



大学计算机基础

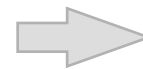


教学课件

北京航空航天大学

第8章 数据库技术基础

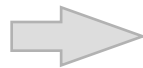
8.1 数据库系统概述



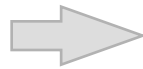
8.2 数据模型



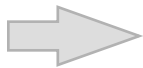
8.3 数据库系统设计



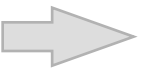
8.4 SQL语言中的数据操作



8.5 简朴数据库应用实例



8.6 Access数据库的建立与维护



本章要点

- 数据库系统的发展和构成
- 数据库系统的基本知识、术语
- 数据模型的概念及E-R模型的使用
- 数据库系统的设计措施与环节
- Access数据库的建立与维护，涉及表、查询、窗体及报表的建立

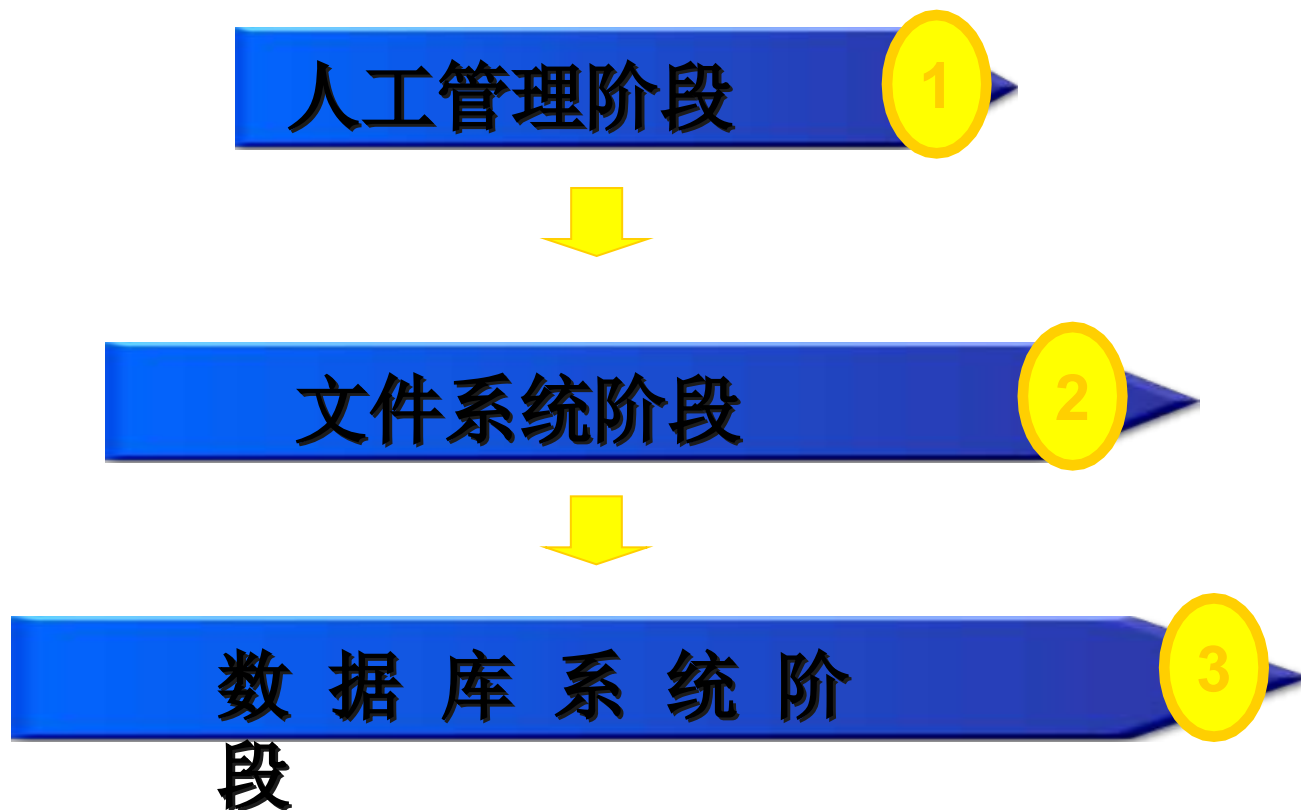


8.1 数据库系统概述

- ◆ 8.1.1 数据库技术的发展阶段
- ◆ 8.1.2 数据库系统的构成
- ◆ 8.1.3 数据库系统的分类

8.1.1 数据库技术的发展阶段

- “**数据管理**”是指怎样对数据进行分类、组织、编码、储存、检索和维护，数据管理技术的发展经历了三个阶段。



1. 人工管理阶段

- 时间：20世纪50年代中期此前
- 背景：
 - 计算机主要用于科学计算。
 - 外存只有磁带、卡片、纸带等，没有磁盘等直接存取设备。
 - 没有操作系统，没有数据管理软件（顾客用机器指令编码）。



人工管理阶段特点

- 数据不在计算机上保存。
- 没有软件系统对数据进行管理。程序规定数据的逻辑构造与物理构造。数据与程序不具有独立性。
- 基本没有文件概念，数据组织方式必须由程序员自行设计。
- 一组数据相应一种程序，数据是面对应用的，程序间不能共享数据。

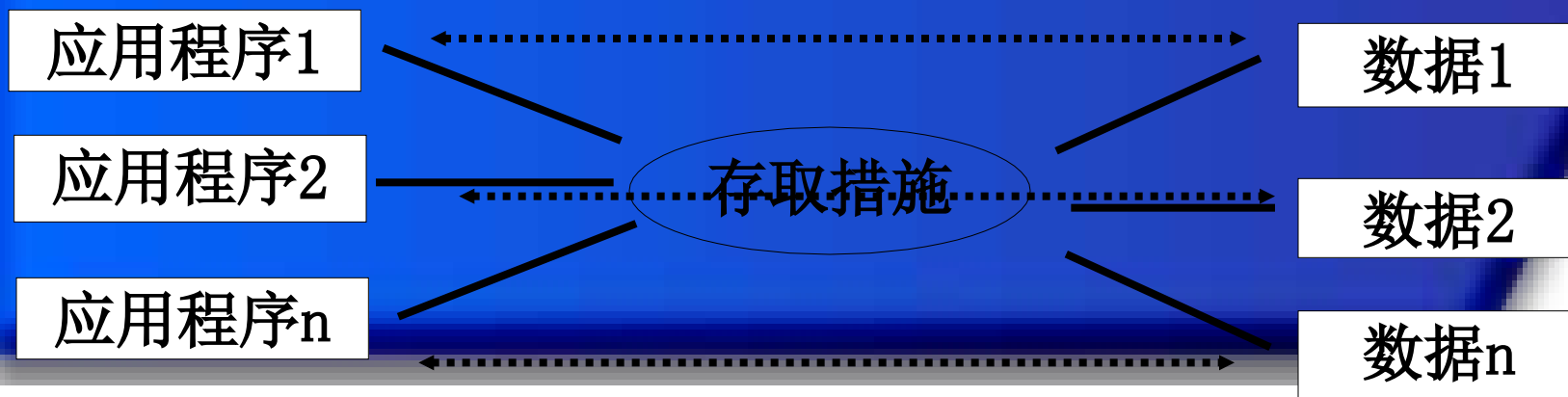


2. 文件系统阶段

- 时间：20世纪50年代后期到60年代中期

- 背景：

- 计算机不但用于科学计算，还用于管理。
- 外存有了磁盘、磁鼓等直接存取设备。
- 有了专门管理数据的软件，一般称为文件系统，包括在操作系统中。



文件系统阶段特点

- 数据以文件形式保存在外存上。
- 程序和数据有一定的独立性。
- 文件多样化。
- 数据的存取基本上以统计为单位。
- 缺陷：
 - 数据冗余度大：挥霍空间并易造成数据的不一致性。
 - 数据和程序缺乏独立性（逻辑独立性）
- 文件系统是不具弹性的无构造的数据集合，数据之间是孤立的，不能反应现实世界事物之间的内在联络。

