

# 数据库系统

哈尔滨工业大学（深圳）

# 数据库设计过程



## 基本内容

1. 数据库设计过程与设计方法
2. E-R图向关系模式的转换
3. 不正确数据库设计引发的问题及其解决

## 重点与难点

- 理解数据库设计的四个过程
- 理解不正确数据库设计引发的问题，为数据库理论的学习奠定问题基础
- 理解不正确数据库设计引发的问题，提升数据建模与数据库设计能力



# 第13讲 数据库设计过程

---

## 13.1 数据库设计过程概述

## 13.2 数据库设计过程之需求分析

## 13.3 数据库设计过程之概念数据库设计

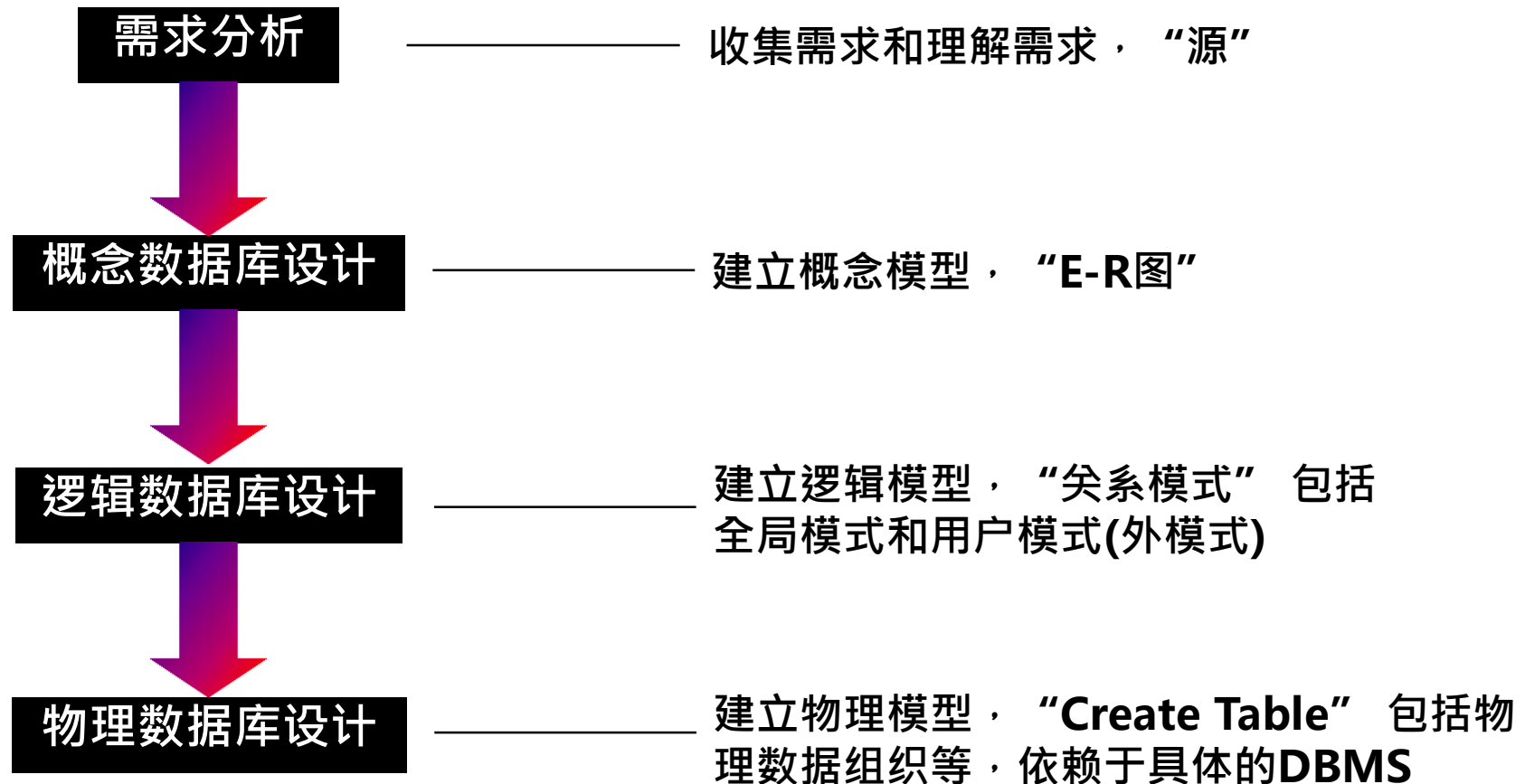
## 13.4 数据库设计过程之逻辑数据库设计

## 13.5 数据库设计过程之物理数据库设计

## 13.1 数据库设计过程概述

---

### (1)数据库设计的四个过程





## 第13讲 数据库设计过程

---

13.1 数据库设计过程概述

13.2 数据库设计过程之需求分析

13.3 数据库设计过程之概念数据库设计

13.4 数据库设计过程之逻辑数据库设计

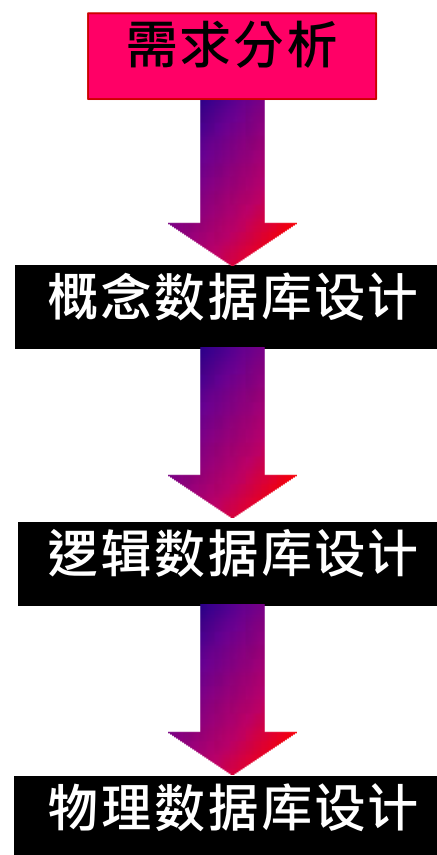
13.5 数据库设计过程之物理数据库设计

## 13.2 数据库设计过程之需求分析



### (1)需求分析

- 目标：理解企业、理解企业业务过程与数据处理流程、理解数据处理的性能需求
- 提交物：需求分析报告
- 使以下内容清楚：
  - ❑ 企业的部门-岗位划分：不同岗位负责的各种日常管理信息表/报表
  - ❑ 形成各种报表的基础数据表
  - ❑ 各种数据表之间的处理关系(What--How)
  - ❑ 围绕数据表的业务处理关系(Who—When--Where)
  - ❑ 尚未实施但未来可能实施的需求
- 形成数据库设计的“源”清单和“属性”清单，及详细描述，尤其注意业务规则与属性处理规则



# 13.2 数据库设计过程之需求分析



## (2)相关结果性内容示意

### “源” 清单(后附源表的格式)

总序号	源表名称	
	处理及存档要求	
	源表属性名表	
总序号	源表名称	
	处理及存档要求	
	源表属性名表	
总序号	源表名称	
	处理及存档要求	
	源表属性名表	

注意收集和理解: 1.业务规则及其在表的处理方面 的体现 2.不仅报表、单据是源，企业的 查询需求与管理需求等也是源



