

第二章 概念模型 --本章内容

- 数据模型
- 概念模型 (ER 模型)-基本概念
- 扩展的 ER 模型
- 实体与联系
- 实例分析

数据模型

- z 作用：对现实世界进行抽象。找到所关心的数据，并将其组织起来，使之可以存放到数据库中。
- z 模型：是描述数据的概念集合。按应用的不同目的分为两类：概念模型、数据模型。
- z 概念模型：也称信息模型，它按用户的观点对数据和信息进行建模，用于信息世界，它强调语义表达能力，是现实世界到信息世界的第一次抽象，是用户和数据库设计人员之间进行交流的语言。

数据模型



- z 数据模型：用于机器世界，是按计算机系统的观点对数据建模，它需要有严格的形式化定义，且常常要加上一些限制和规定，以便机器实现

概念模型 -- ER 模型

- z 是现实世界到机器世界的一个中间层次，概念模型中最常用的是 ER 模型（实体联系）模型
- z ER 模型创始人：Peter Chen（陈品山）
 - 1983 至今任职于 Louisiana State University.
 - Harvard 大学博士，
 - 曾在 MIT, UCLA, Harvard 工作。
 - Entity-Relationship (ER 模型)是计算机软件领域引用率最高的论文之一。
 - 根据对 1,000 个计算机科学领域教授的调查，最初的 ER 模型论文被选为计算机科学 38 篇最具影响力的论文之一。



概念模型 -- ER 模型基本概念

- z ER 模型中的概念
 - 实体 (Entity)：客观存在并可以相互区分的事物叫实体。（例如：一个个学生、一辆辆轿车）
 - 属性 (Attribute)：实体所具有的若干特征，称为实体的属性。例如：学生有学号、姓名等属性
 - 域 (Domain)：一个属性可能取值的范围称为这个属性的域。例如：性别的域值只能为 “男”或“女”

ER 模型基本概念

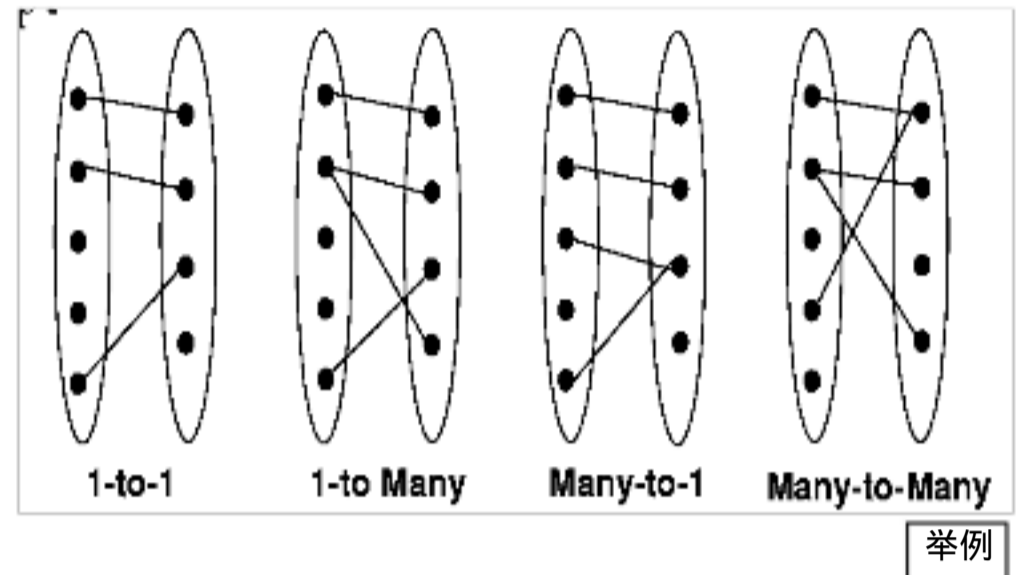
- z 候选码：能够唯一标识实体的属性或最小的属性组称为候选码，可能存在多个候选码，设计者必须指明一个候选码做主码（关键字）
- z 实体型 (Entity type)：具有相同属性的实体具有共同的特征和性质，用实体名及其属性集合来抽象、刻画同类实体，称为实体型。



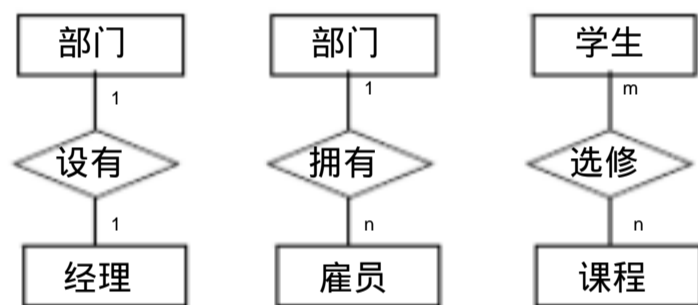
ER 模型基本概念

- z 实体集 (Entity set) : 同型实体的集合
- z 联系 : 现实世界的事物之间是有联系的, 这种联系在信息世界中反映为 : 实体 (型) 内部的联系和实体 (型) 之间的联系。
- z 两个实体型之间的联系
 - ? 一对一联系 (1 : 1) 例如 : 部门、经理
 - ? 一对多联系 (1 : n) 例如 : 部门、雇员
 - ? 多对多联系 (m : n) 例如 : 学生、课程

ER 模型基本概念



ER 模型基本概念

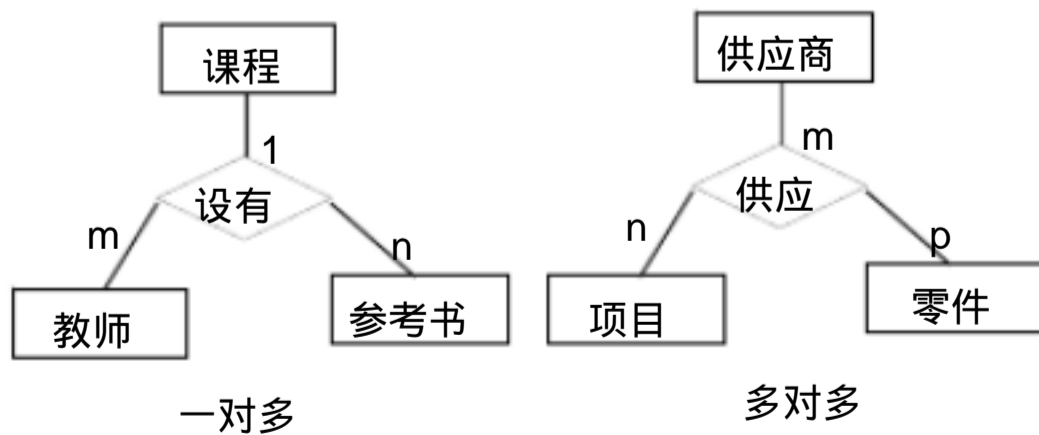


ER 模型基本概念

- z 两个以上实体型之间的联系
 - ? 也存在一对一、一对多和多对多的联系
 - ? 一对多 : 若实体集 E_1, E_2, \dots, E_n 存在联系, 对于实体集 E_j ($j = 1, 2, \dots, i-1, i+1, \dots$) 中的给定实体, 最多只和 E_i 中的一个实体相联系, 则我们说 E_i 与 $E_1, E_2, \dots, E_{i-1}, E_{i+1}, \dots, E_n$ 之间的联系是一对多的。