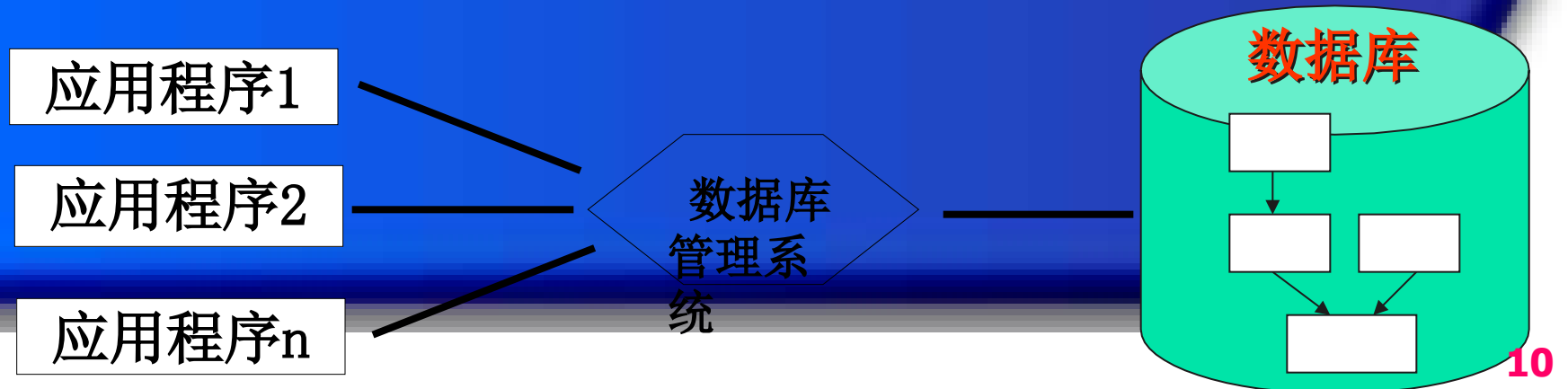


3. 数据库系统阶段

●时间：20世纪60年代后期开始

●背景：

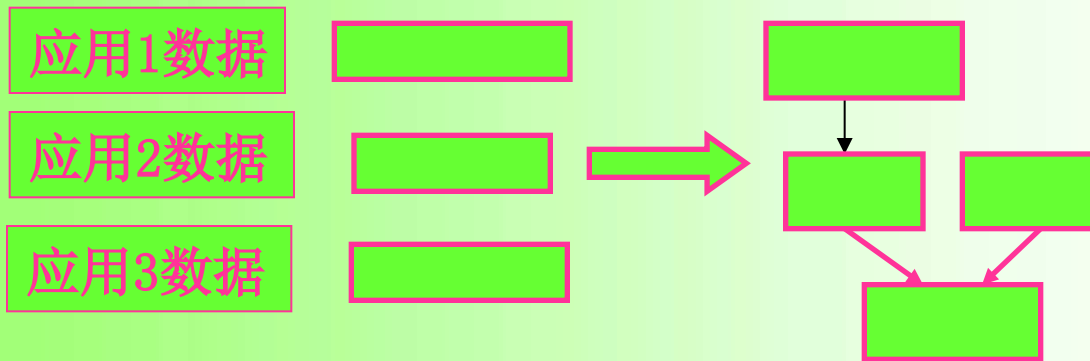
- 计算机管理的数据量大，关系复杂，共享性要求强（多种应用、不同语言共享数据）。
- 外存有大量容量磁盘，光盘。
- 软件价格上升，硬件价格下降，编制和维护软件及应用程序成本相对增长，其中维护的成本更高。



数据库系统数据管理特点 (1/4)

• 面对全组织的复杂的数据构造

• 在描述数据时，不但描述数据本身，还要描述数据之间的联络，使整个组织的数据构造化。



• 数据构造化是数据库主要特征之一，是数据库与文件系统的根本区别。

数据库系统数据管理特点 (2/4)

• 数据冗余度小，易扩充

• 数据库从整体观点描述数据。数据不再面对某个应用，而是面对整个系统，从而大大减小数据的冗余度；

• 数据库数据的应用能够有很灵活的方式，能够取整体数据的多种合理子集用于不同的应用系统，并能够根据应用需求的变化，重新选用不同子集。

数据库系统数据管理特点 (3/4)

• 具有较高的数据和程序的独立性

• 数据独立性

• **数据的物理独立性**: 数据的存储构造 (物理构造) 变化时, 数据的逻辑构造能够不变, 从而应用程序也不必变化;

• **数据的逻辑独立性**: 数据的逻辑构造变化时, 应用程序能够不变;

• 数据库系统提供了两方面的映象 (转换) 功能:

• 数据的存储构造与逻辑构造之间——实现数据的物理独立性

• 数据的总体逻辑构造与某类应用所涉及的局部逻辑构造之间——实现数据的逻辑独立性

数据库系统数据管理特点 (4/4)

• 统一的数据控制功能

• 数据的安全性控制

• 保护数据以预防不正当的使用所造成数据的泄密和破坏。

• 数据的完整性控制

• 指数据的正确性、有效性与相容性。

• 并发控制

• 对多顾客的并发操作进行控制、协调，保护数据的完整性。

• 数据库恢复

• 将数据库从错误状态恢复到某一已知的正确状态



8.1.2 数据库系统的构成

- 数据、数据库、数据库系统、数据库管理系统是与数据库技术亲密有关的基本概念。

1. 数据 (date)

数据是数据库中存储的基本对象。

数据

数字、文字、图形、图象、声音、学生的档案统计等等

将现实世界中的多种信息统计下来的能够辨认的符号；

日常生活

自然语言

计算机

抽象出对这些事物感爱好的特征构成统计

例

(李明, 男, 1972, 江苏, 计算机系, 1990)

注意

数据与其语义的关系 15

2. 数据库 (Database)

简称DB

保存数据：人工方式、文件方式、**数据库方式**

是存储数据的最佳场合

数据库

- ❖ 是长久储存在计算机内、有**组织**、可**共享**的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的**冗余度**，较高的**数据独立性和易扩展性**，并可为多种顾客**共享**。



3. 数据库管理系统

简称DBMS

数据库管理系统

数据库管理系统 (Database Management System) 负责数据库建立, 数据存取、维护和管理**的软件**系统。DBMS是数据库系统的**关键**软件, 在操作系统支持下工作。

如

ORACLE、SYBASE、INFORMIX、DB2、INGRES

ACCESS、FOXPRO、SQLSERVER等等

4. 数据库系统 (1/3)