## 13.2 数据库设计过程之需求分析



### (2) 相关结果性内容示意

#### “属性” 清单

局部序号	属性名称		全局统一序号	属性统一名称		对应“源”序号
	英文	中文		英文	中文	

注意命名: 1.命名要规范, 并且要含义明确  
2.尤其要注意类似于“数量”这样的多含义属性, 比如“计划数量”“采购数量”“到货数量”“装配数量”“完工数量”“销售数量”“发货数量”



## 第13讲 数据库设计过程

---

13.1 数据库设计过程概述

13.2 数据库设计过程之需求分析

13.3 数据库设计过程之概念数据库设计

13.4 数据库设计过程之逻辑数据库设计

13.5 数据库设计过程之物理数据库设计

## 13.3 数据库设计过程之概念数据库设计

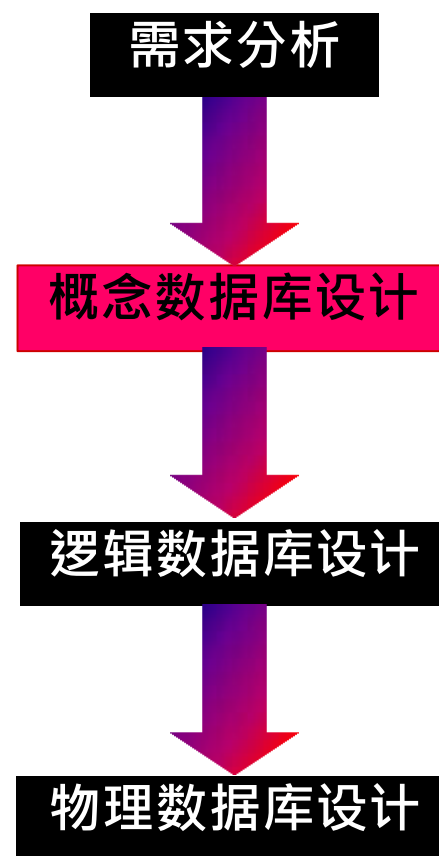


### (1) 概念数据库设计

- 目标：进一步深入理解企业，对信息源进行抽象，发现信息(属性)之间的内在本质联系，这些本质联系可能隐藏于需求分析得到的信息源中。
- 提交物：概念数据库设计报告
- 使以下内容清楚：
  - ❑ 各种实体的发现、划分和定义
  - ❑ 各种实体属性的发现、分析和定义
  - ❑ 各种实体联系的发现、分析和定义
  - ❑ 外部视图(模式)和概念视图(模式)的定义
- 用统一的表达方法，如E-R模型给出描述；不仅绘制出来，而且绘制正确；

用规范的数据模型表达，有助于更好的理解需求

数据模型不仅是自己理解而且要让相关人员理解



## 13.3 数据库设计过程之概念数据库设计



### (3)概念数据库设计的可能冲突

#### 消除冲突

##### ➤属性冲突

—属性域的冲突：属性的类型、取值范围不同

•如不同学校的学号编码方式不同

—属性取值单位冲突

•如重量分别采用磅、千克

##### ➤结构冲突

—同一对象在不同应用中的抽象不同

•如职工在某应用中是实体，在另一应用中则抽象为属性

—同一实体在不同E-R图中属性组成不同

—实体之间的联系在不同E-R图中呈现不同的类型

##### ➤命名冲突

—同名异义：不同意义的对象具有相同的名字

—异名同义：同一意义的对象具有不同的名字



## 第13讲 数据库设计过程

---

13.1 数据库设计过程概述

13.2 数据库设计过程之需求分析

13.3 数据库设计过程之概念数据库设计

13.4 数据库设计过程之逻辑数据库设计

-----Part(I)

-----Part(II)