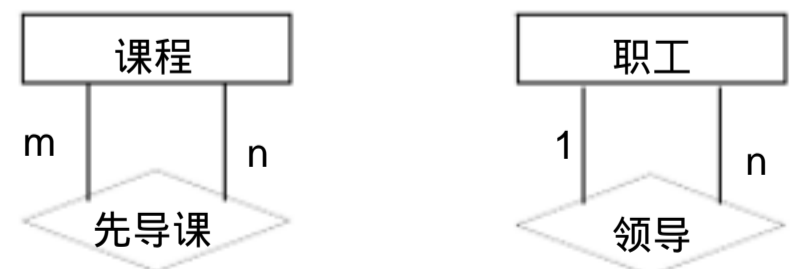
ER 模型基本概念

? 例子 : 三个实体型之间的联系



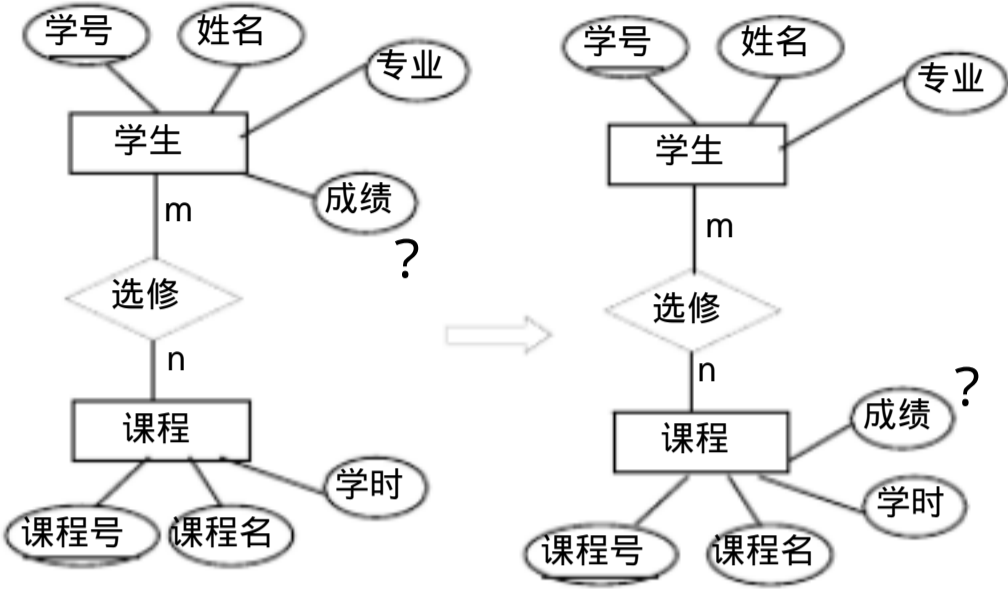
ER 模型基本概念

- z 同一实体集内的各个实体之间的联系
 - 例如 : 职工实体集内的实体有领导和被领导的关系 ; 高数、离散数学是数据结构的先导课, 数据结构是操作系统、数据库原理的先导课。

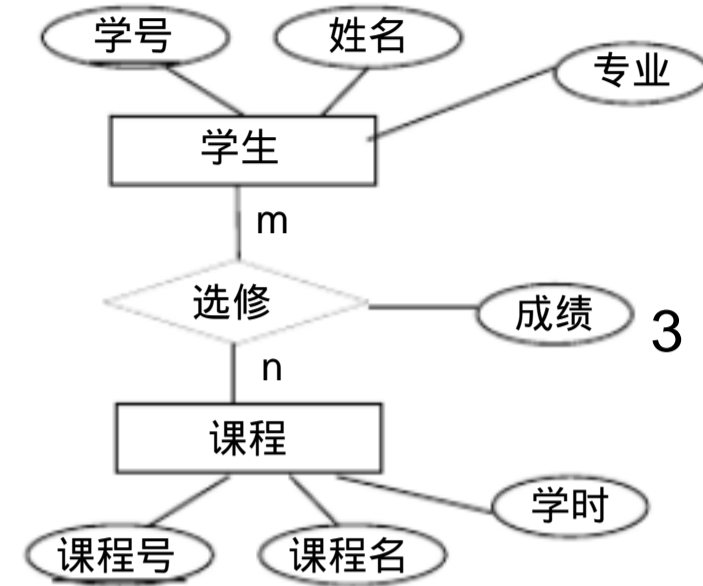


经理也是职工, 经理要领导多个职工, 而一个职工仅被一个经理领导。

联系的属性



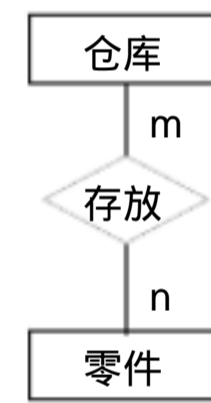
联系的属性



实例分析

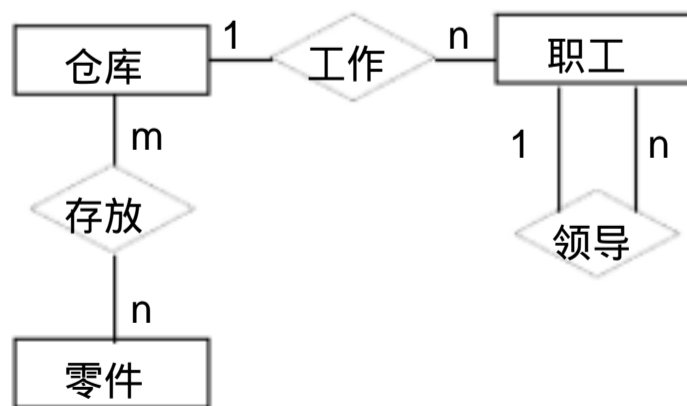
例子 1：假定开发某工厂物资管理系统，业务调查后得到如下的问题描述。在该工厂中，一个仓库可以存放多个零件，一种零件可以存放在多个仓库中。一个仓库有多名职工当保管员，一个职工只能在一个仓库工作，每个仓库有一名主任。仓库存放由许多供应商为一些项目提供的多种零件，一个项目可以由多个供应商提供多种零件，一个供应商可供给多个项目多种零件，每种零件由不同的供应商提供给多个项目。

实例分析 (E-R 图)



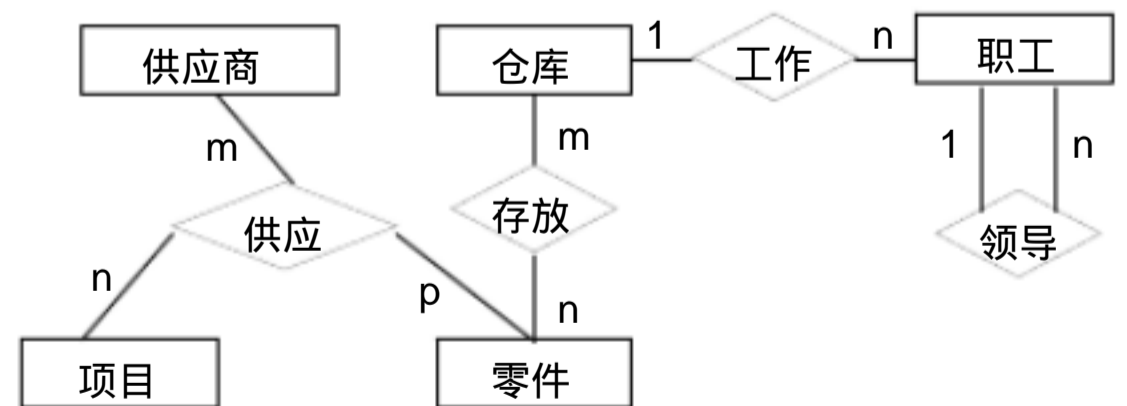
一个仓库可以存放多个零件，一种零件可以存放在多个仓库中。

实例分析 (E-R 图)



