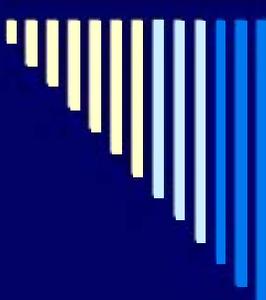


第四讲 数据库设计

- 数据库设计是信息系统的核心组成部分
- 从现实世界到数据世界的转换的过程

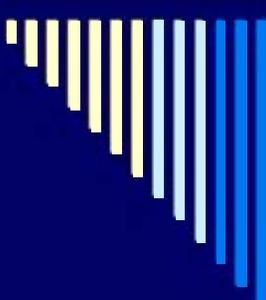




要点

- 数据模型
- E-R图
- 3个范式

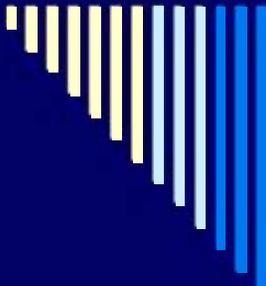




一、数据模型

- 模型是现实世界的特征和抽象。
- 数据模型是工具，是用来抽象、表示和处理现实世界中的数据 and 信息的工具。
- 数据模型应满足三个方面要求：
 - 能够比较真实地模拟现实世界
 - 容易被人理解
 - 便于在计算机系统中实现





三个世界的概念

1. 现实世界

- 现实世界的事物及事物之间存在着联系
- 事物本身的性质

2. 概念世界

- 客观事物及其联系的一种抽象描述
- 实体、属性和联系

3. 数据世界

- 将概念世界中的事物数据转化计算机中数据结构
- 数据库、表、记录



不同世界概念的对应关系

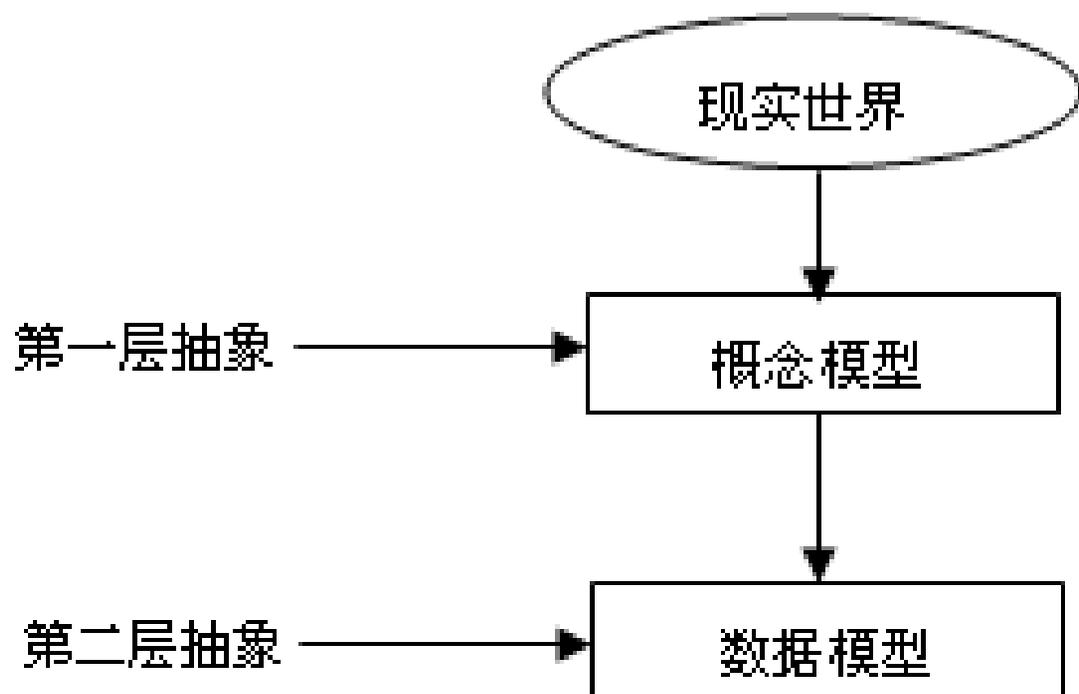
| 客观世界 | 概念世界 | 数据世界 |
|--------|-------|---------|
| 组织 | 实体与联系 | 数据库 |
| 事务类 | 实体 | 文件或表 |
| 事务（对象） | 实体个体 | 记录 |
| 特征（性质） | 属性 | 数据项（字段） |