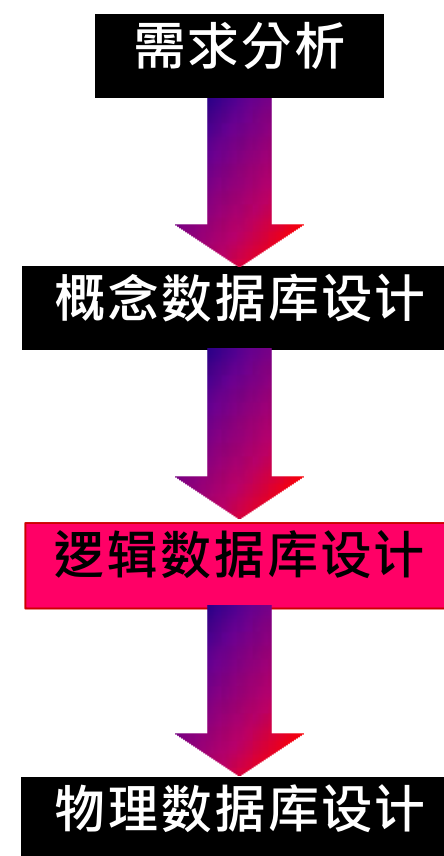
13.5 数据库设计过程之物理数据库设计

## 13.4 数据库设计过程之逻辑数据库设计



### (1)逻辑数据库设计

- 目标：用指定DBMS要求的模式描述方法，给出概念数据库的逻辑模式描述。
- 提交物：逻辑数据库设计报告
- 使以下内容清楚：
  - ❑ 将E-R转换成逻辑模式
  - ❑ 遵循关系范式的设计原则
  - ❑ 也要注意折中，但折中时需要提示应用开发者或使用者可能存在的问题
  - ❑ 外模式和概念模式的定义
- 用关系模型规定的模式描述方法进行描述

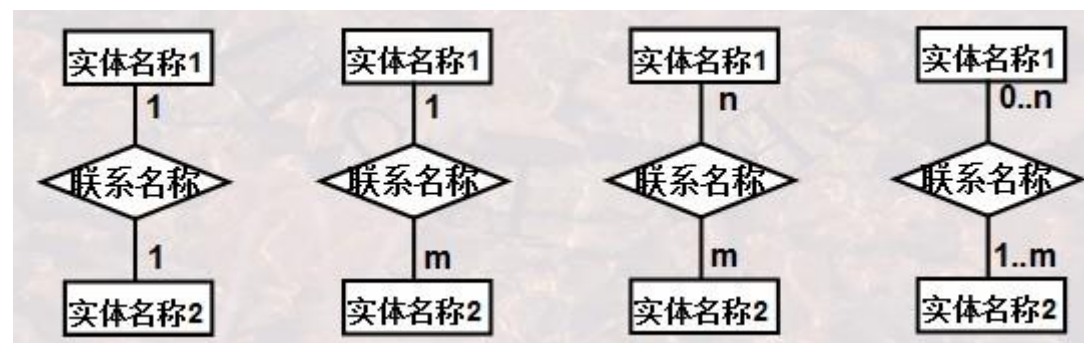
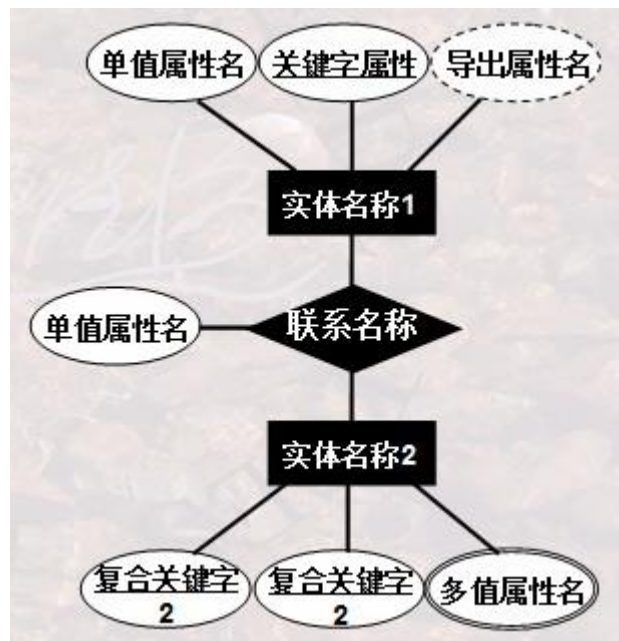


## 13.4 数据库设计过程之逻辑数据库设计



### (2) E-R图向关系模式的转换

#### E-R图(Chen方法)



#### 关系模式(Schema)

$R(A_1:D_1, A_2:D_2, \dots, A_n:D_n)$   
简记为  $R(A_1, A_2, \dots, A_n)$

## 13.4 数据库设计过程之逻辑数据库设计

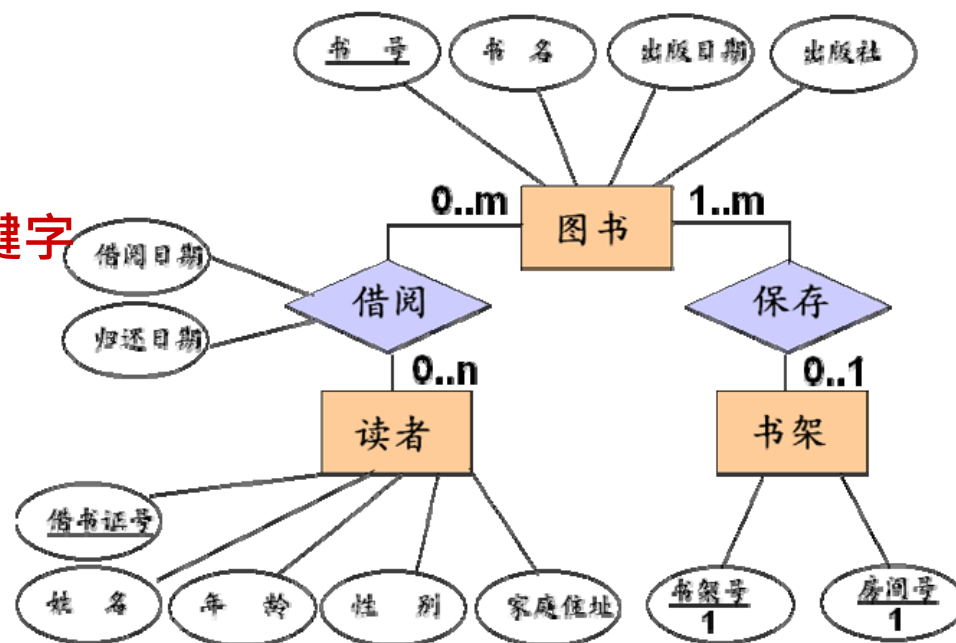


### (2)E-R图向关系模式的转换

基本转换规则：实体-属性-关键字的转换

- E-R图的**实体**转换为**关系**
- E-R图的**属性**转换为关系的**属性**
- E-R图的**关键字**转换为关系的**关键字**

示例



图书(书号,书名,出版日期,出版社)

