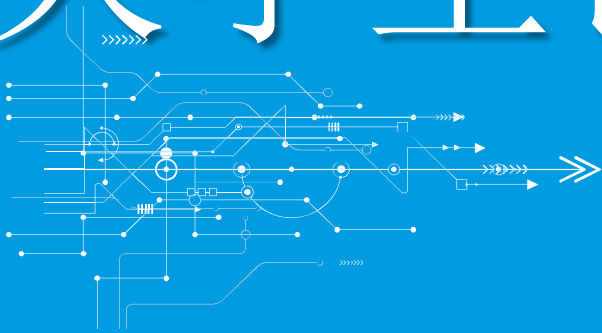




# 大学生计算与信息化素养

## 数据模型



# 数据模型



```
graph LR; A((数据模型)) --- B[客观对象的抽象过程]; A --- C[概念模型]; A --- D[数据模型]; A --- E[关系模型]; B --- F[ ]; C --- F; D --- F; E --- G[ ]
```

客观对象的抽象过程

概念模型

数据模型

关系模型

# 客观对象的抽象过程

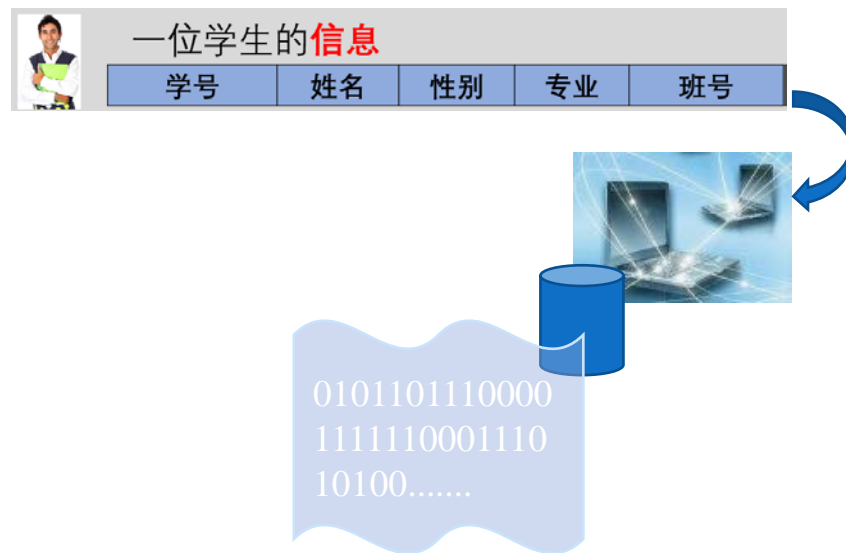


首先对客观事物及事物间的联系进行选择、命名、分类和抽象，并用一种形式化的语义和规则进行描述，即建立模型。

**数据模型**是对现实世界进行抽象和转换的结果。在数据库中用数据模型这个工具来抽象、表示和处理现实世界中的数据和信息。

 数据模型要满足三个要求：

- ✓ 能尽量真实地模拟现实世界
- ✓ 能被人所理解
- ✓ 易于在计算机上处理和存储



# 客观对象的抽象过程



现实世界

客观事物



认识抽象

信息世界

概念模型

用户观点

映射转换

机器世界

DBMS支持的数据模型

机器观点

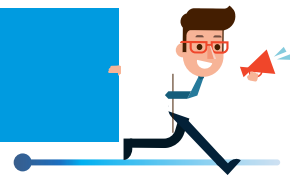


## 1 概念模型的含义

概念模型是按用户的观点将现实世界的信息进行模型化表示后形成的。

概念模型是对整个应用项目涉及的数据的全面描述，需要简洁、恰当地刻画要处理的对象所关注的现实世界中的信息，主要包括两个方面：

- ✓ 筛选出有价值的事物，并对其主要特性进行描述
- ✓ 描述这些事物之间的联系



## 2 概念模型的特点

- ✓ 它是独立于DBMS的数据模型，与具体的计算机平台无关
- ✓ 它具有较强的语义表达能力，能够方便、直接地表达应用中的各种语义知识
- ✓ 它可以辅助数据库设计人员与用户沟通数据的需求
- ✓ 它具有简单、清晰的结构，易于用户理解

