

数据库原理

The Theory of Database System

第五章 数据库设计

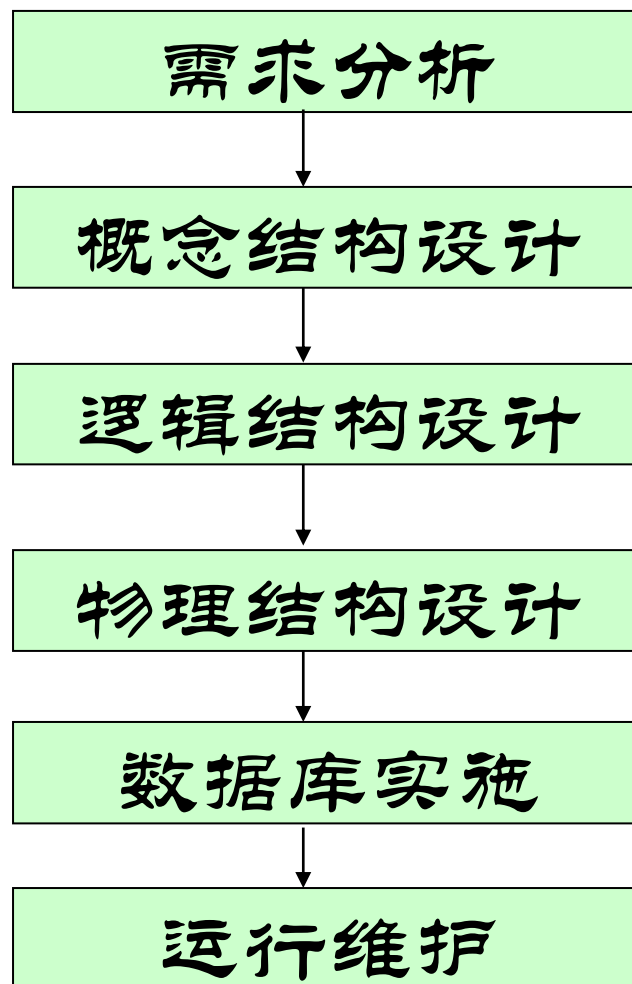


中国矿业大学计算机学院



中国矿业大学数据库原理精品课程

数据库设计的基本步骤



5.3 概念结构设计

- 概念结构
- 设计方法
- 设计步骤
- 局部概念结构设计
- 全局概念结构设计



5.3.1 概念结构设计的定义

- 不同DBMS基于不同的数据模型，而现实应用环境复杂多变，将现实世界中的事物直接转换为机器中的对象，非常不方便。
- 概念模型是现实世界到机器世界的一个中间层次，使用接近计算机存储的方式表示数据，同时又不涉及具体的DBMS。

概念模型的描述工具通常是E-R模型，该模型不依赖于具体的硬件环境和DBMS。



概念结构模型

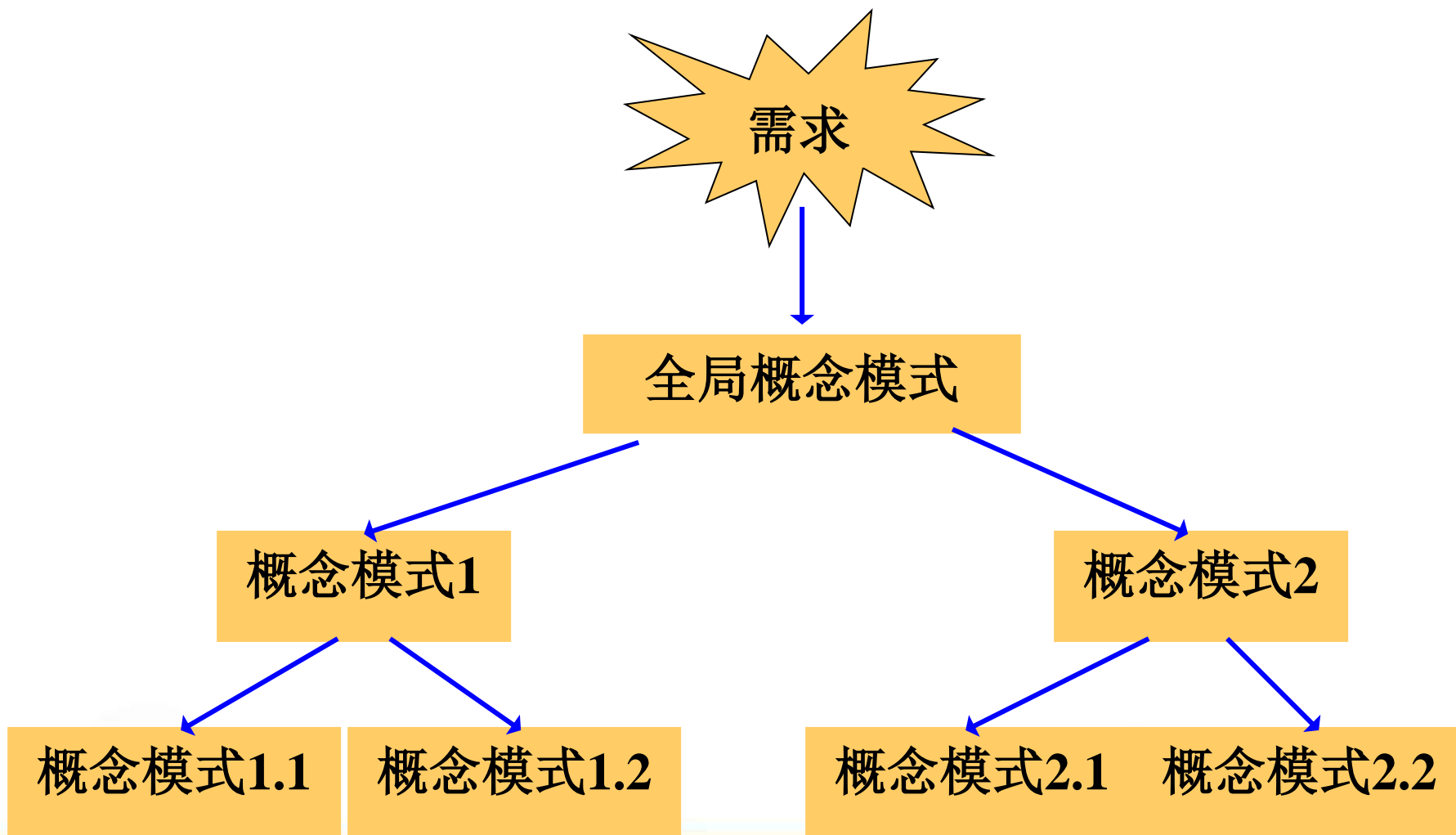
- 1、能真实反映现实世界的模型。
- 2、是各种数据模型的基础。易于向各类数据模型转换。
- 3、易于修改和扩充。
- 4、简单、清晰、用户易于理解。
- 5、是设计人员和用户交流的工具。
- 6、独立于机器。