读者(借书证号,姓名,年龄,性别,家庭住址)

书架(书架号,房间号)



## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

#### 基本转换规则：复合属性的转换

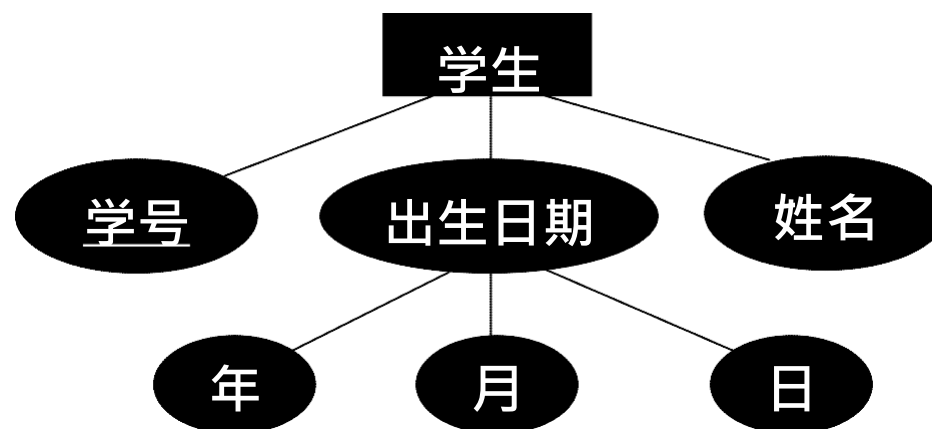
- 将每个分量属性作为复合属性所在实体的属性
- 或者，将复合属性本身作为所在实体的属性

#### 示例

学生(学号, 姓名, 年, 月, 日)

或者

学生(学号, 出生日期, 姓名)



## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

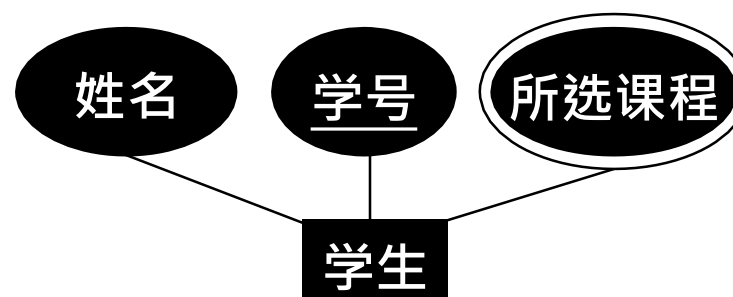
基本转换规则：多值属性的转换

➤将多值属性与所在实体的关键字一起组成一个新的关系

示例

学生(学号,姓名)

选课(学号,所选课程号)



## 13.4 数据库设计过程之逻辑数据库设计

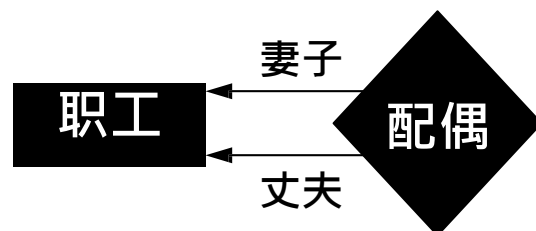


### (2)E-R图向关系模式的转换

#### 基本转换规则：联系的转换

##### ➤ 一对一联系：

□ 若联系双方均部分参与(0..1)，则将联系定义为一个新的关系，属性为参与双方的关键字属性



职工(职工号,...)

配偶(丈夫职工号, 妻子职工号)

□ 若联系一方全部参与(1..1)，则将其关键字作为另一方属性，不需要联系集



职工(职工号, ...)

部门(部门号, 部门名, 职工号)

## 13.4 数据库设计过程之逻辑数据库设计

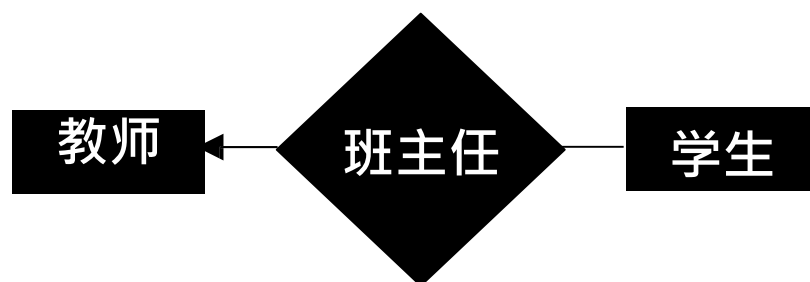


### (2)E-R图向关系模式的转换

#### 基本转换规则：联系的转换

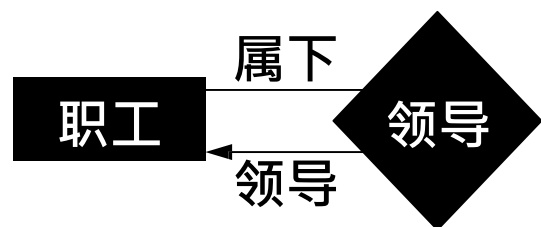
##### ➤一对多联系：

□ 将单方参与实体的关键字，作为多方参与实体对应关系的属性



教师(教工号, ...)