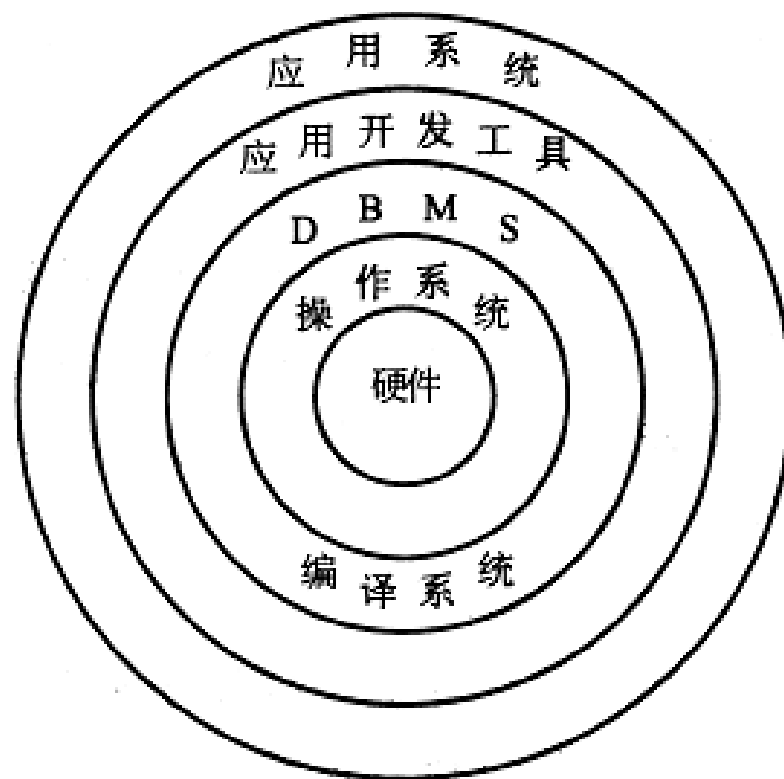
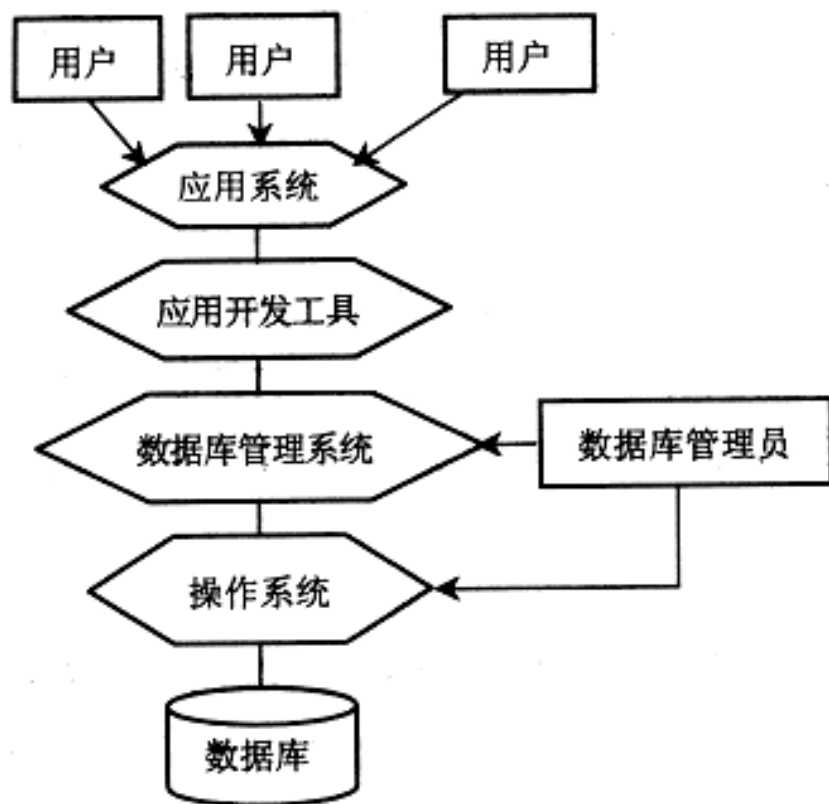
简称DBS

数据库系统 (Database System) ——是指在计算机系统中引入数据库后的系统构成, 一般由**数据库**、**数据库管理系统** (及其开发工具)、**应用系统**、**数据库管理员**和**顾客**构成。应该指出的是, 数据库的建立、使用和维护等工作只靠一种DBMS远远不够, 还要有专门的人员来完成, 这些人称为**数据库管理员** (Database Administrator, 简称DBA)。

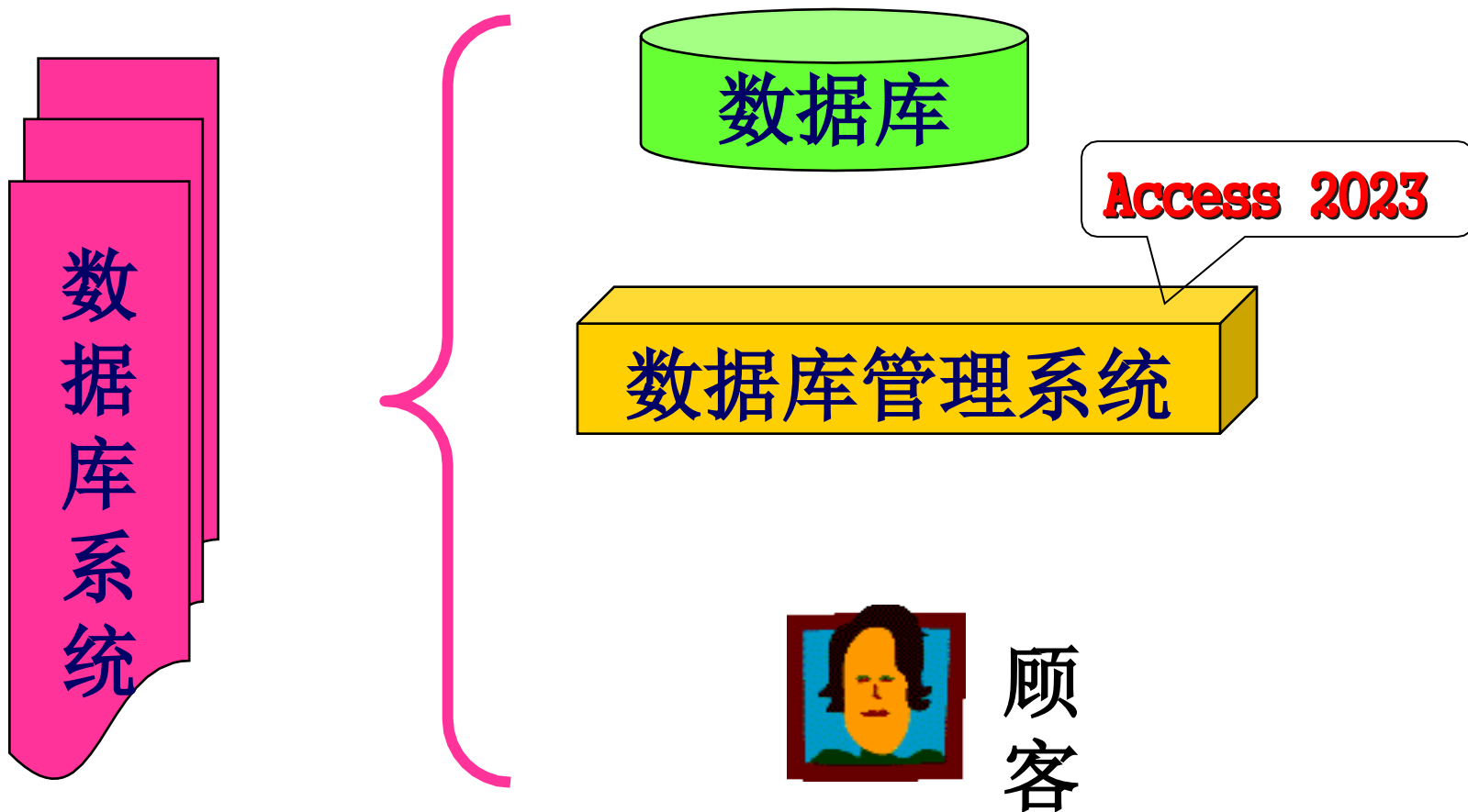
在不引起混同的情况下人们经常把数据库系统简称为数据库。



4. 数据库系统 (2/3)



4. 数据库系统 (3/3)



8.1.3 数据库系统的分类 (1 / 3)

