

数据库系统概论

杨梦

逸夫楼810

m.yang@cumtb.edu.cn



China university of
Mining and Technology-Beijing

Database Applications

- 数据库技术是数据管理的有效技术，是计算机科学的重要分支。
- 数据库技术是信息系统的核心和基础，它的出现极大地促进了计算机应用向各行各业的渗透。
- 数据库已经成为每个人生活中不可缺少的部分。
- Database Applications
 - Banking: all transactions
 - Airlines: reservations, schedules
 - Universities: registration, grades
 - Sales: customers, products, purchases
 - Manufacturing: production, inventory, orders, supply chain
 - Human resources: employee records, salaries, tax deductions
 -



数据库学什么

➤ 数据库的创建和设计

从现实世界抽象出数据模型，并转化成为数据库管理系统 (DBMS) 的存储和管理形式（表，视图等）。



Student Achievement

学号	姓名	专业	DB平时	DB期末	DB总评成绩	OS平时	OS期末	OS总评成绩
3023001093	黄毅照	混合班	85	95	90	85
3011112340	周朝威	计算机科学与技术	80	90	85	88
3020621034	徐鑫	计算机科学与技术	90	90	90			85
3020831035	薄延嵩	计算机科学与技术	70	80	75			90
3021131123	胡俊	计算机科学与技术	70	70	70			75
3022112002	蒋永丽	计算机科学与技术	80	90	85			80
3022112003	顾娉娉	计算机科学与技术	90	90	90	85

.....

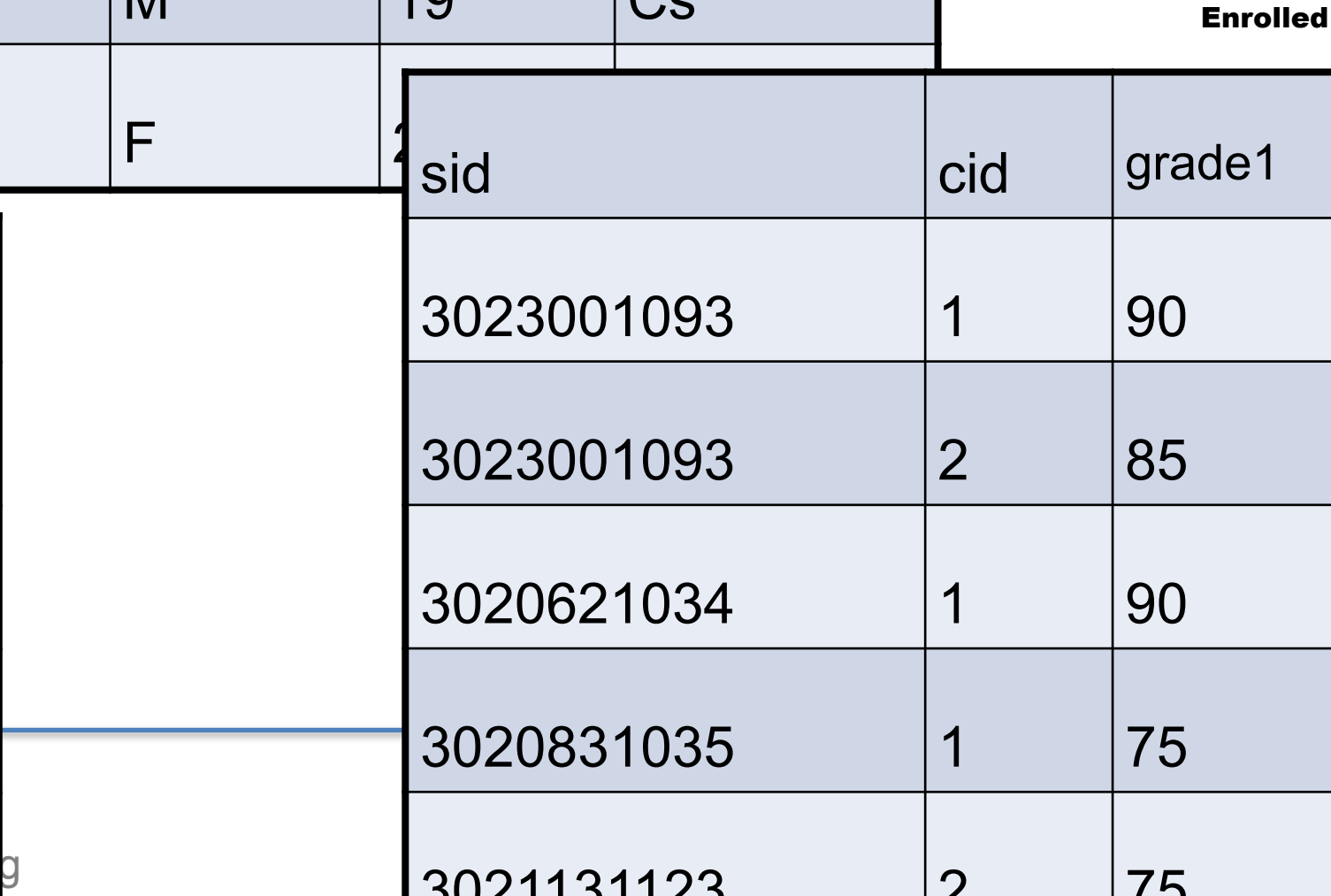
.....



Sid	Sname	Ssex	Sage	Specialty
3023001093	黄毅照	M	21	No
3011112340	周朝威	F	20	Cs
3020621034	徐鑫	M	18	Cs
3020831035	薄延嵩	M	19	Cs
3021131123	胡俊	F	21	Cs

cid	Cname	credit
1	DB	4
2	OS	5
3	English	4
4	Math	4

sid	cid	grade1	grade2	grade3
3023001093	1	90		
3023001093	2	85		
3020621034	1	90		
3020831035	1	75		
3021131123	2	75		



数据库学什么

➤ 数据库的创建和设计

从现实世界抽象出数据模型，并转化成为数据库管理系统(DBMS) 的存储和管理形式（表，视图等）。

➤ 应用编程

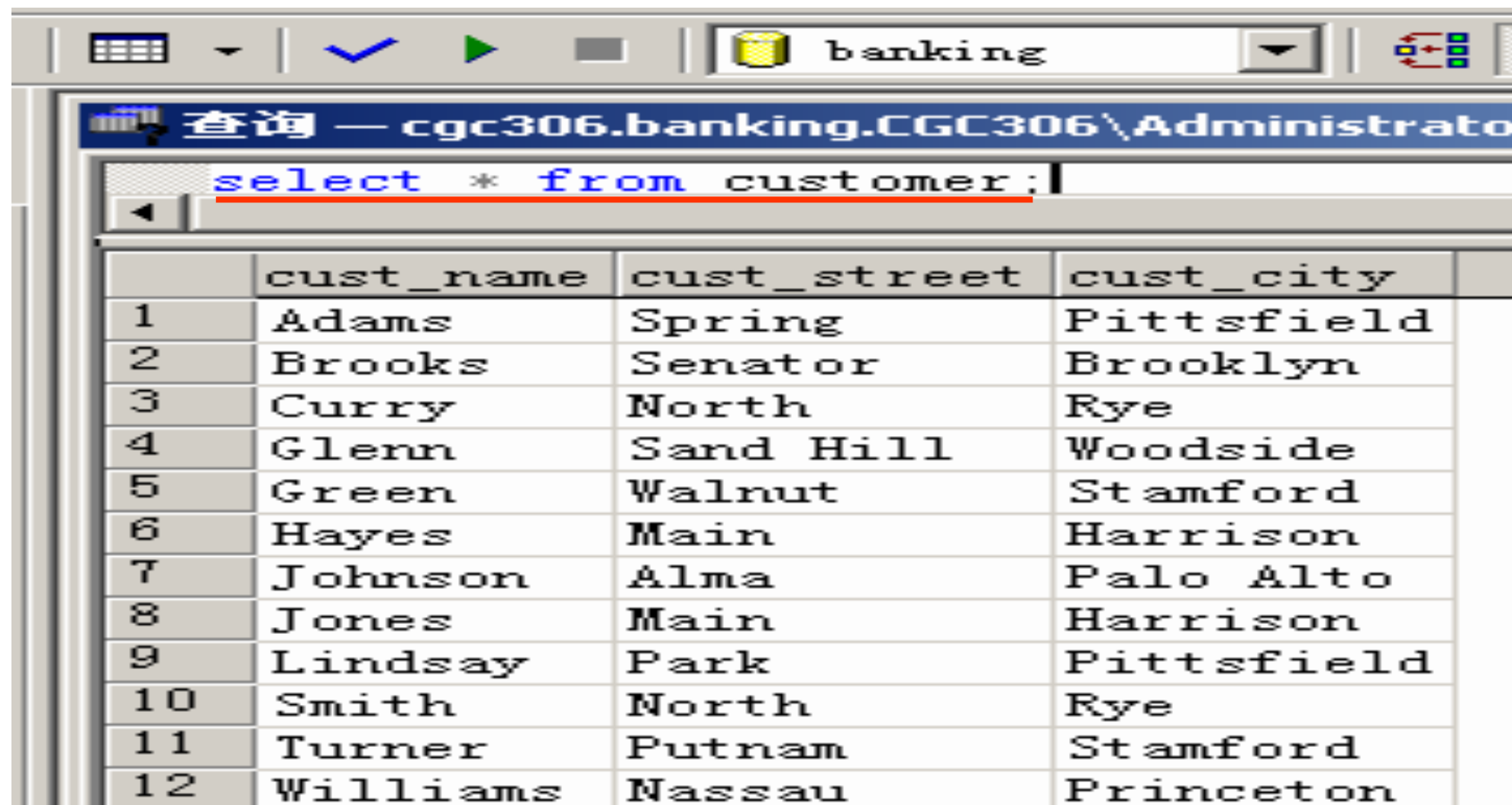
数据的使用和操纵-增删改查.



Database Access Approaches

❑方法1:

- 使用DBMS提供的交互式语言工具访问数据库(e.g., Query Analyzer of SQL Server, Sql*Plus and Work Sheet of ORACLE, etc.)



The screenshot shows a database query tool window titled "查询 — cgc306.banking.CGC306\Administrato". The query entered is "select * from customer;". The result is a table with 12 rows and 4 columns: cust_name, cust_street, and cust_city. The first column is an index from 1 to 12.

	cust_name	cust_street	cust_city
1	Adams	Spring	Pittsfield
2	Brooks	Senator	Brooklyn
3	Curry	North	Rye
4	Glenn	Sand Hill	Woodside
5	Green	Walnut	Stamford
6	Hayes	Main	Harrison
7	Johnson	Alma	Palo Alto
8	Jones	Main	Harrison
9	Lindsay	Park	Pittsfield
10	Smith	North	Rye
11	Turner	Putnam	Stamford
12	Williams	Nassau	Princeton



Database Access Approaches (Cont.)

□方法2:

- 调用 ODBC/JDBC 开发工具访问数据库 (e.g., VC++, PB, Delphi, ASP, JSP, PHP, etc.).

必修课教学计划					
院系选修课					
限定性选修课					
辅修课教学计划					
学院 计算机科学与技术学院 专业 计算机科学与技术 年级 2002 学期 全部					
课程代码	课程名称	学分	周学时	考核方式	课程性
02110010	思想道德修养	2.0	1.0-2.0	考查	必修课
02110020	法律基础	1.5	1.0-1.0	考查	必修课
02110032	毛泽东思想概论(乙)	1.5	1.0-1.0	考试	必修课
03110030	体育 I	1.0	0.0-2.0	考查	必修课
05110010	大学英语 I	3.0	2.0-2.0	考试	必修课
06110042	微积分(甲) I	4.5	4.0-1.0	考试	必修课
06110091	线性代数(甲)	3.0	3.0-0.0	考试	必修课
08110012	工程图学(乙)	2.5	2.0-1.0	考试	必修课
31110010	计算机文化	0.5	0.0-1.0		必修课
03110010	军事理论	1.5	1.0-1.0	考查	必修课
03110040	体育 II	1.0	0.0-2.0	考查	必修课
05110020	大学英语 II	3.0	2.0-2.0	考试	必修课
06110052	微积分(甲) II	4.5	4.0-1.0	考试	必修课
06110200	离散数学	4.0	4.0-0.0	考试	必修课





基础篇

第1章 绪论

第2章 关系数据库

第3章 关系数据库标准语言SQL

第4章 数据库安全性

第5章 数据库完整性



设计与应用开发篇

第6章 关系数据理论

第7章 数据库设计

第8章 数据库编程



Database Applications

- 数据库技术是数据管理的有效技术，是计算机科学的重要分支。
- 数据库技术是信息系统的核心和基础，它的出现极大地促进了计算机应用向各行各业的渗透。
- 数据库已经成为每个人生活中不可缺少的部分。
- Database Applications
 - Banking: all transactions
 - Airlines: reservations, schedules
 - Universities: registration, grades
 - Sales: customers, products, purchases
 - Manufacturing: production, inventory, orders, supply chain
 - Human resources: employee records, salaries, tax deductions
 -



系统篇

第9章 关系查询处理和查询优化

第10章 数据库恢复技术

第11章 并发控制



[教材]

数据库系统概论，萨师煊，王珊。高等教育出版社。

➤ 第 1-11章

[参考书]

- (1) Abraham Silberschatz, Henry F. Korth, and S. Sudarshan. Database System Concepts (6th Edition), 高等教育出版社, 2010
- (2) An Introduction to Database System (Ed.8), C.J. Date, Addison-Wesley, 2003
- (3) 陈根才, 孙建伶, 林怀中, 周波. 数据库课程设计, 浙江大学出版社, 2007.