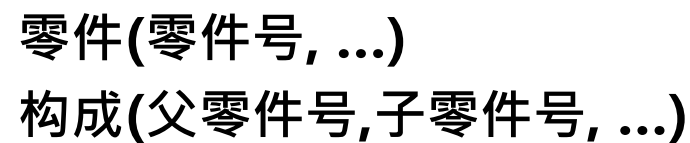
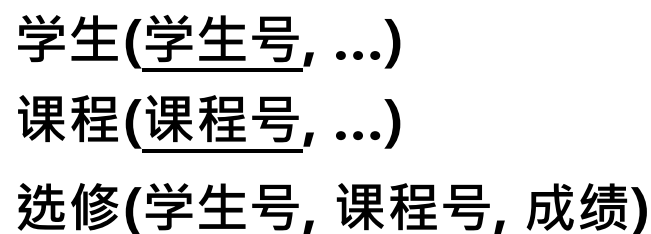
学生(学生号, 学生名, 班主任教工号)



职工(职工号, 职工名, 部门号, 领导职工号)

## 基本转换规则：联系的转换

❑将联系定义为新的关系，属性为参与双方实体的关键字



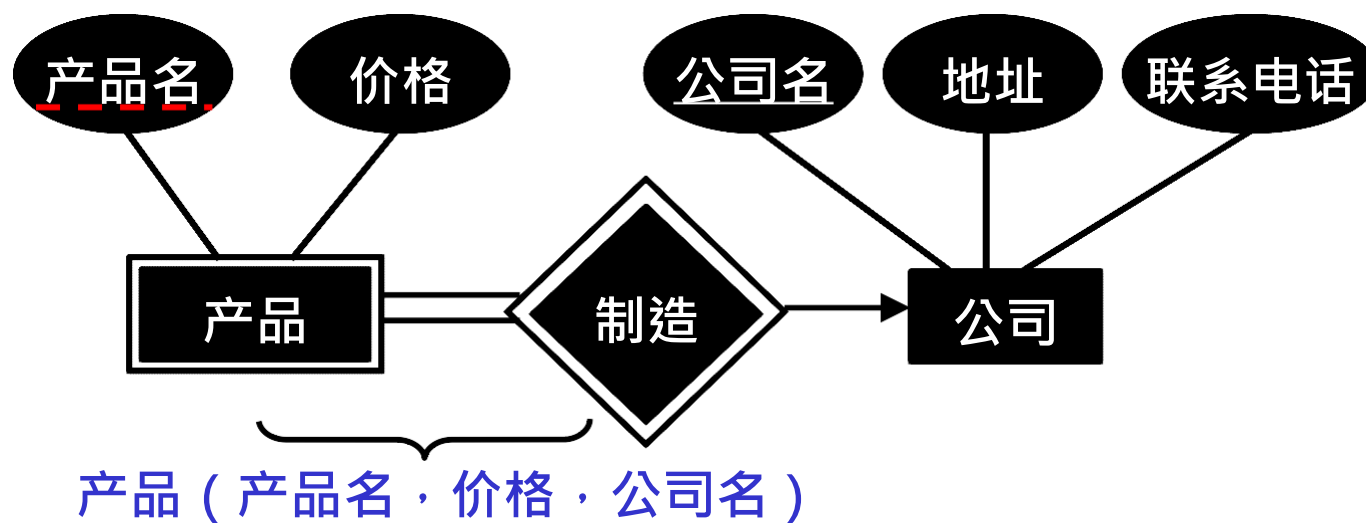
## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

#### 基本转换规则：弱实体的转换

- 所对应关系的关键字由弱实体本身的区分属性再加上所依赖的强实体的关键字构成



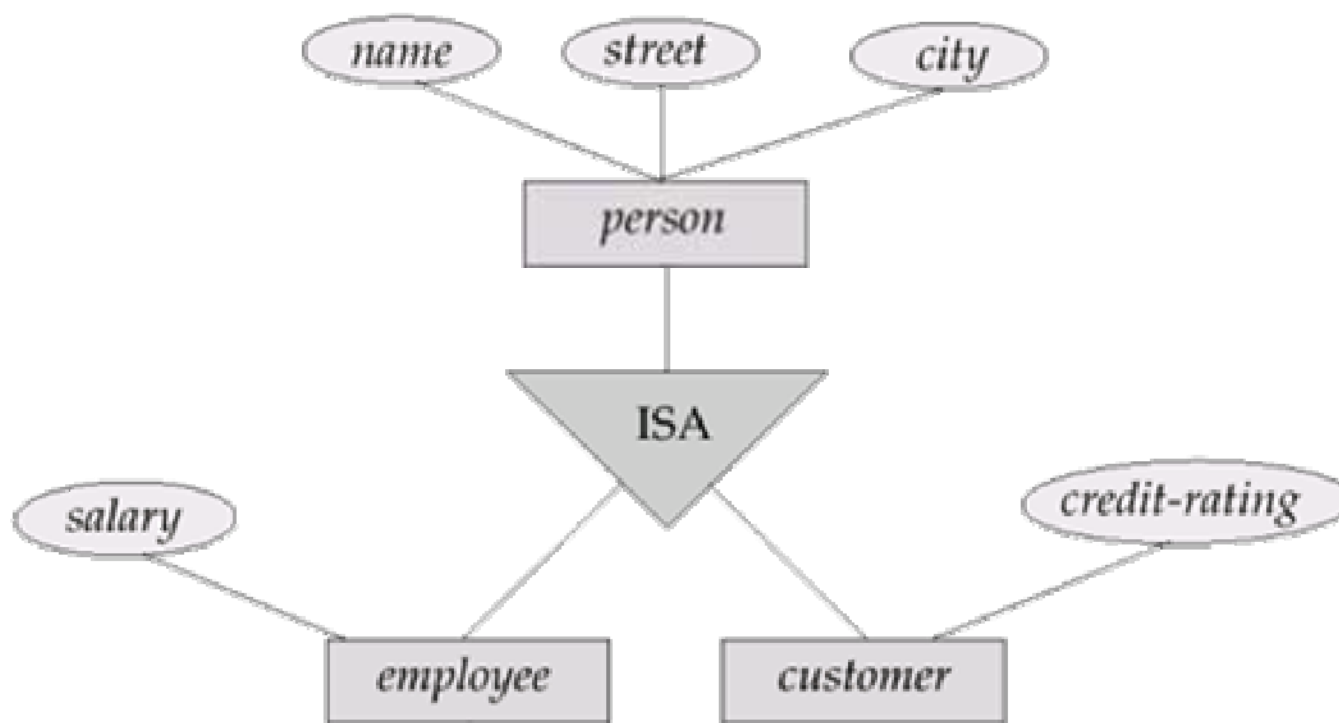
□ 弱实体集(从属实体)与强实体集(独立实体)之间的联系已经在弱实体集所对应的关系中表示出来了

## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

示例：具化/泛化如何转换？



## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

示例：如何转换？

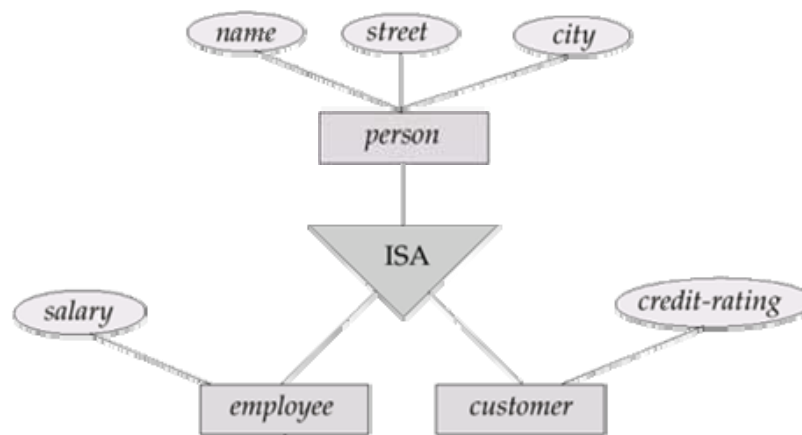
#### □方案1

**person** ( name, street, city )

**customer** ( name, credit-rating )

**employee** ( name, salary )

✓缺点：查询**employee**的地址需要访问两个表



#### □方案2

**person** ( name, street, city )

**customer** ( name, street, city, credit-rating )

**employee** ( name, street, city, salary )

✓如果泛化实体实例是具体化实体实例的全部，可不创建**person**表

✓缺点：地址信息对同时是**customer**和**employee**的人存储两次



## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

基本转换规则：多元联系的转换

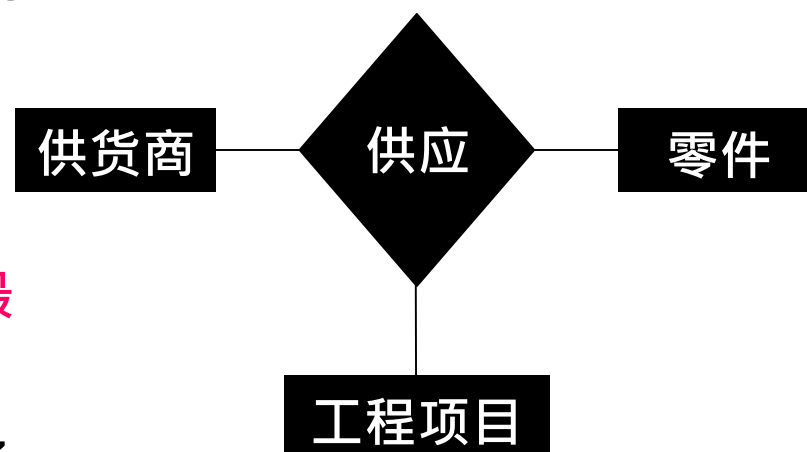
- 多元联系可以通过继承参与联系的各个实体的关键字而形成新的关系
- 这些继承过来的关键字可作为新关系的关键字
- 也可以新增一个区分属性作为关键字

➤注意这两种转换的差异

➤多元联系更需注意分析参与联系的实体的最小基数和最大基数

➤如是否允许参与联系的多实体中有一个或多个实体不参与？

➤多元联系可以转换为多个二元联系进行处理



**1型转换：**供应(工程项目号，供货商号，零件号，数量)

**2型转换：**供应(条目号,工程项目号，供货商号，零件号，数量)

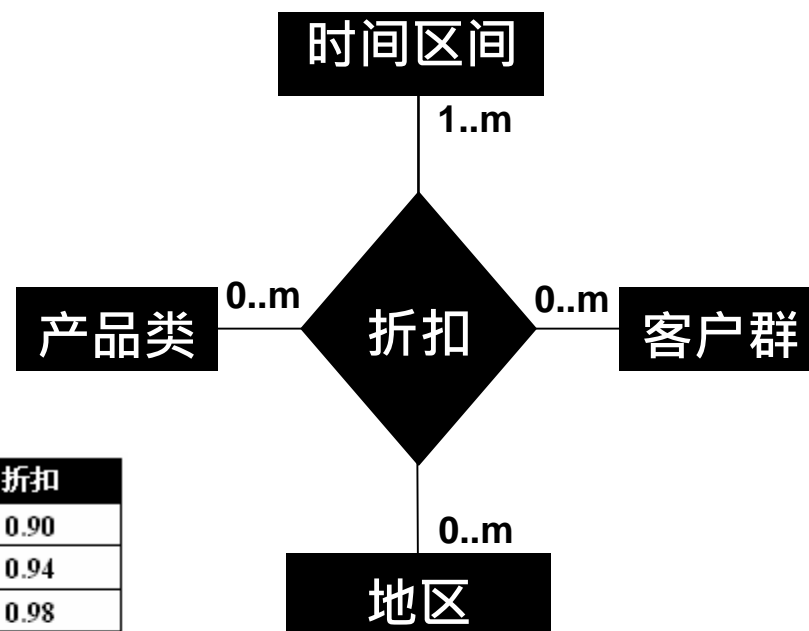
## 13.4 数据库设计过程之逻辑数据库设计



### (2)E-R图向关系模式的转换

#### 示例

➤下表是企业业务规则决定的，如何处理？尤其是检索时，如何注意检索错误？



折扣号	时间区间	地区	产品类	客户群	折扣
Z01	2003/05-2003/07	黑龙江	J1 汽车	经常客户	0.90
Z02	2003/05-2003/07	----	J1 汽车	一般客户	0.94
Z03	2003/05-2003/07	----	----	----	0.98
Z04	2003/05-2003/07	----	----	集团客户	0.945
Z05	2003/05-2003/07	----	J2 汽车	经常客户	0.93
Z06	2003/05-2003/07	黑龙江	----	集团客户	0.94
Z07	2003/05-2003/07	----	J1 汽车	----	0.94
Z08	2003/05-2003/07	----	----	经常客户	0.95