■ 1. 单顾客数据库系统

- 单顾客数据库系统是一种早期的最简朴的数据库系统。在此类系统中，整个数据库系统，涉及DBMS、数据库及应用程序，都装在一台计算机上，由一种顾客独占，不同机器之间不能共享数据。

■ 2. 集中式构造的数据库系统

- 集中式构造，又称为主从式构造，它一般是以一台大型的中心计算机为中心，经过局部网络或区域网络相联的大量终端都能够访问该中心计算机。这种系统的优点是实现中心管理，安全性好，但其缺陷是伴随数据量的增长，系统相当庞大，操作复杂，开销大，因为数据集中存储，大量的通信都要经过主机，造成拥挤。



8.1.3 数据库系统的分类 (2 / 3)

■ 3. 分布式构造的数据库系统

- 分布式数据库系统一般使用较小的计算机系统，每台计算机可单独放在一种地方，每台计算机中都有DBMS的一份完整拷贝副本，并具有自己局部的数据库，位于不同地点的许多计算机经过网络相互连接，共同构成一种完整的、全局的大型数据库。
- 此构造克服了集中式构造的弱点，降低了数据传送代价，系统的可靠性提升了诸多，而且便于系统的扩充，使得增长一种新的局部数据库，或在某个位置扩充一台合适的小型计算机，都很轻易实现。缺陷是实现前面所提到的功能要付出相当高的代价；且因为数据的分布存储，给数据的处理、管理与维护带来困难；当顾客需要经常访问远程数据时，系统效率会明显地受到网络流量的制约。



8.1.3 数据库系统的分类（3 / 3）

4. 客户 / 服务器构造（C/S构造）的数据库系统

- 客户 / 服务器构造把DBMS功能和应用分开，网络中某个结点上的计算机专门用于执行DBMS功能，称为数据库服务器，其他结点上的计算机则只安装DBMS的外围应用开发工具，支持顾客的应用，称为客户机。
- 客户端的顾客祈求被传送到数据库服务器，数据库服务器进行处理后，只将成果返回给顾客（而不是整个数据），从而明显降低了网络上的数据传播量，提升了系统的性能、吞吐量和负载能力。
- 优点是数据库愈加开放。客户与服务器一般都能在多种不同的硬件和软件平台上运营，能够使用不同厂商的数据库应用开发工具，应用程序具有更强的可移植性，同步也能够降低软件维护开销。其缺陷与分布式构造相同，即数据的处理、管理与维护相对比较困难。





8.2 数据模型

- ◆ 8.2.1 数据模型的基本概念
- ◆ 8.2.2 概念模型
- ◆ 8.2.3 构造化数据模型

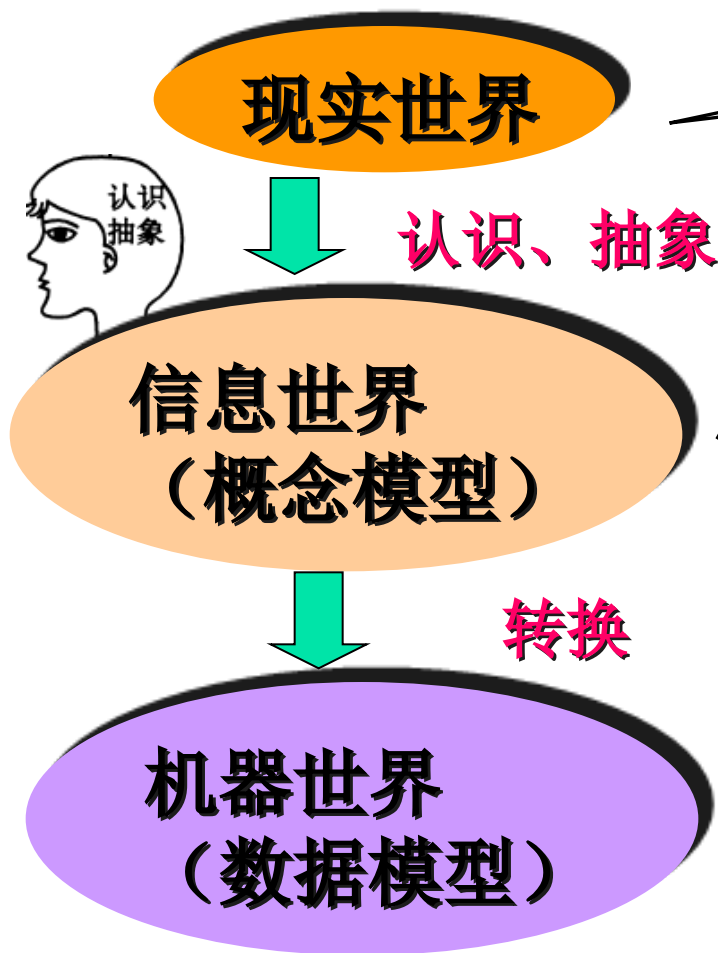
8.2.1 数据模型的基本概念

- 数据模型是一种形式化地描述数据及数据之间的联络措施，是数据库的构造框架，是数据库构造的基础型。一般我们说数据模型有三个要素，即数据构造、数据操作及数据完整性约束条件。
- 数据模型就是对现实世界的模拟。不同的使用对象、不同的应用目的会使用不同的数据模型。
- 在数据库中，根据模型应用的不同目的，将数据模型分为两个层次：概念模型和构造化数据模型。
- 为了把客观存在的复杂的大千世界中的问题，用计算机去处理，即把详细的事物抽象、组织成某一种**DBMS**支持的数据模型，数据库理论中引出了三个世界假设的概念，第一种世界是现实世界，第二个世界是信息世界，第三个世界是机器世界。



8.2.1 数据模型的基本概念 (1 / 3)