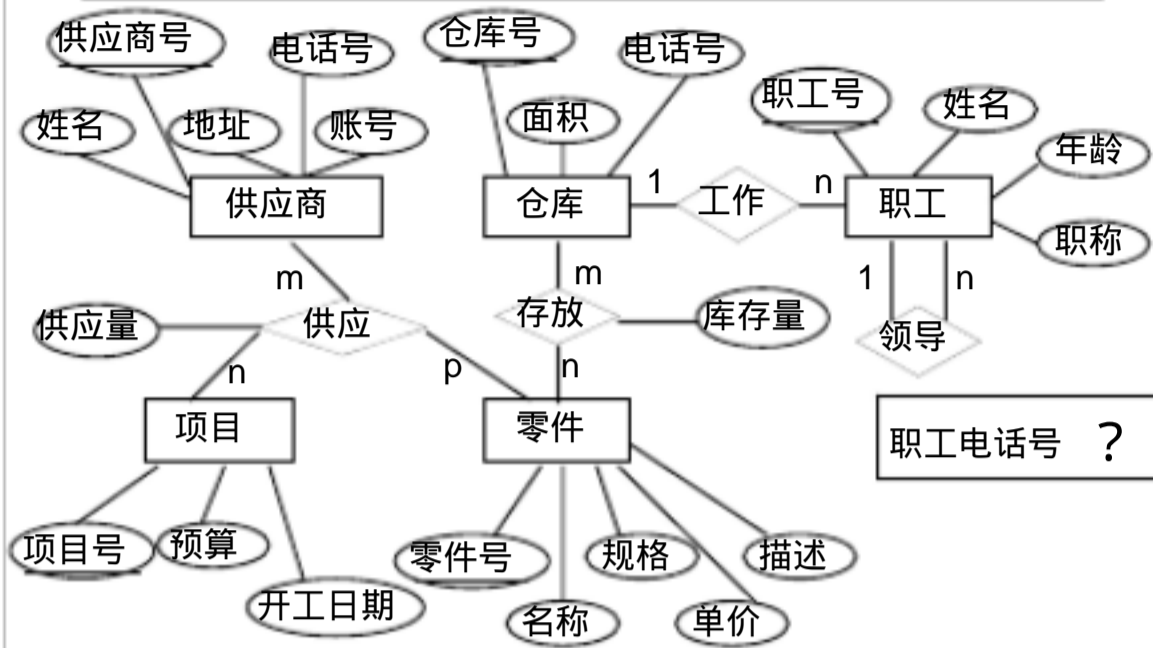
一个仓库有多名职工当保管员，一个职工只能在一个仓库工作，每个仓库有一名主任。

实例分析 (E-R 图)



仓库存放由许多供应商为一些项目提供的多种零件，一个项目可以由多个供应商提供多种零件，一个供应商可供给多个项目多种零件，每种零件由不同的供应商提供给多个项目。

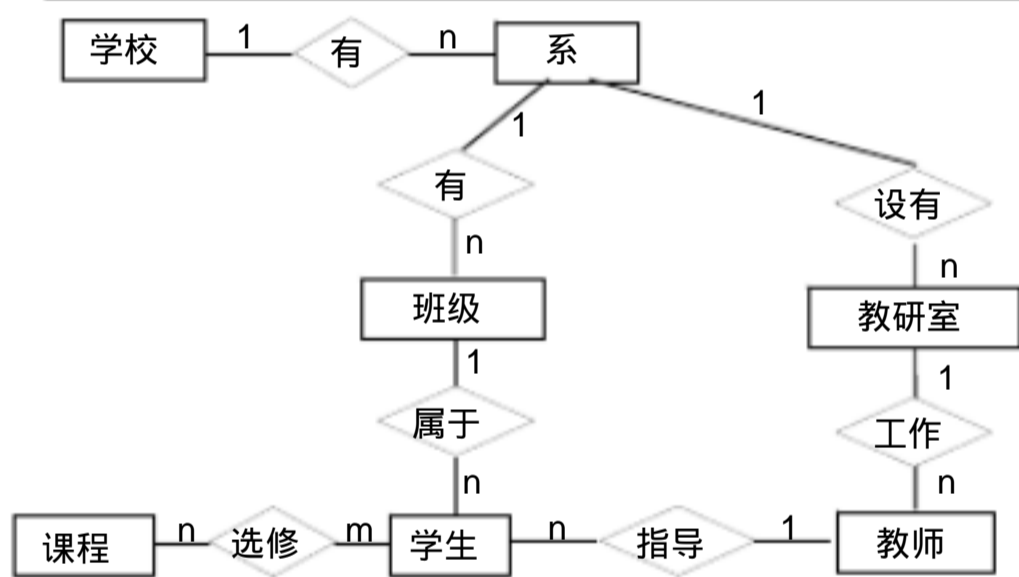
实例分析 (E-R 图)



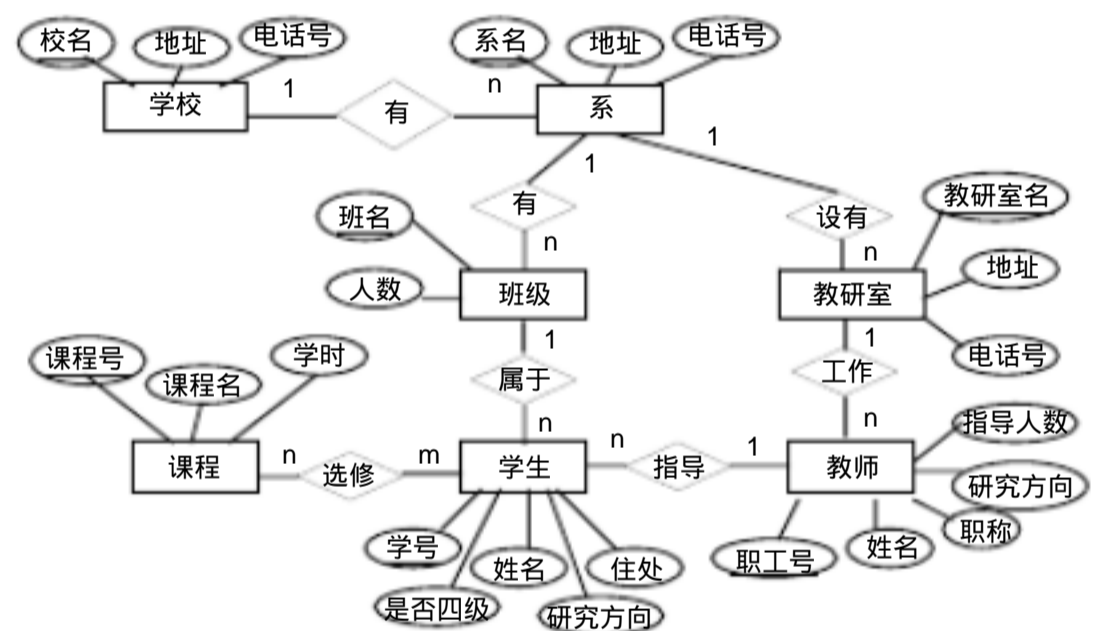
实例分析

z 例子 2：开发学校信息管理系统。学校中有若干系，每个系有若干班级和教研室，每个教研室有若干教师，其中有教授和副教授，每人各带若干名研究生，每个班有若干学生，每个学生选若干课程，每门可由若干学生选修。

实例分析 (E-R 图)

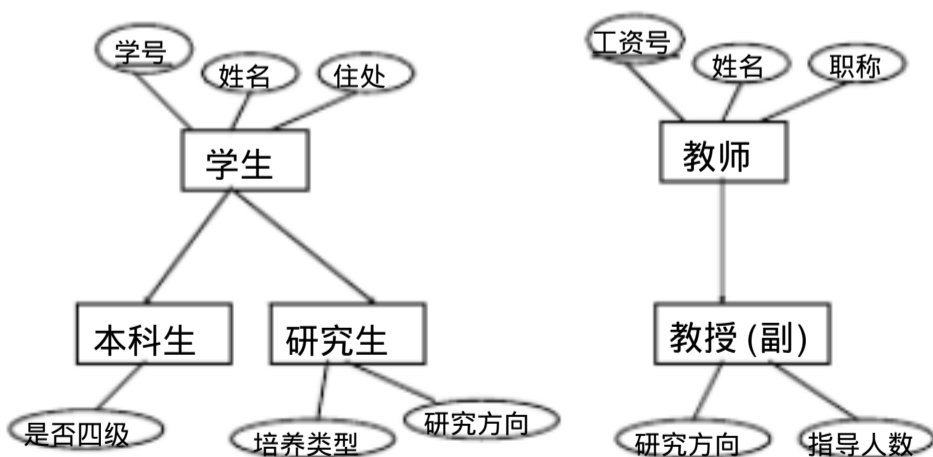


实例分析 (E-R 图)