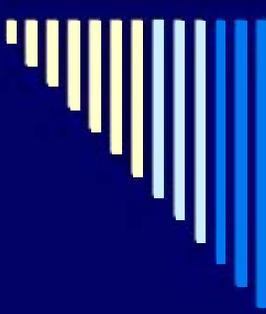


不同世界概念的对应关系（示例）

| 客观世界 | 概念世界 | 数据世界 |
|--------------|--------------|---------------|
| 组织：学校 | 实体与联系：学校ER图 | 数据库：学校数据库 |
| 事务类：学生，教师。。。 | 实体：学生实体、教师实体 | 文件或表：学生表、教师表 |
| 事务（对象）：某个学生 | 实体个体：某个学生个体 | 记录：某个学生记录 |
| 特征（性质）：学号、姓名 | 属性：学号、姓名 | 数据项（字段）：学号、姓名 |

数据抽象的基本过程

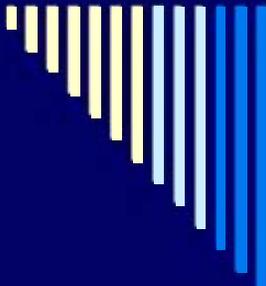




二、E-R方法 (Entity-Relationship Approach)

- 实体-联系方法：通过E-R图表示信息世界中实体、属性、关系的模型
- 反映客观世界事物、事物的特征以及事物之间的关系
- 导出关系数据模型。

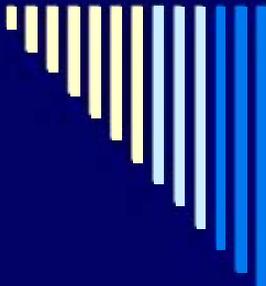




E-R信息模型的设计

- E-R图中三要素：实体、属性、联系
- 实体用方框表示
- 属性用椭圆框表示
- 联系用菱形框表示
- 框内填写相应的实体名、属性名、联系名





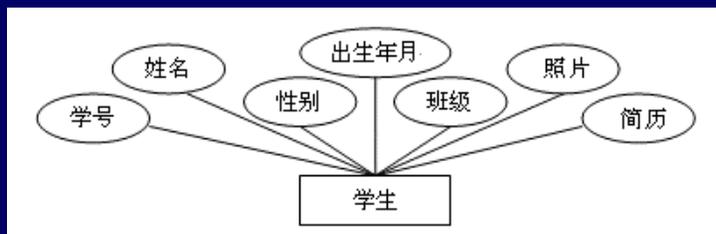
E-R方法

- 画单个的实体图（实体+属性）
- 确定实体之间的联系方式
- 联系方式（1：1、1：m、m：n）
- 如果m：n，通过插入实体的方法改成1：
m\n:1方式
- 画出综合的E-R图



实体例子:

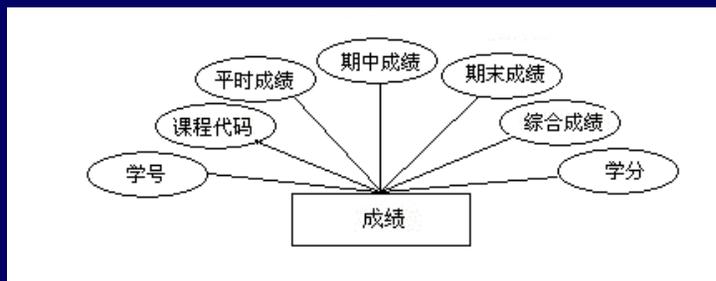
学生



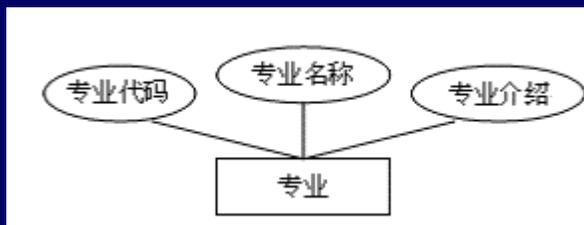
课程

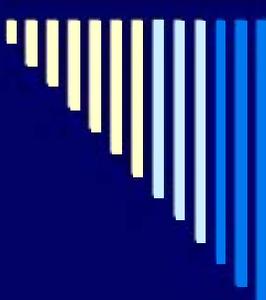


成绩



专业

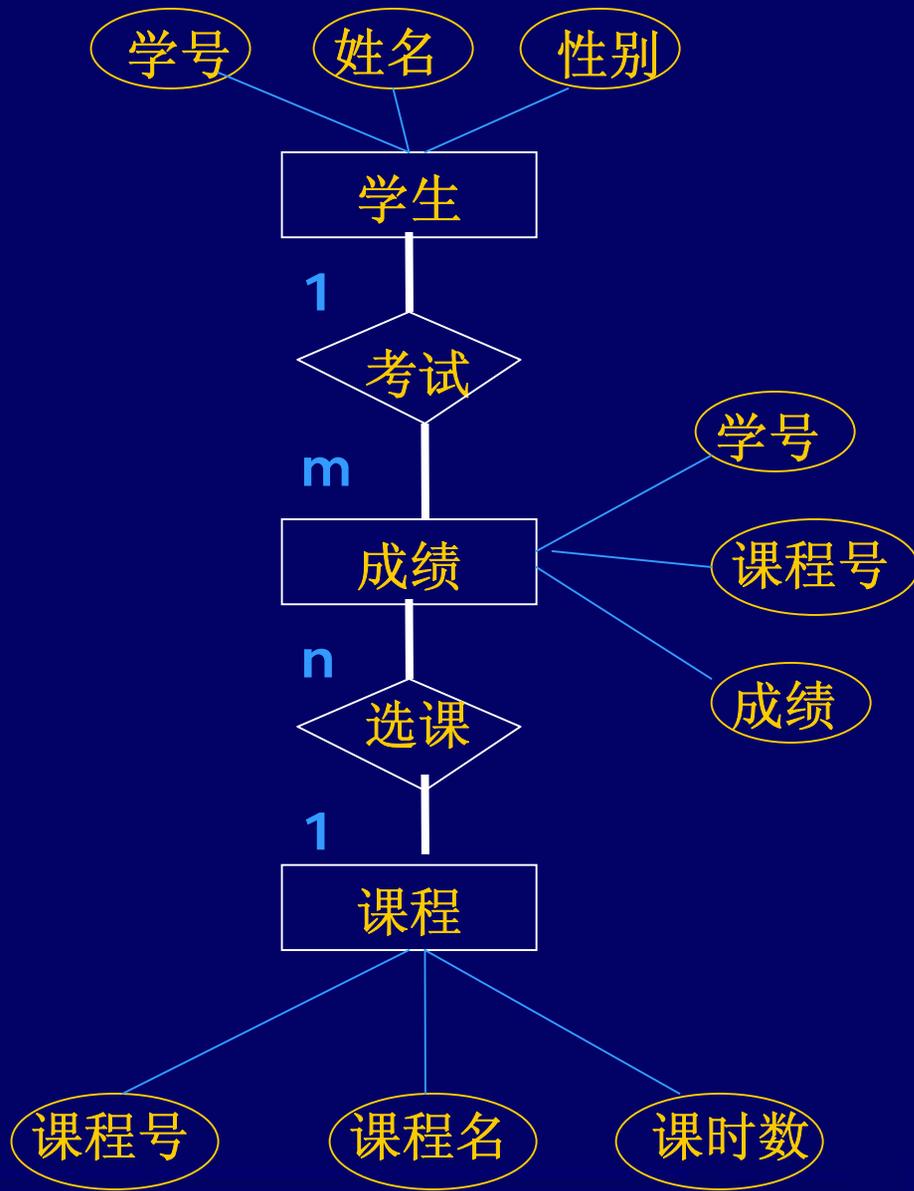
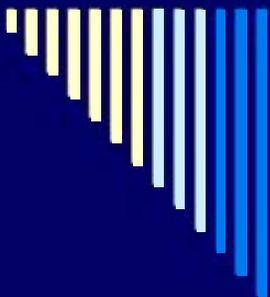


