



数据库原理与应用

教材：

- 王珊 萨师煊, 《**数据库系统概论（第五版）**》 高等教育出版社

参考书籍：

- 王珊 《**数据库系统简明教程**》 高等教育出版社
- 王亚平 《**数据库系统工程师教程**》 清华大学出版社

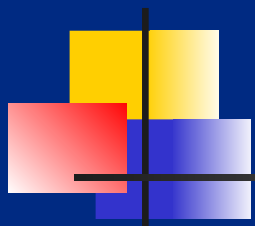


课程时间安排

- 理论课： **16 周** ， 共 **32 学时**
- 实验课： 共八次， 每次一个题目
- 考核办法
 - 闭卷笔试

成绩总结	应参加考试人数	114	人
	实际参加考试人数	110	人
	90-100	29	人 26.4%
	80-90	23	人 20.9%
	70-80	23	人 20.9%
	60-70	17	人 15.5%
	60分以下	22	人 20%

最高分	98
最低分	24
(缺考除外)	



教学内容

基本理论

数据库系统概论 (第 1 章)

关系数据库 (第 2 章)

关系数据库理论 (第 6 章)

应用技术

结构化查询语言 SQL (第 3 章)

安全性与完整性 (第 4、5 章)

数据库设计 (第 7 章)

数据库设计编程、事务管理等
(第 8、9、10 章)



第 1 章 绪论

1.1 数据库系统概述

1.2 数据模型

1.3 数据库系统结构

1.4 数据库系统的组成



数据库技术的重要地位

- ❖ 数据库是数据管理的最新技术，是计算机科学的重要分支
- ❖ 数据库技术在信息化社会中的重要作用
 - ❑ 信息系统的核心和基础技术
 - ❑ 日益广泛的应用

例如：管理信息系统 (MIS)、联机事务处理 (OLTP)、联机分析处理 (OLAP)、计算机辅助设计与制造 (CAD / CAM)、计算机集成制造系统 (CIMS)、办公信息系统 (OIS)、地理信息系统 (GIS).....and ... **Internet !!!**



1.1 数据库系统概述

1.1.1 数据、数据库、数据库管理系统、数据库系统

- 数据 (**Data**):

定义

描述事物的符号记录

种类

文字、图形、图象、音频、视频

特点

数据的表现形式不能完全表达其内容，
与其语义不可分



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ 数据 (Data):

例：学生档案中的学生记录

(李明, 1972 年 5 月, 江苏, 计算机系, 1990)

如果对应语义：学生姓名、出生年月、籍贯、所在系别、
入学时间

则解释为：学生李明, 1972 年 5 月出生, 是江苏人,
1990 年考入计算机系

如果对应语义：学生姓名、毕业年月、工作地点、所
在系别、电子建档日期

则解释为：学生李明, 1972 年 5 月毕业于计算机系,
现在江苏工作, 1990 年建立电子学
籍档案



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□数据库 (DataBase , DB)

定义

长期储存在计算机内、有组织的、可共享的
大量数据集合

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

1.1.1 数据、数据库、 数据库管理系统、数据库系统

数据库管理系统 (DataBase Management System , DBMS)

位于用户与操作系统之间的一层数据管理软件，用于科学组织和存储数据，高效获取和维护数据。

主要功能

1. 数据定义功能
2. 数据组织、存储和管理
3. 数据操纵功能
4. 数据库的事物管理和运行管理
5. 数据库的建立和维护功能
- 6 其他功能



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ 数据库管理系统 (DBMS)

- 1. 数据定义功能：**提供数据定义语言（Data Definition language）来定义数据对象。
- 2. 数据组织、存储和管理：**能分类组织、存储和管理这种数据。确定文件结构、存取方式、数据关系。
- 3. 数据操纵功能：**提供数据操纵语言（Data Manipulation language）实现用户对数据操作，如对数据的增、删、改和查询。



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ 数据库管理系统 (DBMS)