实例分析

思考：
图例中的继承关系如何表示？

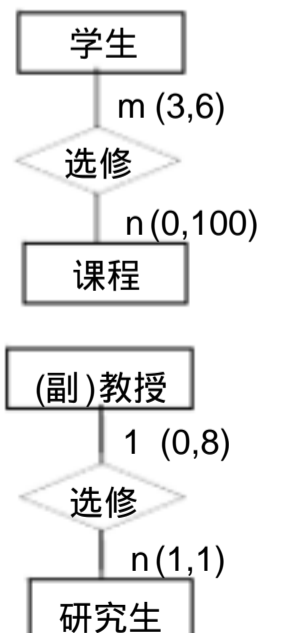


软件开发中，实际使用的 ER 模型有更丰富的语义，介绍扩展 ER 的一些内容，扩展 ER 表示有所不同

ER 模型的其它特征 --参加约束

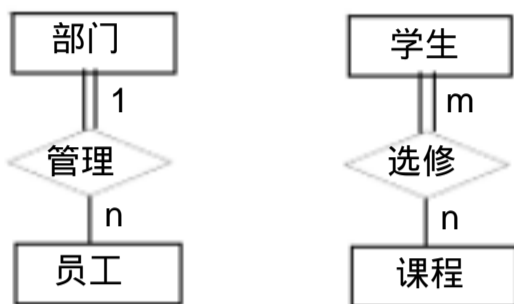
z 实体的参与度 (有些 ER 模型)

? 实体参与联系的最小和最大次数，例如：规定每位学生最少选三门课，最多六门课，则学生在选修联系中的参与度表示为 (3,6)。规定有些课程可以无人选，但任一门课程最多允许 100 人选，则课程的参与度为 (0,100)。表示为 (min,max) , 0<=min<=max , max>=1



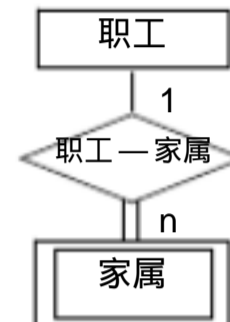
ER模型的其它特征 --弱实体

- ? 部分参与：min=0，实体集中的部分实体参与联系
 - ? 全参与：min>0，实体集中的每个实体都参与联系
- 例如 (1) 每个部门必须有一个经理负责管理该部门
(2) 每名学生必须选修课程



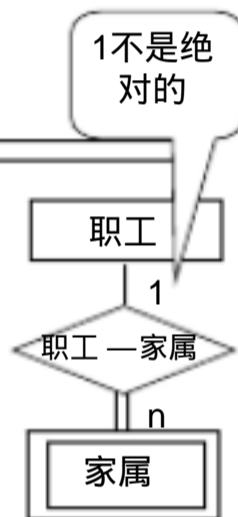
ER模型的其它特征 --弱实体

- z 弱实体
 - ? 现实世界有一种特殊的联系，这种联系代表实体间的所有关系。 例如：职工和家属的关系
 - ? 弱实体是依赖于其它实体存在而存在的实体



ER模型的其它特征 --弱实体

- ? 两个特点：
 - ? 弱实体不能独立存在，它总是依附于某一所有者实体。 owner 实体集与弱实体集之间必须是一对多联系，弱实体必须是全部参与
 - ? 弱实体不一定有自己的码，例如：家属实体集可能有姓名、性别、出生年月等信息，但这些信息不足以标识一个家属 (不同职工的家属有可能会重名)。

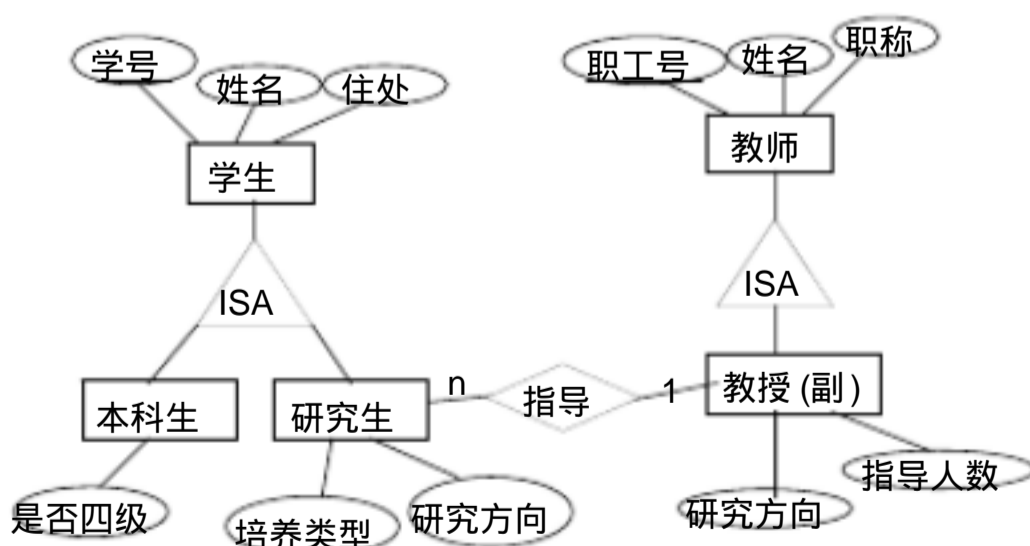


所有者实体的码 + 弱实体的某一属性 = 标识一个弱实体

ER模型的其它特征 -- 类层次

- z 例子：学校中学生，研究生，本科生
- z 实体集研究生和本科生称为实体集学生的子类
- z 从语义上来讲，子类中的每一个实体同时也是超类中的一个实体，子类中实体属性必须包含所有的超类实体集的属性，也就是说，子类将继承超类的属性
- z 超类/子类联系称之为 ISA 联系
- z 表示

ER模型的其它特征 -- 类层次



ER模型的其它特征 -- 类层次

- z ISA 联系可以从两个方面来看
 - ? 可以看作是一个特化，即子类是经过超类特化而得到的
 - ? 也可以看作是一个概括，即超类是对子类进行概括而得到的

ER模型的其它特征 -- 类层次

- z 定义在 ISA 联系上的两种约束关系
 - ? Overlap 约束：Overlap 约束用来指定两个子类之间是否允许包含同一实体
 - ? 例如：合同工和高级雇员
 - ? Covering 约束：Covering 约束用来指定子类中所有的实体是否包含超类中的所有实体
 - ? 例如：全职雇员和兼职雇员

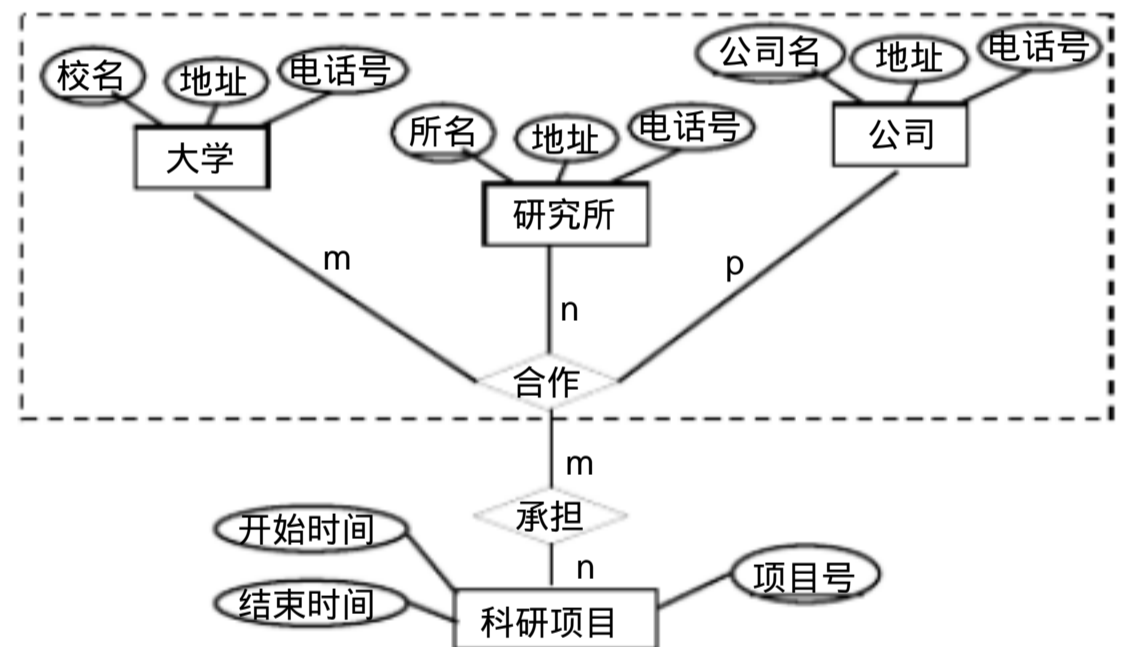
ER模型的其它特征 -- 聚合

- z 联系描述的是多个实体集间的关联，只有实体才能参与联系，不允许联系参与联系，但实际应用中往往有时需要联系参与联系
- z 把联系看成由参与联系的实体组合而成的新实体，这种新实体称为参与联系的实体的聚合
- z 新实体的属性为参与联系的实体的属性和联系的属性的并
- z 新实体参与联系