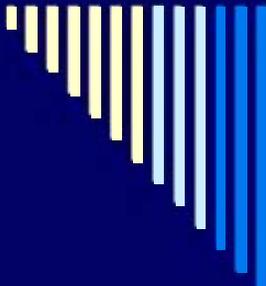


实体之间的联系

- 一对一 (1: 1)
 - 学生与学生家庭
- 一对多 (1: n)
 - 专业与学生
- 多对多 (N: m)
 - 课程与学生



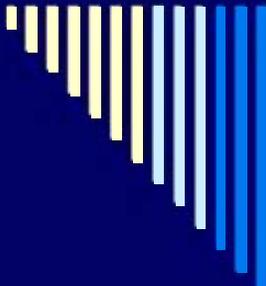




从E-R图导出关系数据模型

- E-R图中的每个实体转换成一个关系，并包括它的所有属性，同时确定关键字
- 实现实体之间的联系
- 如果实体间联系为1:n，“1”方的关键字作为“n”方的外部关键字

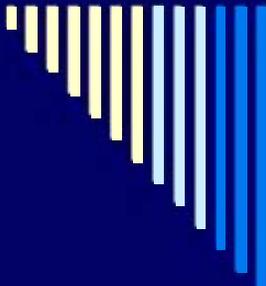




练习1

- 画出学生、班级、课程、教师、成绩的E-R图
- 画出书、出版社、作者的E-R图





三、范式（ Normal Form ）

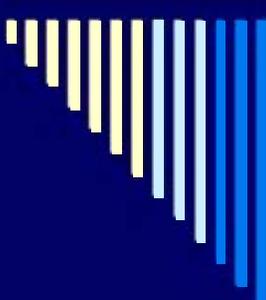
- 对表的数据结构进行规范，规范化的模式称为范式。



发票管理： 发票表的最初形式

- Invoice_no (主键PK)
- Date
- Company_no
- Company_name
- Company_addr
- Company_phone
- Customer_no
- Customer_name
- Customer_addr
- Customer_phone
- Item_no1
- Item_name1
- Item_qty1
- Item_price1
- Item_total1
- Item_no2
- Item_name2
- Item_qty2
- Item_price2
- Item_total2
-
- Item_no?
-





第一范式 (First Normal Form 1NF)

- 如果在一个表中，有一个主键，而且其他字段都是简单的不带重复属性的数据类型时，该表处于第一范式
- 发票的最初形式中，明细字段重复，不满足第一范式



符合第一范式的发票表的形式

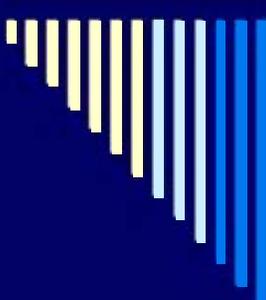
- Invoice_no (主键PK)
- Date
- Company_no
- Company_name
- Company_addr
- Company_phone
- Customer_no
- Customer_name
- Customer_addr
- Customer_phone

发票

- Invoice_no (主键PK)
- Item_no (主键PK)
- Item_name
- Item_qty
- Item_price
- Item_total

明细





第二范式 (Second Normal Form 2NF)

- 如果该表符合第一范式，并且所有非主键的字段都依赖于所有的主键时，该表处于第二范式
- 明细表中，只有item_qty依赖于所有的主键，而item_name等字段只依赖于部分主键（item_no）



符合第二范式的发票表的形式

- Invoice_no (主键PK)
- Date
- Company_no
- Company_name
- Company_addr
- Company_phone
- Customer_no
- Customer_name
- Customer_addr
- Customer_phone

发票

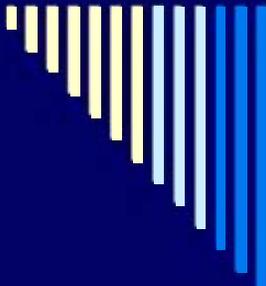
- Invoice_no (主键PK)
- Item_no (主键PK)
- Item_qty