E-R图

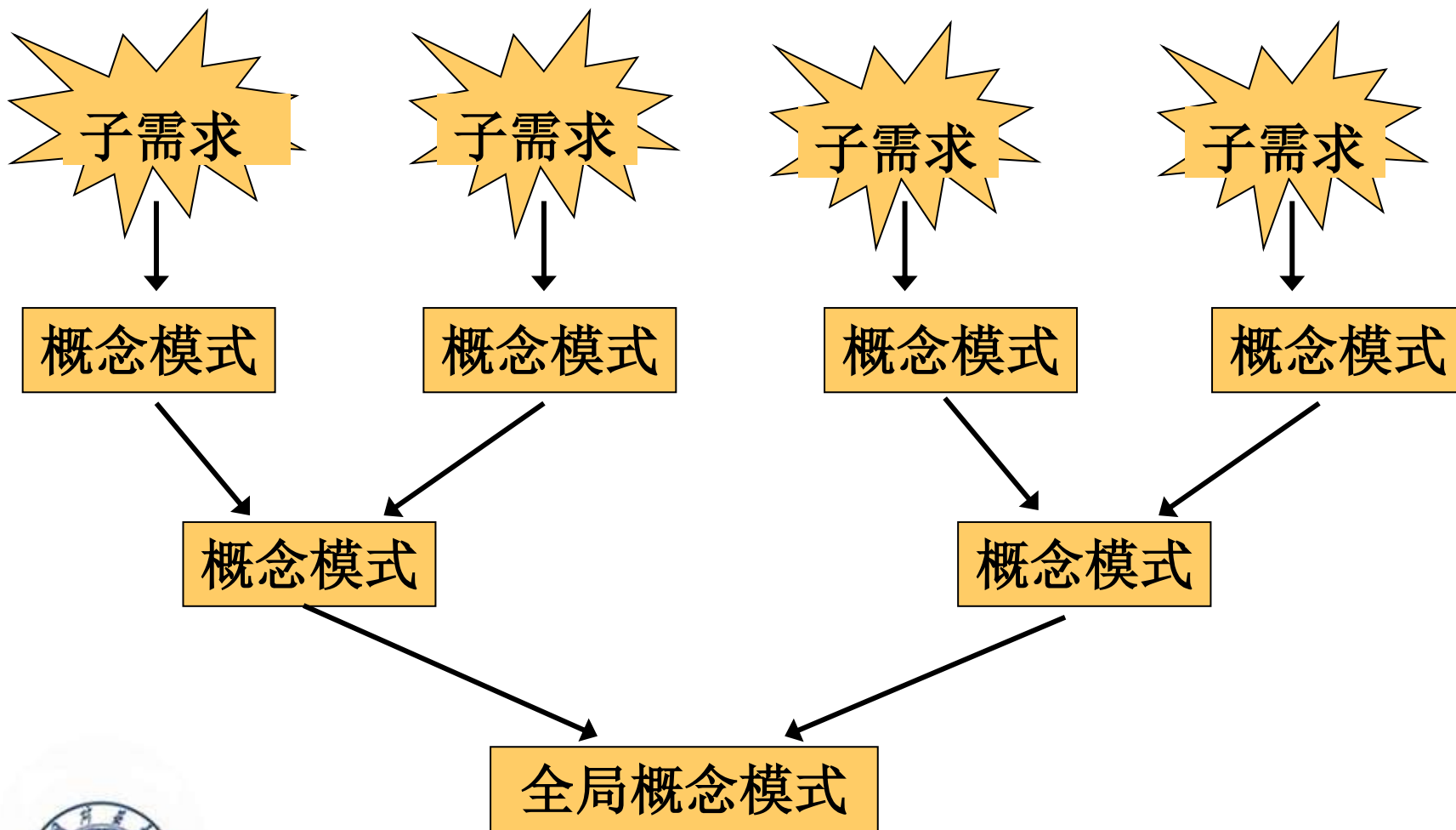


5.3.2 概念结构设计方法

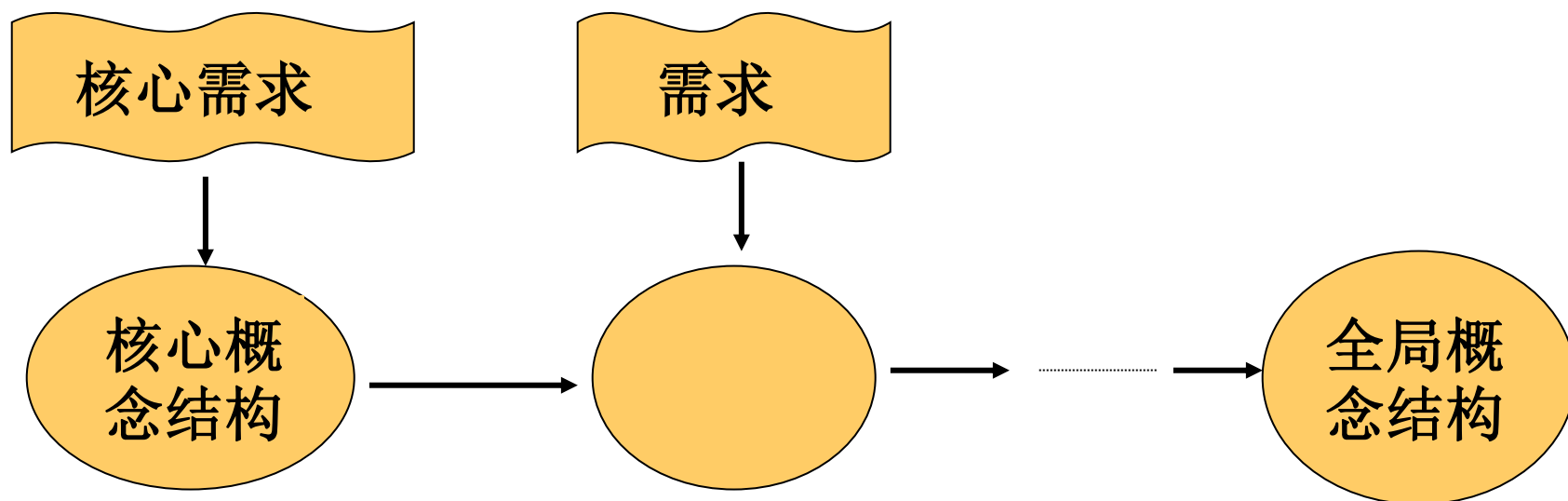
1、自顶向下：先定义全局，然后再逐步细化。



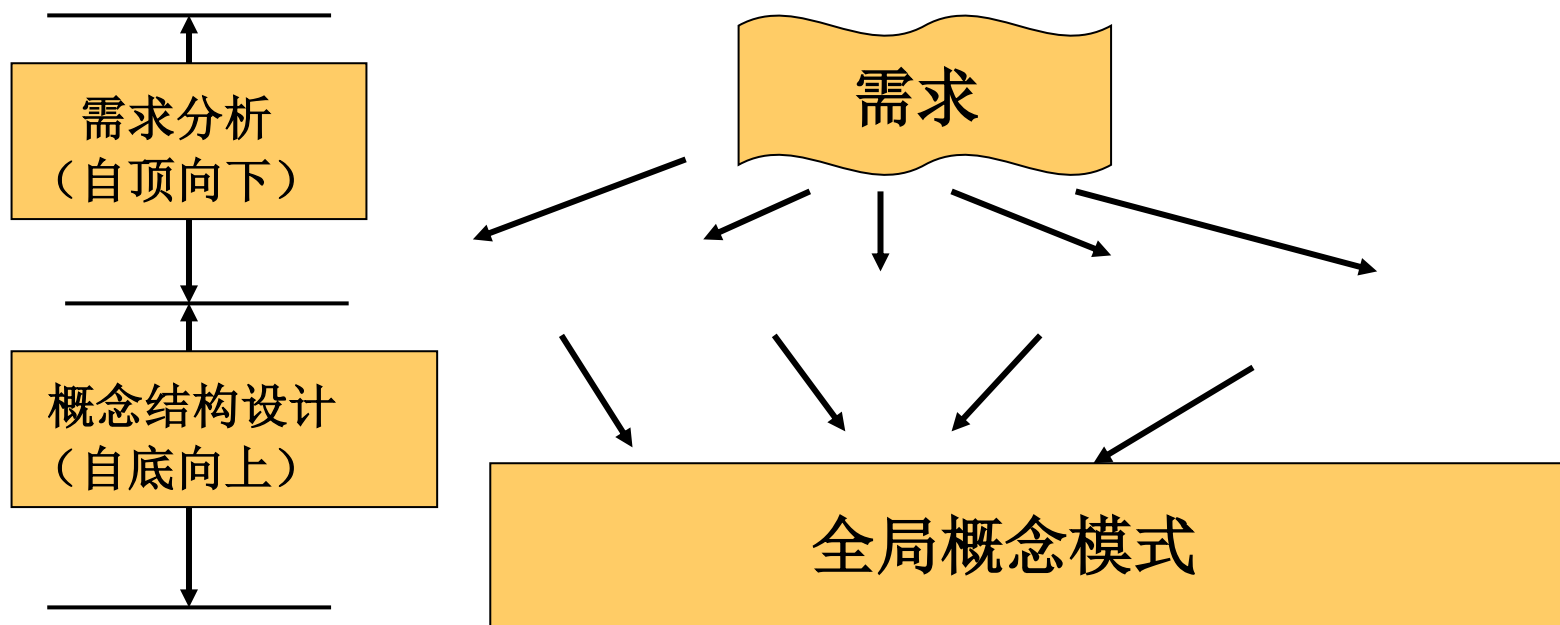
2、自底向上：先定义局部，然后再集成起来。



3、逐步扩张：先定义核心，然后再逐步向外扩充。



4、混合策略：自底向上和自顶向下相结合。



5.3.3 局部（分）E-R图设计

- 选择局部应用。一般而言，中层数据流图能较好反应系统各局部应用的子系统组成，因此通常以中层数据流图作为设计分E-R图的依据。
- 从数据字典中抽取数据。依据数据流图，标定局部应用中的实体、实体属性、标识实体的码，确定联系及类型。