ER模型的其它特征 -- 聚合

- z 存在四个实体集（大学、研究所、公司、科研项目）和两个联系（合作、承担），其中合作是建立在大学、研究所和公司之间的联系
- z 为了定义承担这样的联系，可以对联系合作、实体集大学、研究所和公司进行聚合，将他们当作一个实体集来看待
- z 表示

ER模型的其它特征 -- 聚合



用ER模型进行概念数据库设计

用ER模型来进行概念数据库设计时通常需要注意以下几个问题

- 一个概念应该被建模为一个实体还是一个属性？
- 一个概念应该被建模为一个实体还是一个联系？
- 应该使用多个二元联系还是一个 n元联系？

实体与属性

在进行数据建模时，一个特征建模为属性、实体集，还是联系集，有时并不是十分清楚的

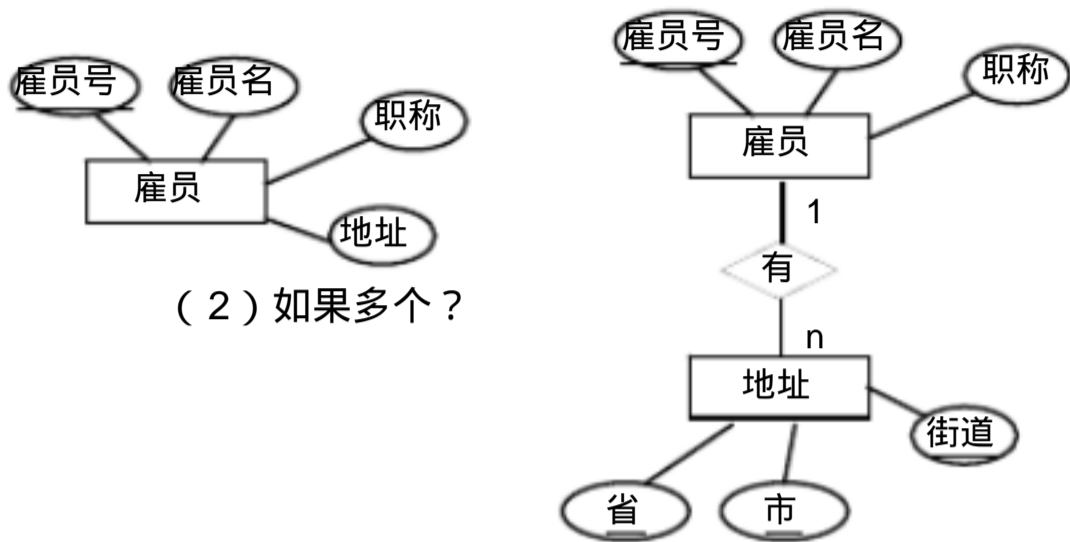
例子：向实体集 Employees 中增加 Address 信息，可能有多种方式

将 Address 作为一个属性：如果我们仅仅是想记载雇员的一个地址，这时这种方式就比较合适

也可以将地址信息建模为一个实体集 Addresses，并在 Employees 和 Addresses 之间建立一个联系 Has_Address，这种设计在以下两种情况是必须的

- ? 必须为雇员记载多个地址
- ? 我们希望结构化雇员的地址信息

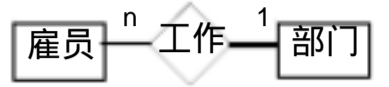
实体与属性



(2) 如果多个?

(1) 结构化的好处?

实体与联系

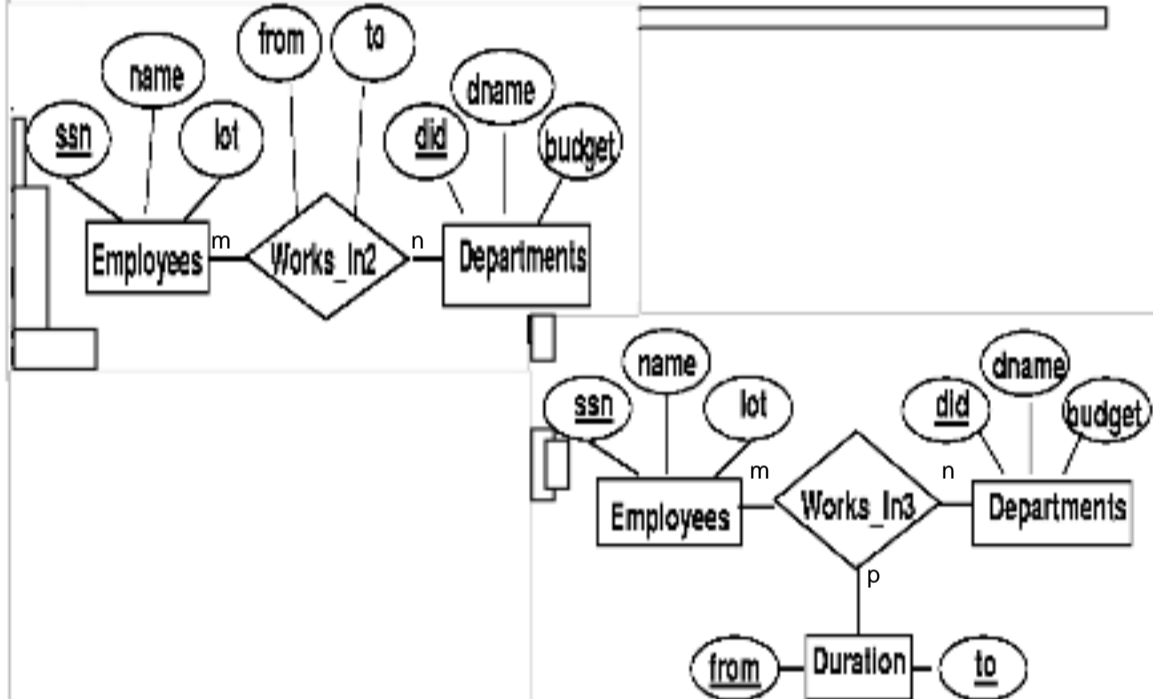


例子 1: 雇员和部门的联系

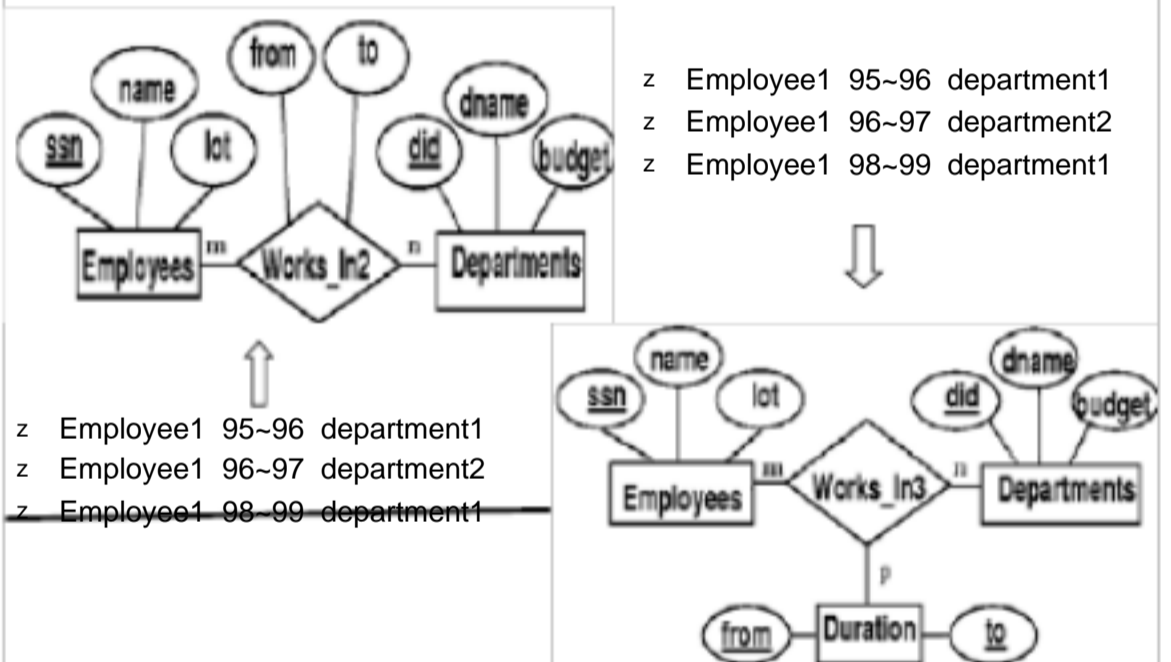
在这个例子中, 如果不允许一个雇员在同一部门工作过两次的话, 则该 ER 图的设计没有任何问题, 但如果允许同一雇员在一个部门多次任职的话, 则上述 ER 图设计存在问题。 Why?