明细

- Item_no (主键PK)
- Item_name
- Item_price

产品





第三范式 (Third Normal Form 3NF)

- 如果该表符合第二范式，而且不包含任何传递依赖时，该表处于第三范式
- 在发票表中，company_name等字段依赖于company_no，customer_name等字段依赖于customer_no

符合第三范式的发票表的形式

- Invoice_no (主键PK)
- Date
- Company_no
- Customer_no

发票

- Invoice_no (主键PK)
- Item_no (主键PK)
- Item_qty

明细

- Company_no (PK)
- Company_name
- Company_addr
- Company_phone

公司

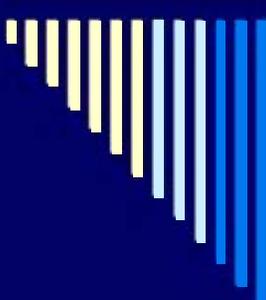
- Customer_no
- Customer_name
- Customer_addr
- Customer_phone

客户

- Item_no (主键PK)
- Item_name
- Item_price

产品



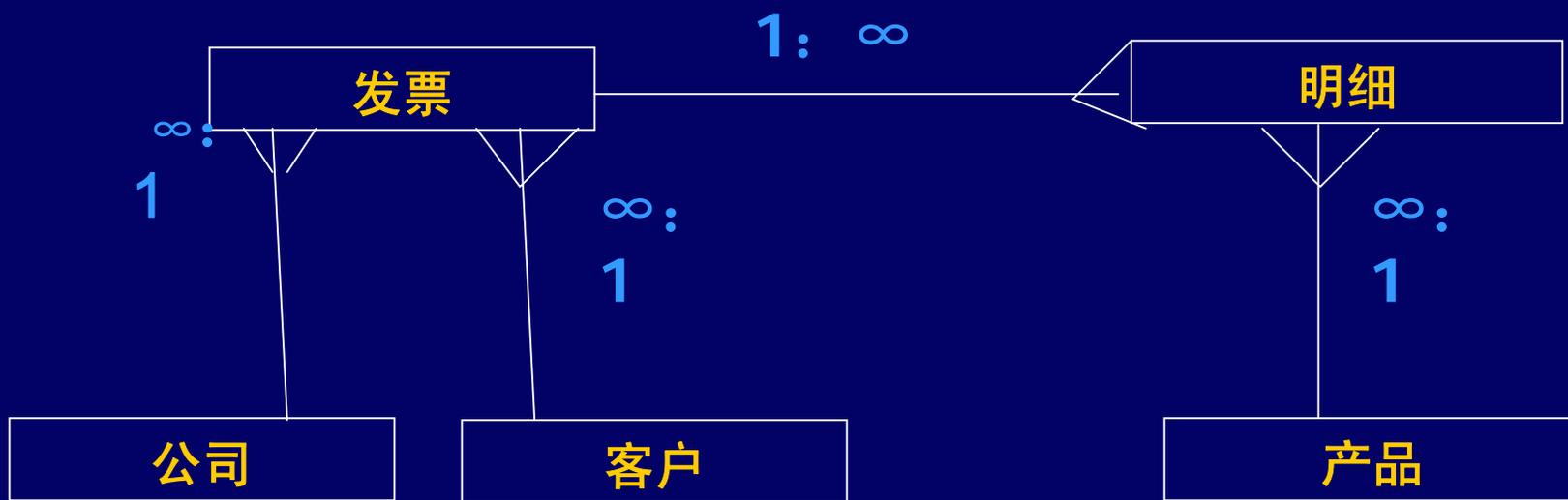


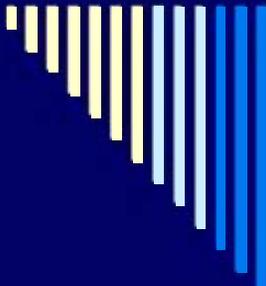
表之间的关系

- 一对一: 1: 1
- 一对多: 1: ∞
- 多对一: ∞ : 1
- 多对多: ∞ : ∞



由发票形成的表之间的关系





练习2

- 将练习1的E-R图设计成满足3rd NF的表
- 将毕业生成成绩单设计成满足3rd NF的表

