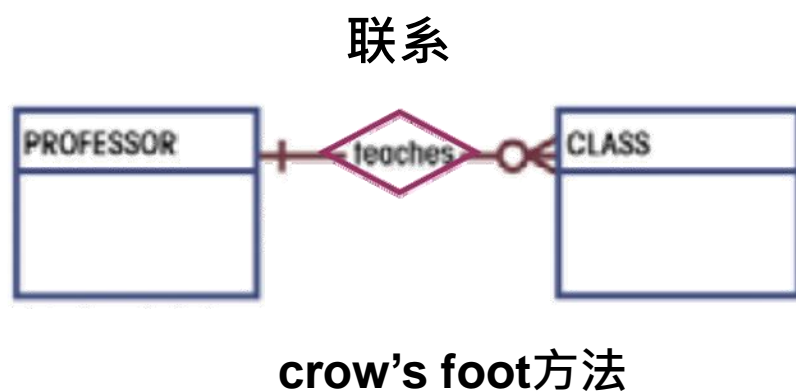


注意与Chen方法区别

# E-R模型—表达方法之Crow's Foot方法



- 联系：菱形框表示，也可以将菱形框省略而直接以联系名来替代

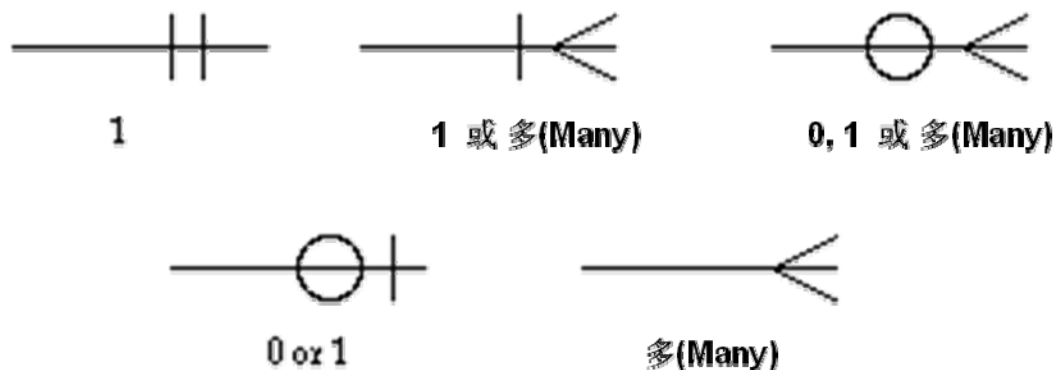


- 注意与chen方法的区别

# E-R模型—表达方法之Crow's Foot方法



## □ 联系的基数表示方法



crow's foot方法

Crow's foot联系基数的表达很形象

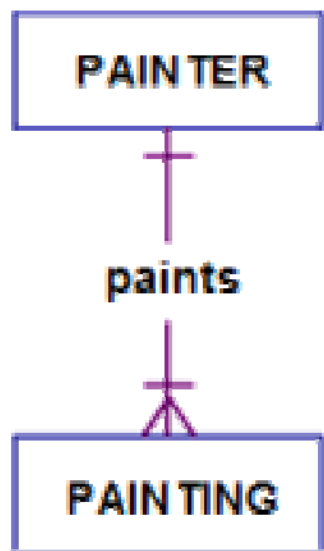
## □ 注意与chen方法的区别

# E-R模型—表达方法之Crow's Foot方法



## Crow's foot方法表达示例：基本实体-联系的表达

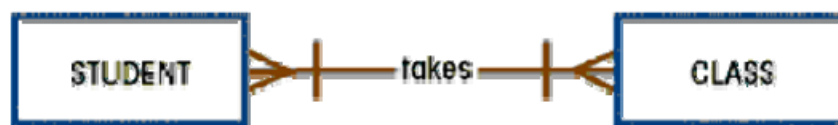
示例：画家与作品之间的关系



**1:m(one-to-many)情况**

**业务规则:**一个painter可以绘制多幅painting, 而一幅painting只能由一个painter来绘制

示例：学生与课程班之间的关系