1、确定局部概念结构的范围

参考原则：

- 1) 联系密切的数据
- 2) 实体(集)数要适中

2、确定实体（集）

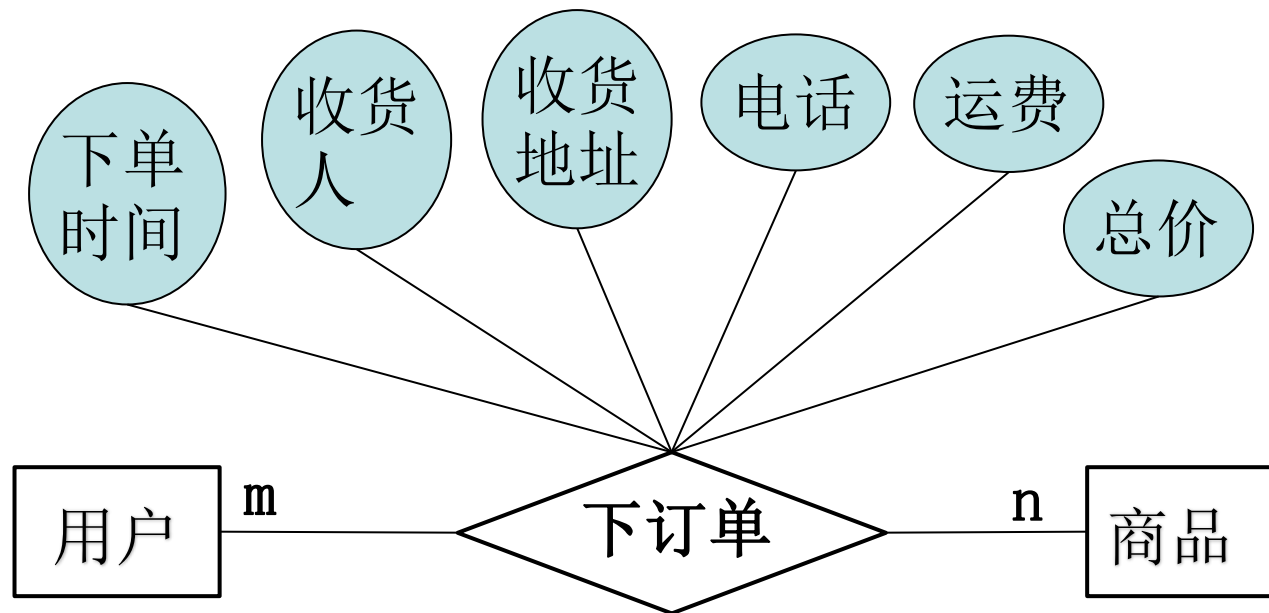
3、确定实体（集）的属性

4、定义实体（集）间的联系



在线订餐系统

用户能够选择多种菜品进行下单，下单时指定收货人、收货地址等信息。请根据日常使用的电商系统，完善相关属性，抽象该E-R模型。



属性方面的讨论：

- 简单/组合属性

- 简单属性：由独立存在的单个部分组成的属性。
- 组合属性：由多个部分组成的属性，每个部分都可以独立存在。

- 例如：

- 姓名，地址



属性方面的讨论：

- 单值/多值属性

- （强实体类型与弱实体类型）

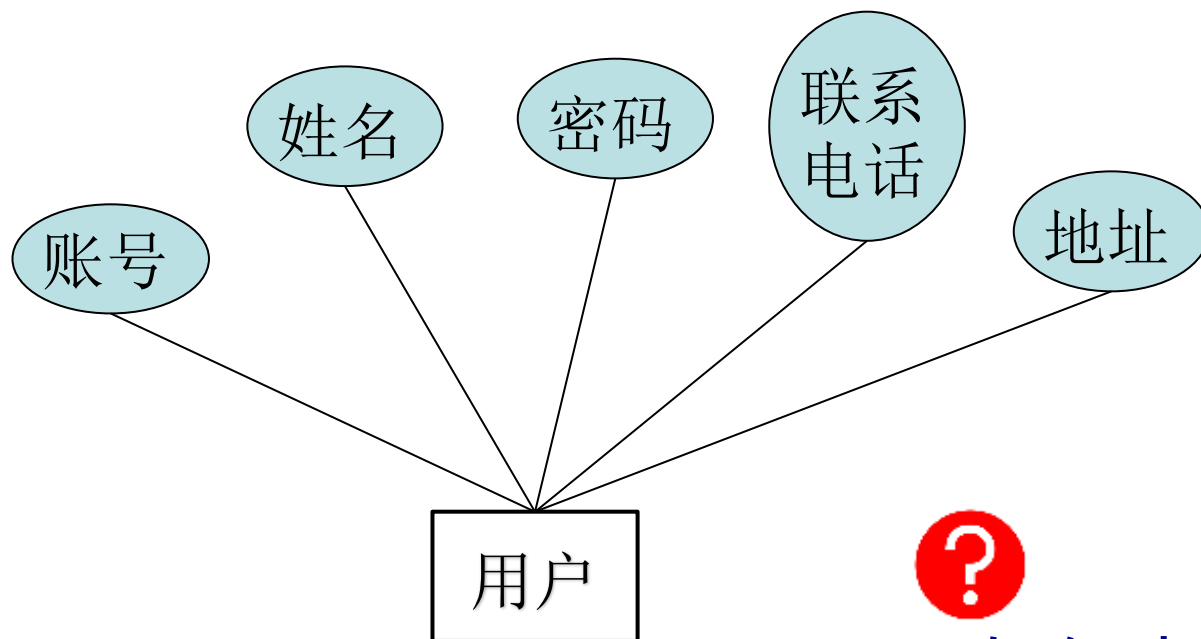
- 单值属性：对于实体类型的每个实例，都只取一个单值的属性。
 - 多值属性：对于实体类型的某些实例，出现可能取多个值的属性。

- 例如：

- 联系方式，简历，订单



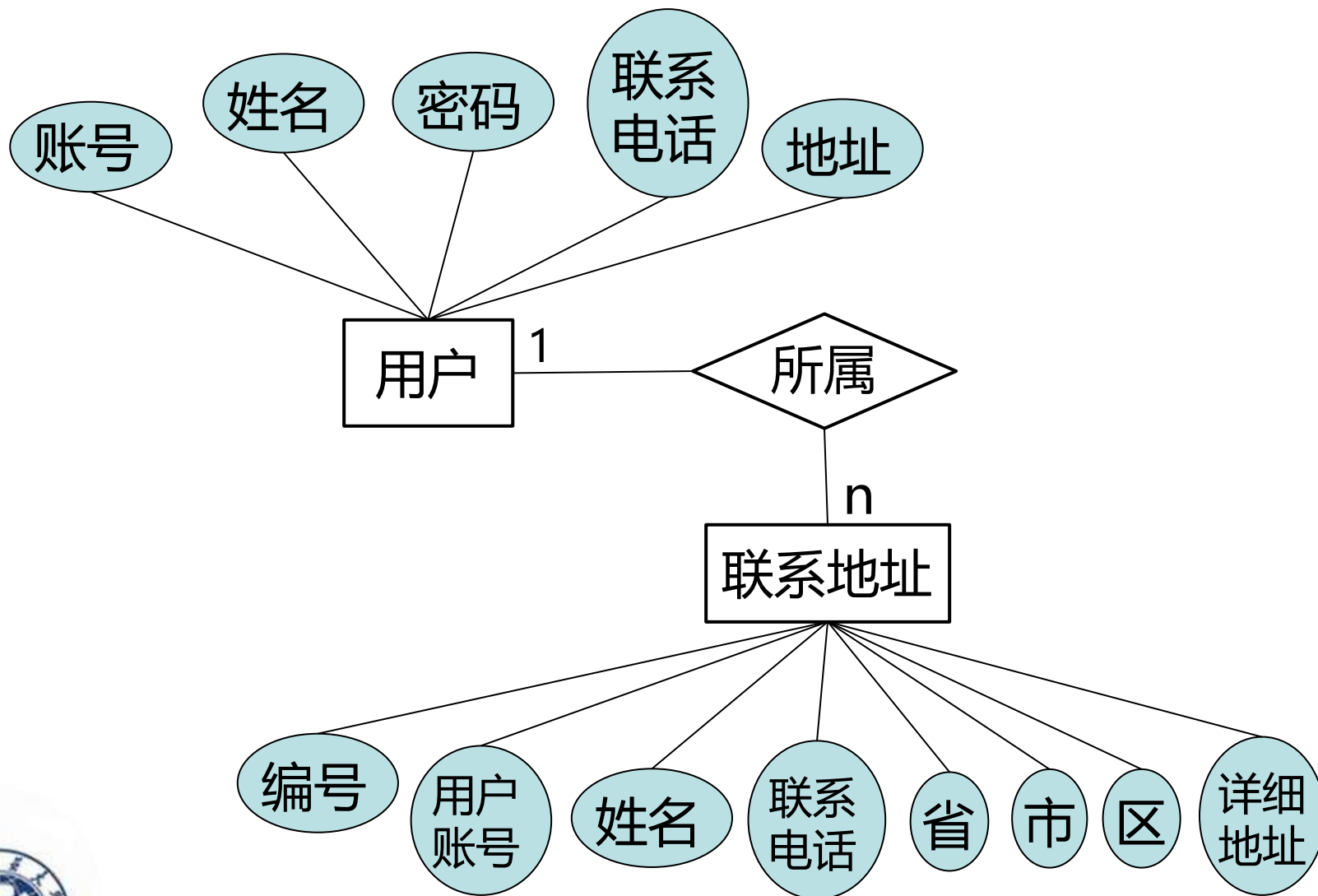
属性方面的讨论:



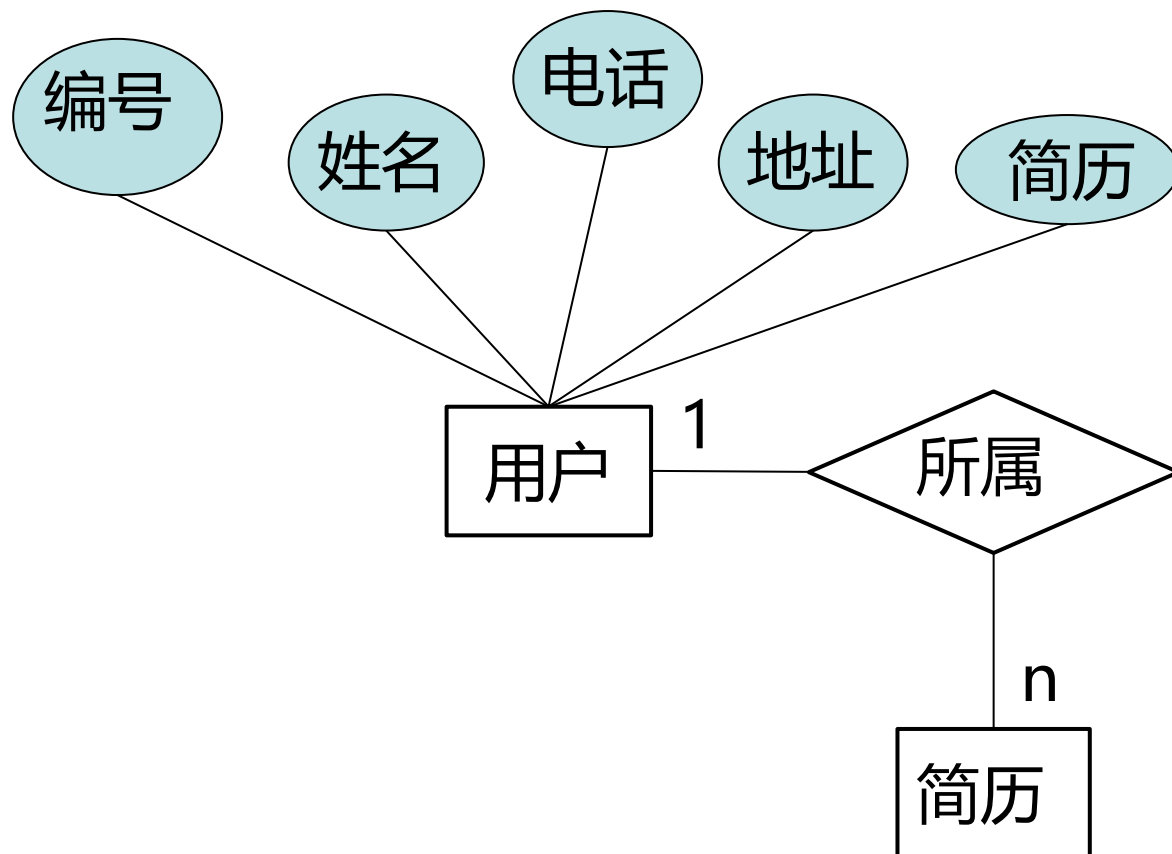
如何支持多地址?