解决策略: 引入一个实体 Duration, 也就是将原来的二元联系建模为一个实体, 并用一个三元联系取代原来的二元联系。

实体与联系 -- 例 1



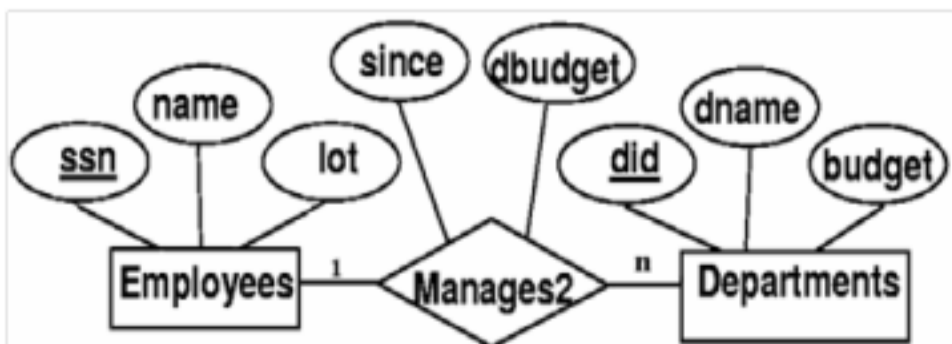
实体与联系 -- 例 1



实体与联系 -- 例 2

例子 2: 一个部门至多只能有一个管理者, 但一个经理可以管理多个部门, 经理管理部门有一个起始时间和所自由支配的财务预算。

如果对于每个部门都有一个财务预算, 则这种表示方式是非常自然的。

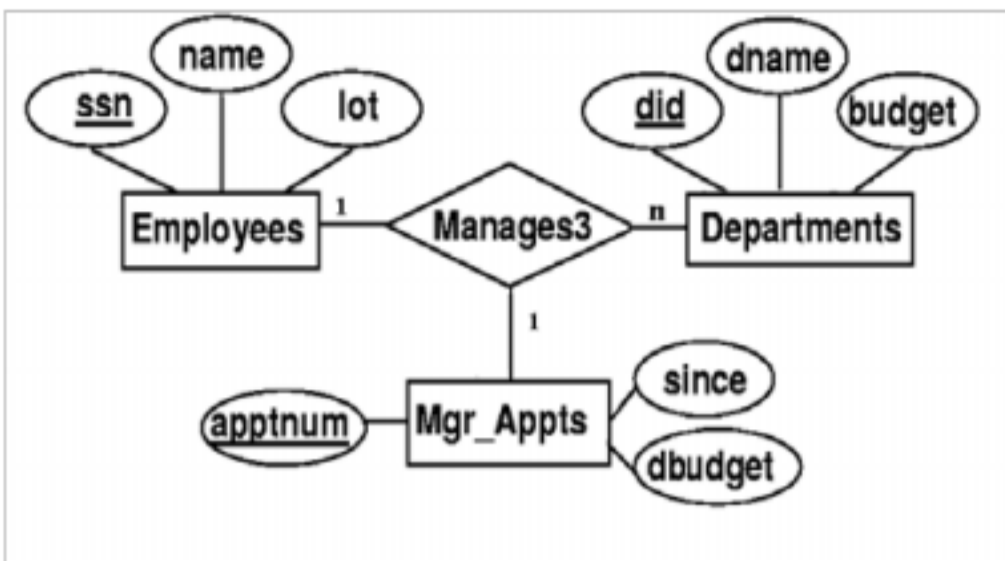


实体与联系 -- 例 2

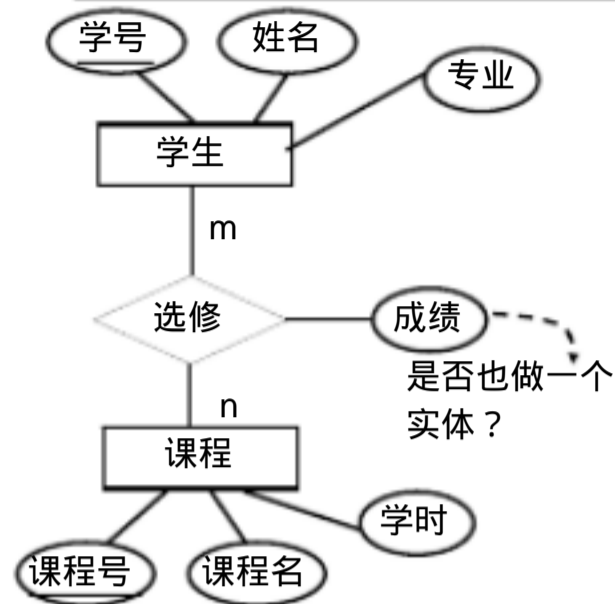
如果属性 dbudget 表示经理对其所管理的所有部门的财务预算之和, 则上述表示 (1) 会引起 dbudget 信息的冗余; (2) 会误导系统设计人员。

解决策略: 引入一个实体 Mgr_Appts 来表示一组部门的经理的委派, 也就是将原来的二元联系建模为一个实体, 并用一个三元联系取代原来的二元联系

实体与联系 -- 例2



实体与联系 -- 例3



是否存在：

王立 数据库 98
王立 数据库 78

特点：不是多值！！

结论：只能做属性

实体与联系 -- 例3

成绩作为实体的情况

是否存在：

姓名	课名	平时成绩	实验成绩	考试成绩	总成绩
王立	数据库	8	18	60	86
李梅	数据库	7	19	53	79

特点：是多属性值！！

结论：可以做实体

实体与联系 -- 例3

成绩作为实体的情况

是否存在：

姓名	课名	平时成绩	实验成绩	考试成绩	总成绩
王立	数据库	8	18	60	86
王立	数据库	7	19	53	79

特点：是多属性值！！

结论：可以做实体

思考：职工地址是否是同样的道理？

实体与联系 -- 例3

成绩作为实体的情况

是否存在：

姓名	课名	平时成绩	实验成绩	考试成绩	总成绩
王立	数据库	8	18	60	86
王立	数据库	7	19	53	79

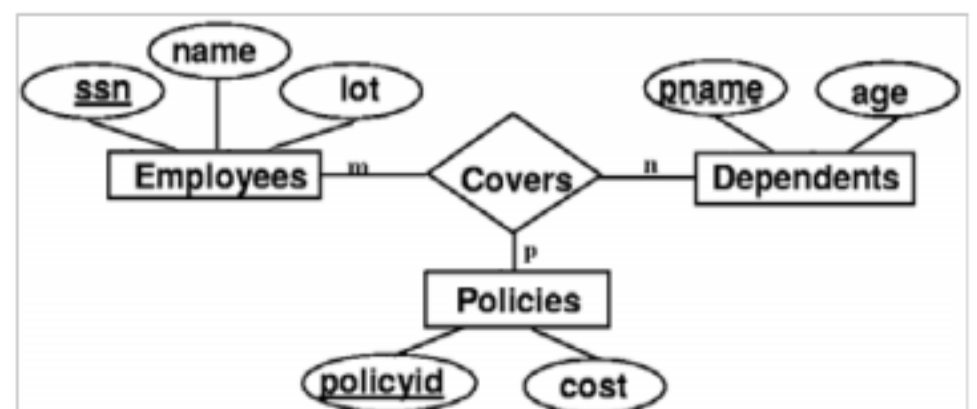
特点：是多值！！

结论：可以做实体

缺点：语义的正确性需要系统自行维护

二元与三元联系

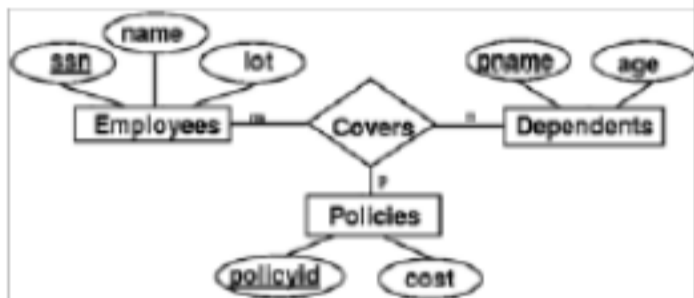
例1：每个雇员可以购买几份保险，每份保险可以被多个雇员所拥有且每个家属可以被多个保险单所涵盖，...三元多对多联系。