**m:n(many-to-many)情况 业务规则:**

一个student可以选择多个class,  
而一个class又可由多个  
student来组成

强调：一种图形代表着一类业务规则

强调：不是在画图，而是在表达业务规则

# E-R模型—表达方法之Crow's Foot方法



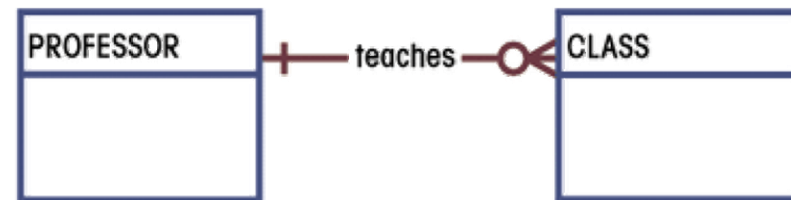
## Crow's foot方法表达示例：完全参与和部分参与联系的表达

示例：教师与课程班之间的联系—不同的图表达了不同的联系



**1:m(one-to-many)完全参与联系** 业务规则:

一个professor至少要教一个class( $\geq 1$ ), 而一个class只能由一个professor来教



**1:m(one-to-many)部分参与联系** 业务规则:

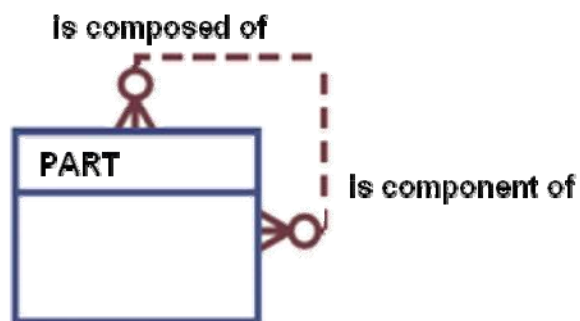
一个professor可以教多个class, 但也可以不教; 一个class只能由一个professor来教