- 4. **数据库的事务管理和运行管理：**数据库在建立、运用和维护时统一由 DBMS 控制和管理。
- ✓ **数据的安全性保护：**保护数据，防止泄密和破坏；用户只能按照规定，对某些数据以某些方式进行使用和处理
- ✓ **数据的完整性检查：**检查数据的正确性、有效性和相容性。将数据控制在有效的范围内，保证数据之间满足一定的关系。
- ✓ **并发控制：**对多用户的并发操作加以控制和协调
- ✓ **数据库恢复：**具有将数据库从错误状态恢复到某一已知正确状态的功能。



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ 数据库管理系统 (DBMS)

5. 数据库的建立和维护功能:

- ✓ 数据库原始数据的输入、转换功能
- ✓ 数据库的转储、恢复功能
- ✓ 数据库的重组织和性能监视、分析功能



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ 数据库管理系统 (DBMS)

6. 其他功能

- ✓ DBMS 与网络中其他软件系统的通信功能
- ✓ 一个 DBMS 与另一个 DBMS 或文件系统的数据转换功能
- ✓ 异构数据库之间的互访和互操作功能



1.1.1 数据、数据库、 数据库管理系统、数据库系统

□ **数据库系统 (DataBase System, DBS)**

定义：在计算机系统中引入数据库后的系统

组成：数据库

DBMS(以及应用开发工具)

应用系统

数据库管理员 (DBA)