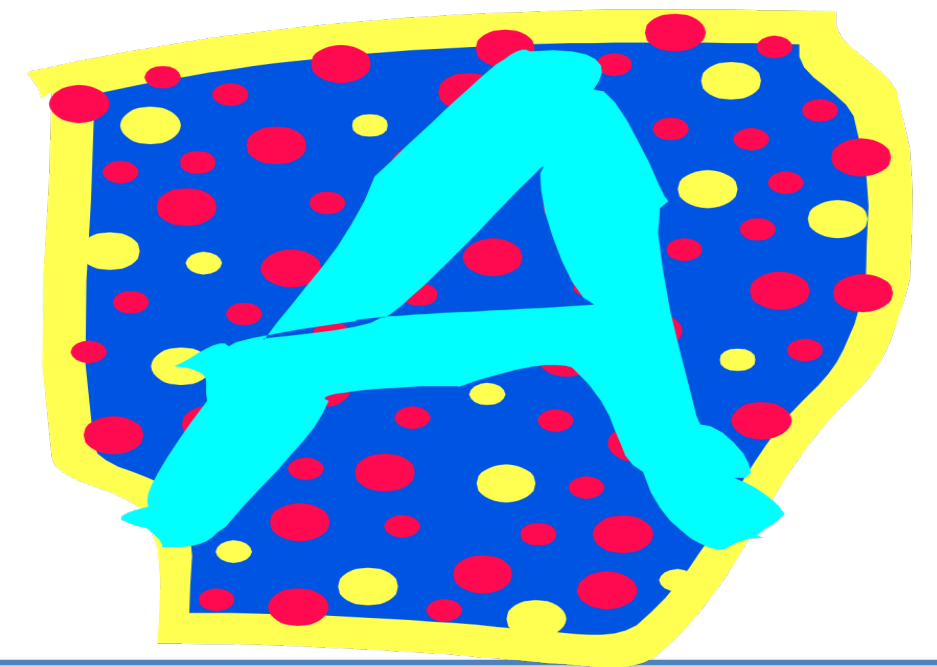


考试成绩

☀平时成绩

（出勤率，书面作业，上机练习）

☀期末考试



数据库系统概论

第一章 绪论



第一章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的结构

1.4 数据库系统的组成

1.5 小结



学习要求

- (1) 了解学习数据库系统的目的
- (2) 掌握数据库中有关的基本概念、数据库系统的组成及各部分的主要功能。
- (3) 了解数据库管理技术的三个发展阶段及各阶段的主要特点。
- (4) 掌握数据库系统的三级模式结构和两级映象功能
- (5) 了解常用的数据模型及特点



第一章 绪论

.....

1.1 数据库系统概述

- ◆ 数据库的4个基本概念
- ◆ 数据库系统的特点



第一章 绪论

.....

1.1 数据库系统概述

◆ 数据库的4个基本概念

数据 (Data)

数据库 (Database)

数据库管理系统 (DBMS)

数据库系统 (DBS)

◆ 数据库系统的特点



1.1 数据库系统概述

信息与数据

一、信息的定义：信息是关于现实世界事物的存在方式或运动状态的反映的综合。

信息的特征：

- 信息源于物质和能量。
- 信息是可以感知的，不同的信息源有不同的感知形式。
- 信息是可以存储、加工、传递、再生。



二、数据

数据的定义：数据是对信息的一种符号化表示，即用一定的符号表示信息。数据是数据库中存储的基本对象。

描述事物的符号可以是数字、也可以是文字、图形、音频、视频等，数据有多种表现形式，都可以经过数字化后存入计算机。

三、数据与信息的联系

数据是信息的载体，信息是数据的内涵。

同一信息可以有不同的数据表示形式；而同一数据也可能有不同的解释。



数据举例

◆学生档案中的学生记录

(李明, 男, 199505, 江苏南京市, 计算机系, 2013)

语义: 学生姓名、性别、出生年月、出生地、所在院系、入学时间

解释: 李明是个大学生, 1995年5月出生, 江苏南京市人, 2013年考入计算机系



二、数据

数据的定义：数据是对信息的一种符号化表示，即用一定的符号表示信息。数据是数据库中存储的基本对象。

描述事物的符号可以是数字、也可以是文字、图形、音频、视频等，数据有多种表现形式，都可以经过数字化后存入计算机。

三、数据与信息的联系

数据是信息的载体，信息是数据的内涵。

同一信息可以有不同的数据表示形式；而同一数据也可能有不同的解释。



数据举例

◆数据的含义称为数据的语义，数据与其语义是不可分的。

例如 93是一个数据

语义1：学生某门课的成绩

语义2：某人的体重

语义3：计算机系2013级学生人数

语义4：请同学给出.....



.....

- 数据库的定义

数据库（Database，简称DB）是长期储存在计算机内、有组织的、可共享的大量数据的集合。

- 数据库的基本特征

数据按一定的数据模型组织、描述和储存

可为各种用户共享

冗余度较小

数据独立性较高

易扩展



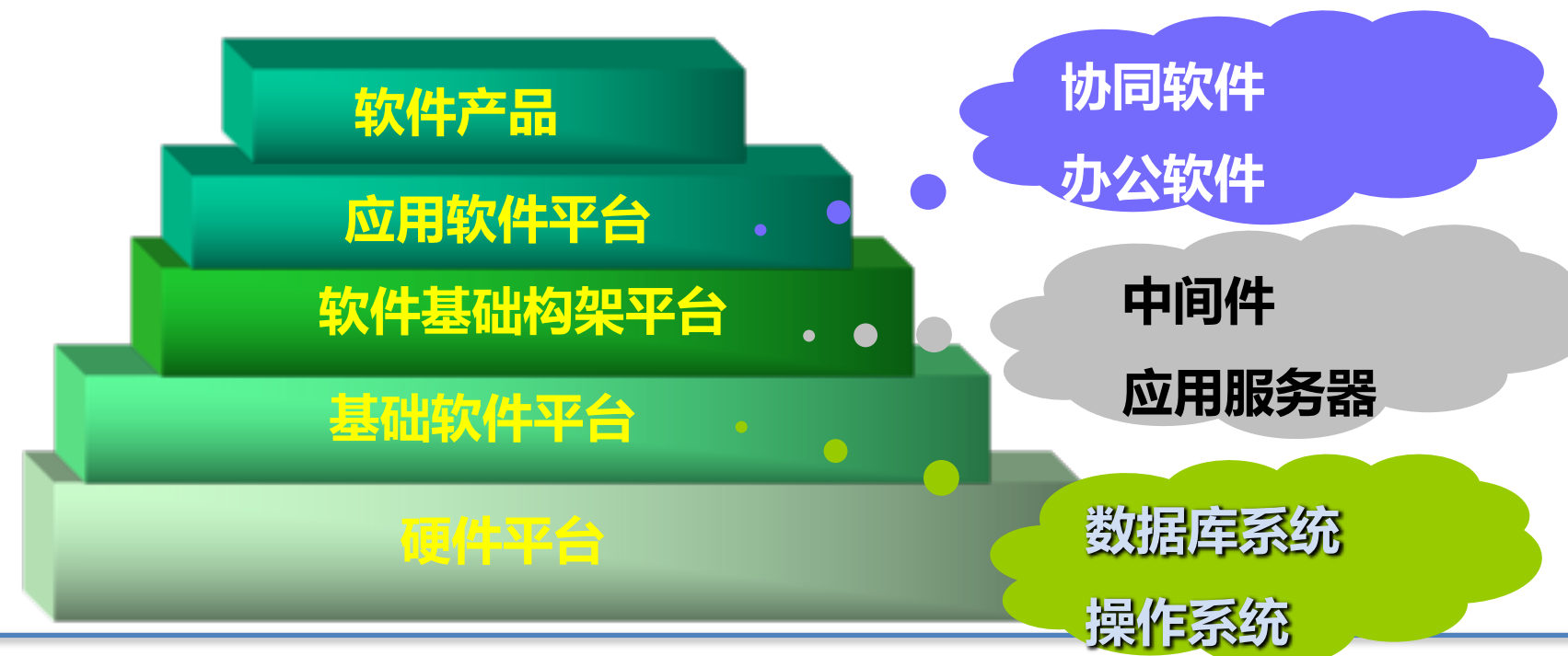
2. 数据库管理系统

- ◆ 数据库管理系统 Database Management System (DBMS)
(数据库) + 一系列软件，用于科学的组织和存储，
高效的获取和维护数据。



2. 数据库管理系统

- 数据库管理系统
 - 位于用户与操作系统之间的一层数据管理软件，是基础软件，是一个大型复杂的软件系统
 - 数据库管理系统的用途：科学地组织和存储数据、高效地获取和维护数据



.....数据库管理系统的主要功能.....

1. 数据定义功能

提供数据定义语言（DDL），定义数据库中的数据对象的组成和结构

2. 数据组织、存储和管理

分类组织、存储和管理各种数据

确定组织数据的文件结构和存取方式

提供多种存取方法提高存取效率



数据库管理系统的主要功能.....

3. 数据操纵功能

提供数据操纵语言（DML），实现对数据库的基本操作（查询、插入、删除和修改）

4. 数据库的事务管理和运行管理

数据库在建立、运行和维护时由数据库管理系统统一管理和控制

保证数据的安全性、完整性、多用户对数据的并发使用

发生故障后的系统恢复



数据库管理系统的主要功能.....

5.数据库的建立和维护功能

数据库初始数据的装载和转换

数据库转储、恢复功能

数据库的重组织

性能监视、分析等

6.其它功能

数据库管理系统与网络中其它软件系统的通信

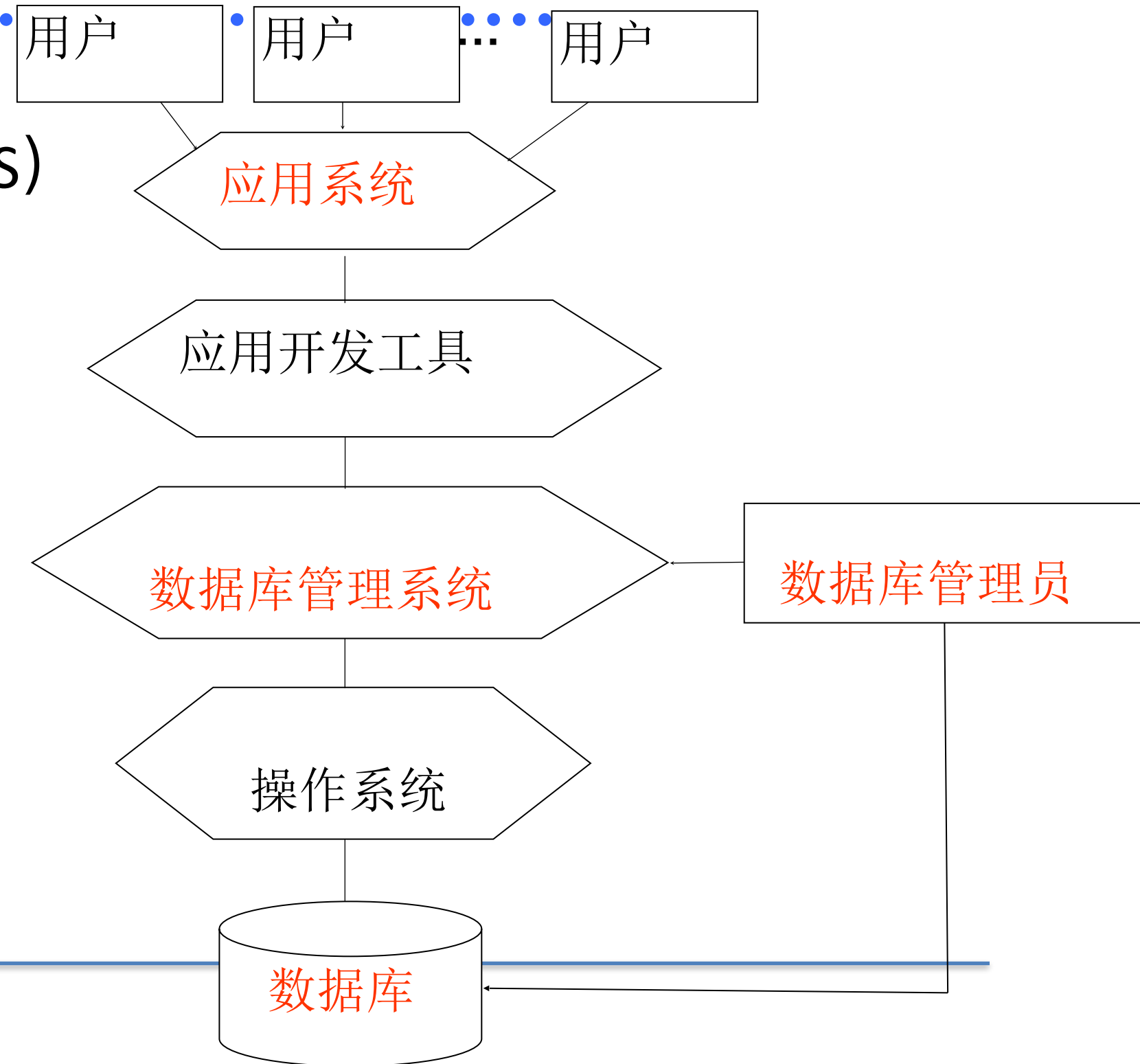
数据库管理系统系统之间的数据转换

异构数据库之间的互访和互操作



4.数据库系统

- 数据库系统 (Database System, 简称DBS)
- 数据库系统的构成
 - 数据库
 - 数据库管理系统 (及其应用开发工具)
 - 应用程序
 - 数据库管理员



第一章 绪论

.....