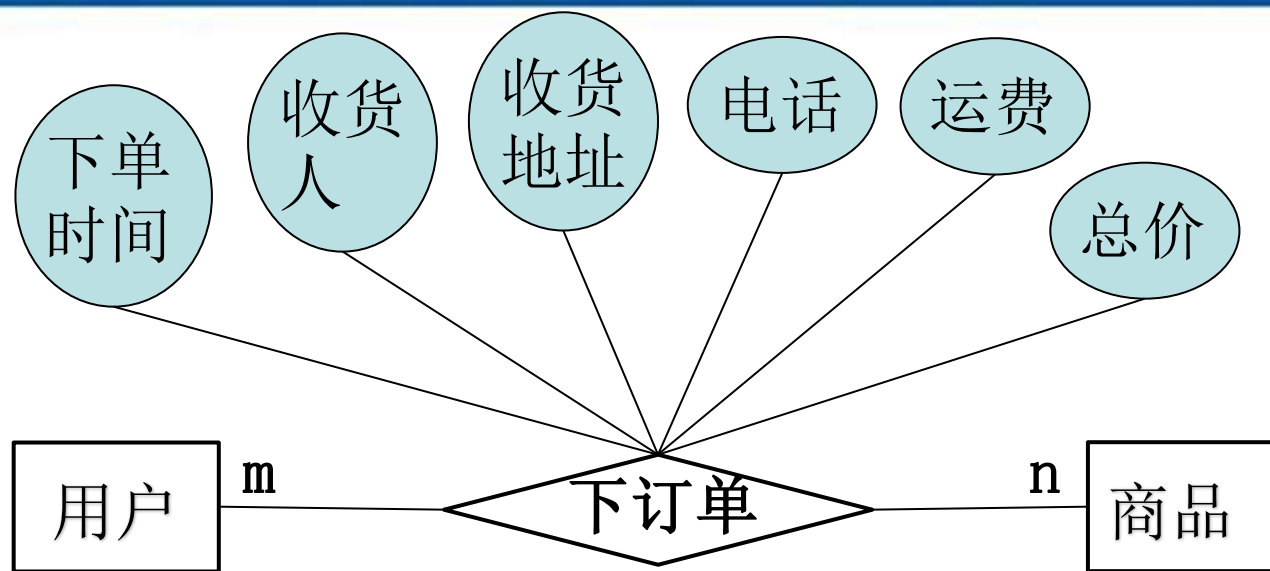
属性方面的讨论:



属性方面的讨论：



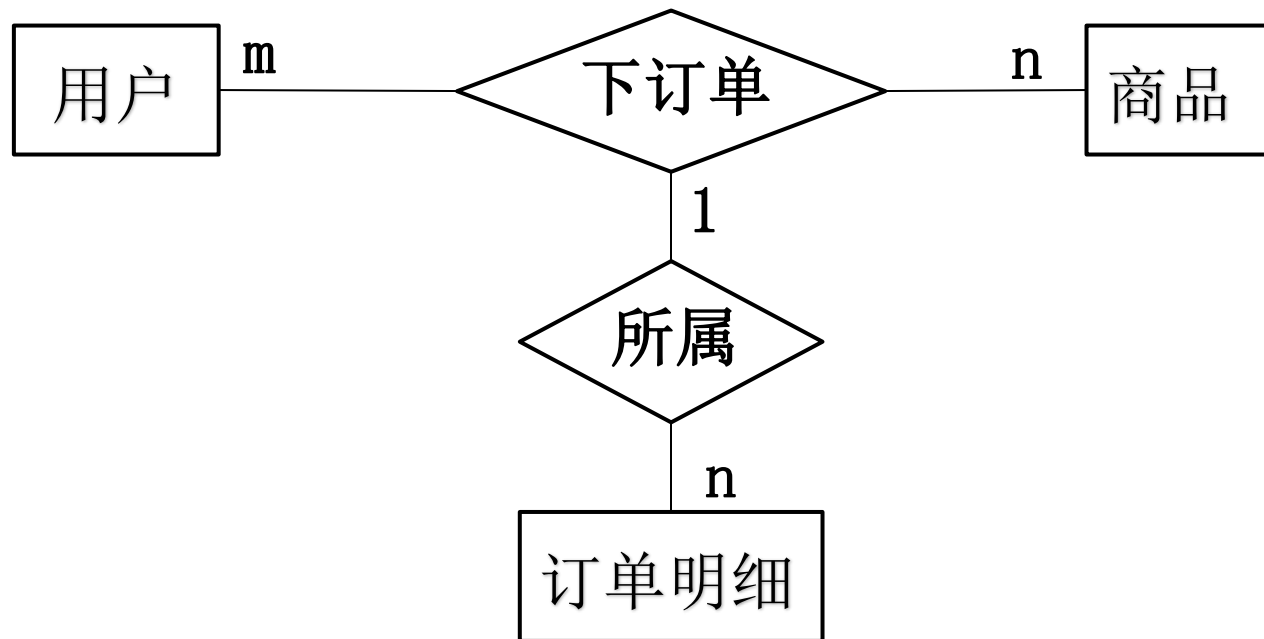
属性方面的讨论:



订单编号	用户id	下单时间	收货人	地址	电话	运费	总价



订单



属性方面的讨论：

- 导出属性

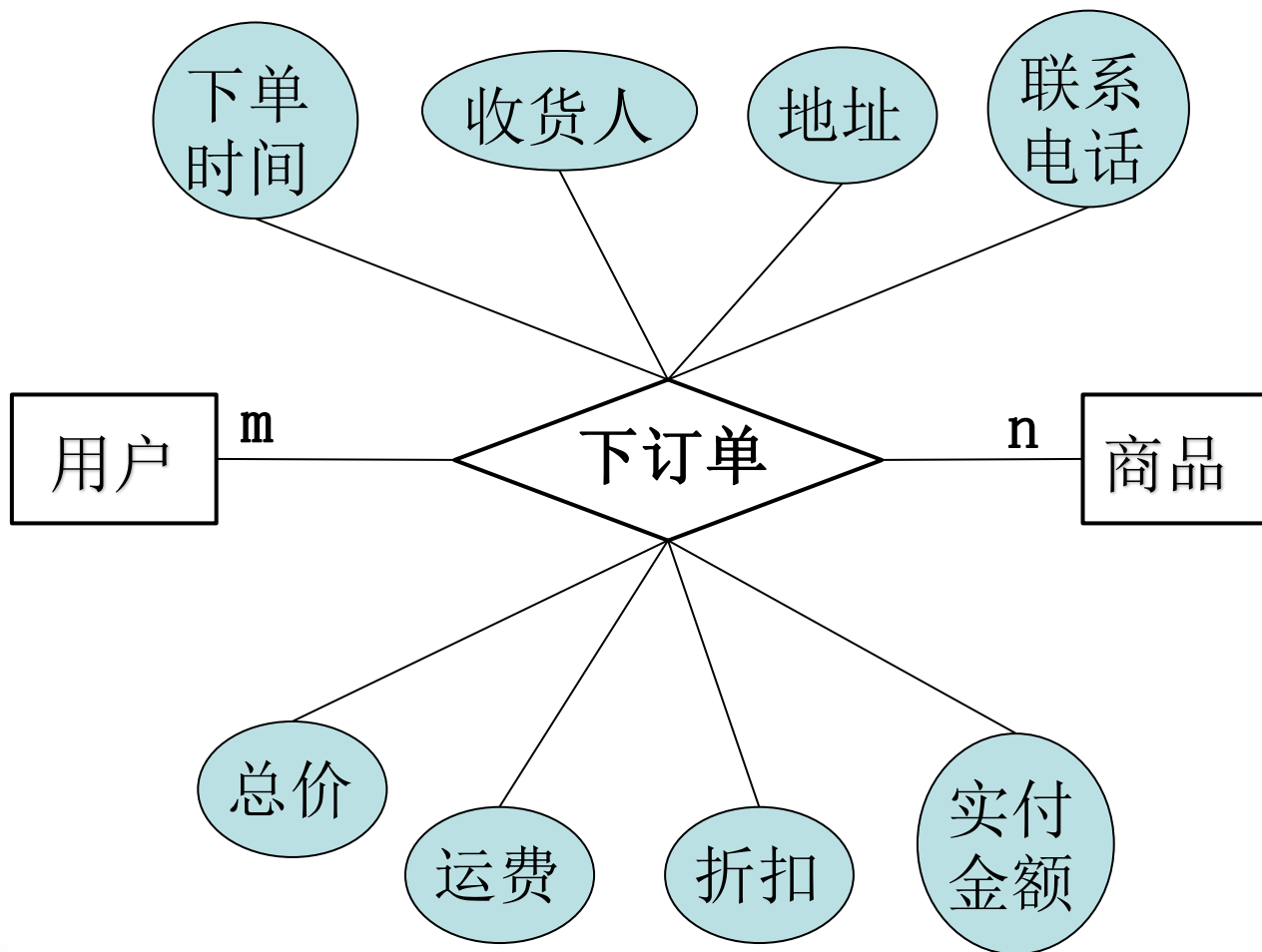
- 属性的值是从相关的一个或者一组属性（不一定来自同一个实体类型）的值导出出来的属性。

- 例如：

- 租用时间，总人数，成绩

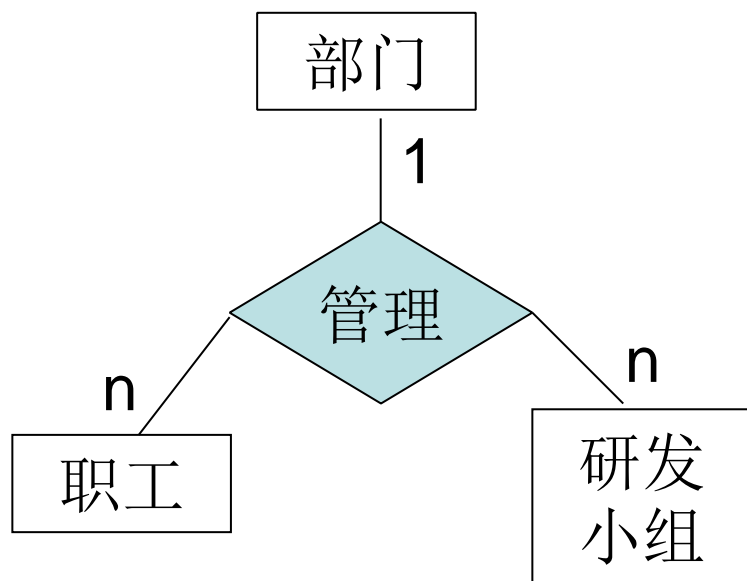


属性方面的讨论:



联系方面的讨论：

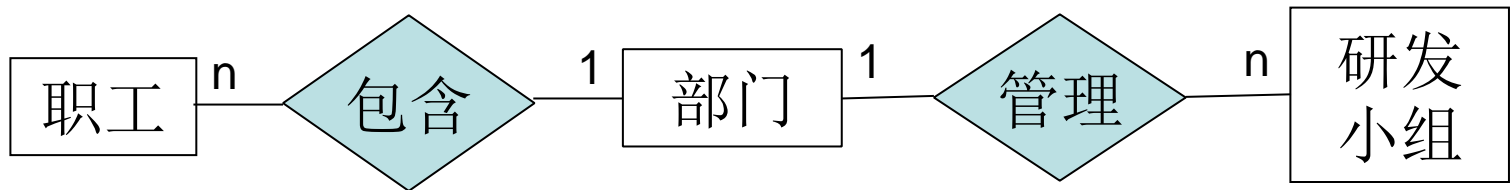
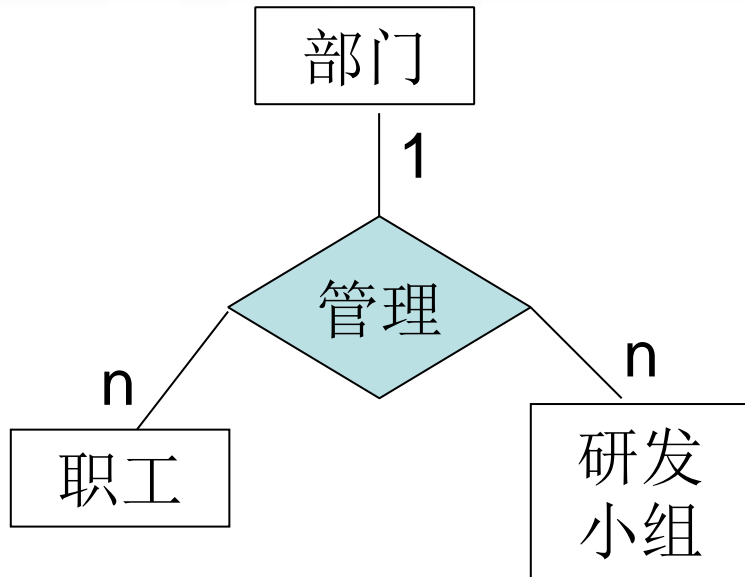
【例】一个部门包含多名职工以及多个研发小组，每个职工和每个研发小组只属于一个部门。



扇形陷阱：

模型给出了两实体类型之间的一种联系，但在某些实体出现之间存在着多条通路。





联系方面的讨论：

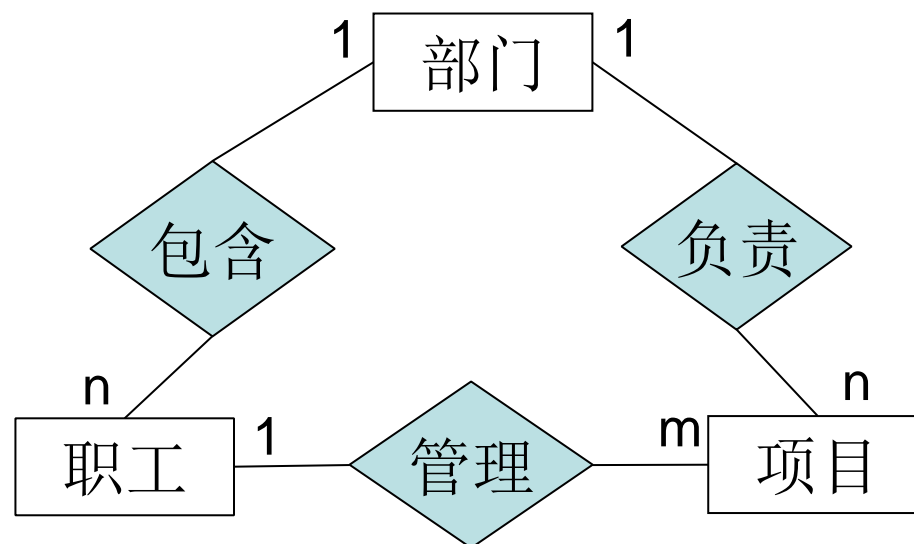
【例】一个部门包含多名职工并承接多个项目，部门会给每个项目指派相应的负责人。



• 断层陷阱

- 模型表明某些实体类型之间存在着联系，但某些实体出现之间却不存在通路。





5.3.4 集成全局视图

视图集成要解决的问题

(1) 确定模式之间的对应和冲突

- 属性冲突 (属性域冲突 取值单位冲突)
- 命名冲突 (同名异议 异名同义)
- 结构冲突

(2) 修改视图使得相互一致

(3) 合并视图

(4) 重构(可选)



1、属性冲突

- 两类属性冲突
 - 属性域冲突：属性值的类型、取值范围不同。
 - 属性取值单位冲突。
- 属性冲突的解决方法
 - 通常用讨论、协商等行政手段加以解决



2、命名冲突

- 两类命名冲突
 - 同名异义：不同意义的对象在不同的局部应用中具有相同的名字
 - 异名同义（一义多名）：同一意义的对象在不同的局部应用中具有不同的名字



2、命名冲突

➤ 属性层面

s(sno, sname,)

s(sid, sname,)

➤ 实体层面

学生(学号, 姓名,)

用户(学号, 姓名,)

• 命名冲突的解决方法

- 通过讨论、协商等行政手段加以解决



3、结构冲突

三类结构冲突

- 同一实体在不同局部视图所包含的属性不完全相同。

- 同一对象在不同应用中具有不同的抽象。

在应用A中作为实体，在应用B中作为联系

例如：开课



- 实体之间的联系在不同局部视图中呈现不同的类型。



结构冲突 (1)

毕业设计系统

学生（学号，姓名，密码，电话，学院，
专业，班级）

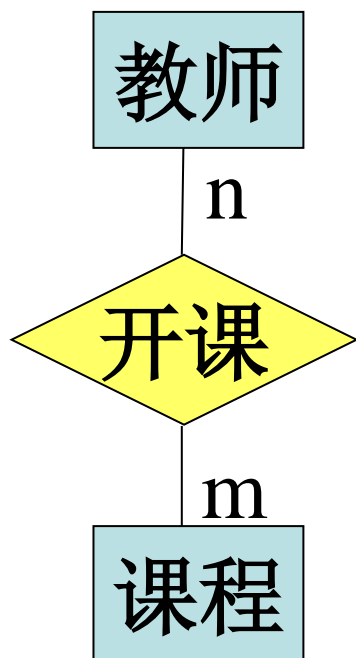
学生信息管理系统

学生（学号，姓名，密码，电话，年龄，入学
年份，籍贯，学院，专业，班级，宿舍）

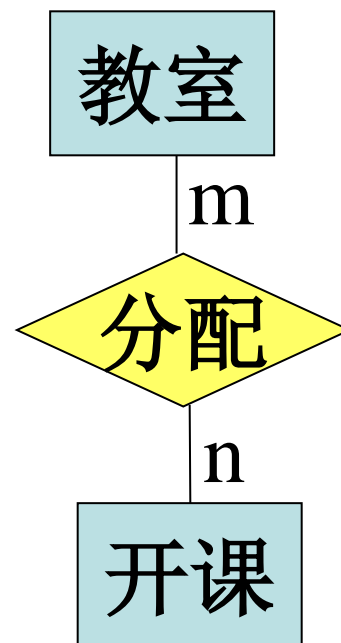


结构冲突 (2)

局部应用1:

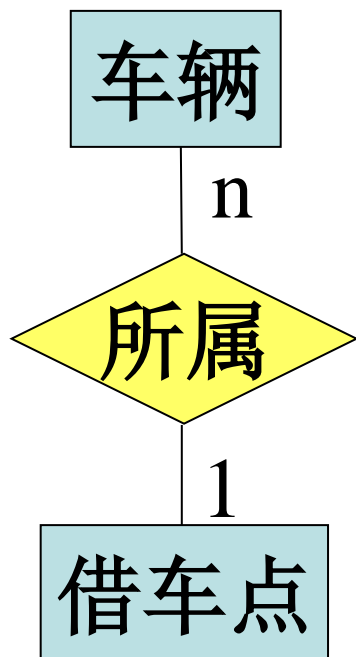


局部应用2:

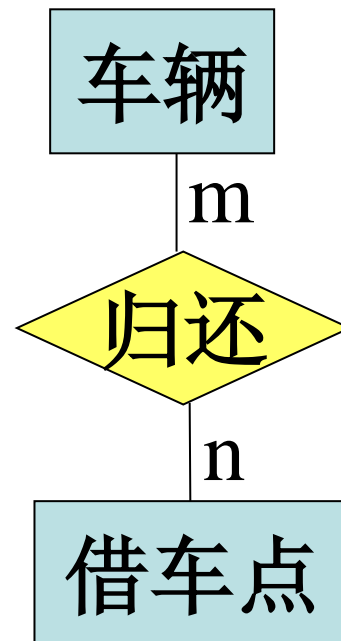


结构冲突 (1)

局部应用1:

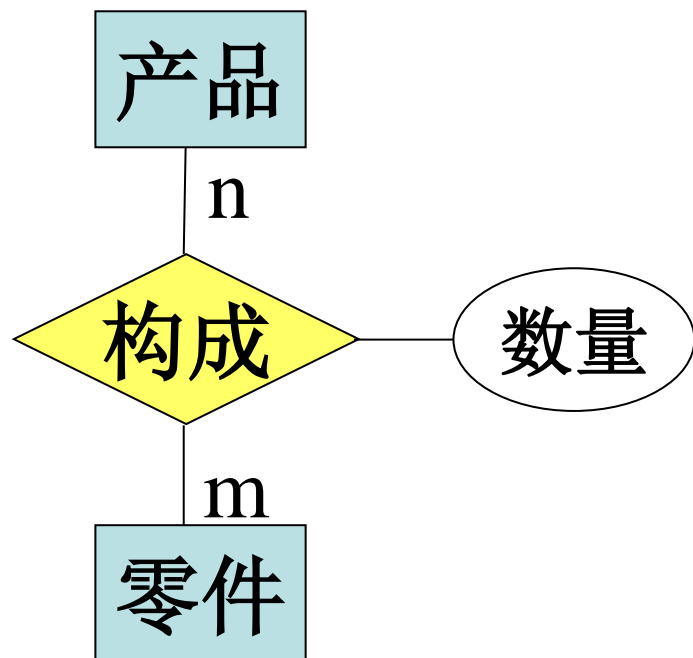


局部应用2:

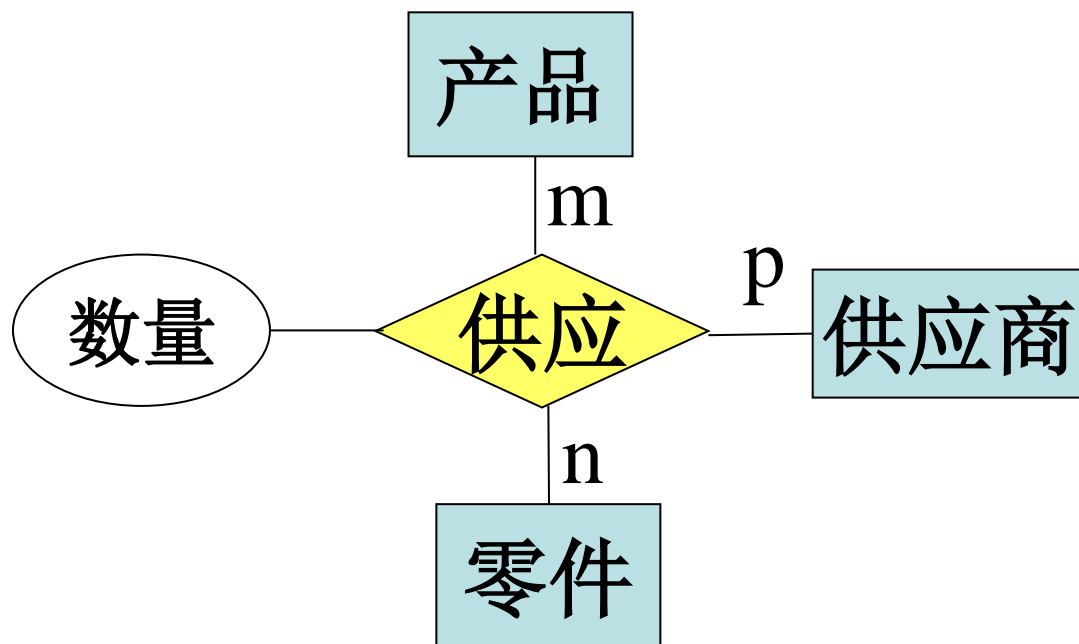


结构冲突 (1)

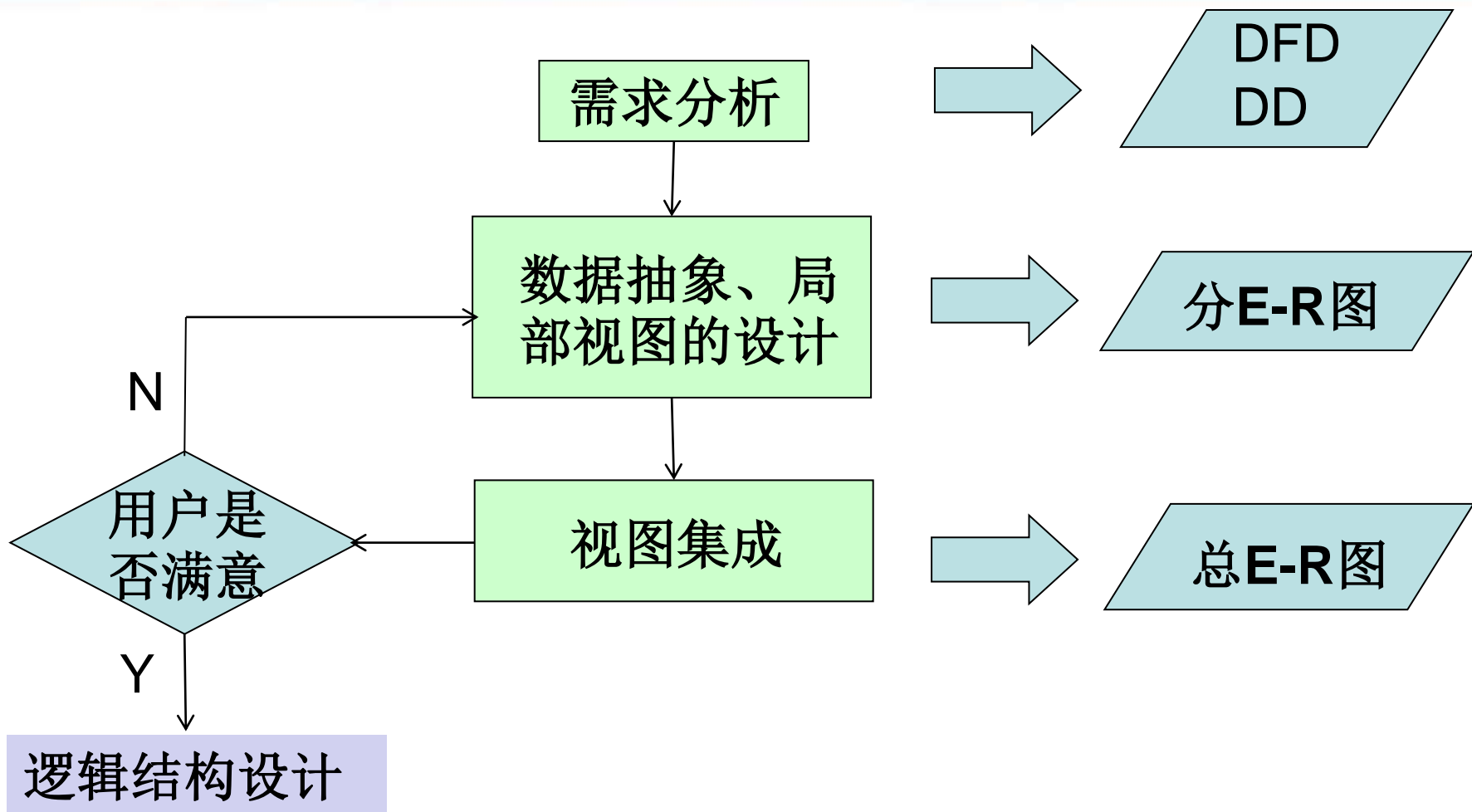
局部应用1:



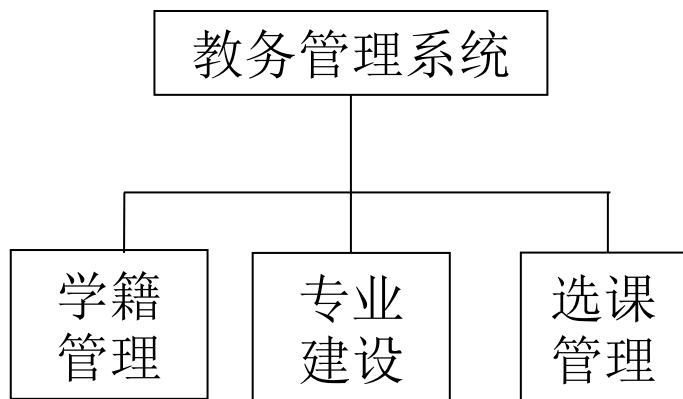
局部应用2:



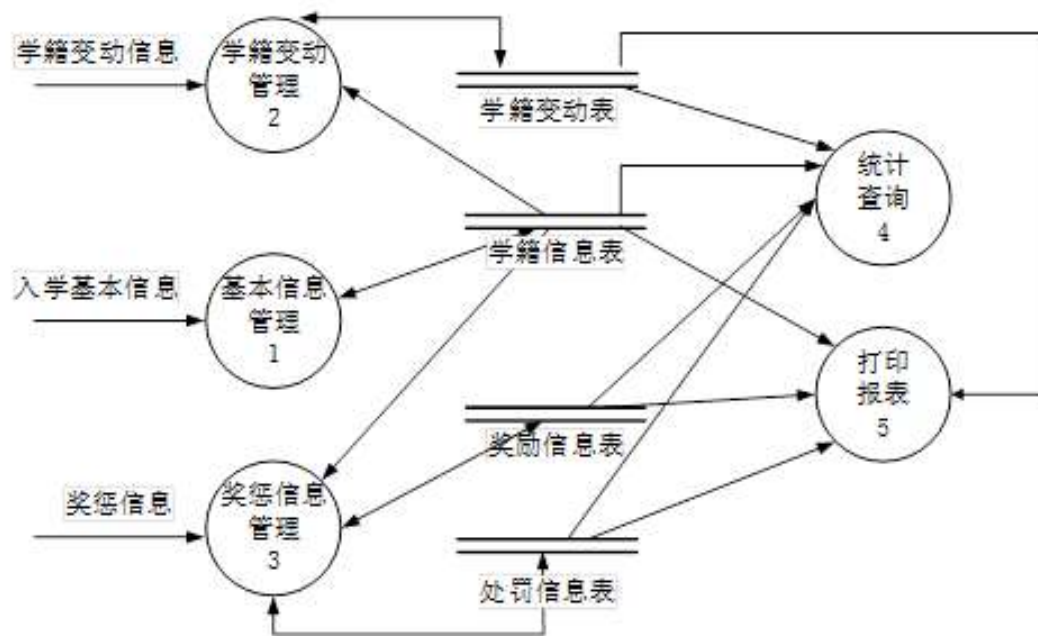
5.3.4 集成全局视图



学校管理信息系统



教务管理系统功能结构图



学籍管理子系统一层数据流图



局部视图设计

➤在“学籍管理模块”中，涉及到的实体有哪些？

学生，班级，学院，专业，宿舍，班主任

➤实体包含哪些属性？实体之间有什么联系？

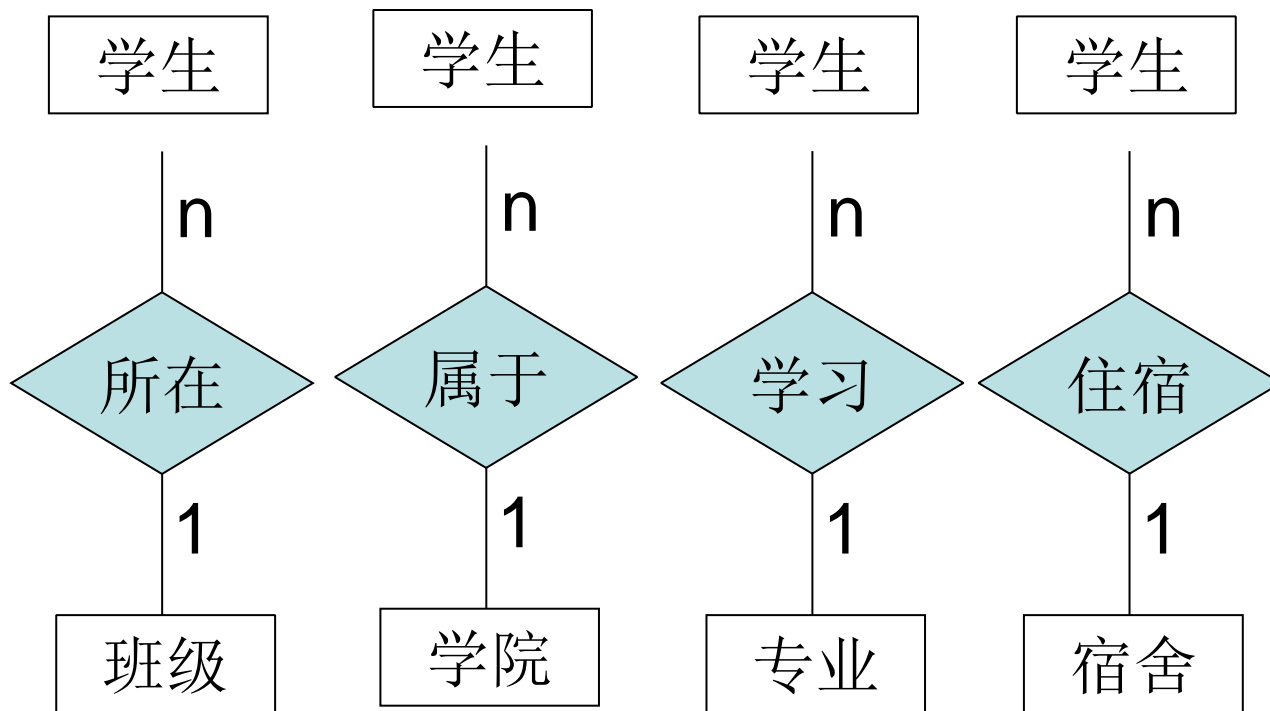


➤ 局部应用中的实体及属性

- 学生：{学号，姓名，性别，年龄，出生日期，院系，专业，班级，宿舍}
- 班级：{班级编号，班级备注信息}
- 宿舍：{宿舍编号，地址，电话}
- 学院：{学院编号，学院名称，院长，办公电话}
- 专业：{专业编号，专业名称}
- 班主任：{职工号，姓名，性别}
- 教师：{职工号，姓名，性别，职称}

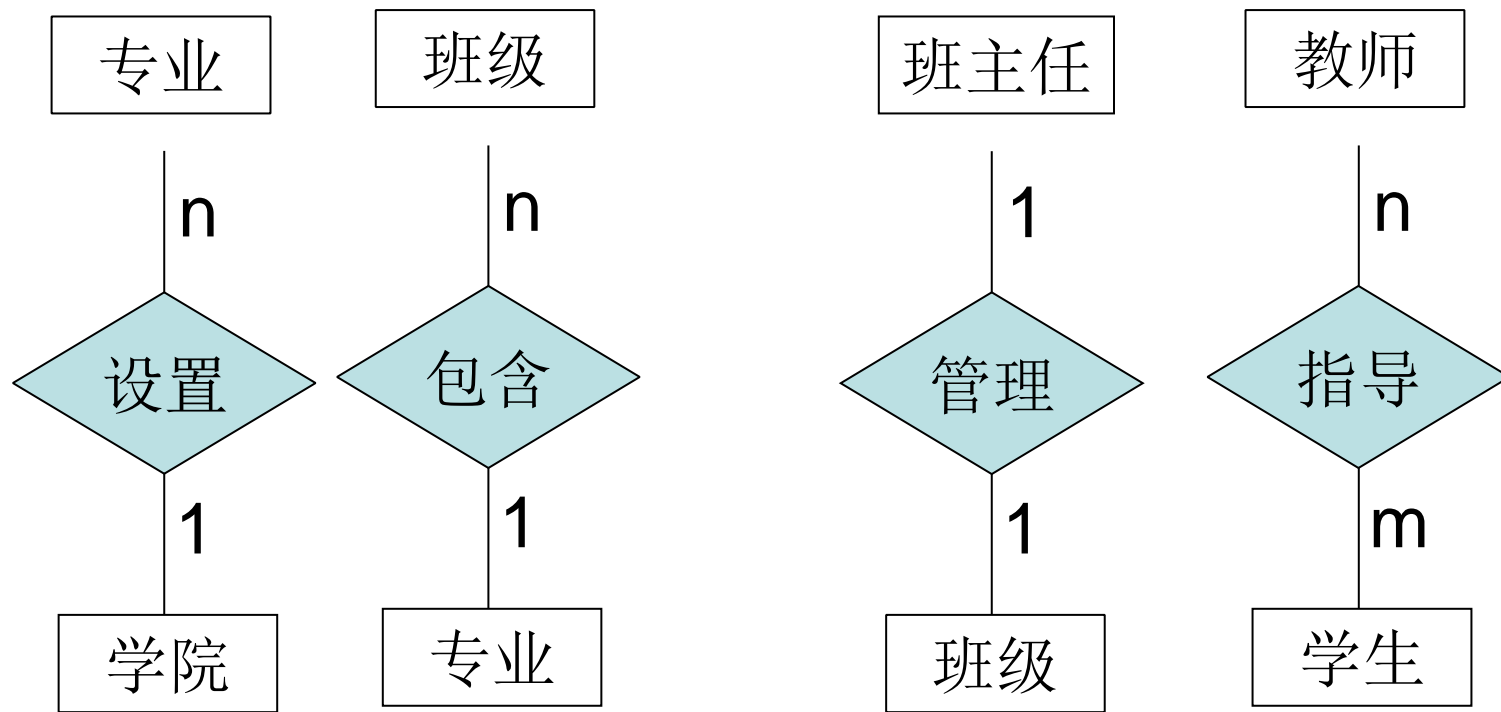


➤ 实体之间的联系及类型



学籍管理子系统分E-R图





学籍管理子系统分E-R图



应用举例——消除冲突

学籍管理子系统

- 学生：{学号，姓名，性别，年龄，出生日期，院系，专业，**班级**，宿舍}
- **班级**：{**班级编号**，班级信息备注}
- 宿舍：{宿舍编号，地址，电话}
- 学院：{**编号**，学院名称，院长，办公电话}
- 专业：{**编号**，专业名称}
- 班主任：{职工号，姓名，性别}