二元与三元联系 -- 例 1

考虑下列附加的需求

- 多个雇员不能同时购买同一份保险
- 每份保险必须由雇员购买
- 一份保险可以涉及到多位家属
- 家属 (Dependents) 是一个弱实体集
- 每位家属只对应一份保险



二元与三元联系

根据上述需求，给出如下方案：

对应 1，将 m 改成 1

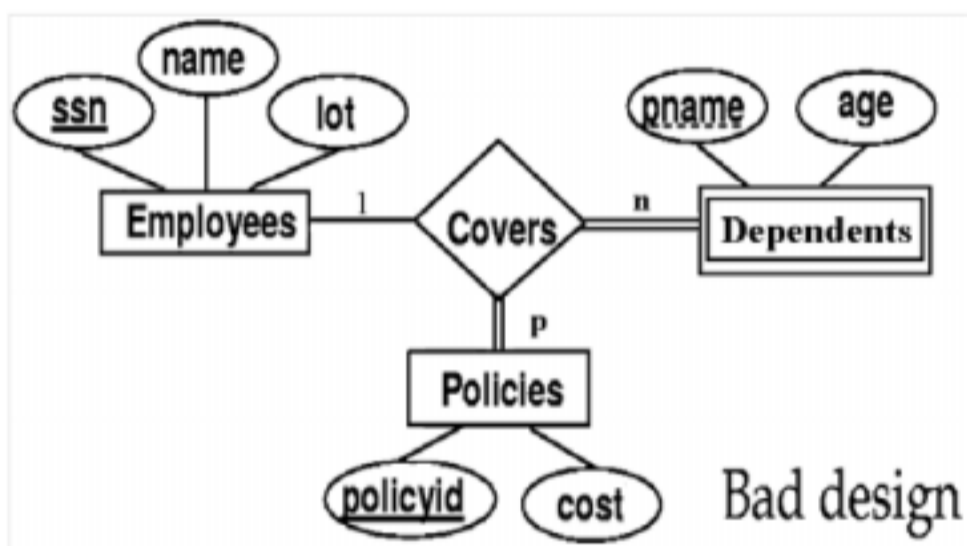
不能解决一位家属只对应一份保险
在 Policies 上施加一个 Total 参加约束

产生一个副作用：每份险单至少涉及到一位家属
将 p 改为 1

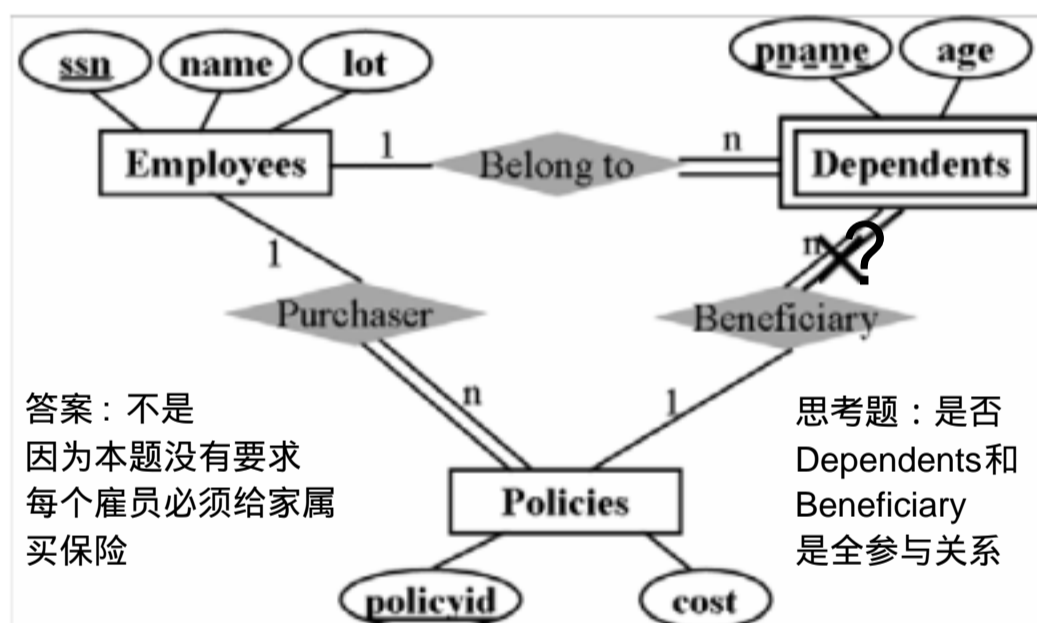
(不满足：因为要求每位家属只对应一份保险)

引入一个弱实体集

二元与三元联系 -- 例 1



二元与三元联系 -- 例 1

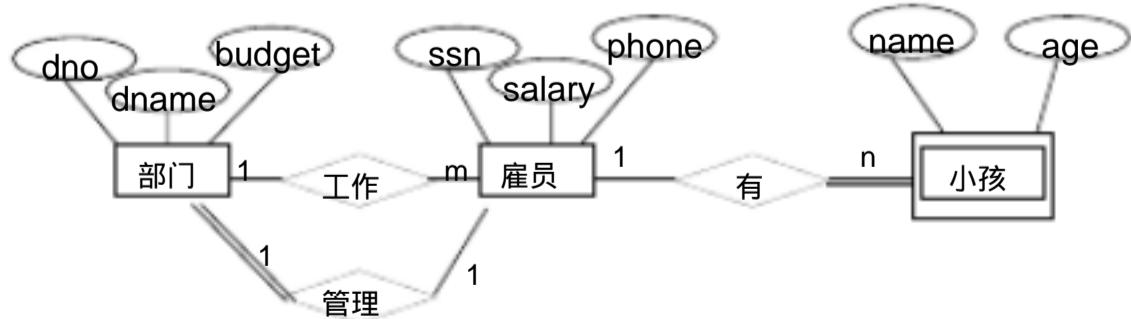


答案：不是
因为本题没有要求
每个雇员必须给家属
买保险

思考题：是否
Dependents 和
Beneficiary
是全参与关系

实例分析

一个公司数据库需要存储雇员、部门和雇员小孩的信息。雇员工作在部门（一个雇员只能工作在一个部门），每个部门由一个雇员管理，每个雇员小孩的名字是唯一的，假定小孩只有一个家长工作在这个公司，而且我们不关心那些已经调离雇员的小孩情况。请画出 ER 图捕获这些信息。



实例分析

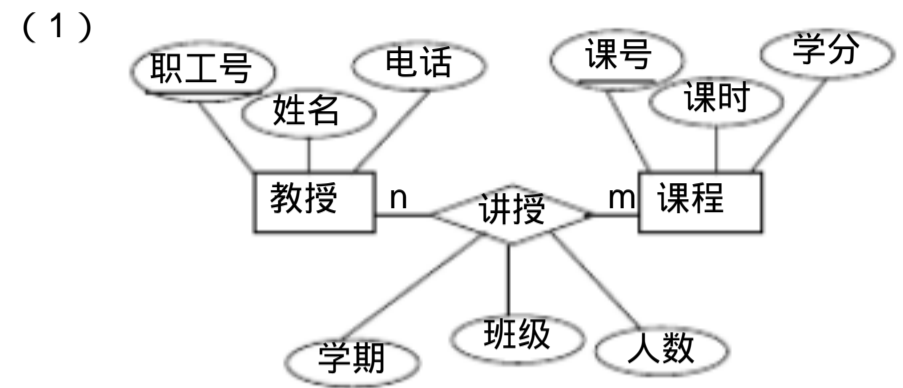
一个大学数据库包括教授、课程信息。教授讲授课程，下面几种情况都是描述有关讲授联系集的，对于每一种情况画 ER 图描述