## 第13讲 数据库设计过程

---

13.1 数据库设计过程概述

13.2 数据库设计过程之需求分析

13.3 数据库设计过程之概念数据库设计

13.4 数据库设计过程之逻辑数据库设计

-----Part(I)

-----Part(II)

13.5 数据库设计过程之物理数据库设计

## 13.4 数据库设计过程之逻辑数据库设计



### (3)不正确设计数据库会引发什么问题呢？

#### 不正确设计数据库引发的问题

➤冗余：数据库中存在大量冗余

➤非受控冗余

●例如，右侧数据库设计

➤非受控冗余问题

●当数据发生改变时，如何使冗余数据同步更新？

学号	姓名	班级	班主任	班主任职称
2003510101	张三	035101	张林	讲师
2003510102	李四	035101	张林	讲师
2003510103	王五	035101	张林	讲师
2003510104	李六	035101	张林	讲师
2003510105	张四	035101	张林	讲师
2003510106	张五	035101	张林	讲师
2003510107	张小三	035101	张林	讲师
2003510108	张小四	035101	张林	讲师
2003510109	李小三	035101	张林	讲师
2003510110	李小四	035101	张林	讲师
2003520201	周三	035202	郑东	副教授
2003520202	赵四	035202	郑东	副教授
2003520203	赵五	035202	郑东	副教授
2003520204	赵六	035202	郑东	副教授
2003520205	钱四	035202	郑东	副教授
2003520206	强五	035202	郑东	副教授
2003520207	梁小三	035202	郑东	副教授
2003520208	梁小四	035202	郑东	副教授
2003520209	王小三	035202	郑东	副教授
2003520210	王小四	035202	郑东	副教授



## 13.4 数据库设计过程之逻辑数据库设计

### (3)不正确设计数据库会引发什么问题呢?

➤受控冗余：如Table中的外键(继承其他Table中的键值)

班级	班主任	班主任职称
035101	张林	讲师
035202	郑东	副教授

学号	姓名	班级
2003510101	张三	035101
2003510102	李四	035101
2003510103	王五	035101
2003510104	李六	035101
2003510105	张四	035101
2003510106	张五	035101
2003510107	张小三	035101
2003510108	张小四	035101
2003510109	李小三	035101
2003510110	李小四	035101
2003520201	周三	035202
2003520202	赵四	035202
2003520203	赵五	035202
2003520204	赵六	035202
2003520205	钱四	035202
2003520206	强五	035202
2003520207	梁小三	035202
2003520208	梁小四	035202
2003520209	王小三	035202
2003520210	王小四	035202



## 13.4 数据库设计过程之逻辑数据库设计

### (3)不正确设计数据库会引发什么问题呢?

#### 不正确设计数据库引发的问题

➤ **删除异常**：当四系所有同学被删除后，则四系的有关信息则随之丢失。

学号	姓名	所属系别	系名称	系主任
2003510101	张三	三系	计算机系	徐晓飞
2003510102	李四	三系	计算机系	徐晓飞
2003510103	王五	三系	计算机系	徐晓飞
2003510104	李六	三系	计算机系	徐晓飞
2003510105	张四	三系	计算机系	徐晓飞
2003510106	张五	三系	计算机系	徐晓飞
2003510201	张小三	四系	自动控制系	李教授
2003510202	张小四	四系	自动控制系	李教授
2003510203	李小三	四系	自动控制系	李教授
2003510204	李小四	四系	自动控制系	李教授
2003510205	周三	四系	自动控制系	李教授
2003510206	赵四	四系	自动控制系	李教授
2003510207	赵五	四系	自动控制系	李教授



# 13.4 数据库设计过程之逻辑数据库设计

## (3)不正确设计数据库会引发什么问题呢?

如何避免?

➤设计满足规范性，由DBMS或数据库本身来保证

➤设计不满足规范性，由使用者或应用程序员使用过程中加以注意

学号	姓名	班级	班主任	班主任职称
2003510101	张三	035101	张林	讲师
2003510102	李四	035101	张林	讲师
2003510103	王五	035101	张林	讲师
2003510104	李六	035101	张林	讲师
2003510105	张四	035101	张林	讲师
2003510106	张五	035101	张林	讲师
2003510107	张小三	035101	张林	讲师
2003510108	张小四	035101	张林	讲师
2003510109	李小三	035101	张林	讲师
2003510110	李小四	035101	张林	讲师
2003520201	周三	035202	郑东	副教授
2003520202	赵四	035202	郑东	副教授
2003520203	赵五	035202	郑东	副教授
2003520204	赵六	035202	郑东	副教授
2003520205	钱四	035202	郑东	副教授
2003520206	强五	035202	郑东	副教授
2003520207	梁小三	035202	郑东	副教授
2003520208	梁小四	035202	郑东	副教授
2003520209	王小三	035202	郑东	副教授
2003520210	王小四	035202	郑东	副教授

班级	班主任	班主任职称
035101	张林	讲师
035202	郑东	副教授

学号	姓名	班级
2003510101	张三	035101
2003510102	李四	035101
2003510103	王五	035101
2003510104	李六	035101
2003510105	张四	035101
2003510106	张五	035101
2003510107	张小三	035101
2003510108	张小四	035101
2003510109	李小三	035101
2003510110	李小四	035101
2003520201	周三	035202
2003520202	赵四	035202
2003520203	赵五	035202
2003520204	赵六	035202
2003520205	钱四	035202
2003520206	强五	035202
2003520207	梁小三	035202
2003520208	梁小四	035202
2003520209	王小三	035202
2003520210	王小四	035202