# E-R模型—表达方法之Crow's Foot方法



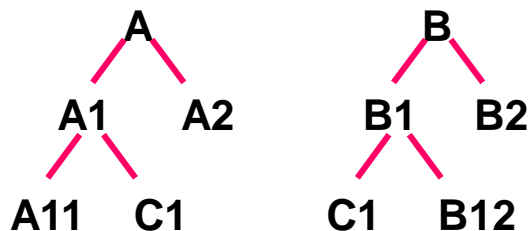
## Crow's foot方法表达示例：一元联系的表达

示例：产品构成暨物项之间的联系

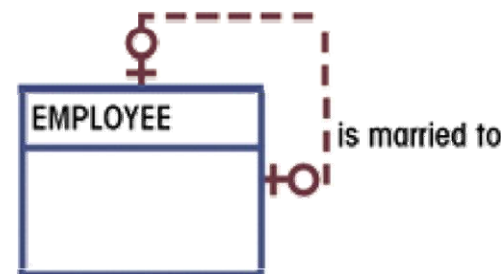


**m:n(many-to-many)部分参与联系**

**业务规则:**一个part可以由多个其他的part构成, 而一个part也可以构成多个其他的part(但都可以不参与)



示例：婚姻暨人与人之间的联系



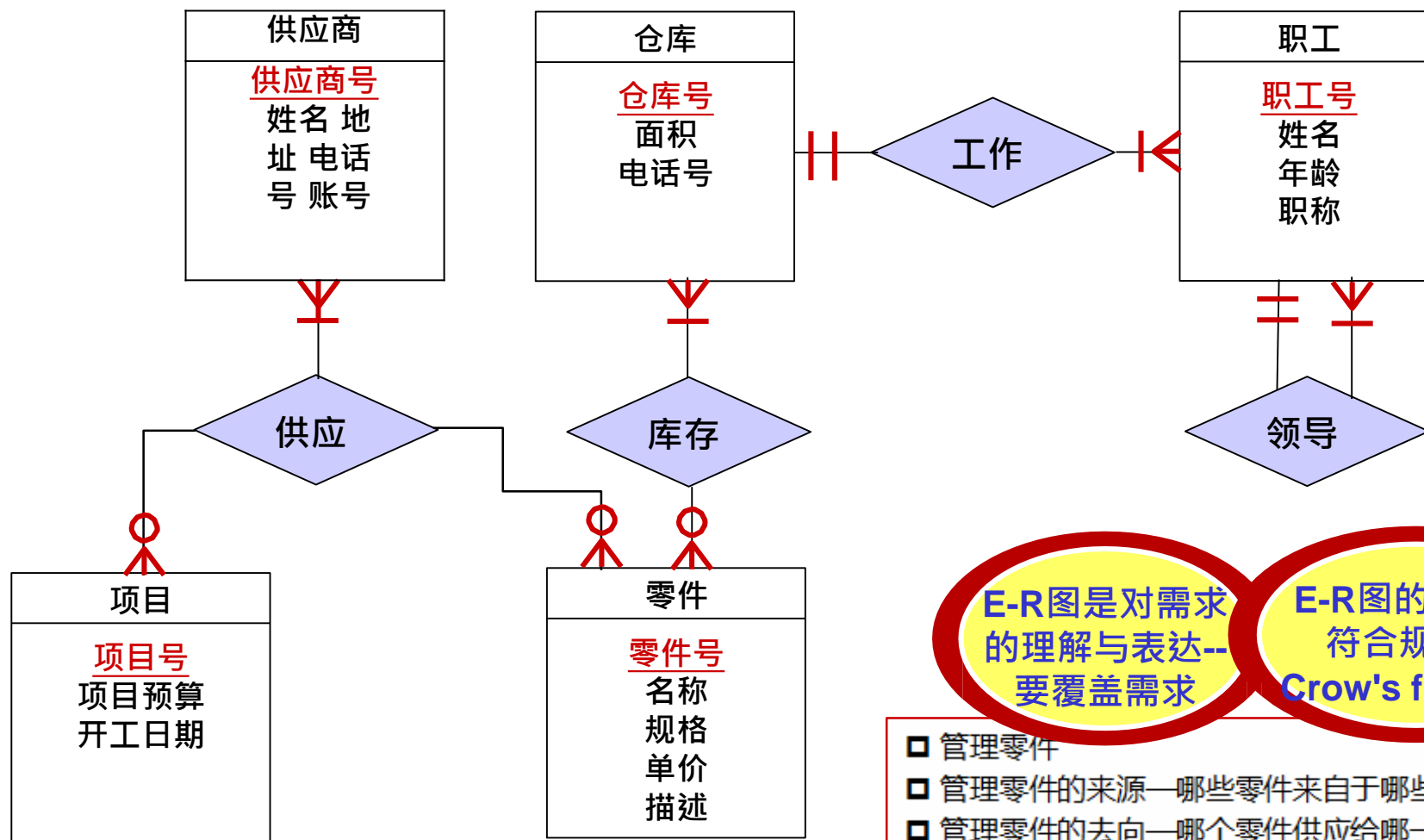
**1:1(one-to-one)部分参与联系 业务**

**规则:**一个person可以与另一个且只能一个person结婚, 但其可以不结婚

# E-R模型—表达方法之Crow's Foot方法



## 示例：仓储管理的E-R Diagram



E-R图是对需求的理解与表达--要覆盖需求

E-R图的绘制要符合规范--Crow's foot方法

- ❑ 管理零件
- ❑ 管理零件的来源—哪些零件来自于哪些供应商
- ❑ 管理零件的去向—哪个零件供应给哪一个项目使用
- ❑ 管理多个仓库---哪个零件存在哪个仓库中
- ❑ 管理职工---哪个职工管理哪个仓库