三个世界的假设



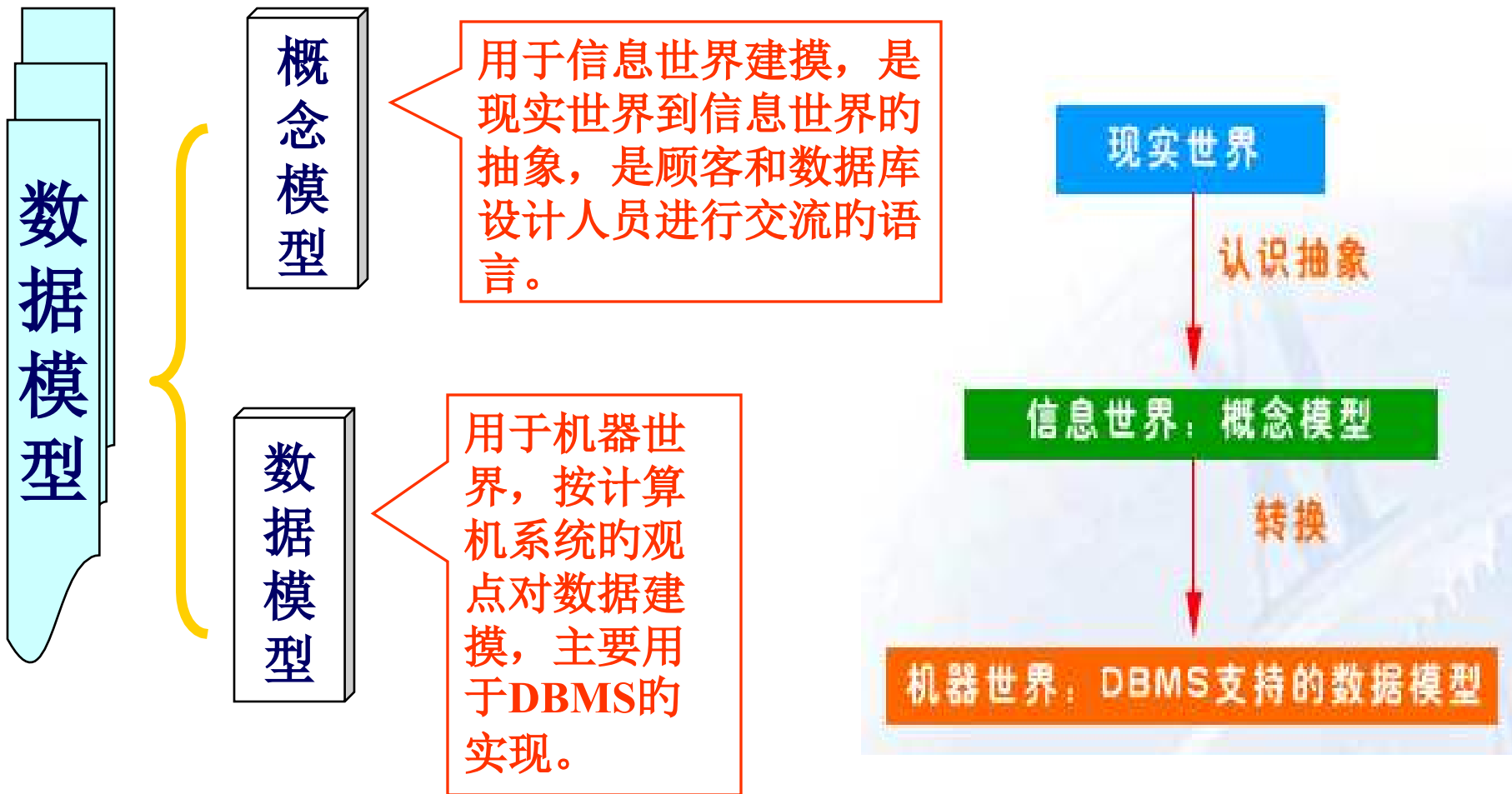
存在于人们头脑之外的客观世界，称为现实世界。

是现实世界在人们头脑中的反应，人们用文字、图形和符号等表达它们，构成信息世界。在信息世界中，数据库技术一般用到下列术语：实体、实体集，属性、联络和实体标识等。

计算机只能处理数据化的信息，对信息世界中的信息必须进行数据化，数据化后的信息称为数据。机器世界中数据描述的术语主要有：字段、统计、文件、键等

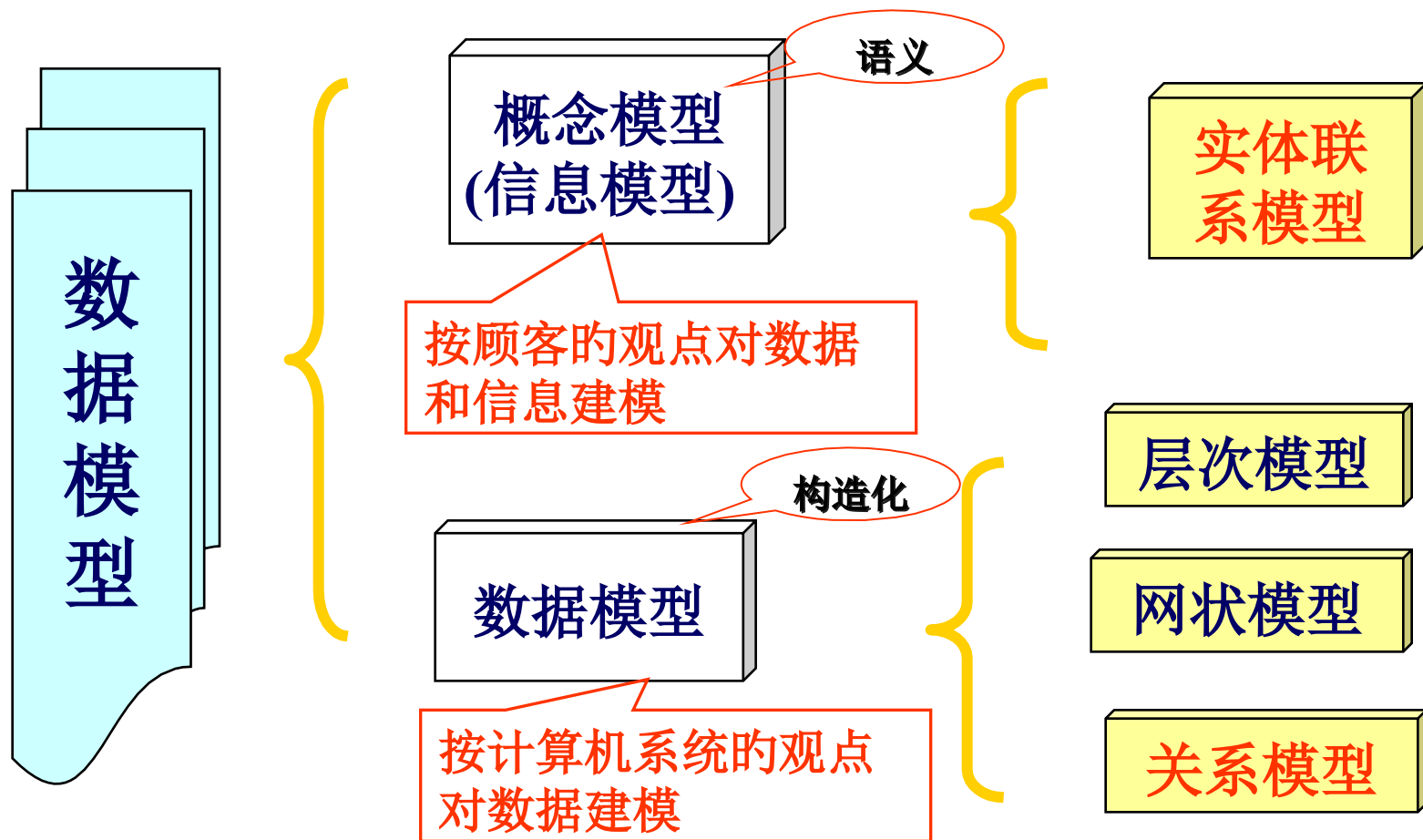
8.2.1 数据模型的基本概念 (2 / 3)

在数据库中，根据模型应用的不同目的，将模型分为两类或两个层次：



8.2.1 数据模型的基本概念 (3 / 3)

在数据库中，根据模型应用的不同目的，将模型分为两类或两个层次



8.2.2 概念模型

概念模型是按顾客的观点对数据和信息建模，概念模型是基于信息世界的主要模型，又叫信息模型，它具有较强的语义体现能力，能够以便、直接体现应用中的多种语义。表达概念模型能够使用多种不同的措施，这里简介的是1976年P. S. Chen提出的实体-联络措施（Entity—Relationship），简称E-R措施。



1. 信息世界中的基本术语——实体、属性

(1) 实体 (Entity)

客观存在并可相互区别的事物

能够是详细的人、事、物，也能够是抽象的概念或联络，例如，一种职员、一种学生、一种部门、一门课、学生的一次选课、部门的一次订货，老师与系的工作关系(即某位老师在某系工作)等都是实体。

(2) 属性 (Attribute)

实体所具有的某一特征；
属性有属性名、属性类型、
属性定义域和属性值之分。

一种实体能够由若干个属性来刻画。例如，学生实体能够由学号、姓名、性别、出生年份、系、入课时间等属性构成(94002268, 张山, 男, 976, 计算机系, 1994)。这些属性组合起来表征了一种学生。

单值属性、多值属性； NULL属性； 派生属性

1. 信息世界中的基本术语——码、域

(3) 码 (Key)