通常采用的概念模型工具是实体联系模型。



## 3 实体联系模型（又称E-R模型）

实体联系模型（Entity-Relationship Model）是 Peter Chen 于 1976 年首先提出。它属于一种概念化的模型。由于 E-R 模型提供了不受任何 DBMS 约束的、面向用户的表达方法，因此在数据库概念设计阶段常被用来进行数据建模。

E-R 模型能够将现实世界中的问题对象转化成实体、属性、联系等几个基本元素，并用一种较为简单的图形表示，即 **E-R 图**。



## 1 实体 (Entity)

客观存在并可相互区别的事物称为实体。

实体可以是实际的事物，如一个学生、一个课程 ……

也可以是具体的事件，如一场比赛、一个展销会 ……



学生



课程



比赛



展销会





## 2 属性 (Attribute)

实体所具有的某一特性称为属性。一个实体可以由若干个属性来刻画其特征。

属性一般由**属性名**和**属性值**组成。

例如，学生实体 —— 学号、姓名、性别、专业 等属性

属性的值：（180834125，张海，男，计算机）  
表示一个名字为“张海”的学生实体。



学生

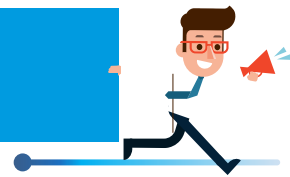


## 3 实体型 (Entity Type)

用实体名及其属性名的集合来抽象和刻画同类实体，称为实体型。

例如，“学生”作为实体名，用小括号括起其全部的属性名来表示一个学生实体型：

学生（学号，姓名，性别，专业）



## 4 实体集 (Entity Set)

同一类型实体的集合称为实体集，即具有同一类属性的事物的集合。例如，一个班的学生在上述属性值的集合就是学生实体的一个实体集。



## 5 关键字（Keyword）

如果某个属性或属性组合能够唯一地标识出实体集中的各个实体，这个属性或属性组合就称为**关键字**，或称为**码**。

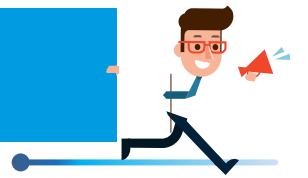
例如，学生的学号可以作为学生实体的关键字，但学生的姓名（可能会有重名）则不一定适宜作学生实体的关键字。



