数据库的几个基本概念，数据，对象，实体，属性，数据模型

如何将实体抽象成数据模型这类的

人们把客观存在的事物以数据的形式存储到计算机中，经历了对现实e5a48de588b662616964757a686964616f31333233666230生活中事物特性的认识、概念化到计算机数据库里的具体表示的逐级抽象过程，即现实世界－概念世界－机器世界三个领域。有时也将概念世界称为信息世界；将机器世界称为存储或数据世界。  
一、三个世界  
1、现实世界 人们管理的对象存于现实世界中。现实世界的事物及事物之间存在着联系，这种联系是客观存在的，是由事物本身的性质决定的。例如学校的教学系统中有教师、学生、课程，教师为学生授课，学生选修课程并取得成绩。  
2、概念世界 概念世界是现实世界在人们头脑中的反映，是对客观事物及其联系的一种抽象描述，从而产生概念模型。概念模型是现实世界到机器世界必然经过的中间层次。涉及到下面几个术语：  
实体：我们把客观存在并且可以相互区别的事物称为实体。实体可以是实际事物，也可以是抽象事件。如一个职工、一场比赛等。  
实体集：同一类实体的集合称为实体集。如全体职工。注意区分"型"与"值"的概念。如每个职工是职工实体"型"的一个具体"值"。  
属性：描述实体的特性称为属性。如职工的职工号，姓名，性别，出生日期，职称等。  
关键字：如果某个属性或属性组合的值能唯一地标识出实体集中的每一个实体，可以选作关键字。用作标识的关键字，也称为码。如"职工号"就可作为关键字。  
联系：实体集之间的对应关系称为联系，它反映现实世界事物之间的相互关联。联系分为两种，一种是实体内部各属性之间的联系。另一种是实体之间的联系。  
3、机器世界 存入计算机系统里的数据是将概念世界中的事物数据化的结果。为了准确地反映事物本身及事物之间的各种联系，数据库中的数据必须有一定的结构，这种结构用数据模型来表示。数据模型将概念世界中的实体，及实体间的联系进一步抽象成便于计算机处理的方式。  
二、E－R模型  
E－R模型（实体联系模型）简称E－R图。它是描述概念世界，建立概念模型的实用工具。E－R图包括三个要素：  
实体（型）――用矩形框表示，框内标注实体名称。  
属性――用椭圆形表示，并用连线与实体连接起来。  
实体之间的联系――用菱形框表示，框内标注联系名称，并用连线将菱形框分别与有关实体相连，并在连线上注明联系类型。  
联系归结为三种类型：  
1）一对一联系（1:1）  
设A、B为两个实体集。若A中的每个实体至多和B中的一个实体有联系，反过来，B中的每个实体至多和A中的一个实体有联系，称A对B或B对A是1:1联系。注意，1:1联系不一定都是一一对应的关系。可能存在着无对应。如一个公司只有一个总经理，一个总经理不能同时在其它公司再兼任总经理，某公司的总经理也可能暂缺。  
2）一对多联系（1:n）  
如果A实体集中的每个实体可以和B中的几个实体有联系，而B中的每个实体至我和A中的一个实体有联系，那么A对B属于1:n联系。如一个部门有多名职工，而一名职工只在一个部门就职，部门与职工属于一对多的联系。  
3）多对多联系（m:n）  
若实体集A中的每个实体可与和B中的多个实体有联系，反过来，B中的每个实体也可以与A中的多个实体有联系，称A对B或B对A是m:n联系。如一个学生可以选修多门课程，一门课程由多个学生选修，学生和课程间存在多对多的联系。  
必须强调指出,有时联系也有属性,这类属性不属于任一实体只能属于联系。  
三、数据模型简介  
数据模型由三部分组成，即模型结构、数据操作和完整性规则。这里主要介绍模型结构。DBMS所支持的数据模型分为四种：层次模型、网状模型、关系模型、面向对象模型。其中第四种数据模型目前并未成熟，因此传统的说法有前三种数据模型。  
1、基本层次联系  
层次模型和网状模型有共同点，可以把它们统称为格式化模型。基本层次联系就是包括两结点和一边的基本单元，两个实体间的有向边代表的基本层次联系表示一对多的联系。通常把表示1的实体放在上方,称为父结点，把表示N的实体放在下面，称为子结点。  
2、层次数据模型  
用村形结构表示实体及其之间的联系的模型称为层次模型。该模型的实际存储数据由链接指针来体现联系。特点：有且仅有一个结点无父结点，此结点即为根结点；其它结点有且仅有一个父结点。适合用表示一对多的联系。  
3、网状模型  
用网状结构表示实体及其之间的联系的模型称为网状模型。允许结点有多于一个的父结点，可以有一个以上的结点无父结点。适合用于表示多对多的联系。  
层次模型和网状模型从本质上都是一样的。存在的缺陷：难以实现系统扩充，插入或删除数据时，涉及到大量链接指针的调整。