实体标识或键

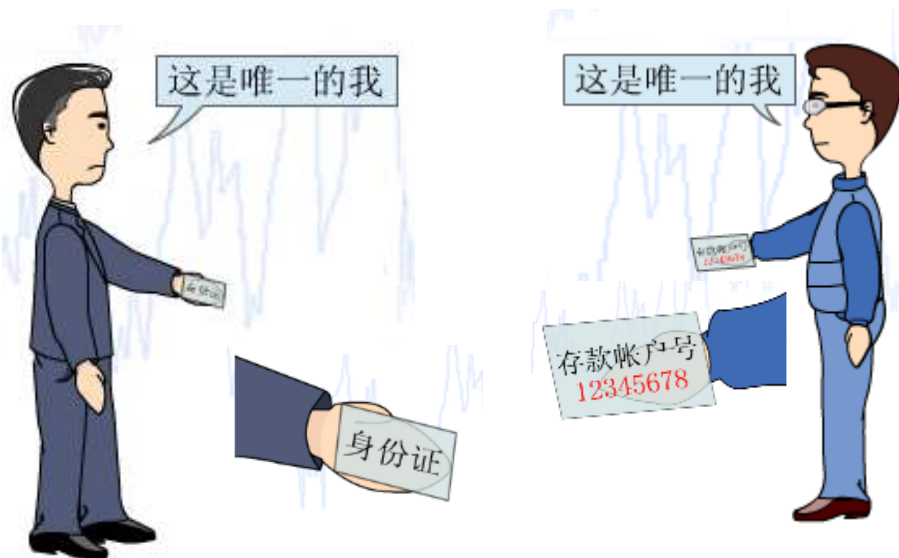
唯一标识实体的属性集

主码、候选码和超码，
是**实体集**的性质

(4) 域 (Domain)

某个（些）属性的取值范围；

例如，学号的域为8位整数（或**字符**），姓名的域为字符串集合，年龄的域为不小于1且不大于120的整数，性别的域为(男，女)等等。



1. 信息世界中的基本术语—实体

(5) 实体型 (Entity Type)

用实体名及其属性名集合来抽象、刻画同类实体；

描述实体集的信息构造,一般涉及类型名和实体的属性名集合。
例如, 学生 (学号, 姓名, 性别, 年龄, 班级)。

(6) 实体集 (Entity Set)

同型实体的集合称为实体集；

构造和特征相同的实体构成的集合称为实体集。
例如, 学生实体集, 班级实体集等。

1. 信息世界中的基本术语-联络

(7) 联络Relation

例如，每个学生隶属一种学校，每个学生和其隶属的一种学校之间有一种隶属联络。

现实世界中，事物之间的相互联络是客观存在的，**联络反应实体间的相互关系**。这种联络必然要在信息世界中进行描述。而且，联络除用作属性外，联络本身也能够有属性值。

实体型之间的联络

一对一联络 (1: 1)

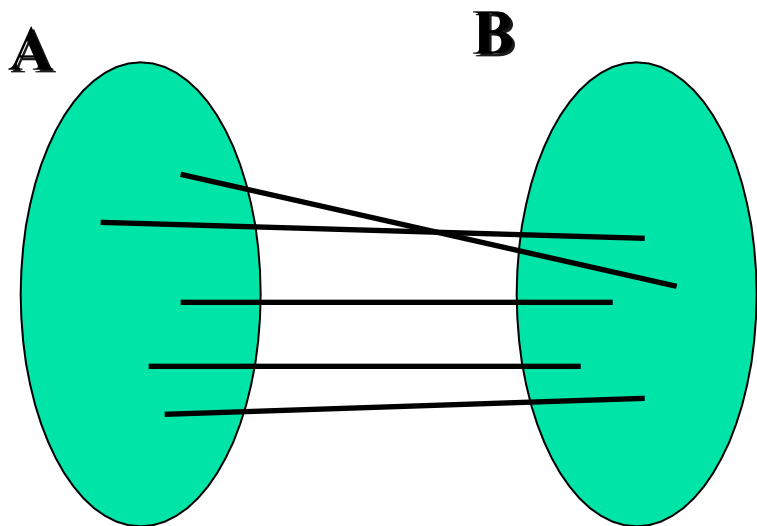
一对多联络 (1: n)

多对多联络 (m: n)

1. 信息世界中的基本术语——一对一联络

一对一联系(1:1)

实体集A中的每一种实体，实体集B中至多有一种实体与之联络，反之亦然，则称实体集A与实体集B具有一对一联络。记为1:1。



例

学校里面，一种班级只有一种正班长，而一种班长只在一种班中任职，则班级与班长之间具有一对一联络。

例？

例？

例？

1. 信息世界中的基本术语——一对多联络

一对多联系(1:n)

例

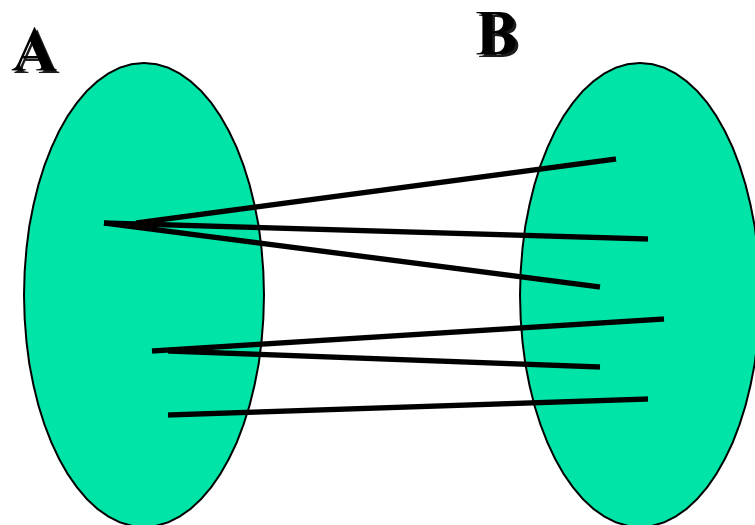
一种班级中有若干名学生，而每个学生只在一种班级中学习，则班级与学生之间具有一对多联络。

例?

例?

例?

实体集A中的每一种实体，实体集B中有n个实体($n \geq 0$)与之联络，反之，对于实体集B中的每一种实体，实体集A中至多只有一种实体与之联络，则称实体集A与实体集B有一对多联络。记为1:n。



1. 信息世界中的基本术语-多对多联络

多对多联系(m:n)

实体集A中的每一种实体，实体集B中有n个实体 ($n \geq 0$) 与之联络，反之，对于实体集B中的每一种实体，实体集A中也有m个实体 ($m \geq 0$) 与之联络，则称实体集A与实体集B具有多对多联络。记为m:n。

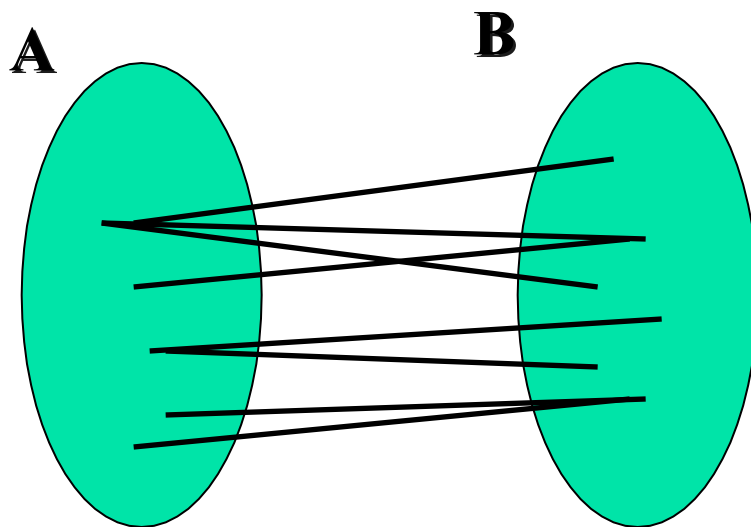
例

一门课程同步有若干个学生选修，而一种学生能够同步选修多门课程，则课程与学生之间具有多对多联络。

例?