## 13.4 数据库设计过程之逻辑数据库设计



### (3)不正确设计数据库会引发什么问题呢?

什么是规范的数据库设计

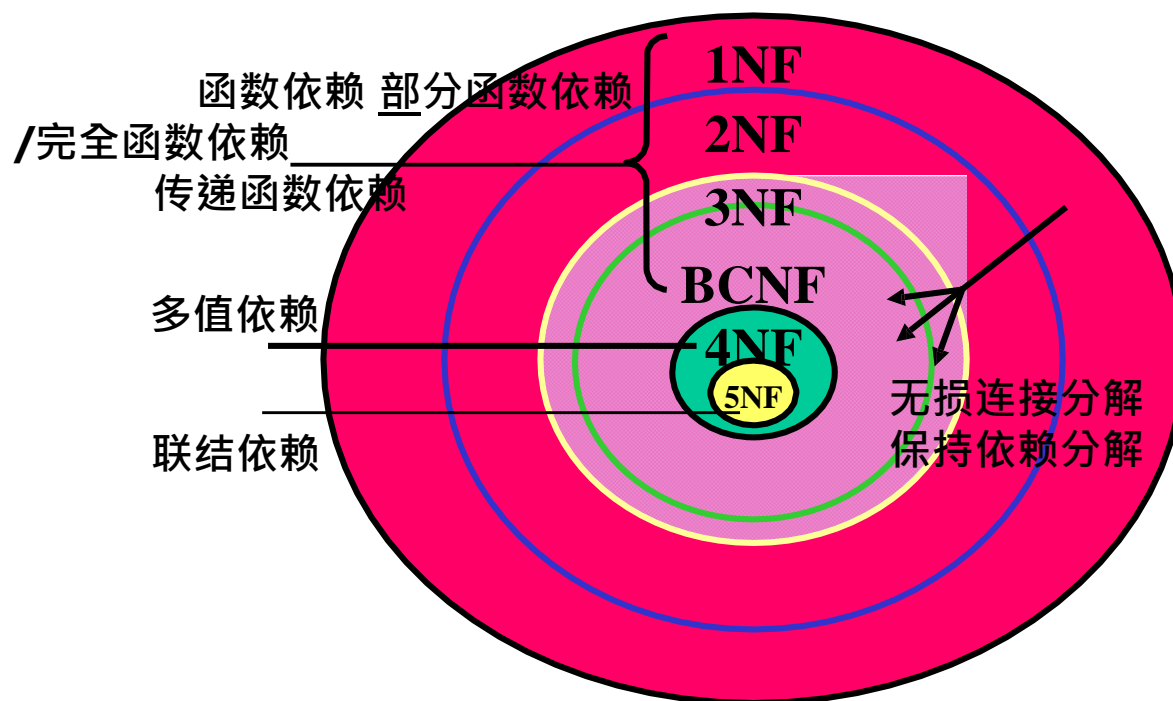
➤数据库的规范性设计需要分析数据库Table中的属性在取值方面有什么依存关系？数据库设计过程中应遵循什么样的原则

➤数据库设计理论

□数据依赖理论

□关系范式理论

□模式分解理论







## 第13讲 数据库设计过程

---

13.1 数据库设计过程概述

13.2 数据库设计过程之需求分析

13.3 数据库设计过程之概念数据库设计

13.4 数据库设计过程之逻辑数据库设计

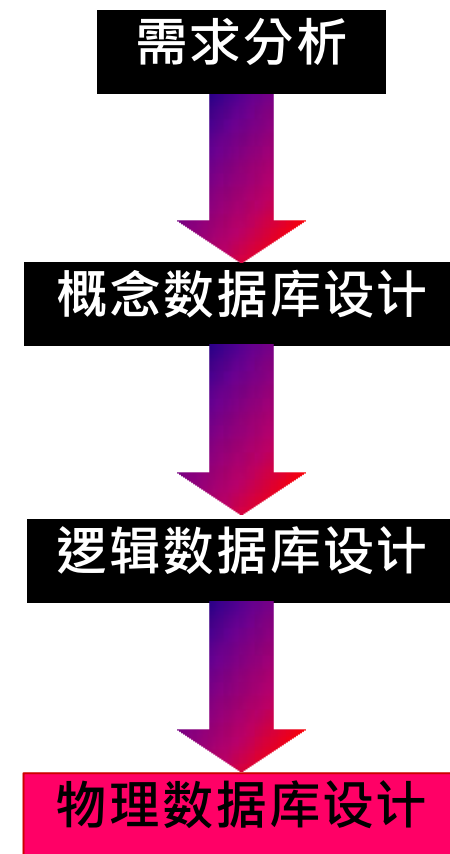
13.5 数据库设计过程之物理数据库设计

# 13.5 数据库设计过程之物理数据库设计



## (1)物理数据库设计

- 目标：结合指定**DBMS**物理数据库管理方法，给出概念数据库的物理模式描述。
- 提交物：物理数据库设计报告
- 使以下内容清楚：
  - ❑ **DBMS**选型
  - ❑ 确定数据库的存储结构，文件类型：如定长文件、不定长文件；堆文件、散列文件或**B-Tree**文件等
  - ❑ 用**Triggers**, 设计一些完整性控制约束
  - ❑ 确定数据库的高效访问方式(索引访问，直接访问... ..)
  - ❑ 评估和设置磁盘空间需求



## 13.5 数据库设计过程之物理数据库设计



### (1)物理数据库设计

- ❑ 设计用户视图及访问控制规则，以进行安全性控制
  - ❑ 建立索引
  - ❑ 设计使数据库运行达到最佳效率的一些措施
  - ❑ 设计备份**Backup**和恢复**Recovery**的步骤
- 理解**Oracle**、**Sybase**或其他**DBMS**的物理数据库管理方式，这是数据库管理员(**DBA**)的基本责任。