- ◆ 数据库管理系统在数据库建立、运用和维护时对数据库进行统一控制，以保证数据的完整性、安全性，并在多用户同时使用数据库时进行并发控制，在发生故障后对数据库进行恢复

1.1 数据库系统概述

- ◆ 数据库的4个基本概念
- ◆ DBMS技术的发展和数据库系统的特点



数据管理技术

- ◆ 数据管理技术的发展过程
 - 人工管理阶段（20世纪50年代中之前）
 - 文件系统阶段（20世纪50年代末--60年代中）
 - 数据库系统阶段（20世纪60年代末--现在）



1.2 数据库管理技术的发展

一、人工管理阶段

硬件中的外存只有卡片、纸带、磁带，没有磁盘等直接存储设备；软件只有汇编语言，没有操作系统和专门的数据管理软件，数据由人来管理。计算机主要用于科学计算，数据量不大。

人工管理数据的特点：

- (1) 数据不保存。
- (2) 系统没有专门的软件对数据进行管理。
- (3) 数据不共享。
- (4) 数据不具有独立性。

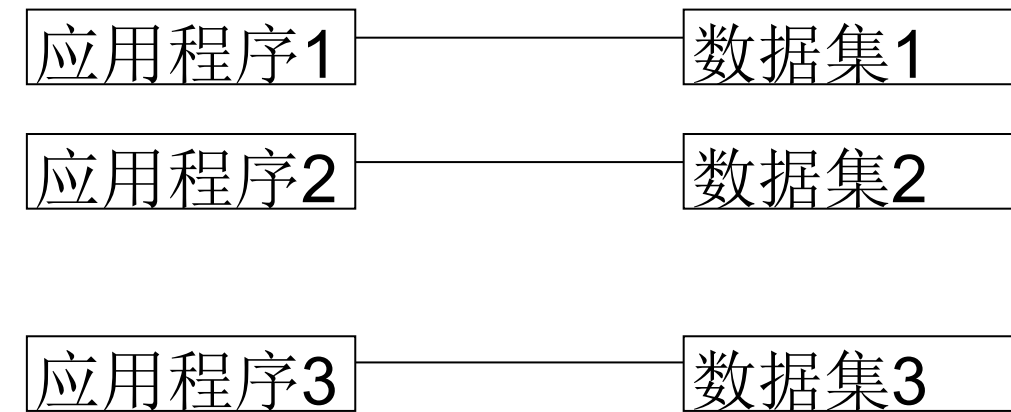


图1-1 人工管理阶段应用程序与数据之间的对应关系



二、文件系统阶段

这一时期，计算机外存已有了磁鼓、磁盘等存储设备，软件有了**操作系统**。人们在操作系统的支持下，设计开发了一种专门管理数据的计算机软件，称为文件系统。这时，计算机不仅用于科学计算，而且大量用于数据处理。

用文件系统管理数据具有以下特点：

- (1) 数据以文件的形式长期保存。
- (2) 由文件系统管理数据。

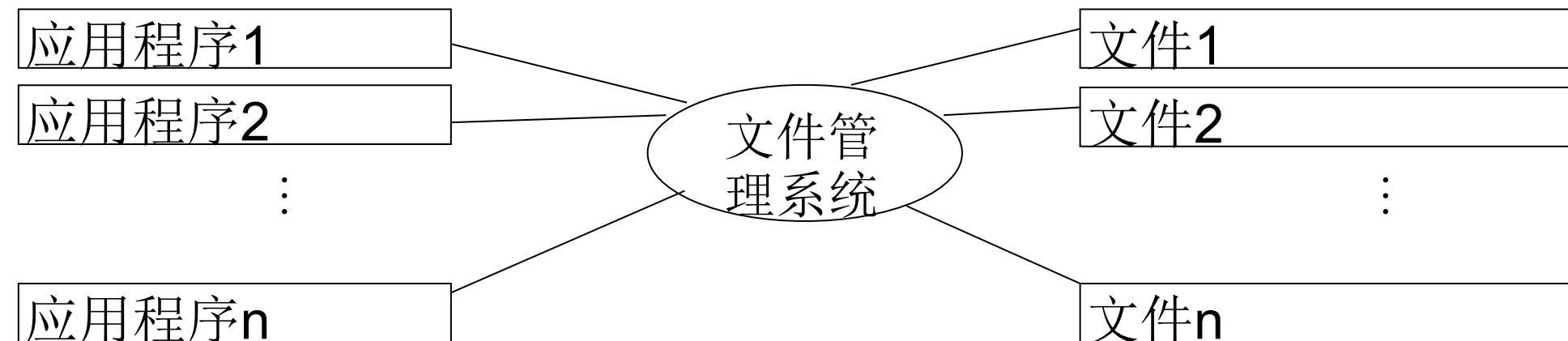


图1-2 文件系统阶段应用程序与数据之间的对应关系 34



文件系统阶段（续）

❖缺点

- 数据面向的对象：某一应用
- 数据的共享程度：共享性差、冗余度大
- 数据的独立性：独立性差
- 数据控制能力：应用程序自己控制



3. 数据库系统阶段

◆时期

20世纪60年代末以来

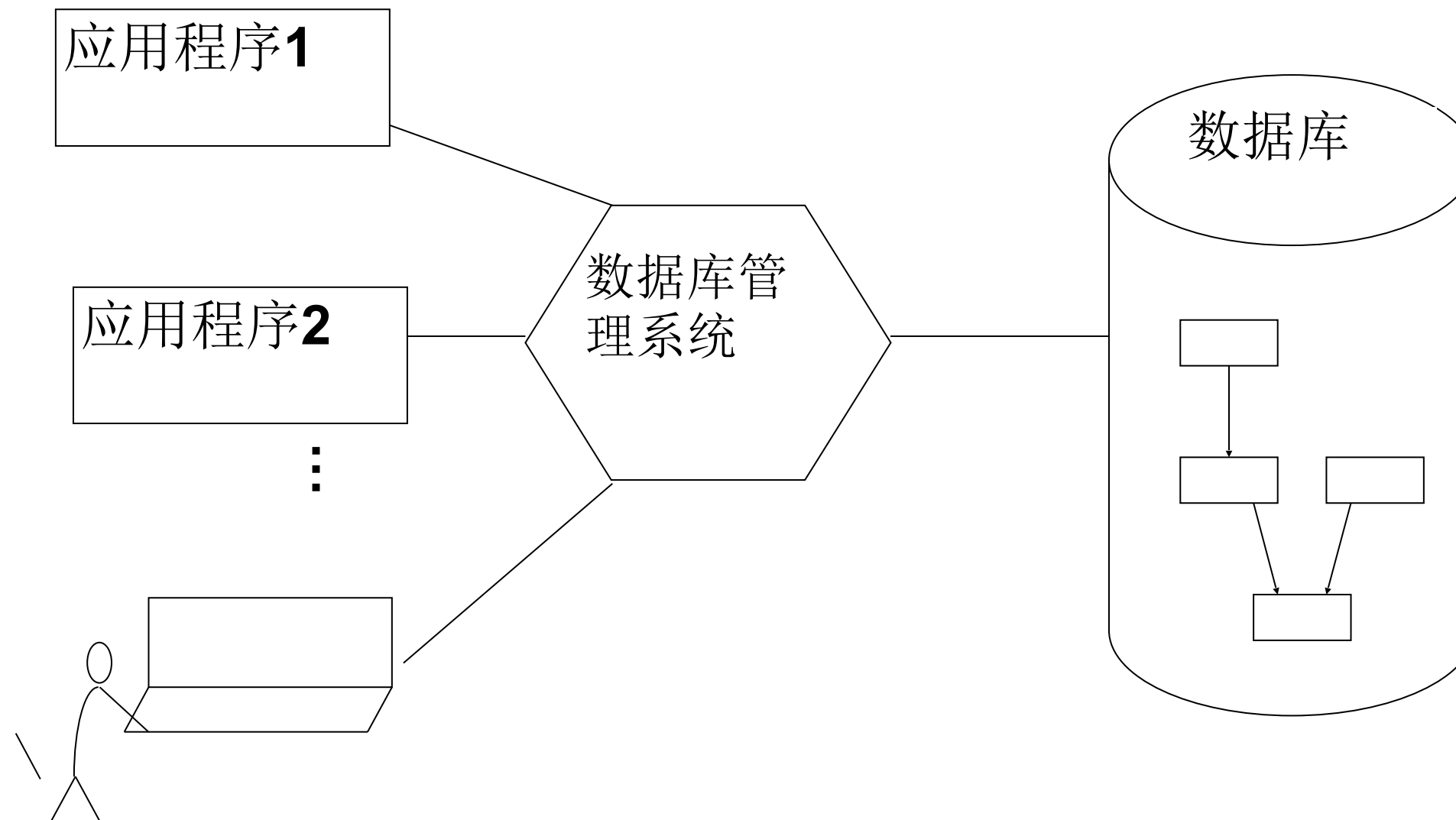
◆产生的背景

计算机管理的对象规模越来越大，应用范围越来越广，数据量急剧增长，并且多应用互相共享数据集合的要求越来越强烈。

为了解决多用户，多应用共享数据的需要，使数据尽可能多的应用服务，数据库技术便应用而生，出现了统一管理数据的专门软件系统—DBMS



应用程序与数据的对应关系（数据库系统阶段）



数据库系统阶段——应用程序与数据之间的对应关系



1.1.3 数据库系统的特点

.....

- 1.数据结构化
- 2.数据的共享性高，冗余度低且易扩充
- 3.数据独立性高
- 4.数据由数据库管理系统统一管理和控制



一个例子

◆学生的信息包括学号、姓名、性别、年龄、专业和奖励

◆用文件系统实现学籍管理

- 数据存储

- 定长记录 存储在“学生基本信息”文件中

- 变长记录 存放在另一个“奖励”文件

- “学生基本信息”表中的位置和长度描述“奖励”文件中记录的开始

位置和长度

- 查询数据

- 编写应用程序，实现数据的录入和查找

◆缺点：程序员必须关注记录结构和不同文件中记录之间的联系，工作量大，编程复杂，开发速度慢



一个例子（续）

❖数据库系统管理

- 存储数据

建立两张表:

STUDENT表-存放学生的基本信息,

AWARD表-存放学生的奖励情况

使用两条插入命令 完成学生基本信息和奖励情况的数据录入功能

- 查询功能

可以用一条查询语句实现



1.数据结构化

.....

◆数据的整体结构化是数据库的主要特征之一

整体结构化：

不再仅仅针对某一个应用，而是面向全组织

不仅数据内部结构化，整体是结构化的，数据之间具有联系



2.数据的共享性高，冗余度低且易扩充

.....

◆数据面向整个系统，可以被多个用户、多个应用共享使用。

•数据共享的好处

减少数据冗余，节约存储空间

避免数据之间的不相容性与不一致性

使系统易于扩充



3.数据独立性高

- 物理独立性

指用户的应用程序与数据库中数据的物理存储是相互独立的。当数据的物理存储改变了，应用程序不用改变。

- 逻辑独立性

指用户的应用程序与数据库的逻辑结构是相互独立的。数据的逻辑结构改变了，应用程序不用改变。

数据独立性由数据库管理系统的系统结构——二级映像功能来保证。



4.数据由数据管理系统统一管理和控制

.....