例?

例?



2. 实体—联络措施 (1 / 3)

- 概念层次的数据模型称为概念数据模型，简称概念模型。概念模型离机器最远，从机器的立场上看是抽象级别的最高层。目的是按顾客的观点或认识来对现实世界建模。

(1) 实体型

用矩形表达，矩形框内写明实体名。

(2) 属性

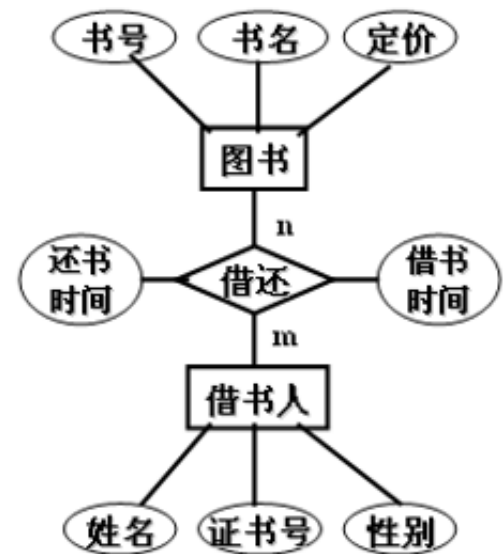
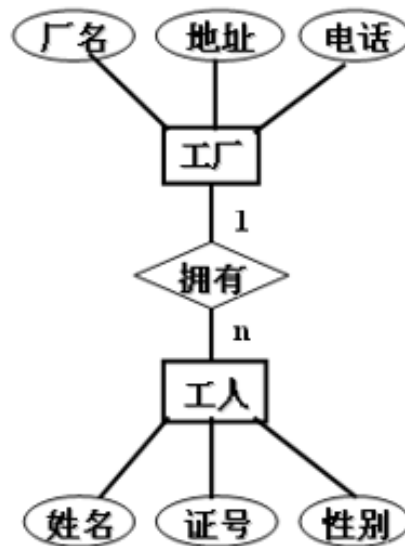
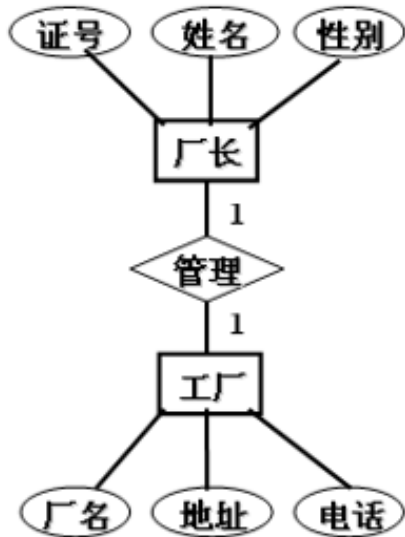
用椭圆形表达，并用无向边将其与相应的实体连接起来。

(3) 联络

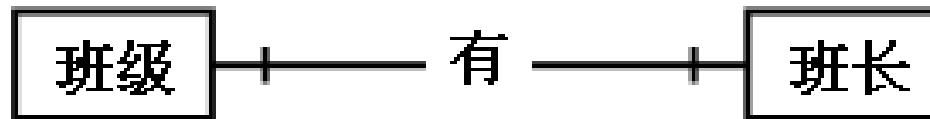
用菱形表达，菱形框内写明联络名，并用无向边分别与有关实体连接起来，同步在无向边旁标上联络的类型(1:1, 1:n或m:n)。

2. 实体—联络措施 (2 / 3)

- 概念层次的数据模型称为概念数据模型，简称概念模型。概念模型离机器最远，从机器的立场上看是抽象级别的最高层。目的是按顾客的观点或认识来对现实世界建模。



2. 实体—联络措施 (3 / 3)



(a) 两个实体型间的1:1联系



(b) 两个实体型间的1:n联系



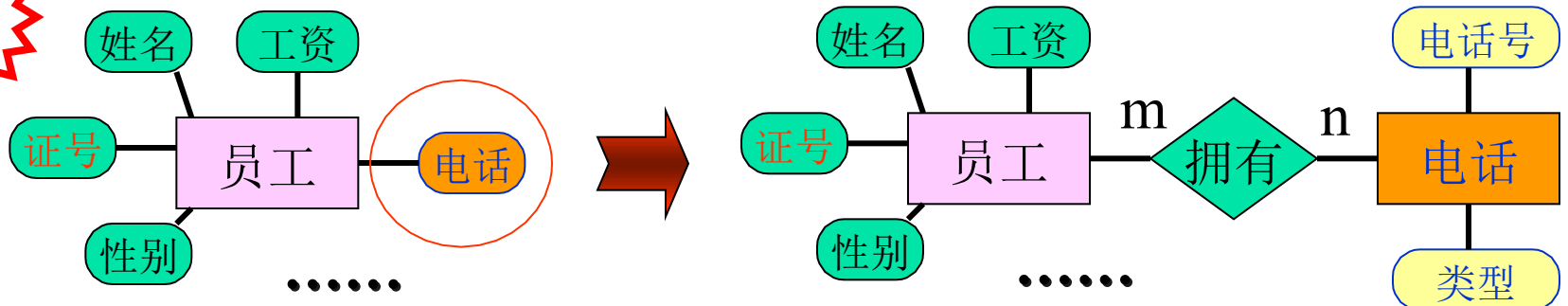
(c) 两个实体型间的n:m联系

E-R图实例 (鸭掌模型)

3. E-R模型设计中常出现的问题

- E-R模型设计中常遇到的问题之一是把所关心的对象表达成实体集好呢？还是用属性表达它更加好？而把对象描述成实体集和把它描述成某个实体集的属性是完全不同的，应该根据实际需求情况决定。
- 因为属性较简朴，假如能用属性表达清楚，就尽量用属性，但注意属性不能再用其他的属性来修饰，且它不能与其他实体集或属性再发生联络。当用简朴的属性表达不能满足实际需求时，就需用实体集表达。

在E-R模型设计中，把所关心的对象描述成实体集还是描述成某个实体集的属性是完全不同的。



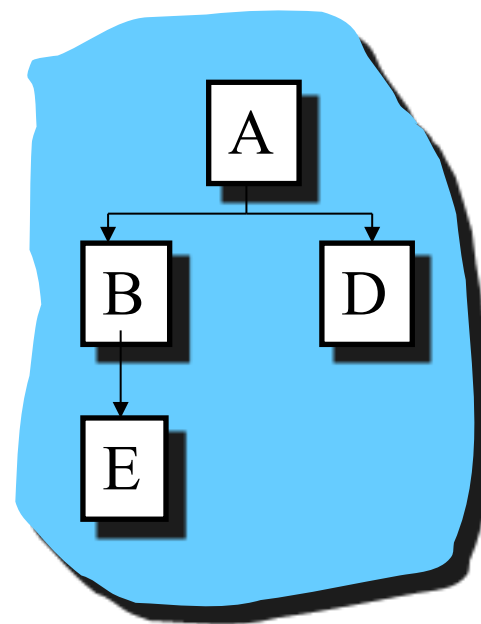
8.2.3 构造化数据模型

构造化的数据模型用于描述从信息世界到机器世界的模型，是对现实世界的第二层次的抽象，它是直接面对数据库的逻辑构造的，因为它涉及到计算机系统和数据库管理系统，所以又称为“基本数据模型”，如层次数据模型、网状数据模型、关系数据模型及面向对象数据模型等。它们都着重于描述数据的组织构造，极少考虑数据的语义和顾客对数据的了解，主要关心的是提供一种一致的、高效的数据库存储和检索所需要的物理构造和信息构造，提升数据存取和检索的速度。目前用得最多的是关系数据模型，这里也着重简介关系数据模型，简朴了解层次及网状数据模型。