## 6 域 (Domain)

属性的值的取值范围称为该属性的域。

例如：学生的成绩如果作为学生的属性，那么“成绩”属性的域设定为 $[0, 100]$ 。



## 7 联系 (Relationship)

现实世界的事物内部和事物之间都存在关联，这也可以被抽象和反映，并在概念模型中表示出来——**联系**。

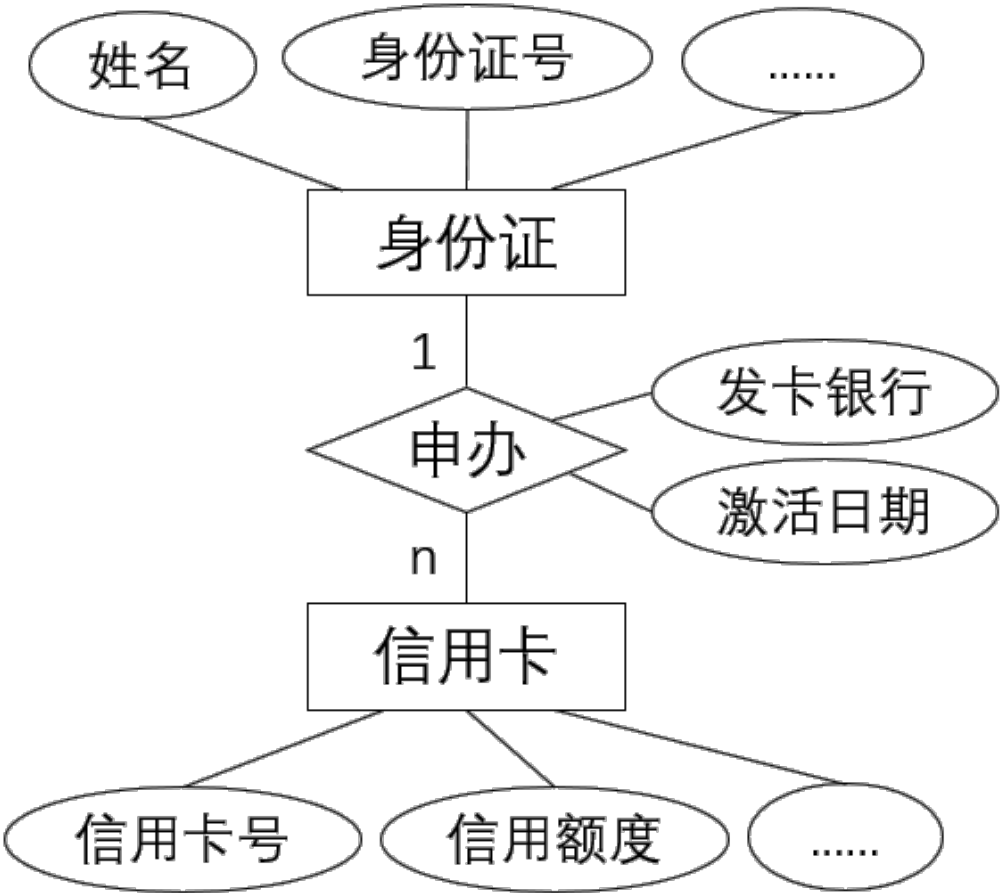
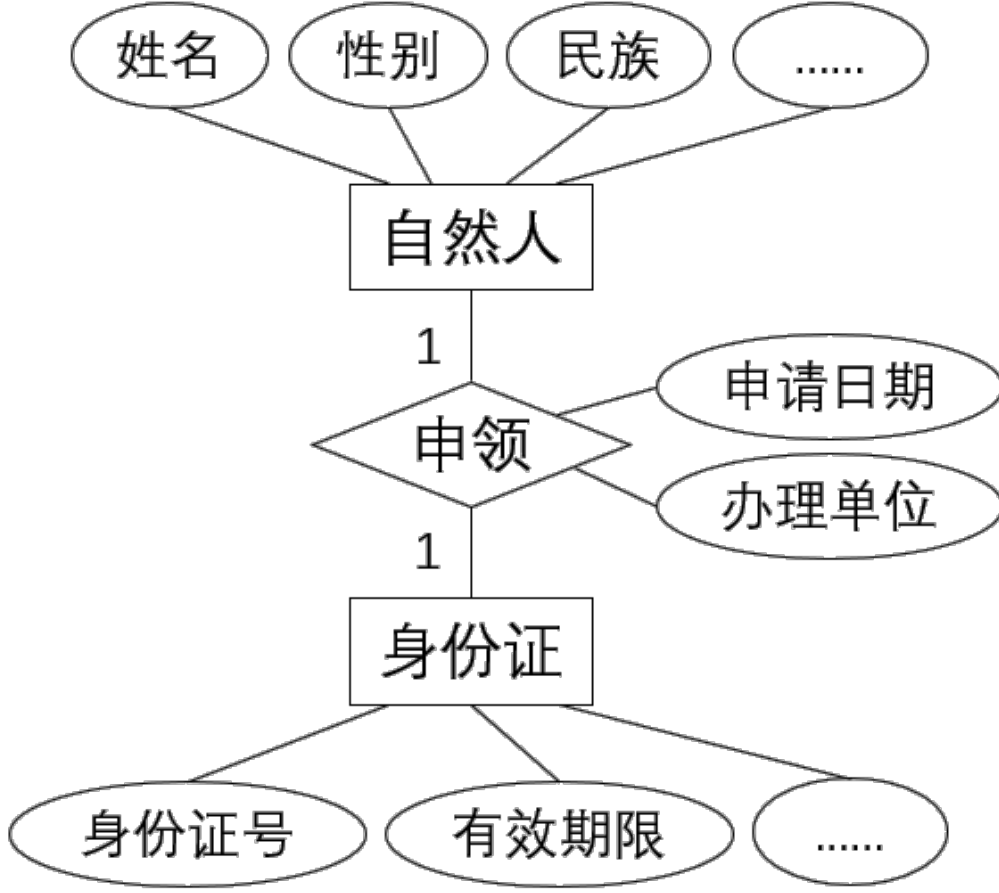
事物内部的关联属于实体内部的联系，是指组成实体的属性之间的联系。事物之间的关联属于实体（集）之间的外在的联系，这种联系归纳为三种类型：

- 一对一联系 (1:1)
- 一对多联系 (1:n)
- 多对多联系 (n:m)