◆数据库管理系统提供的数据控制功能

(1) 数据的安全性 (Security) 保护

保护数据以防止不合法的使用造成的数据的泄密和破坏。

(2) 数据的完整性 (Integrity) 检查

保证数据的正确性、有效性和相容性。

(3) 并发 (Concurrency) 控制

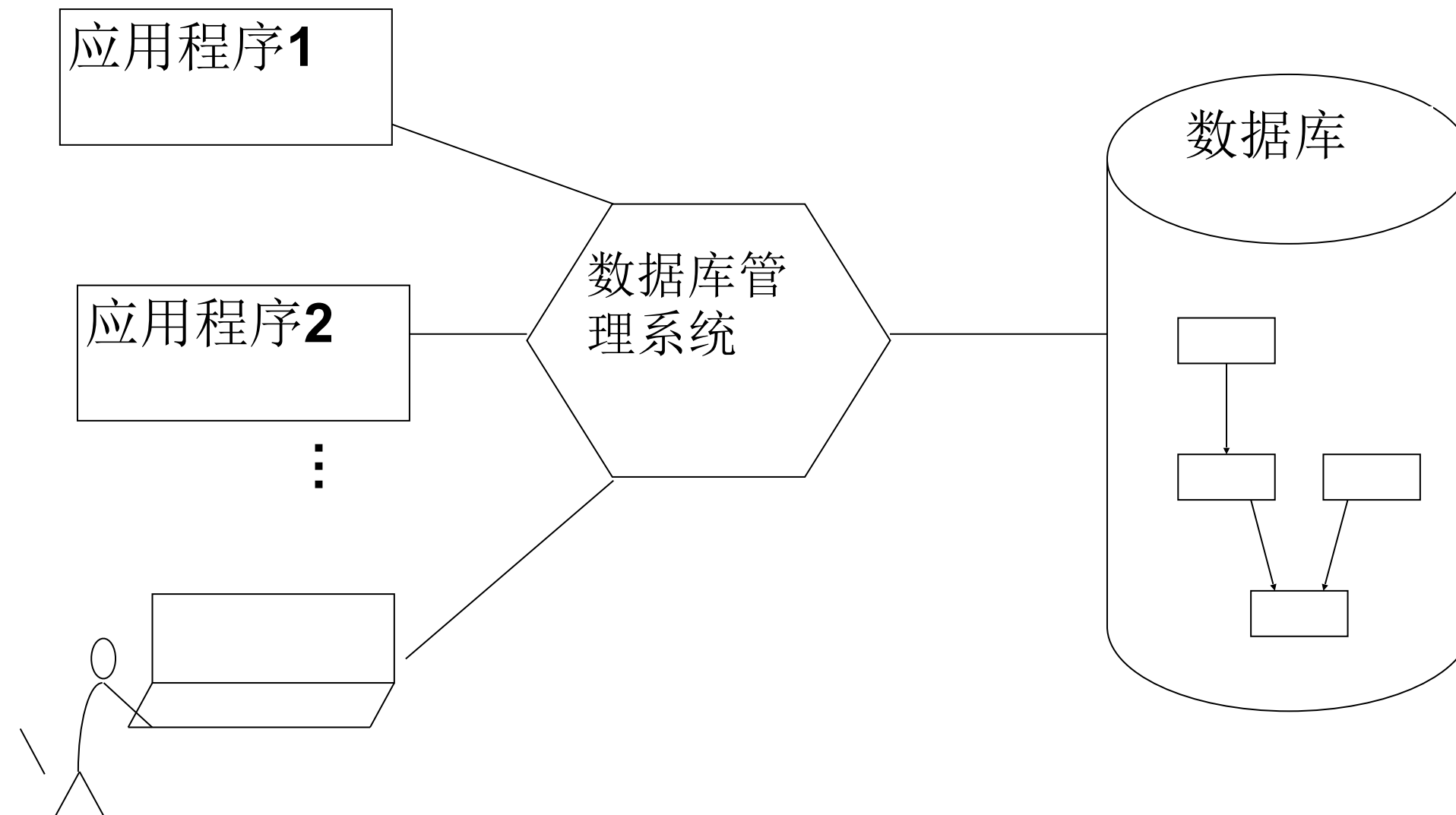
对多用户的并发操作加以控制和协调，防止相互干扰而得到错误的结果。

(4) 数据库恢复 (Recovery)

将数据库从错误状态恢复到某一已知的正确状态。



应用程序与数据的对应关系（数据库系统阶段）



数据库系统阶段 应用程序与数据之间的对应关系



数据库系统分类

.....

从最终用户角度，数据库系统分为单用户结构、主从式结构、分布式结构、客户/服务器结构。

1、单用户结构的数据库系统

又称为桌面型数据库系统

特点：将应用程序、**DBMS**、数据库都装在一台计算机上，由一个用户独占使用，不同计算机间不能共享数据。如**Access**。

优点：工作在单机环境，侧重在可操作性、易开发和简单管理等方面。

适用：未联网用户、个人用户等。



2、主从式结构的数据库系统

是指大型主机带多终端的多用户结构的系统。

特点：将应用程序、**DBMS**、数据库都集中存放在大型主机上，所有处理任务由主机完成，而连于主机上的终端，只作为主机的输入/输出设备，各个用户通过主机的终端并发地存取数据库，共享数据资源。

优点：结构简单、易于管理和维护。

缺点：所有处理任务由主机完成，对主机的性能要求较高，当终端数量太多时，主机的处理任务过重，易形成瓶颈，使系统性能下降；另外当主机出现故障时，整个系统无法使用。



3、客户/服务器结构的数据库系统

特点：**DBMS**和数据库存放在数据库服务器上，应用程序和相关开发工具存放于客户机上。客户机负责管理用户界面，接收用户数据，处理应用逻辑，生成数据库服务请求，将该请求发送到服务器。数据库服务器进行处理后，将处理结果返回给客户机，并按一定格式显示给用户。这种模式和主从式模式比，又称为胖客户机模式。是一种二层结构，即客户机、数据库服务器。

客户/服务器结构优点：

（1）网络运行效率提高。因为服务器只将处理结果返回给客户机，降低了网络上数据的传输量。

（2）应用程序的运行和计算处理工作由客户机完成，减少了客户机与服务器间不必要的通讯开销，减轻了服务器的处理工作。

客户/服务器结构缺点：

维护升级不方便，需要在每个客户机上安装前端客户应用程序，当应用程序修改后，需要在每台客户机上重新安装。



4、分布式结构的数据库系统

是分布式网络技术与数据库技术相结合的产物，是分布在计算机网络上的多个逻辑相关的数据库的集合。

特点：

(1) 数据在物理上分布：数据库中的数据不集中存放在一台服务器上，而是分布在不同地域的服务器上，每台服务器被称为节点。

(2) 所有数据在逻辑上是整体的：数据库中的数据物理分布，在逻辑上相互关联，是相互联系的整体。

(3) 节点上分布存储的数据相对独立：对用户而言，整个数据库系统仍然是集中的整体，用户不关心数据的分片存储，也不关心物理数据的具体分布，完全由网络数据库在网络操作系统的支持下完成。用户既可以存取本地节点上的数据库，也可以存取异地节点上的数据库。



第一章 绪论

.....

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的结构

1.4 数据库系统的组成

1.5 小结



1.2 数据模型

1.2.1 什么是数据模型

1.2.2 概念模型

1.2.3 数据模型的组成要素

1.2.4 常用的数据模型

1.2.5 层次模型

1.2.6 网状模型

1.2.7 关系模型

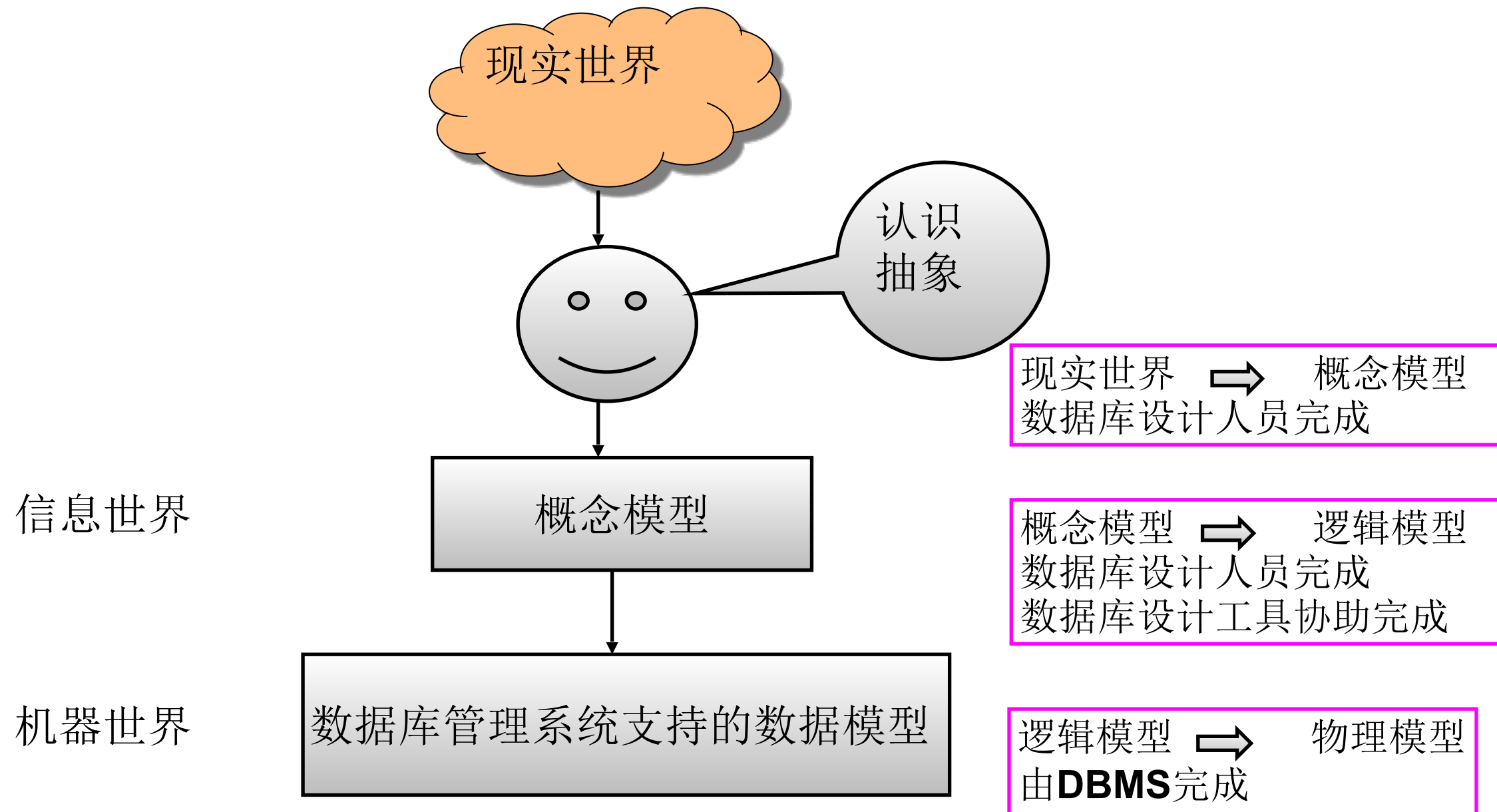


1.2 数据模型

- ◆数据模型是数据库系统的核心和基础，数据模型是对现实世界数据特征的抽象。
- ◆通俗地讲数据模型就是现实世界的模拟。
- ◆数据模型应满足三方面要求
 - 能比较真实地模拟现实世界
 - 容易为人所理解
 - 便于在计算机上实现



两类数据模型（续）



现实世界中客观对象的抽象过程



1.2.1 两类数据模型

数据模型分为两类（两个不同的层次）

(1) **概念模型** 也称信息模型，它是按用户的观点来对数据和信息建模，用于数据库设计。

(2) **逻辑模型和物理模型**

- 逻辑模型主要包括网状模型、层次模型、关系模型、面向对象数据模型、对象关系数据模型、半结构化数据模型等。按计算机系统的观点对数据建模，用于DBMS实现。

- 物理模型是对数据最底层的抽象，描述数据在系统内部的表示方式和存取方法，在磁盘或磁带上的存储方式和存取方法。



1.2 数据模型

1.2.1 两大类数据模型

1.2.2 概念模型

1.2.3 数据模型的组成要素

1.2.4 最常用的数据模型

1.2.5 层次模型

1.2.6 网状模型

1.2.7 关系模型



1.2.2 概念模型

◆概念模型的用途

概念模型用于信息世界的建模

是现实世界到机器世界的一个中间层次

是数据库设计的有力工具

数据库设计人员和用户之间进行交流的语言

◆对概念模型的基本要求

较强的语义表达能力

简单、清晰、易于用户理解



1. 信息世界中的基本概念

.....

(1) 实体 (Entity)

客观存在并可相互区别的事物称为实体。

可以是具体的人、事、物或抽象的概念。

(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。

(3) 码 (Key)

唯一标识实体的属性集称为码。



信息世界中的基本概念（续）

(4) 实体型 (Entity Type)

用实体名及其属性名集合来抽象和刻画同类实体称为
实体型

(5) 实体集 (Entity Set)

同一类型实体的集合称为实体集



信息世界中的基本概念（续）

（6）联系（Relationship）

在现实世界中事物内部以及事物之间的联系在信息世界中反映为实体（型）内部的联系和实体（型）之间的联系。

实体内部的联系通常是指组成实体的各属性之间的联系

实体之间的联系通常是指不同实体集之间的联系

实体之间的联系有一对一、一对多和多对多等多种类型



联系：

(1) 一对一联系（简记为1:1）

如果对于实体集A中的每一个实体，实体集B中至多有一个（也可以没有）实体与之联系，反之亦然，则称实体集A与实体集B具为1:1联系。

例如：

✓ 观众与座位，病人与床位。



(2) 一对多联系 (1:n)

如果对于实体集A中的每一个实体，实体集B中有n个实体 ($n \geq 0$) 与之联系，反之，对于实体集B中的每一个实体，实体集A中至多只有一个实体与之联系，则称实体集A与实体集B具有一对多联系，记为1:n。

一对多联系包含了一对一联系。

例如：

✓ 一个班级中有若干名学生，而每个学生只在一个班级中学习，则班级与学生之间具有一对多联系。

✓ 公司与职员、省与市



(3) 多对多联系 (m:n)

如果对于实体集A中的每一个实体，实体集B中有n个实体 ($n \geq 0$) 与之联系，反之，对于实体集B中的每一个实体，实体集A中也有m个实体 ($m \geq 0$) 与之联系，则称实体集A与实体集B具有多对多联系，记为m:n。

多对多联系包含了一对多联系。

例如：

✓ 学生选课，一门课同时有若干个学生选修，而一个学生可以同时选修多门课程，则课程与学生之间具有多对多联系。

✓ 教师与学生，工厂与产品



1.2.2 概念模型

.....