命名冲突

结构冲突

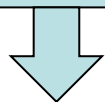


应用举例——消除冲突

教师：{职工号，姓名，性别，职称}

班主任：{职工号，姓名，性别}

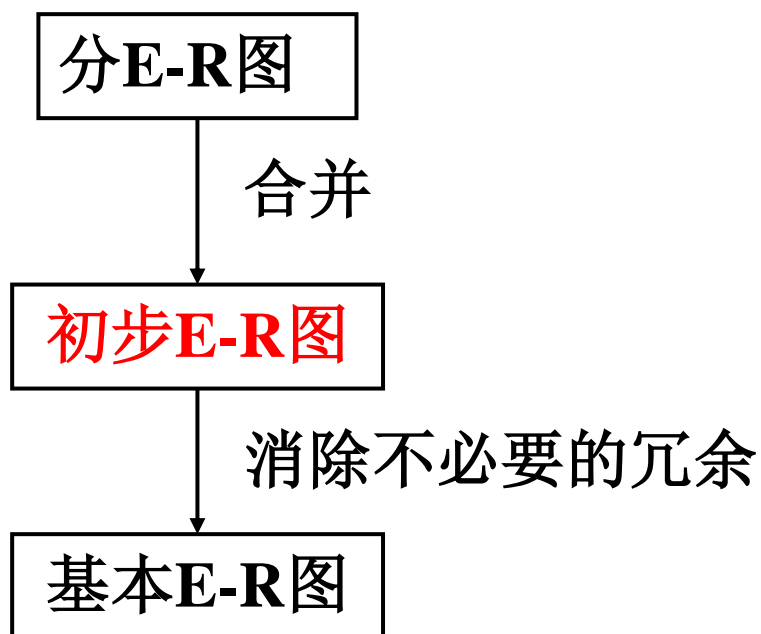
合并



教师：{职工号，教师姓名，性别，职称}



消除不必要的冗余，设计基本E-R图



可能存在冗余的数据
和冗余的实体间联系



消除冗余的方法

1. 分析法

- 以数据字典和数据流图为依据，根据数据字典中关于数据项之间逻辑关系的说明来消除冗余。

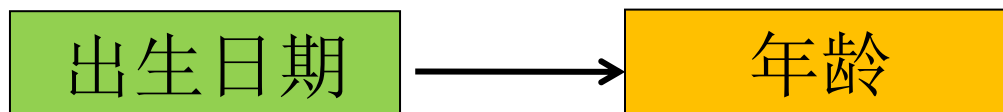
2. 规范化理论法

- 函数依赖的概念提供了消除冗余联系的形式化工具

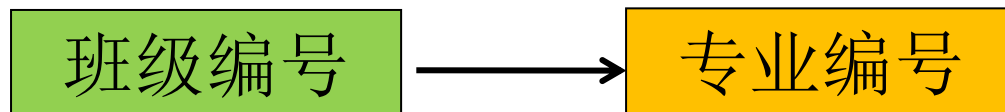


消除冗余的属性

➤ 学生：{学号，姓名，性别，出生日期，学院编号，专业编号，班级编号，宿舍编号}



➤ 学生：{学号，姓名，性别，出生日期，学院编号，专业编号，班级编号，宿舍编号}



➤ 学生：{学号，姓名，性别，出生日期，学院编号，班级编号，宿舍编号}



消除冗余的联系

➤ 根据语义分析

数据流图和数据字典

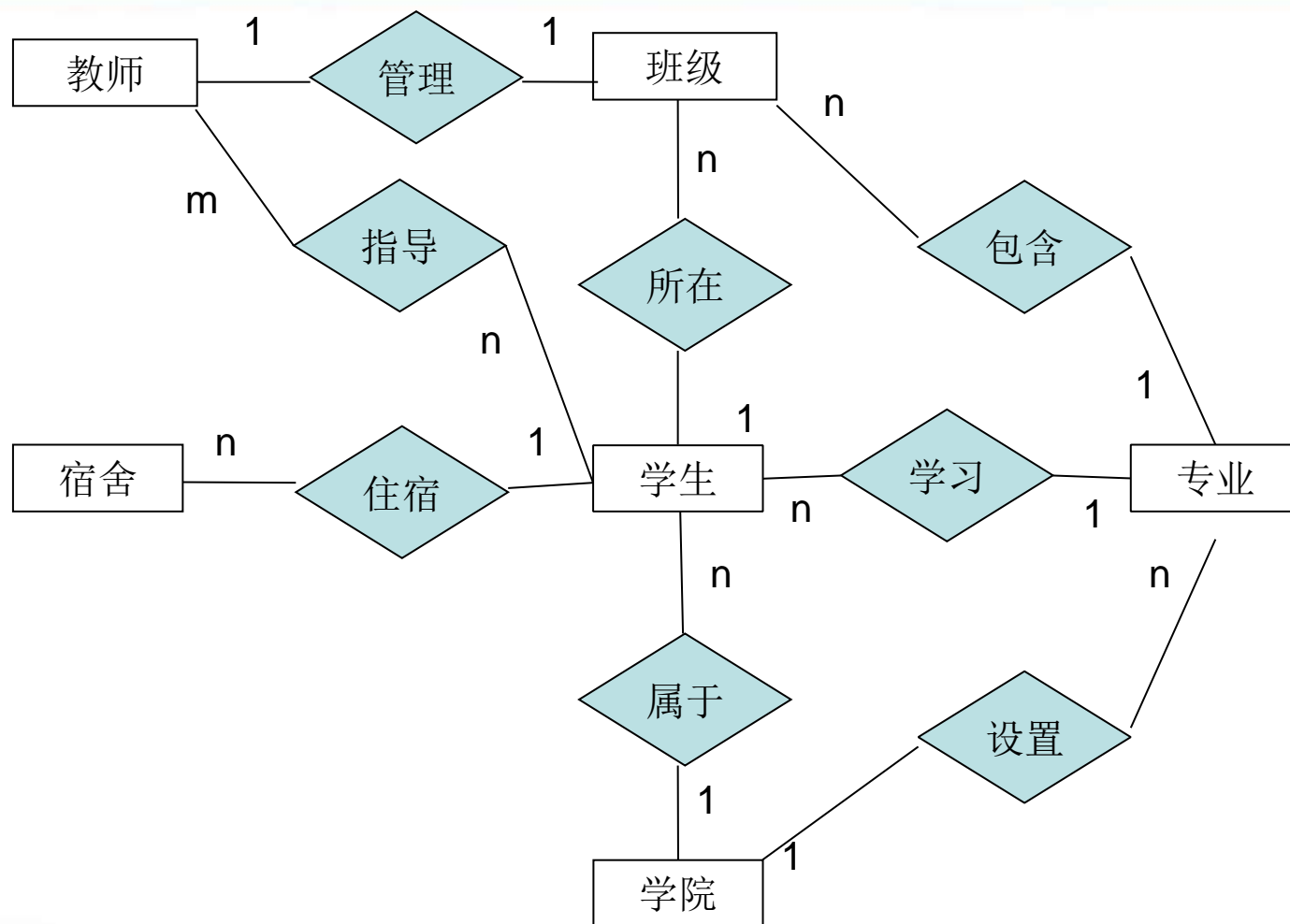
➤ 规范化

求 F 的最小覆盖 G ，差集为 $D = F - G$ 。

逐一考察 D 中的函数依赖，确定是否是冗余的联系。



视图集成



学籍管理子系统总E-R图



第一章课后题

设计一个学生档案管理系统，学生的信息主要包括学生的学号、姓名、性别，入学年份、出生日期、联系电话，宿舍等，和学生相关的信息还有学生所在学院，所学专业，所在班级，班主任等。其中一个学院可以有多个专业，一个专业可以有多个班级，一个专业只属于某个学院，一个班级也只属于某个专业，一个班级只能有一个班主任，一个班主任也只能带一个班级，学院信息包括学院代号、学院名称、学院负责人，专业信息包括专业代号、专业名称、专业负责人，班级信息包括班级代号、班级名称，班主任信息包括工号、姓名、职称，和班主任有关信息包括所在学院和所在系部，一个教师只能在一个学院下的一个系部工作。

要求：(1) 确定有哪些实体，每个实体包括哪些属性。

(2) 找出实体间的联系，并画出E-R图。

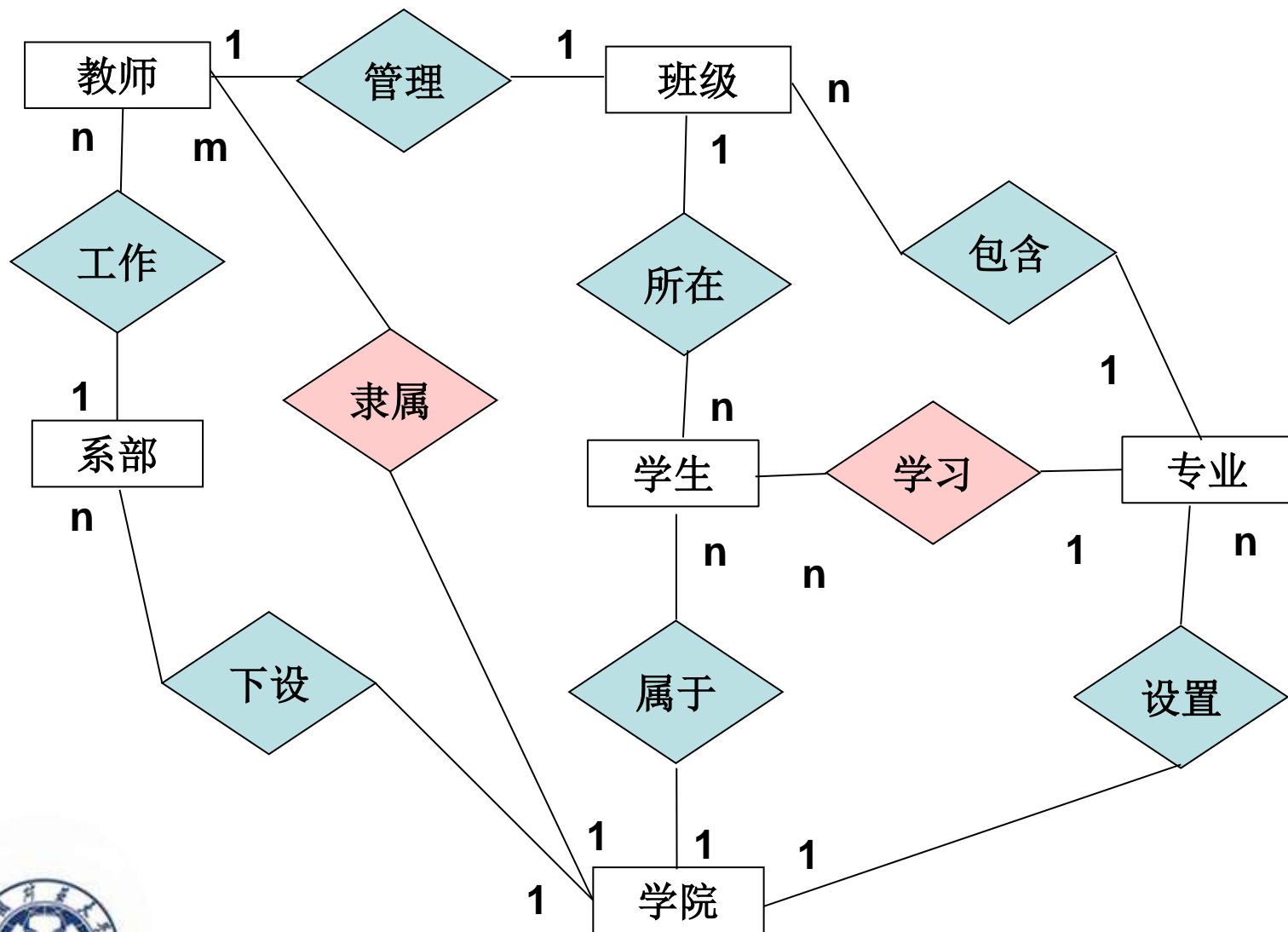


➤ 实体及其属性

- 学生 {学号, 姓名, 性别, 入学年份, 出生日期, 联系电话, 宿舍}
- 学院 {学院代号, 学院名称, 学院负责人}
- 专业 {专业代号, 专业名称, 专业负责人}
- 班级 {班级代号, 班级名称}
- 班主任 {工号, 姓名, 职称}
- 教师 {职工号, 姓名, 性别, 职称}
- 系部 {系部代号, 系部名称}



视图集成



个人博客系统

- 用户发表博客
- 管理员审核博客
- 用户浏览、收藏喜欢的博客
- 用户可以评论博客或者