- ? 教授可以在几个学期讲授同一门课程，但仅最近一次的讲授活动需被记录下来
- ? 教授可以在几个学期讲授同一门课程，每次讲授活动需被记录下来
- ? 每个教授必须讲授课程
- ? 每个教授只讲授一门课程

实例分析

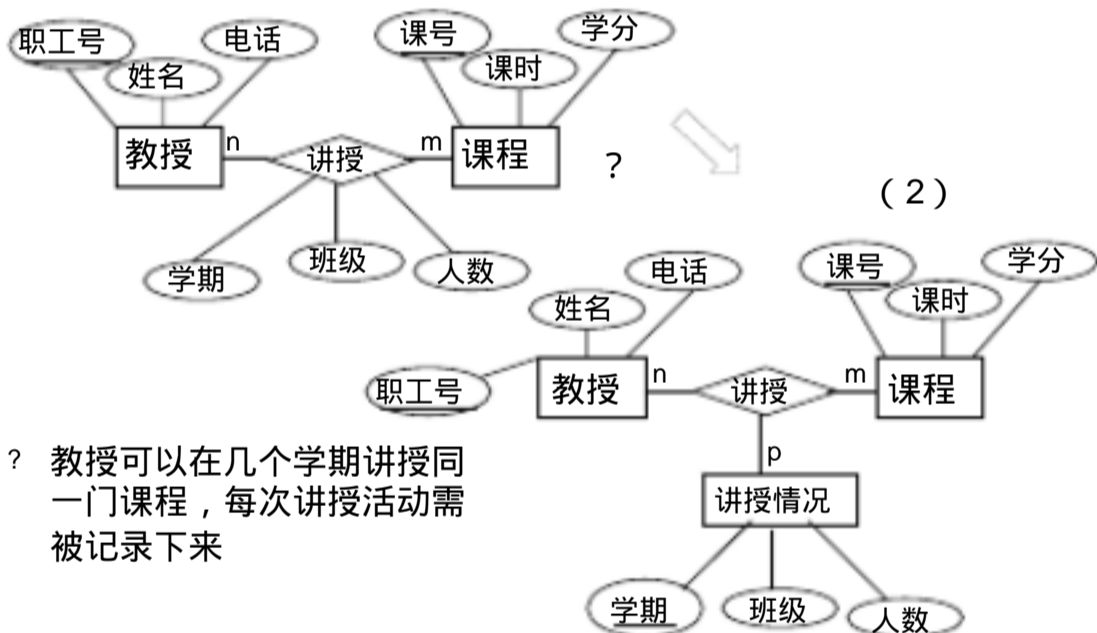
- ? 每个教授只讲授一门课程，每门课程可有几位教授讲授
- ? 假定一些课程可由一组教授联合讲授
- ? 假定一些特定课程只能由一组教授联合讲授，且这些教授中的任一位不可能独立讲授这门课程（思考题）

实例分析



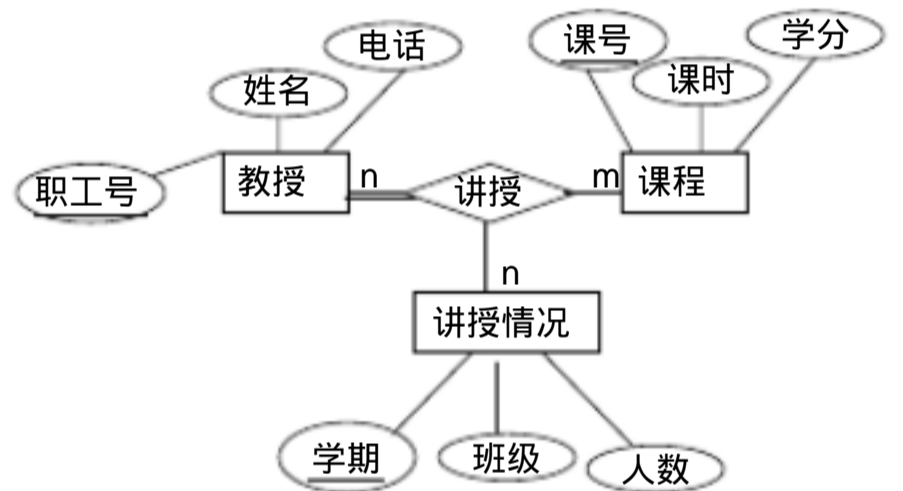
- ? 教授可以在几个学期讲授同一门课程，但仅最近一次的讲授活动需被记录下来

实例分析



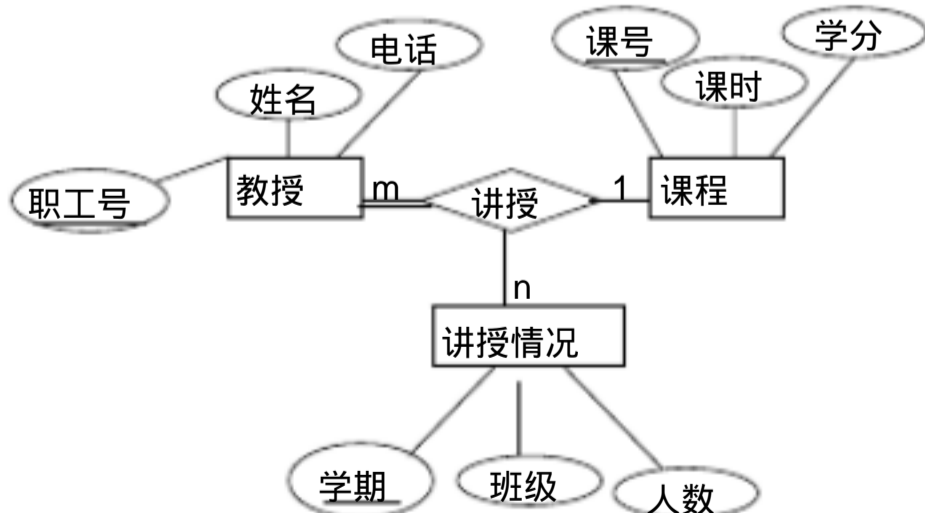
- ? 教授可以在几个学期讲授同一门课程，每次讲授活动需被记录下来

实例分析 (3)



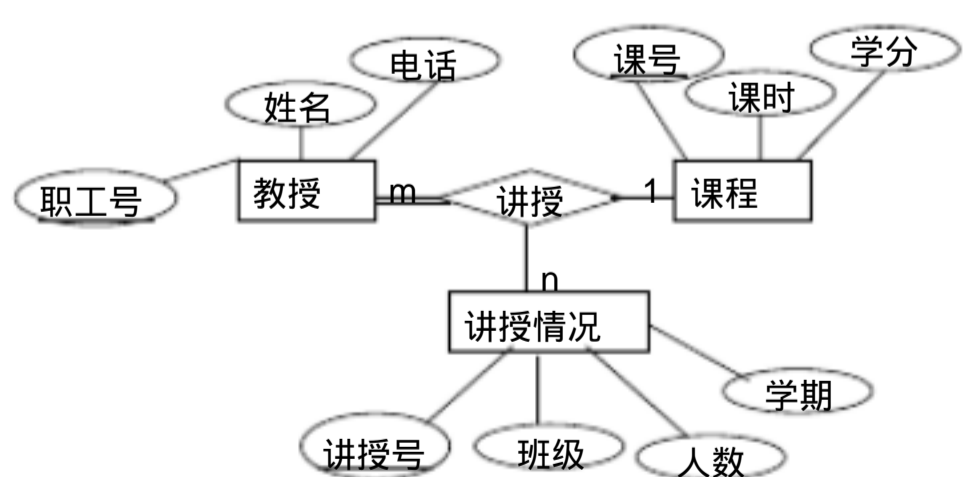
- ? 每个教授必须讲授课程

实例分析 (4, 5)



- ? 每个教授只讲授一门课程
- ? 每个教授只讲授一门课程，每门课程可有几位教授讲授

实例分析 (6)



- ? 假定一些课程可由一组教授联合讲授

实例分析 (7)

? 假定一些特定课程只能由一组教授联合讲授，且这些教授中的任一位不可能独立讲授这门课程