◆概念模型的用途

概念模型用于信息世界的建模

是现实世界到机器世界的一个中间层次

是数据库设计的有力工具

数据库设计人员和用户之间进行交流的语言

◆对概念模型的基本要求

较强的语义表达能力

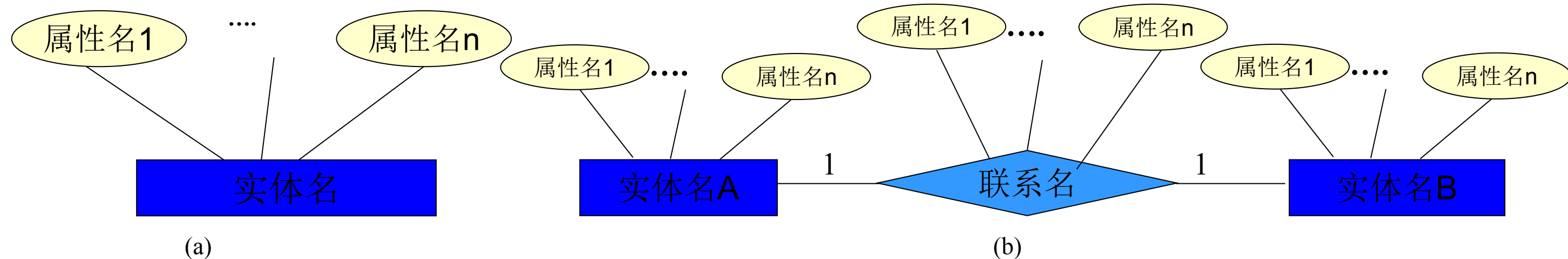
简单、清晰、易于用户理解



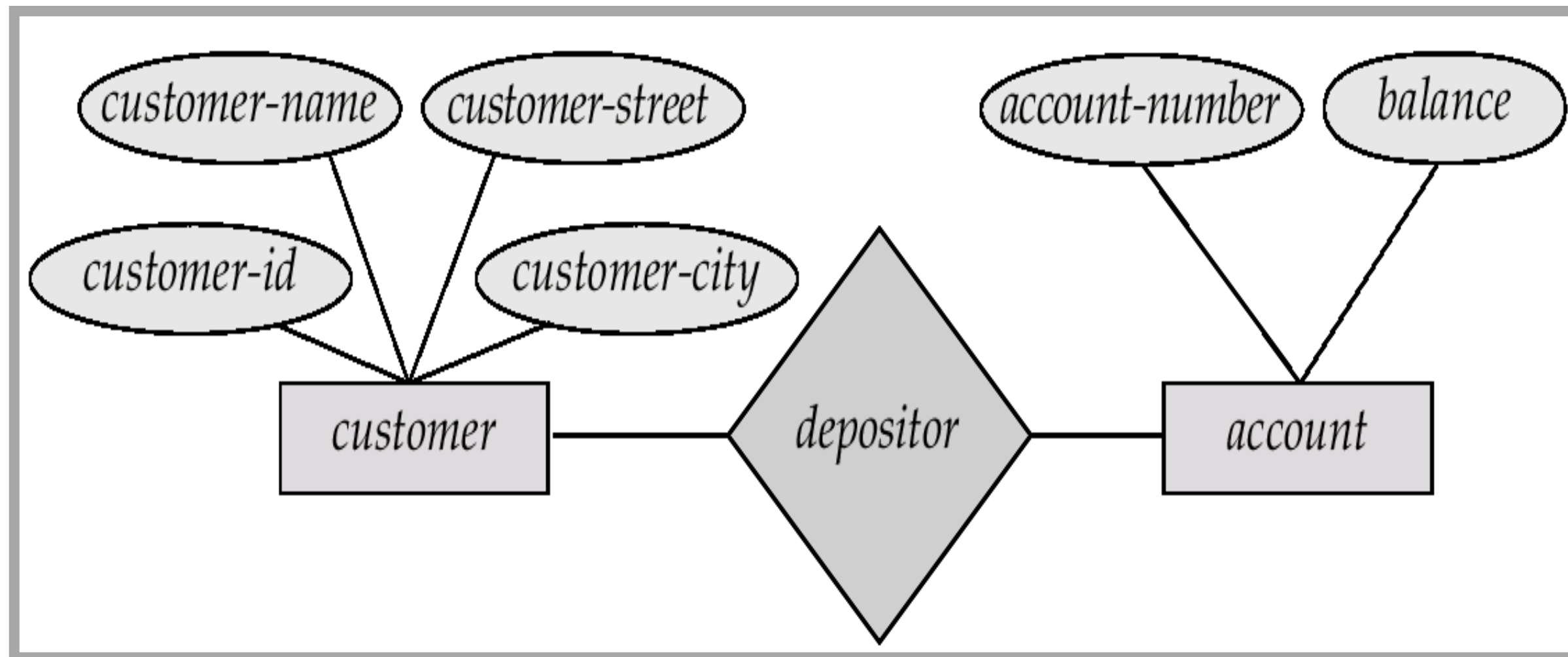
概念模型的表示方法：实体联系表示法——E-R图

E-R图的表示方法：

- ❖ 实体用方框表示。
- ❖ 联系用菱形表示，并且用边将其与有关的实体连接起来，并在边上标有联系的类型；
- ❖ 属性用椭圆表示，并且用边将其与相应的实体连接起来。
- ❖ 对于有些联系，其自身也会有某些属性，同实体与属性的连接类似，将联系与其属性用边连接起来。



Entity-Relationship (E-R) Model (Cont.)



1.2 数据模型

1.2.1 两大类数据模型

1.2.2 概念模型

1.2.3 数据模型的组成要素

1.2.4 最常用的数据模型

1.2.5 层次模型

1.2.6 网状模型

1.2.7 关系模型



❖数据模型的组成要素

一般说，数据模型是严格定义的一组概念的集合，这些概念精确地描述了数据库系统的静态特性、动态特性和完整性约束条件。因此，数据模型通常由数据结构、数据操作和数据的完整性约束三部分组成。

- 数据结构描述了数据库的组成对象以及对象间的关系，即数据结构一方面描述数据对象的类型、内容、性质等，另一方面描述数据对象间的联系。数据结构是对数据库系统静态特性的描述。
- 数据操作是指对数据库中各种对象（型）的实例（值）允许执行的操作的集合，是对数据库系统动态特性的描述。数据库有查询和更新（插入、修改、删除）两类操作。
- 数据的完整性约束是一组完整性规则的集合。完整性规则是给定的数据模型中数据及其联系所具有的制约和依存规则，用来限定符合数据模型的数据库状态以及状态的变化，以保证数据的正确、有效、相容。



1.2 数据模型

1.2.1 两大类数据模型

1.2.2 数据模型的组成要素

1.2.3 概念模型

1.2.4 常用的数据模型

1.2.5 层次模型

1.2.6 网状模型

1.2.7 关系模型



.....1.2.4 常用的数据模型.....

层次模型 (Hierarchical Model)

网状模型 (Network Model)

关系模型 (Relational Model))

面向对象数据模型 (Object Oriented Data Model)

对象关系数据模型 (Object Relational Data Model)

半结构化数据模型 (Semistruture Data Model)



1.2 数据模型

1.2.1 两大类数据模型

1.2.2 数据模型的组成要素

1.2.3 概念模型

1.2.4 常用的数据模型

1.2.5 层次模型

1.2.6 网状模型

1.2.7 关系模型

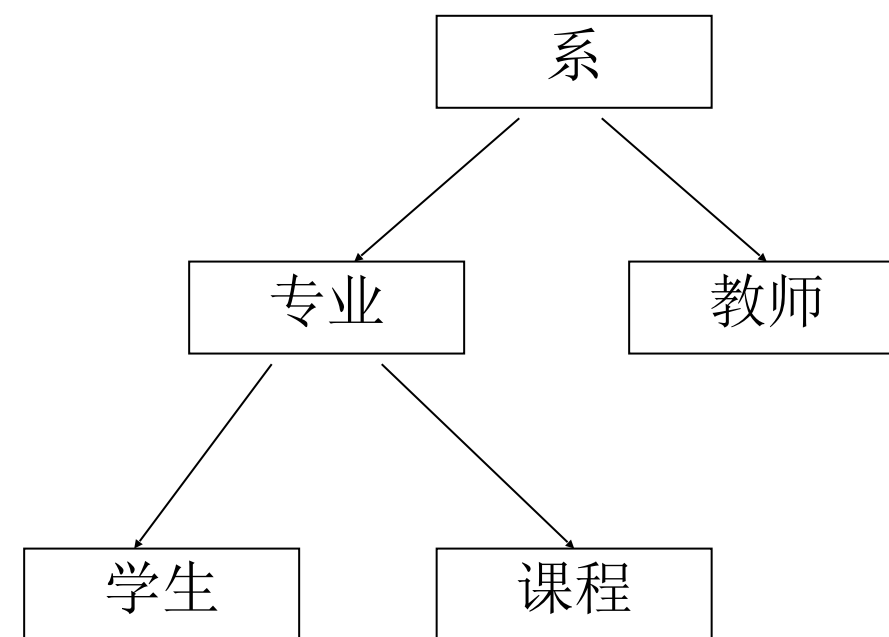


1.7.1 层次模型

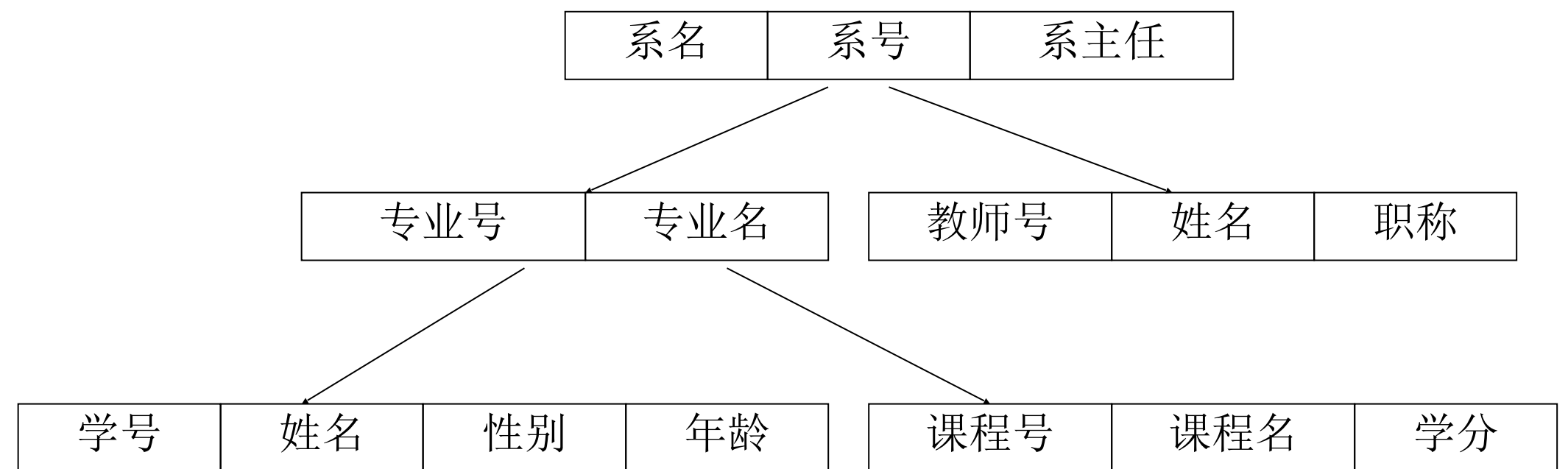
定义：用树型（层次）结构表示实体类型及实体间联系的数据模型称为层次数据模型。

(1) 数据结构

有且仅有一个结点没有父结点，它就是根结点；根以外的其他结点有且仅有一个双亲结点。父子节点之间的联系是一对多的联系。



(a) 实体之间的联系



(b) 实体型之间的联系

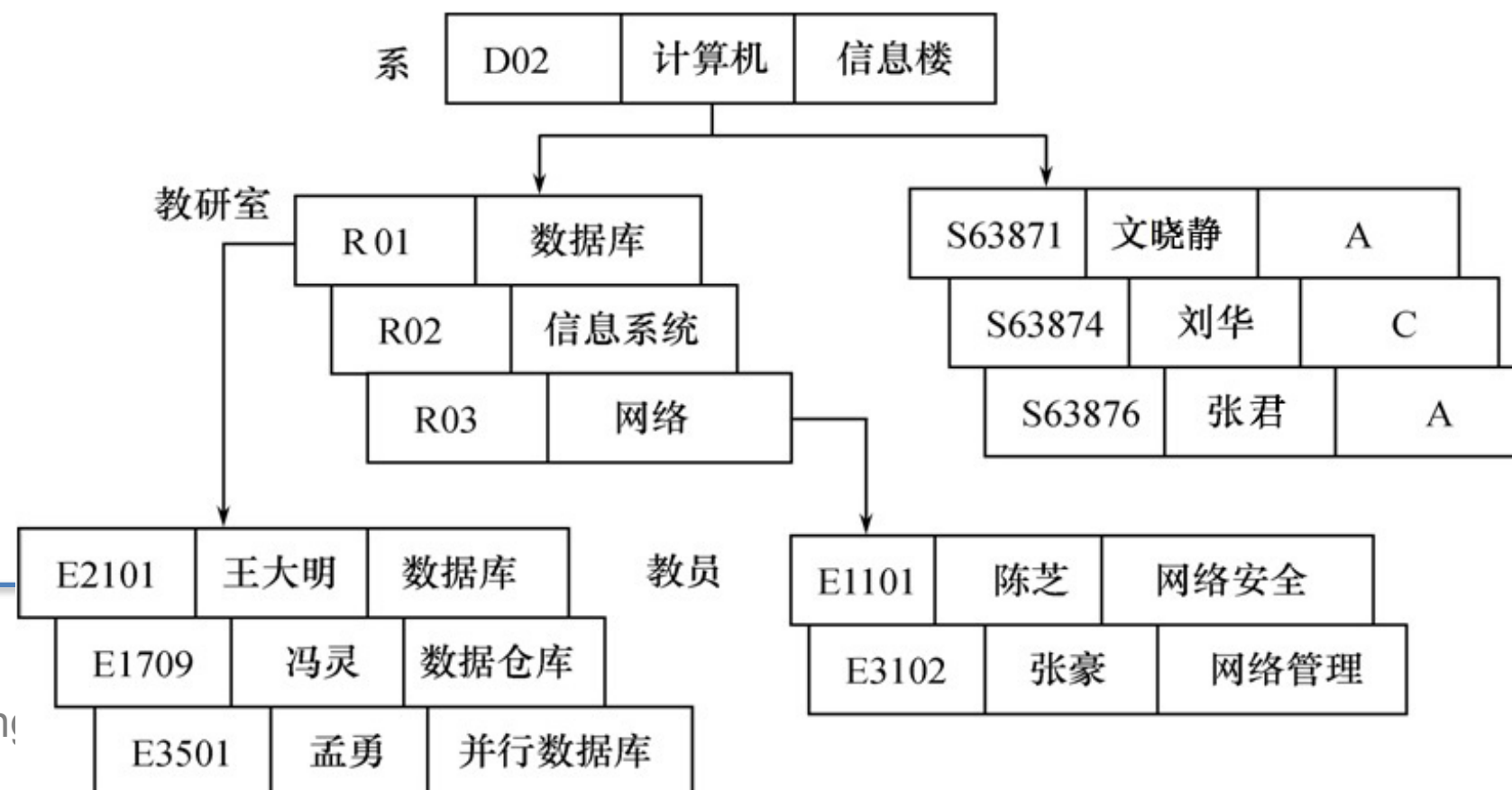


(2) 层次模型的数据操纵与完整性约束

层次模型的数据操纵主要包括查询、插入、删除和修改，进行插入、删除和修改操作时要满足层次模型的完整性约束条件。

■ 进行插入操作时，如果没有相应的双亲结点值就不能插入子女结点值。

■ 进行删除操作时，如果删除了双亲结点值，则相应的子女结点值也被同时删除。



(3) 层次模型的优缺点

优点:

结构简单，层次分明，便于在计算机内实现；

结点间联系简单，从根结点到树中任一结点均存在一条唯一的层次路径，当要存取某个结点的记录值时，沿着这条路径很快就能找到该记录值，因此，以该种模型建立的数据库系统查询效率很高。提供了良好的数据完整性支持。

缺点:

不能直接表示两个以上的实体型间的复杂联系和实体型间的多对多联系，只能通过引入冗余数据或创建虚拟结点的方法解决，容易产生不一致性。

对数据插入和删除的操作限制太多，比如，查询子女结点必须通过双亲结点。



1.7.2 网状模型

定义：用有向图（网状结构）表示实体及实体之间联系的数据模型称为网状数据模型。

（1）数据结构

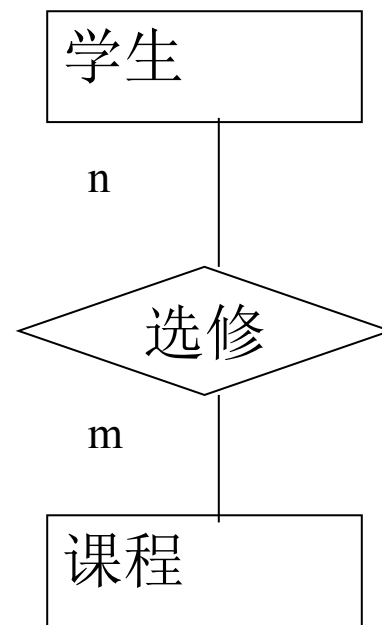
有一个以上的结点没有双亲结点。

允许结点有多个双亲结点。

允许两个结点之间有多种联系。如学生和课程间的联系。



用有向图表示的网状模型中，每个结点是记录类型，箭头表示从箭尾的记录类型到箭头的记录类型间的联系是1:n联系。



学生选课E-R图

学生记录类型S

学号	姓名	性别	年龄
----	----	----	----

课程记录类型C

课程号	课程名	学分
-----	-----	----

选课联系记录类型L

学号	课程号	学分
----	-----	----

SL

CL

学生选课网状模型

在网状模型中，每个n: m联系用两个1:n的联系来实现。学生选课网状模型图有3个结点两条边组成，学生和课程间的n: m联系用边SL、CL实现。



(2) 网状模型的数据操纵与完整性约束

网状模型的数据操纵主要包括查询、插入、删除和修改，进行插入、删除和修改操作时要满足网状模型的完整性约束条件。

- 插入数据时，允许插入尚未确定双亲结点值的子女结点值。如，可增加一名尚未分配到某个教研室的新教师。
- 删除数据时，允许只删除双亲结点值。如可删除一个教研室，而该教研室所有老师的信息仍保留在数据库中。
- 修改数据时，只需更新指定记录即可。
- 网状模型没有像层次模型那样有严格的完整性约束条件，只提供一定的完整性约束。



(3) 网状模型的优缺点

优点:

能更为直接地描述客观世界，可表示实体间的多种复杂联系。

具有良好的性能和存储效率。

缺点:

数据结构复杂，并且随着应用环境的扩大，数据库的结构变得越来越复杂，不便
于终端用户掌握。

其数据定义语言（DDL）、数据操纵语言（DML）极其复杂，不易使用户掌握。

由于记录间的联系本质上是通过存取路径实现的，应用程序在访问数据库时要指
定存取路径，即用户要了解系统结构的细节，加重了编写应用程序的负担。



1.7.3 关系模型

定义：用二维表格结构来表示实体及实体之间联系的数据模型称为关系模型。

(1) 数据结构及有关概念

关系模型的数据结构是一张规范化的二维表，它有表名、表头、表体。表名是二维表的名称，表头决定了二维表的结构，表体即二维表中的数据。每个二维表称为关系。关系模型与层次模型、网状模型不同，它是建立在严格的数学概念之上。

学 号	姓 名	年 龄	性 别	系 名	年 级
2013004	王小明	19	女	社会学	2013
2013006	黄大鹏	20	男	商品学	2013
2013008	张文斌	18	女	法律	2013
...



S (学生关系)

SNo 学号	SN 姓名	Sex 性别	Age 年龄	Dept 系别
S1	赵亦	女	17	计算机
S2	钱尔	男	18	信息
S3	孙珊	女	20	信息
S4	李思	男	21	自动化
S5	周武	男	19	计算机
S6	吴丽	女	20	自动化

SC (选课关系)

SNo 学号	CNo 课程号	Score 成绩
S1	C1	90
S1	C2	85
S2	C5	57
S2	C6	80
S2	C7	
S2	C4	70
S3	C1	75
S3	C2	70
S3	C4	85
S4	C1	93
S4	C2	85
S4	C3	83
S5	C2	89

C (课程关系)

CNo 课程号	CN 课程名	CT 课时
C1	程序设计	60
C2	微机原理	80
C3	数字逻辑	60
C4	数据结构	80
C5	数据库	60
C6	编译原理	60
C7	操作系统	60



关系必须是规范化的，满足一定的规范条件

最基本的规范条件：关系的每一个分量必须是一个不可分的数据项，不允许表中还有表

图1.15中工资和扣除是可分的数据项，不符合关系模型要求

职工号	姓名	职称	工 资			扣 除		实 发
			基本工资	岗位津贴	业绩津贴	三险	个人所得税	
86051	陈平	讲师	1305	1200	1850	160	112	4083

图1.15 一个工资表（表中有表）实例



关系模型中基本概念:

- **关系**: 一个关系就是一张二维表。(对应信息世界的实体集)
- **元组**: 表中的一行即为一个元组, 也称为一个记录。(对应信息世界的实体)
- **属性**: 表中的列称为属性, 或称为字段, 每一列有一个属性名。(对应信息世界的属性。)
- **域**: 属性的取值范围。(对应信息世界的域)
- **分量**: 每一行对应的列的属性值, 即为元组中的一个属性值。
- **关系模式**: 对关系的描述称为关系模式, 关系模式是关系的型, 是关系的框架。格式为:

关系名 (属性名1, 属性名2, ..., 属性名n)



关系模型的数据结构（续）

表1.5 术语对比

关系术语	一般表格的术语
关系名	表名
关系模式	表头（表格的描述）
关系	（一张）二维表
元组	记录或行
属性	列
属性名	列名
属性值	列值
分量	一条记录中的一个列值
非规范关系	表中有表（大表中嵌有小表）



2. 关系模型的操纵与完整性约束

◆数据操作是集合操作，操作对象和操作结果都是关系

查询

插入

删除

更新

•存取路径对用户隐蔽，用户只要指出“干什么”，不必详细说明

“怎么干”



3. 关系模型的优缺点

◆优点

- 建立在严格的数学概念的基础上
- 概念单一

实体和各类联系都用关系来表示

对数据的检索结果也是关系

- 关系模型的存取路径对用户隐蔽

具有更高的数据独立性，更好的安全保密性

简化了程序员的工作和数据库开发建立的工作



关系模型的优缺点（续）

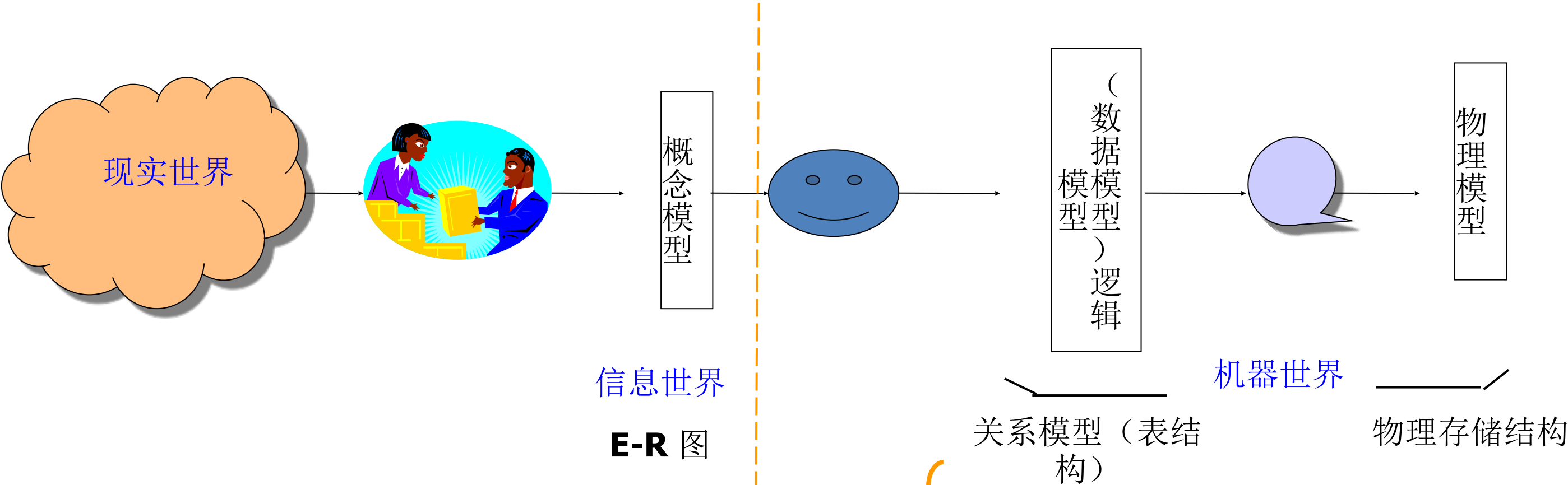
.....

◆缺点

- 存取路径对用户隐蔽，查询效率往往不如格式化数据模型
- 为提高性能，必须对用户的查询请求进行优化，增加了开发数据库管理系统的难度



Steps of Database Design



86

DBTG模型 | CODASYL系统
网状结构 (一个结点可有多多个双亲, 允许多个结点无双亲)



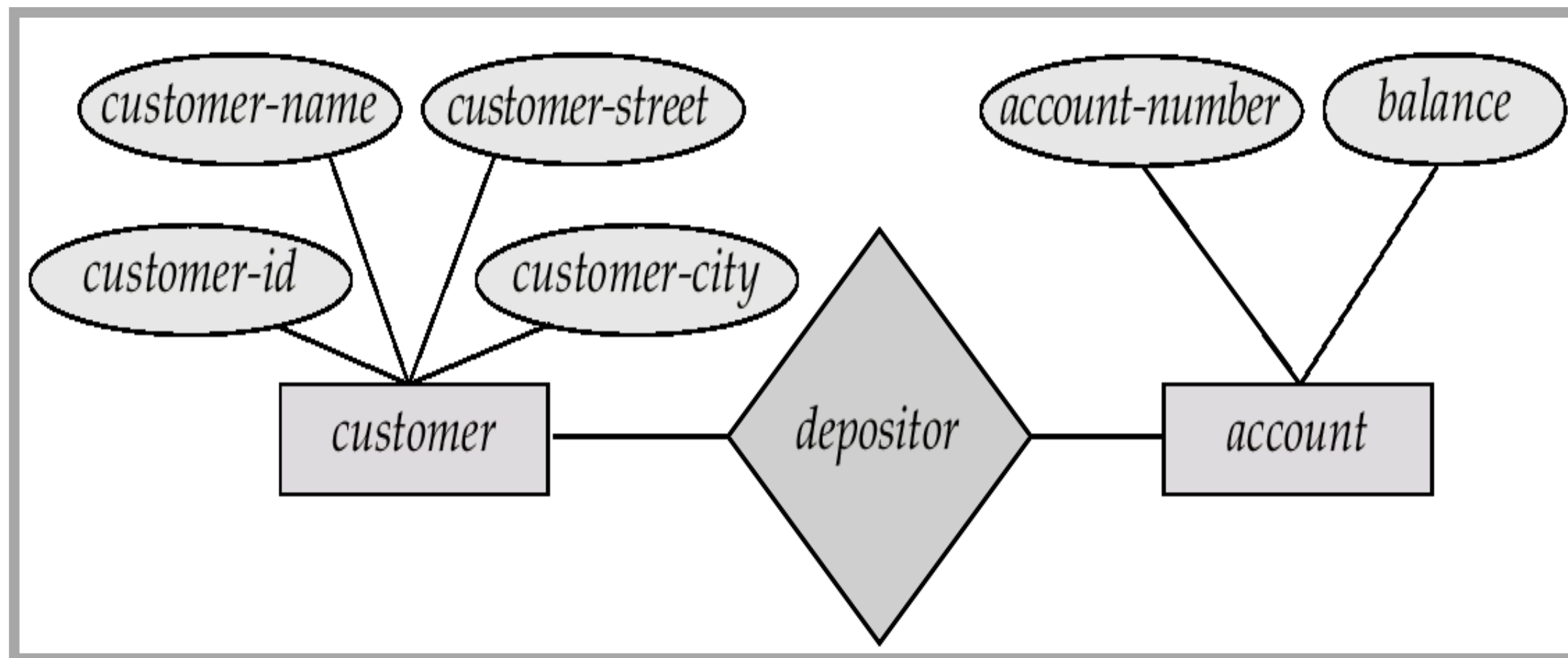
China university of Mining and Technology-Beijing

Honeywell IDS/2, HP image

面向对象模型
--非关系模型--
层次模型
网状模型

1968, IBM IMS
树型结构
(除根结点外, 每个结点有且只有一个双亲结点)

Entity-Relationship (E-R) Model (Cont.)