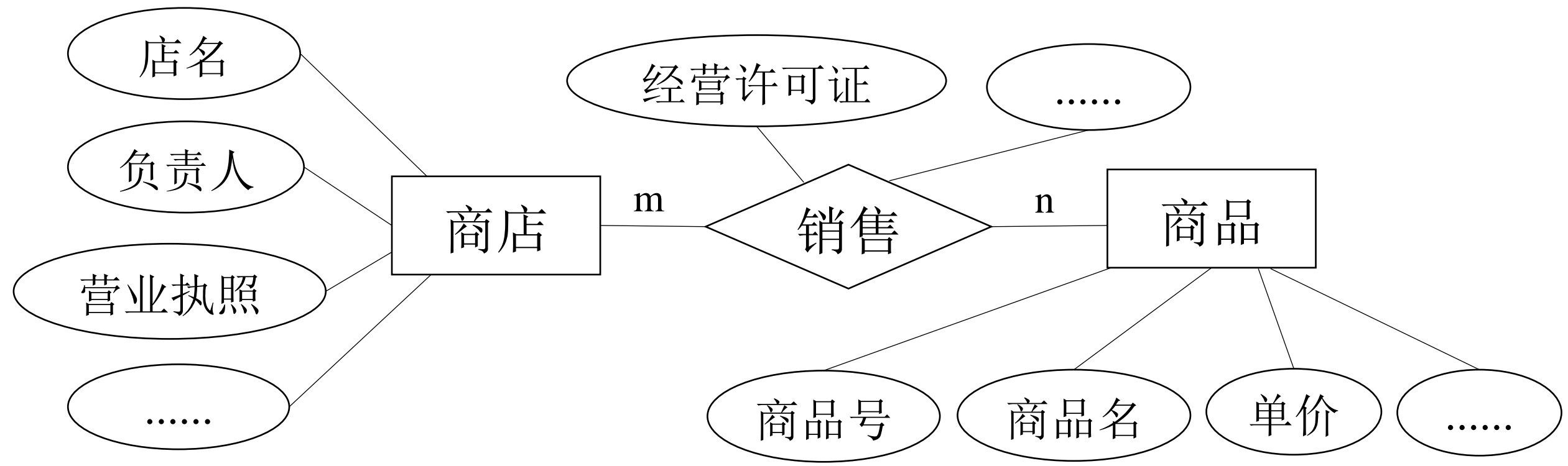


E-R图是一种直观反映现实世界的实体及其联系的模型工具，是体现概念模型的一种直观图。直观图的构成元素如下：

- 矩形框——表现**实体**
- 椭圆——表示实体的**属性**
- 无向线段——起连接作用
- 菱形框——表示**联系**





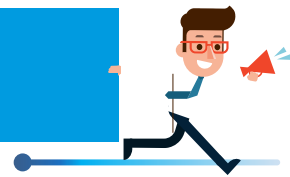




数据模型是信息世界到机器世界的抽象，将信息世界中的概念模型进一步抽象，转换成适于计算机处理的模型形式，也被称为逻辑模型。

逻辑模型（Logical Model）是按计算机实现的观点对数据进行建模，主要服务于具体的DBMS的实现过程。

数据库中的数据存储由DBMS完成，并自动实现逻辑模型向物理模型的转换。物理模型（Physical Model）是对数据最底层的抽象，它描述数据在计算机系统内部的表示方式和存取方法。



## 逻辑模型的构成要素有三个：

- ✓ 数据结构
- ✓ 定义在数据结构上的操作
- ✓ 数据约束



## 1 数据结构

是逻辑模型中最重要的部分，描述了逻辑模型的静态特征，它描述处理对象的类型、性质、内容，以及数据对象之间的联系。常用的3种数据结构有：网状结构、层次结构和关系结构。