1. 层次数据模型

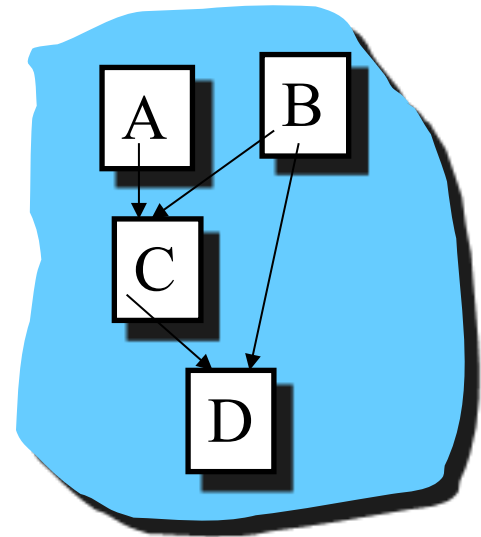
- 层次模型用树型构造来表达各类实体及实体间的联络，其特点是有且仅有一种结点无父结点，此结点称为根结点；其他结点有且仅有一种父结点。它适于表达一对多的联络。
- 层次模型可自然地体现数据间具有层次规律的关系，构造简朴，但因它只适于表达一对多的联络，所以在构造上有一定的不足。



层次数据模型

2. 网状数据模型

- 网状数据模型是一种比层次模型更具普遍性的构造，它去掉了层次模型的两个限制，允许多种结点没有双亲结点，允许结点有多种双亲结点，另外它还允许两个结点之间有多种联络，所以网状数据模型能够更直接地描述现实世界。
- 而层次构造实际上是网状构造的一种特例。



网状数据模型

3. 关系数据模型

- 用二维表格构造（关系）表达实体及实体之间的联络。

学生

学号	姓名	年龄	系别
s1	张红	18	CS
s2	李明	18	CS
s3	张红	18	MA

课程

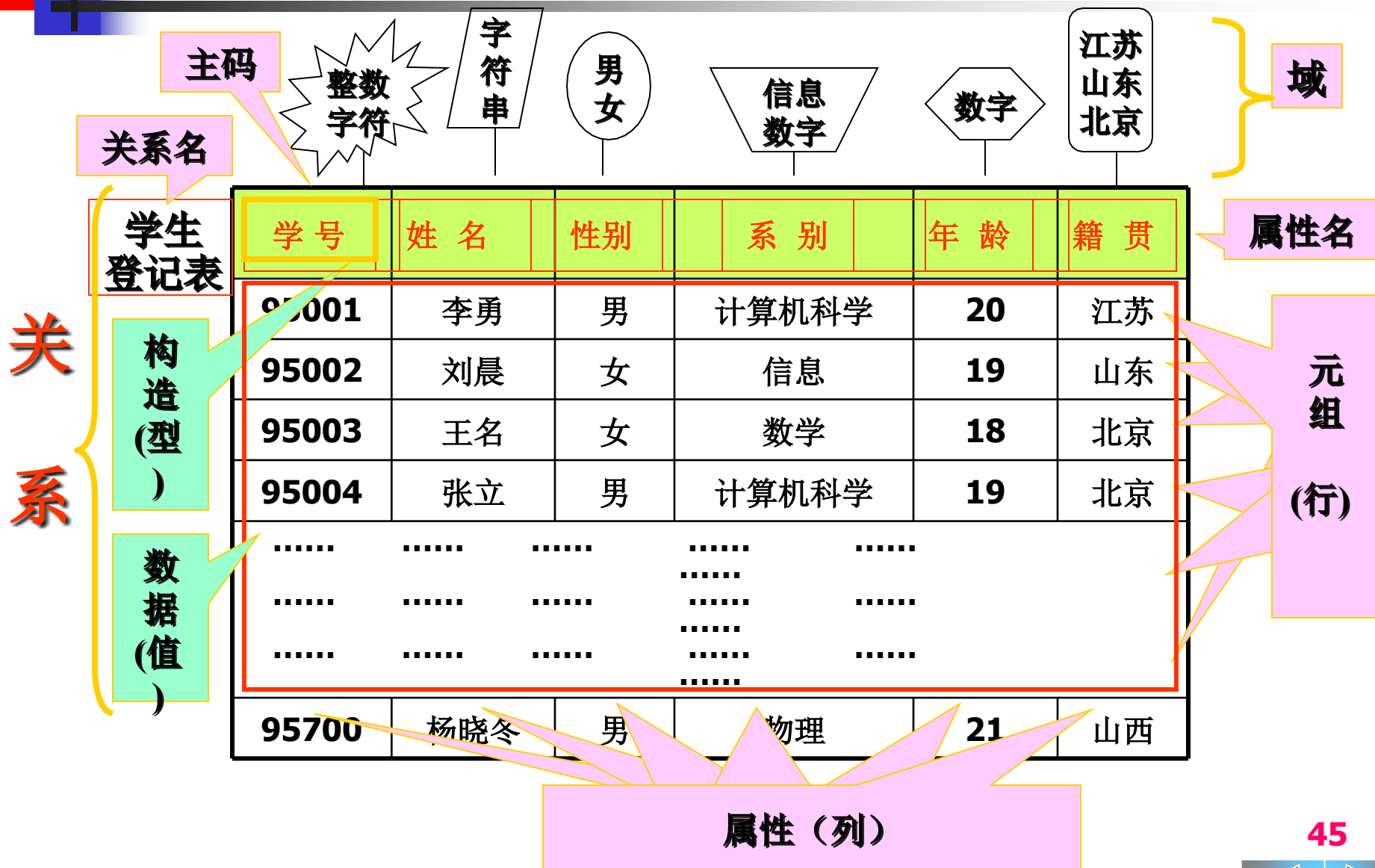
课号	课名	学分
c1	C语言	5
c2	高数	6
c3	数据库	6

学生选课

学号	课号	成绩
s1	c1	80
s1	c2	90
s2	c1	95



4. 关系数据模型的数据构造



关系数据模型中的概念及术语

- (1) 关系：即二维表，如图所示是一种图书表，也就是一种关系。
- (2) 元组：图中的每一行即称为一种元组或叫一种统计。
- (3) 属性：图中的每一列即称为一种属性，也可称为字段。(4) 主码：某个属性组，它能够唯一拟定一种元组，如图图书表中的书号属性。
- (5) 域：属性的取值范围，如书的定价一般在1至几十元之间，不可能有负数。
- (6) 分量：元组中的一种属性值，如图中的“S1”、“高等数学”等等。
- (7) 关系模式：即对关系的描述，一般表达为：关系名（属性1，属性2，……，属性n），例如图中的图书（书号，书名，出版社，定价，作者）。

5. 关系数据模型的操纵与完整性约束

- **实体完整性**：若属性**A**是基本关系**R**的主属性，则属性**A**不能取空值。
- **参照完整性**：若属性（或属性组）**F**是基本关系**R**的外码、它与基本关系**S**的主码**Ks**相相应（基本关系**R**和**S**不一定是不同的关系），则对于**R**中每个元组在**F**上的值必须为：
 - 或者取空值（**F**的每个属性值均为空值）
 - 或者等于**S**中某个元组的主码值
- **顾客定义完整性**：顾客定义的完整性就是针对某一详细关系数据库的约束条件，它反应某一详细应用所涉及的数据必须满足的语义要求，例如，某个属性必须取唯一值、某个非主属性也不能取空值、某个属性的取值范围在**0~100**之间等。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/005112202310011303>