A Sample Relational Database

<i>customer-id</i>	<i>customer-name</i>	<i>customer-street</i>	<i>customer-city</i>
192-83-7465	Johnson	12 Alma St.	Palo Alto
019-28-3746	Smith	4 North St.	Rye
677-89-9011	Hayes	3 Main St.	Harrison
182-73-6091	Turner	123 Putnam Ave.	Stamford
321-12-3123	Jones	100 Main St.	Harrison
336-66-9999	Lindsay	15	
019-28-3746	Smith	72	

(a) The *customer* table

<i>account-number</i>	<i>balance</i>
A-101	500
A-215	700
A-102	400
A-305	350
A-201	900
A-217	750
A-222	700

(b) The *account* table

<i>customer-id</i>	<i>account-number</i>
192-83-7465	A-101
192-83-7465	A-201
019-28-3746	A-215
677-89-9011	A-102
182-73-6091	A-305
321-12-3123	A-217
336-66-9999	A-222
019-28-3746	A-201

(c) The *depositor* table

第一章 绪论

.....

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的结构

1.4 数据库系统的组成

1.5 小结



1.3 数据库系统的结构

.....

从数据库应用开发人员角度看，数据库系统通常采用三级模式结构，是数据库系统内部的系统结构

1.3.1 数据库系统模式的概念

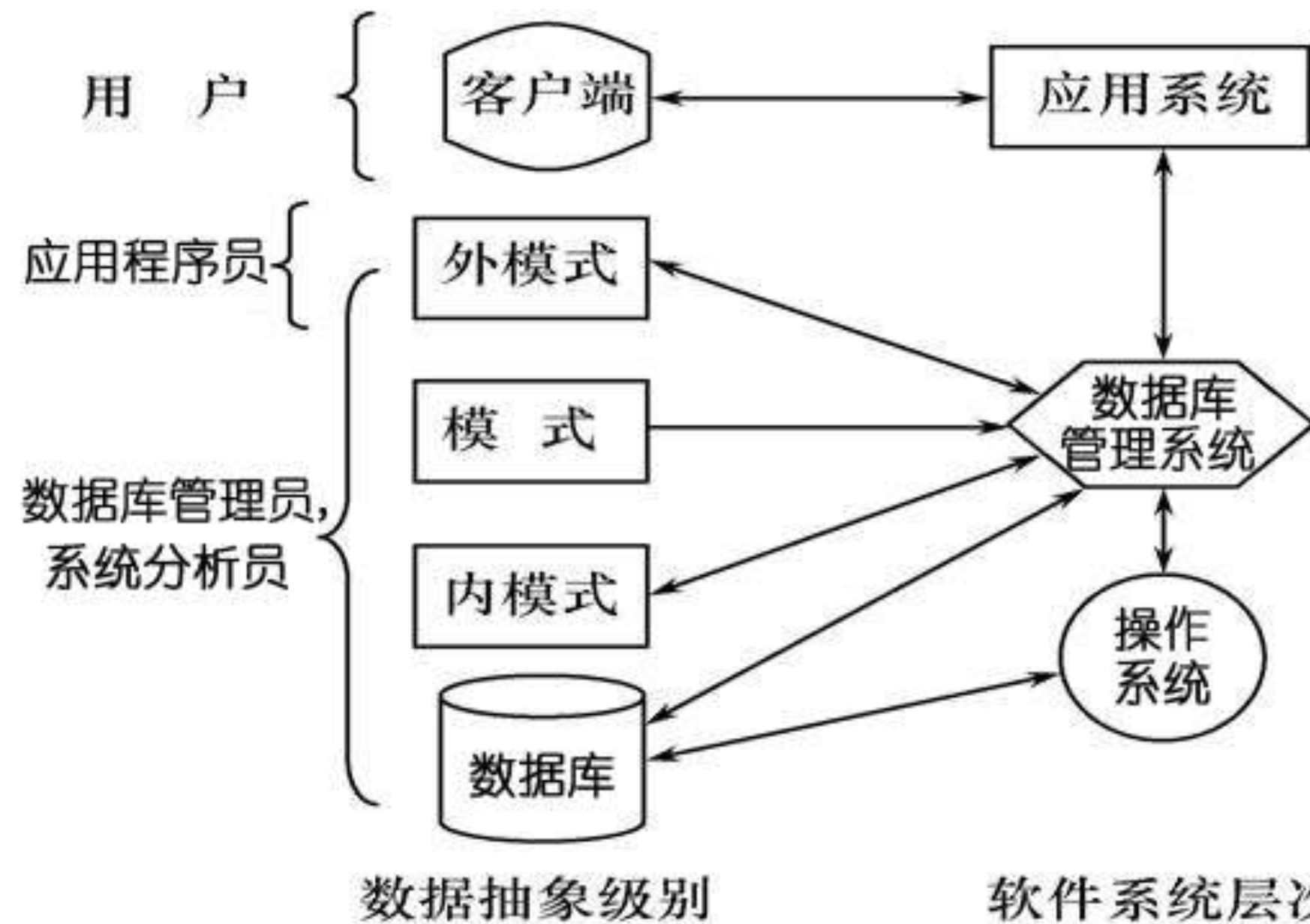
1.3.2 数据库系统的三级模式结构

1.3.3 数据库的二级映像功能与数据独立性



人员 (续)

- 不同的人员涉及不同的数据抽象级别，具有不同的数据视图，如下图所示



1.3.1 数据库系统模式的概念

一、“型”和“值”的概念

- 型 (Type)

对某一类数据的结构和属性的说明

- 值 (Value)

是型的一个具体赋值

- 例如

学生记录:

(学号, 姓名, 性别, 系别, 年龄, 籍贯)

一个记录值:

(201315130, 李明, 男, 计算机系, 19, 江苏南京市)



数据库系统模式的概念（续）

二、模式（Schema）和实例（Instance）

- 模式（Schema）

数据库逻辑结构和特征的描述

是型的描述，不涉及具体值

反映的是数据的结构及其联系

模式是相对稳定的

- 实例（Instance）

模式的一个具体值

反映数据库某一时刻的状态

同一个模式可以有很多实例

实例随数据库中的数据的更新而变动



数据库系统结构（续）

.....

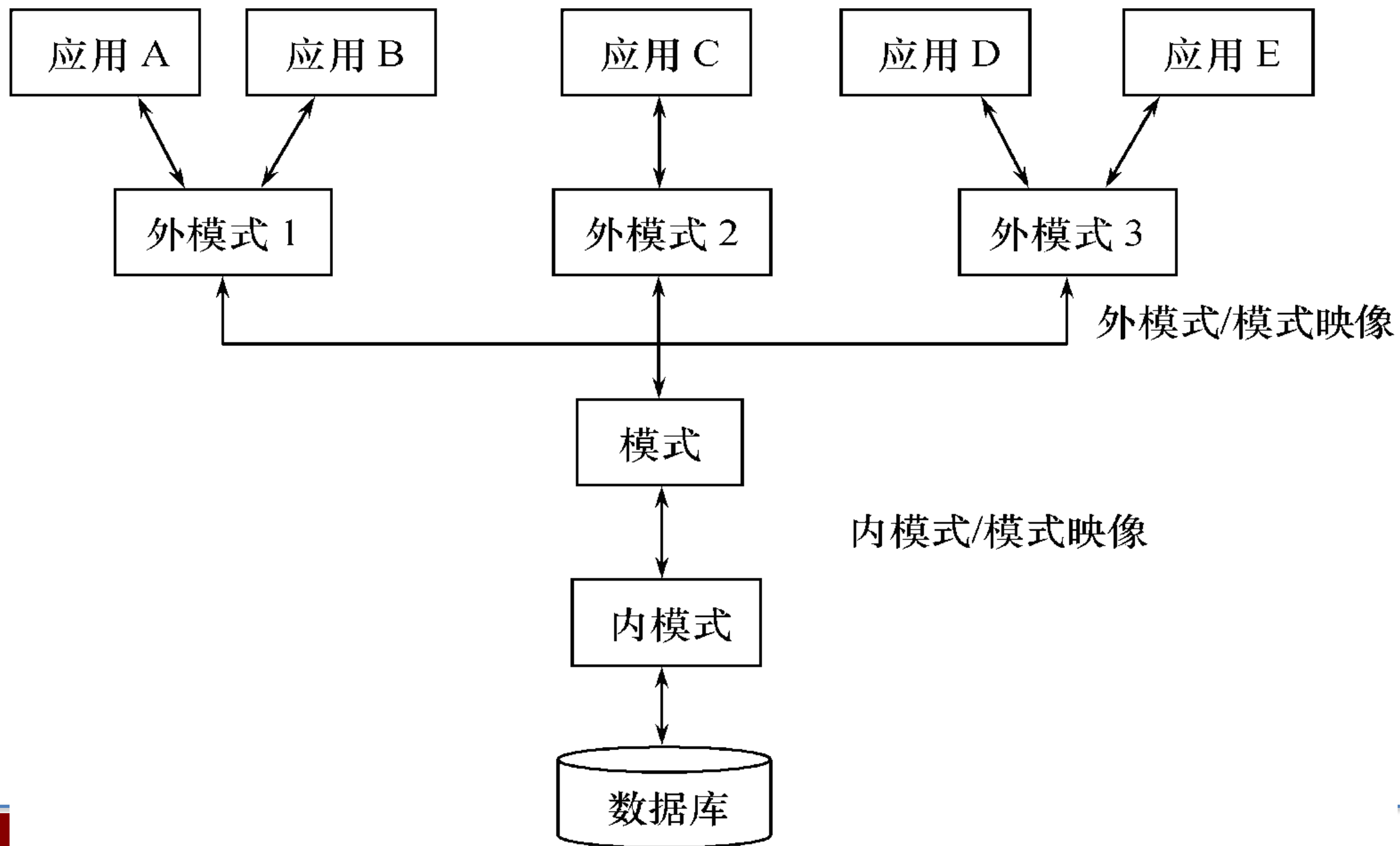
1.3.1 数据库系统模式的概念

1.3.2 数据库系统的三级模式结构

1.3.3 数据库的二级映像功能与数据独立性



1.3.2 数据库系统的三级模式结构



1. 模式 (Schema)

- 模式 (也称逻辑模式)

数据库中**全体**数据的逻辑结构和特征的描述；所有用户的公共数据视图；一个数据库只有一个模式

- 模式的地位：是数据库系统模式结构的中间层

与数据的物理存储细节和硬件环境无关

与具体的应用程序、开发工具及高级程序设计语言无关

- 模式的定义

数据的逻辑结构 (数据项的名字、类型、取值范围等)

数据之间的联系

数据有关的安全性、完整性要求



China University of
Mining and Technology-Beijing

- 外模式（也称子模式或用户模式）

数据库用户（包括应用程序员和最终用户）使用的局部数据的逻辑结构和特征的描述；数据库用户的数据视图，是与某一应用有关的数据的逻辑表示

- 外模式的地位：介于模式与应用之间

- 模式与外模式的关系：一对多。外模式通常是模式的子集；一个数据库可以有多个外模式。反映了不同的用户的应用需求、看待数据的方式、对数据保密的要求；对模式中同一数据，在外模式中的结构、类型、长度、保密级别等都可以不同

- 外模式与应用的关系：一对多。同一外模式也可以为某一用户的多个应用系统所使用，但一个应用程序只能使用一个外模式

- 外模式的用途

保证数据库安全性的一个有力措施；每个用户只能看见和访问所对应的外模式中的



Another example: University Database

Achievement

Sid	Sname	Cname	credit	grade
3023001093	Tom	DB	4	92
3023001093	Tom	OS	5	88
3020621034	Jack	DB	4	70
3020831035	Smith	DB	4	85
3021131123	Alane	OS	5	95

view level
(subschema)