## 2 数据结构上的操作

是描述数据模型中允许各种数据对象所做操作的集合，如查询、删除、更改等等，以及相关的操作规则。它描述了逻辑模型的动态特征。



### 3 数据约束

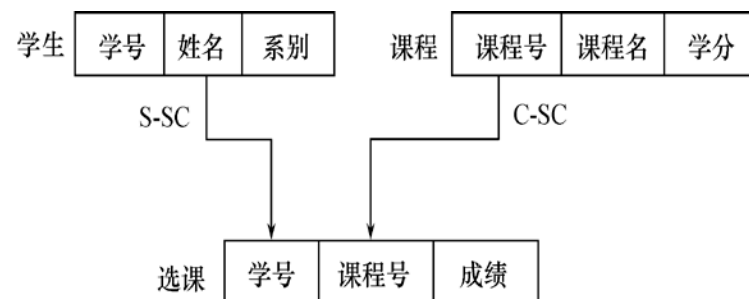
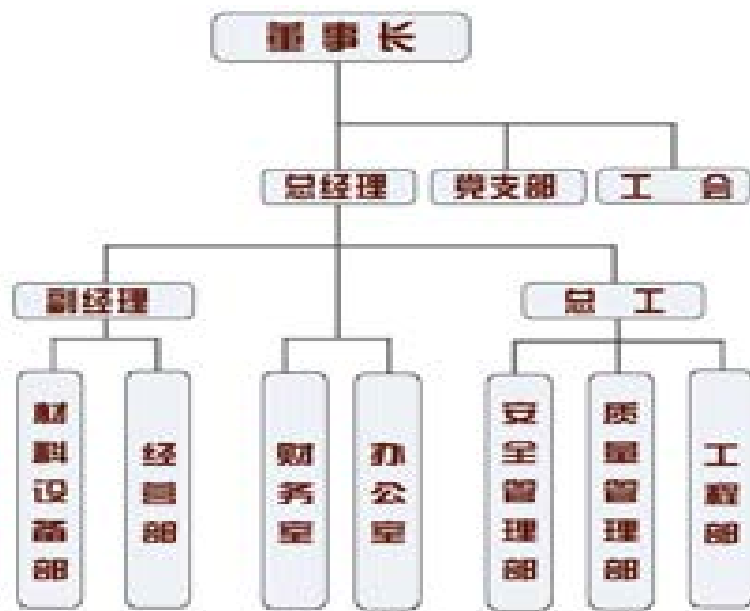
是对数据模型中数据取值或数据取值变化时的制约关系的描述。它是一组数据完整性约束规则的集合。能确保数据动态变化时，数据库依然保持数据的正确性、有效性和相容性。



## 逻辑模型的种类：

目前成熟的应用于DBMS中的逻辑模型有三种：

层次模型、网状模型和关系模型。



学号	姓名	成绩
160224105	洪云	82
160224206	张心语	71

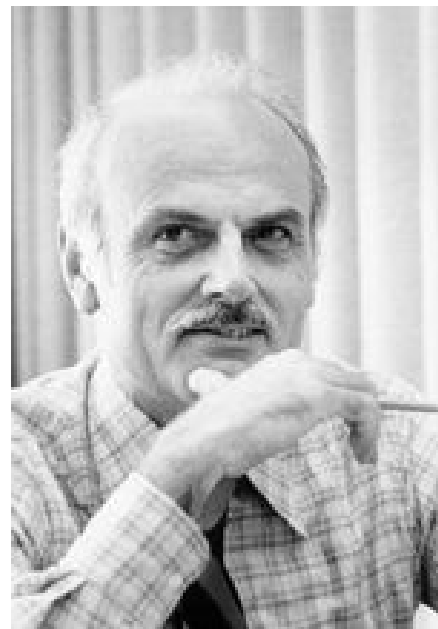


关系模型是逻辑模型种类中的一个，它采用简单易懂的二维表来表示数据间的联系，是构造关系型数据库的模型。

关系模型的思想是IBM公司的E.F.Codd于1970年提出。他于1981年获得图灵奖。

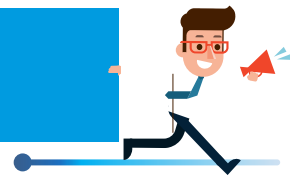
## 关系模型的优点

- ✓ 理念成熟
- ✓ 结构简单
- ✓ 使用方便





## 小结



现实世界的客观事物转变为计算机世界里数据库存储的数据是一个逐步抽象的过程，首先将现实世界的客观事物抽象为信息世界的概念模型，再将概念模型抽象转换为逻辑模型（数据模型），逻辑模型再次转换为物理模型，实现信息世界转化为机器世界。

概念模型的工具是E-R图。

逻辑模型种类中应用比较广泛的是关系模型，它是构造关系型数据库Access的模型。

下一节，我们将详细介绍关系模型的数据结构。