Logical level
(schema)

Sid	Sname	Ssex	Sage	sdept
3023001093	Tom	M	21	Cs
3011112340	Mary	F	20	Cs
3020621034	Jack	M	18	Cs
3020831035	Smith	M	19	Ma
3021131123	Alane	F	22	Is

student

course

cid	Cname	credit
1	DB	4
2	OS	5
3	English	4
4	Math	4

Enrolled

sid	cid	grade
3023001093	1	92
3023001093	2	88
3020621034	1	70
3020831035	1	85
3021131123	2	95



3. 内模式 (Internal Schema)

- 内模式 (也称存储模式)
 - 是数据物理结构和存储方式的描述
 - 是数据在数据库内部的表示方式
 - 记录的存储方式 (例如, 顺序存储, 按照B树结构存储, 按hash方法存储等)
 - 索引的组织方式
 - 数据是否压缩存储
 - 数据是否加密
 - 数据存储记录结构的规定
- 一个数据库只有一个内模式



1.3.2 数据库系统的三级模式结构

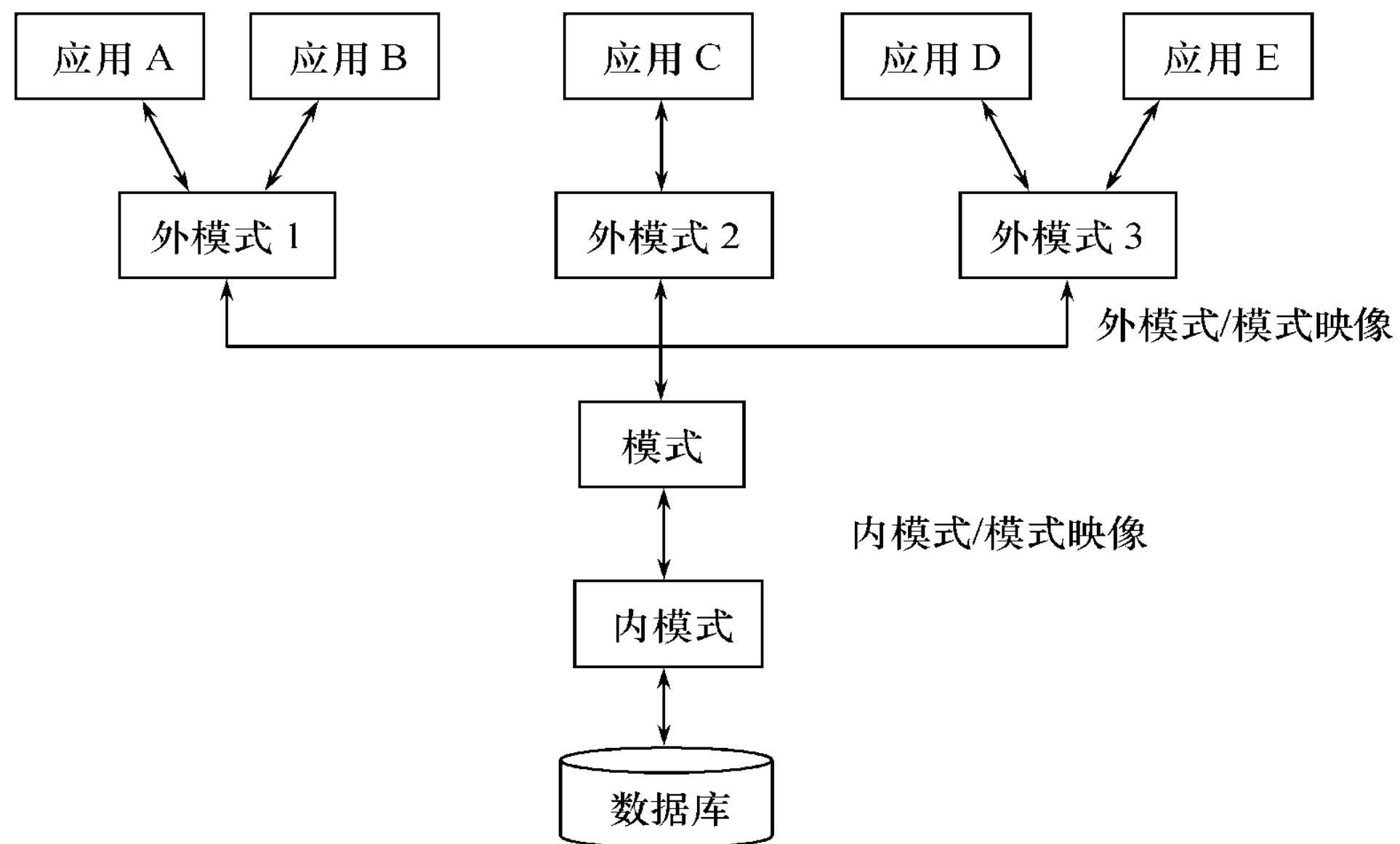


图1.16 数据库系统的三级模式结构



数据库系统结构（续）

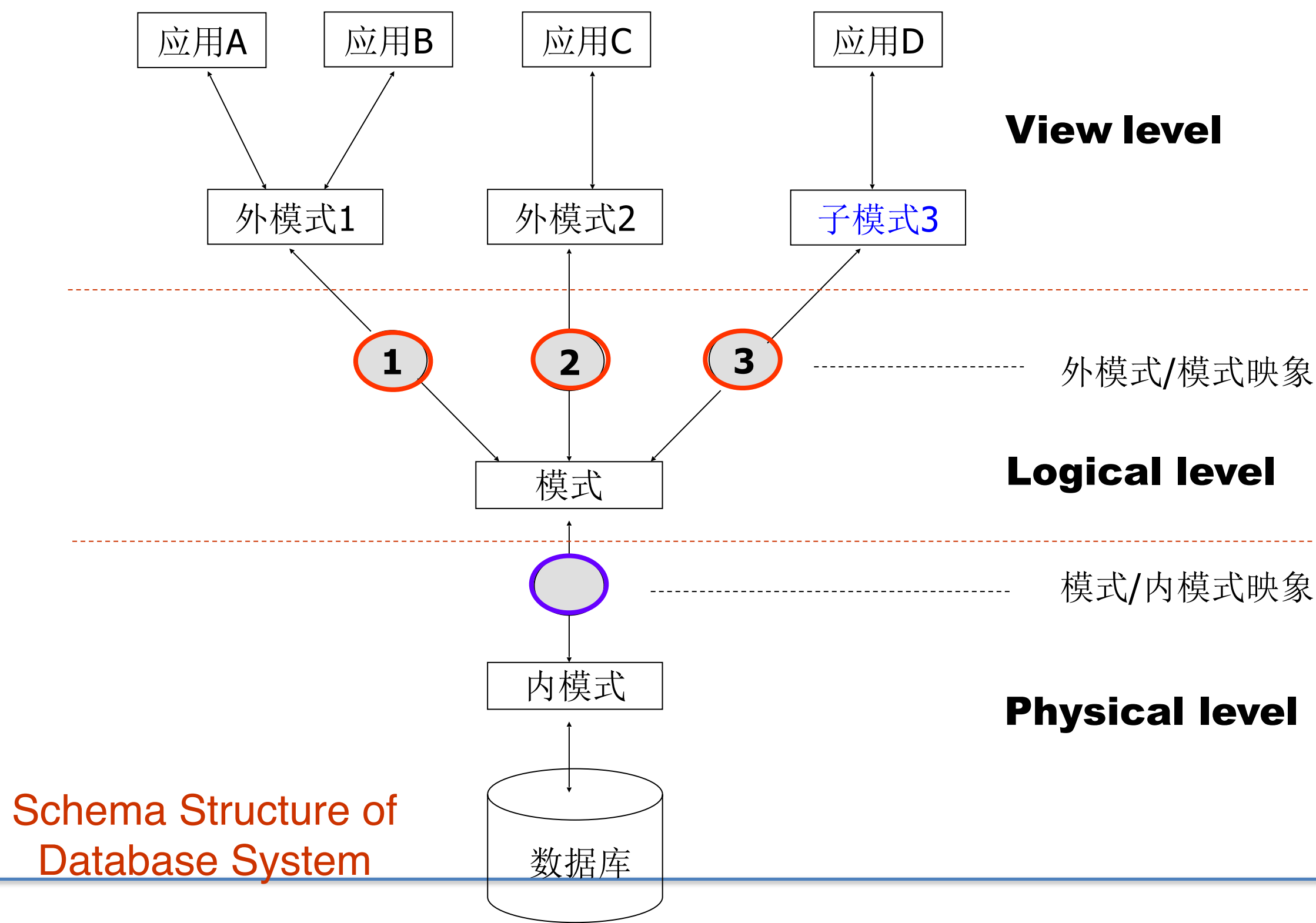
1.3.1 数据库系统模式的概念

1.3.2 数据库系统的三级模式结构

1.3.3 数据库的二级映像功能与数据独立性



数据库的二级映像功能



Schema Structure of Database System



1. 外模式 / 模式映像

- 模式：描述的是数据的全局逻辑结构
- 外模式：描述的是数据的局部逻辑结构
- 同一个模式可以有任意多个外模式
- 每一个外模式，数据库系统都有一个外模式 / 模式映像，定义外模式与模式之间的对应关系
- 映像定义通常包含在各自外模式的描述中



外模式 / 模式映像（续）

◆保证数据的逻辑独立性

当模式改变时，数据库管理员对外模式 / 模式映像作相应改变，使外模式保持不变

应用程序是依据数据的外模式编写的，应用程序不必修改，保证了数据与程序的逻辑独立性，简称数据的逻辑独立性



Another example: University Database

Achievement

Sid	Sname	Cname	credit	grade
3023001093	Tom	DB	4	92
3023001093	Tom	OS	5	88
3020621034	Jack	DB	4	70
3020831035	Smith	DB	4	85
3021131123	Alane	OS	5	95

view level
(subschema)

Logical level
(schema)

Sid	Sname	Ssex	Sage	sdept
3023001093	Tom	M	21	Cs
3011112340	Mary	F	20	Cs
3020621034	Jack	M	18	Cs
3020831035	Smith	M	19	Ma
3021131123	Alane	F	22	Is

student

course

cid	Cname	credit
1	DB	4
2	OS	5
3	English	4
4	Math	4

Enrolled

sid	cid	grade
3023001093	1	92
3023001093	2	88
3020621034	1	70
3020831035	1	85
3021131123	2	95

105



2. 模式 / 内模式映像

- 模式 / 内模式映像定义了数据全局逻辑结构与存储结构之间的对应关系。

例如，说明逻辑记录和字段在内部是如何表示的

- 数据库中模式 / 内模式映像是唯一的
- 该映像定义通常包含在模式描述中



模式 / 内模式映像（续）

◆保证数据的物理独立性

当数据库的存储结构改变了（例如选用了另一种存储结构），数据库管理员修改模式 / 内模式映像，使模式保持不变。

应用程序不受影响。保证了数据与程序的物理独立性，简称数据的物理独立性。



数据独立性

- 数据与程序之间的独立性，使得数据的定义和描述可以从应用程序中分离出去
- 数据的存取由数据库管理系统管理
 - 简化了应用程序的编制
 - 大大减少了应用程序的维护和修改



第一章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的结构

1.4 数据库系统的组成

1.5 小结



数据库系统的组成（续）

.....

1.硬件平台及数据库

- (1) 足够大的内存
- (2) 足够的大的磁盘或磁盘阵列等设备
- (3) 较高的通道能力，提高数据传送率

2.软件

- (1) 数据库管理系统
- (2) 支持数据库管理系统运行的操作系统
- (3) 与数据库接口的高级语言及其编译系统
- (4) 以数据库管理系统为核心的应用开发工具

3.人员

- (1) 数据库管理员
- (2) 系统分析员和数据库设计人员
- (3) 应用程序员
- (4) 最终用户

- (5) 为特定应用环境开发的数据库应用系统



1.数据库管理员 (DBA)

具体职责：

1. 决定数据库中的信息内容和结构
2. 决定数据库的存储结构和存取策略
3. 定义数据的安全性要求和完整性约束条件
4. 监控数据库的使用和运行

周期性转储数据库

系统故障恢复

介质故障恢复



监视审计文件

2. 系统分析员和数据库设计人员

系统分析员

负责应用系统的需求分析和规范说明

与用户及数据库管理员结合，确定系统的软硬件配置

参与数据库系统的概要设计

数据库设计人员

参加用户需求调查和系统分析

确定数据库中的数据

设计数据库各级模式

3. 应用程序员

设计和编写应用系统的程序模块

进行调试和安装



4. 用户

用户是指最终用户（End User）。最终用户通过应用系统的用户接口使用数据库。

1. 偶然用户

不经常访问数据库，但每次访问数据库时往往需要不同的数据库信息

企业或组织机构的高中级管理人员

2. 简单用户

主要工作是查询和更新数据库

银行的职员、机票预定人员、旅馆总台服务员

3. 复杂用户

工程师、科学家、经济学家、科技工作者等

直接使用数据库语言访问数据库，甚至能够基于数据库管理系统的应用程序接口编制自己的应用程序



China university of
Mining and Technology-Beijing

第一章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统的结构

1.4 数据库系统的组成

1.5 小结



1.5 小结

(1) 数据库系统概述

数据库的基本概念

数据管理的发展过程

数据库系统的特点

(2) 数据模型

数据模型的三要素

三种主要数据库模型



小结（续）

（3）数据库系统内部的系统结构

数据库系统三级模式结构

数据库系统两层映像系统结构

（4）数据库系统的组成