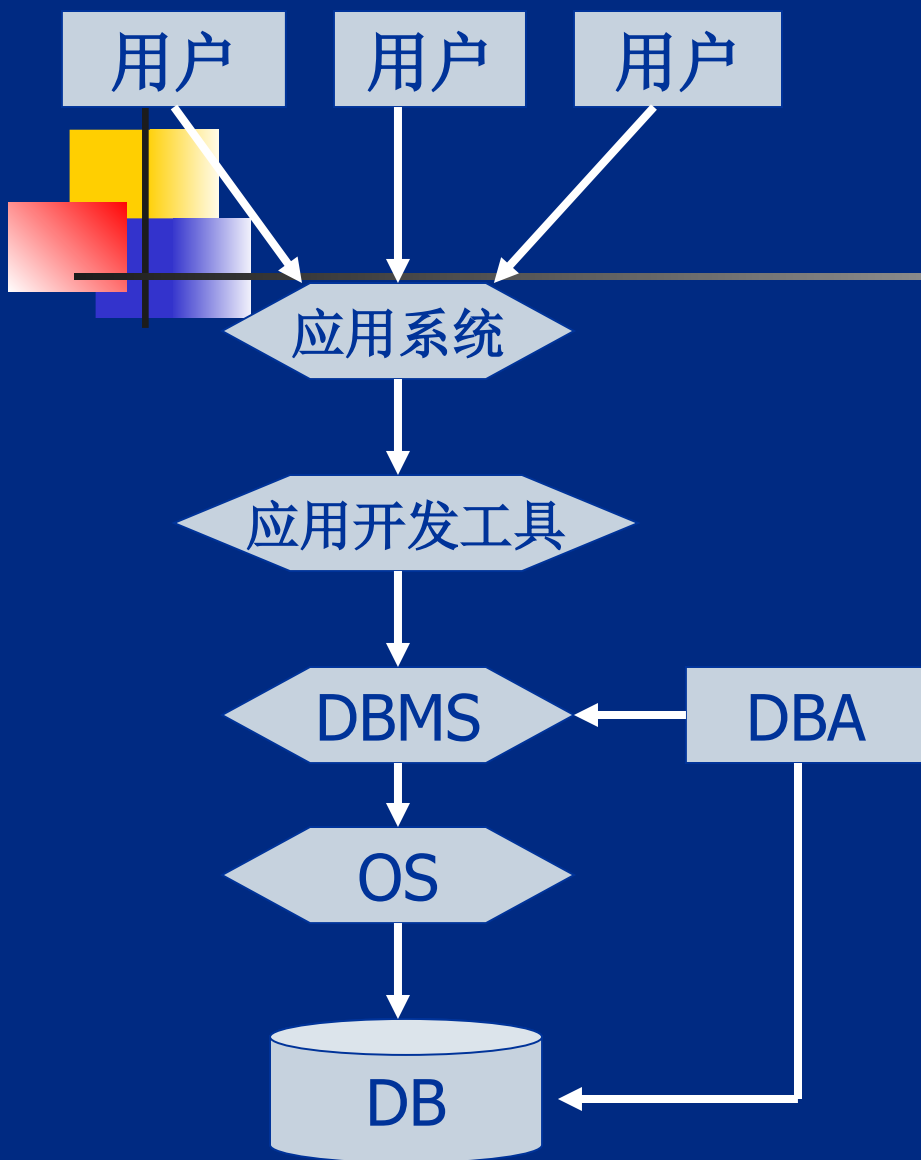


图 1 数据库系统 (DBS)

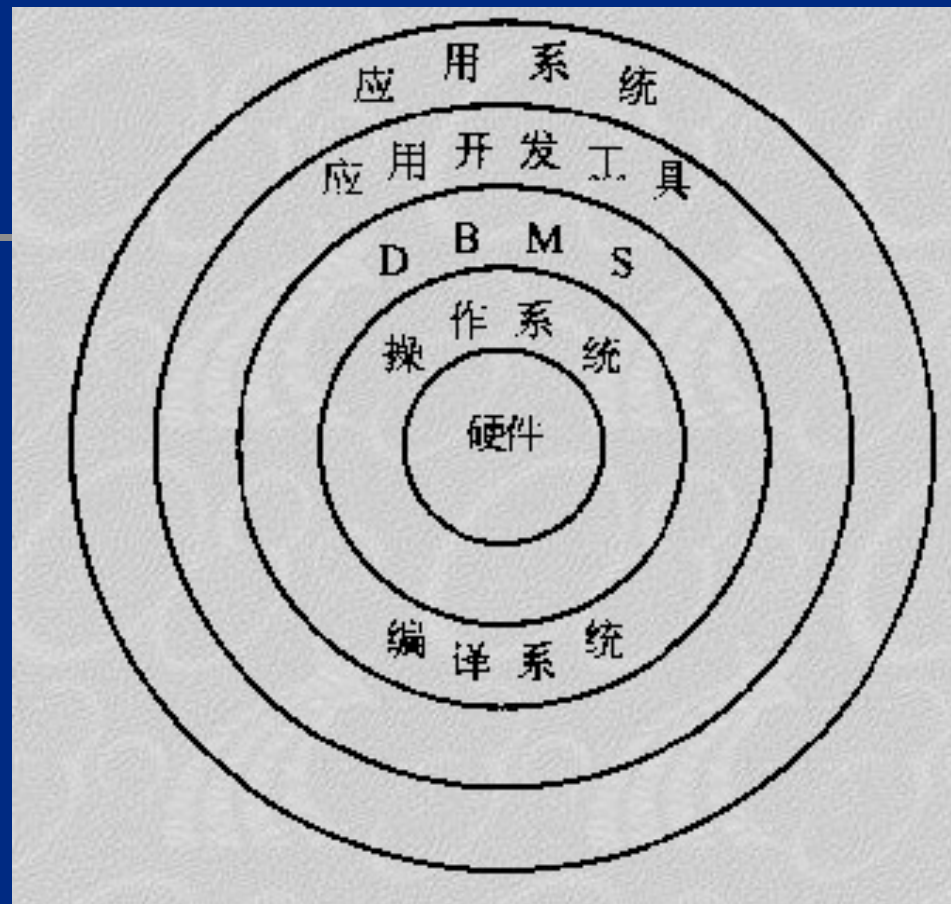


图 2 数据库在计算机系统中的地位



1.1.2 数据管理技术的产生和发展

❖ 数据库技术是应数据管理任务的需要而产生的。

数据处理是指对各种数据进行收集、存储、加工和传播的一系列活动的总和。**数据管理**则是指对数据进行分类、组织、编码、存储、检索和维护，它是数据处理的中心问题。

❖ 数据管理技术经历了人工管理、文件系统、数据库系统三个阶段。



1.1.2 数据管理技术的产生和发展

□ 人工管理阶段 (20 世纪 50 年代中期以前)

背景:

应用需求: 科学计算

硬件水平: 无直接存取存储设备

软件水平: 没有操作系统和管理数据的专门软件

处理方式: 批处理



1.1.2 数据管理技术的产生和发展

□ 人工管理阶段 (20 世纪 50 年代中期以前)

特点:

1) 数据不保存

2) 应用程序管理数据

数据需要由应用程序自己管理，没有相应的软件系统负责数据的管理工作。应用程序中不仅要规定数据的逻辑结构，而且要设计物理结构，包括存储结构、存取方法、输入方式等。因此程序员负担很重。

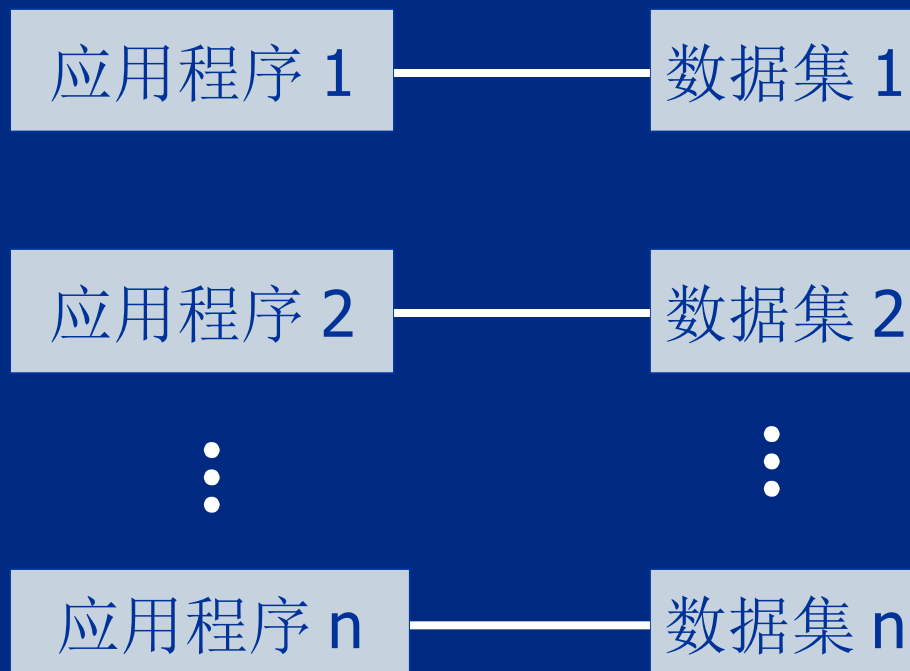
3) 数据不共享

数据是面向应用的，一组数据只能对应一个程序。

4) 数据不具有独立性

数据的逻辑结构或物理结构发生变化后，必须对应用程序做相应的修改。

1.1.2 数据管理技术的产生和发展



人工管理阶段应用程序与数据之间的对应关系



1.1.2 数据管理技术的产生和发展

□ 文件系统阶段 (50 年代后期— 60 年代中期)

背景:

应用需求	科学计算、管理
硬件水平	磁盘、磁鼓
软件水平	有文件系统
处理方式	联机实时处理、批处理

1.1.2 数据管理技术的产生和发展

□ 文件系统阶段 (50 年代后期— 60 年代中期)

特点:

(1) 数据可长期保存

(2) 由文件系统管理数据

按文件名访问，按记录进行存取，实现记录内的结构性而整体无结构。

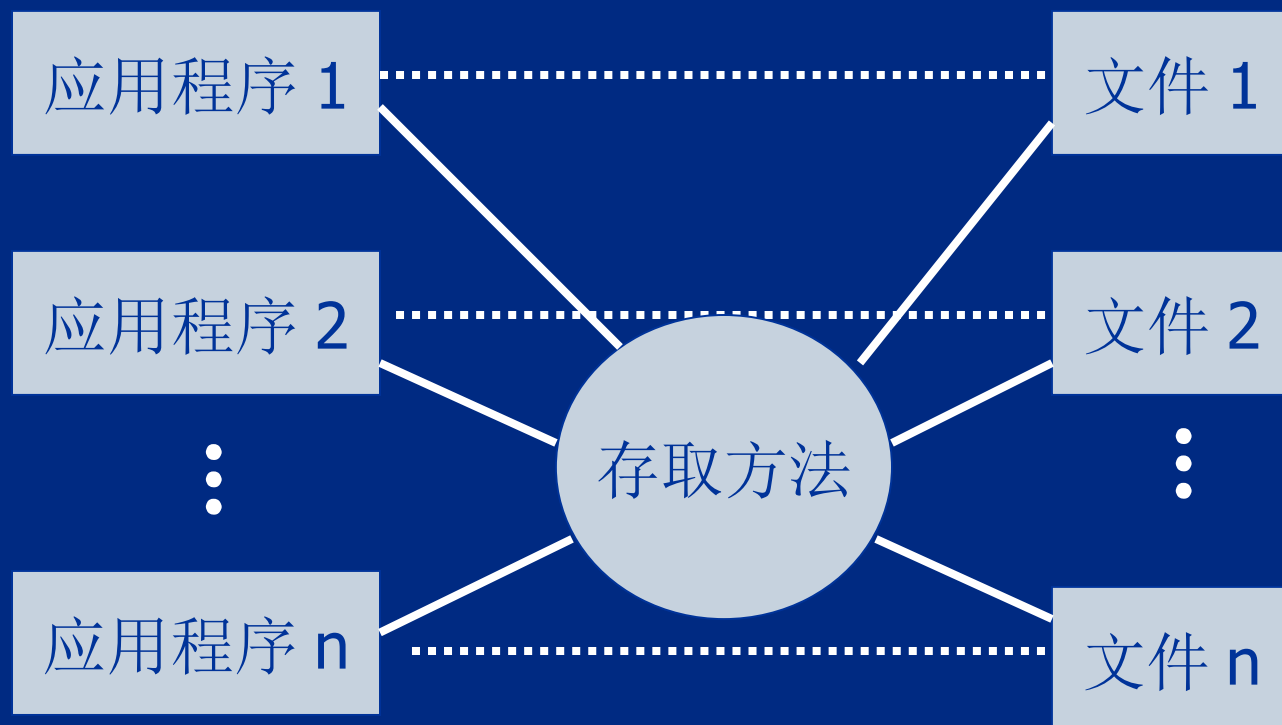
(3) 数据共享性差，冗余度大

文件仍然是面向应用的。当不同的应用程序具有部分相同的数据时，也必须建立各自的文件，不能共享相同的数据。

(4) 数据独立性差

一旦数据的逻辑结构改变，必须修改应用程序，修改文件结构定义；应用程序的改变，也要引起文件数据结构的变化；文件之间互相独立，不能反映出现实世界的内在联系。

1.1.2 数据管理技术的产生和发展



文件系统阶段应用程序与数据之间的对应关系



1.1.2 数据管理技术的产生和发展

□ 数据库系统阶段 (20 世纪 60 年代后期—)

背景:

应用背景

大规模管理

硬件背景

大容量磁盘

软件背景

有数据库管理系统

处理方式
理

联机实时处理、分布处理、批处

1.1.2 数据管理技术的产生和发展

□ 数据库系统阶段 (20 世纪 60 年代后期—)
特点:

1) 数据结构化

数据结构化是数据库与文件系统的根本区别。

- 文件系统中，尽管其记录内部已有了某些结构，但记录之间没有联系。数据的最小存取单位是记录。
- 数据库系统实现整体数据的结构化，描述数据时不仅要描述数据本身，还要描述数据之间的联系。数据不再针对其应用，是面向全组织，具有整体的结构化。数据的最小存取单位是数据项。

1.1.2 数据管理技术的产生和发展

学生文件 student 的记录结构

学生编号	姓名	性别	年龄	系	家庭住址	联系电话
------	----	----	----	---	------	------

课程文件 course 的记录结构

课程编号	课程名称	学时数	教材名称
------	------	-----	------

学生选课文件 sc 的记录结构

学生编号	课程编号	学期	成绩
------	------	----	----

学生、课程、学生选课文件结构

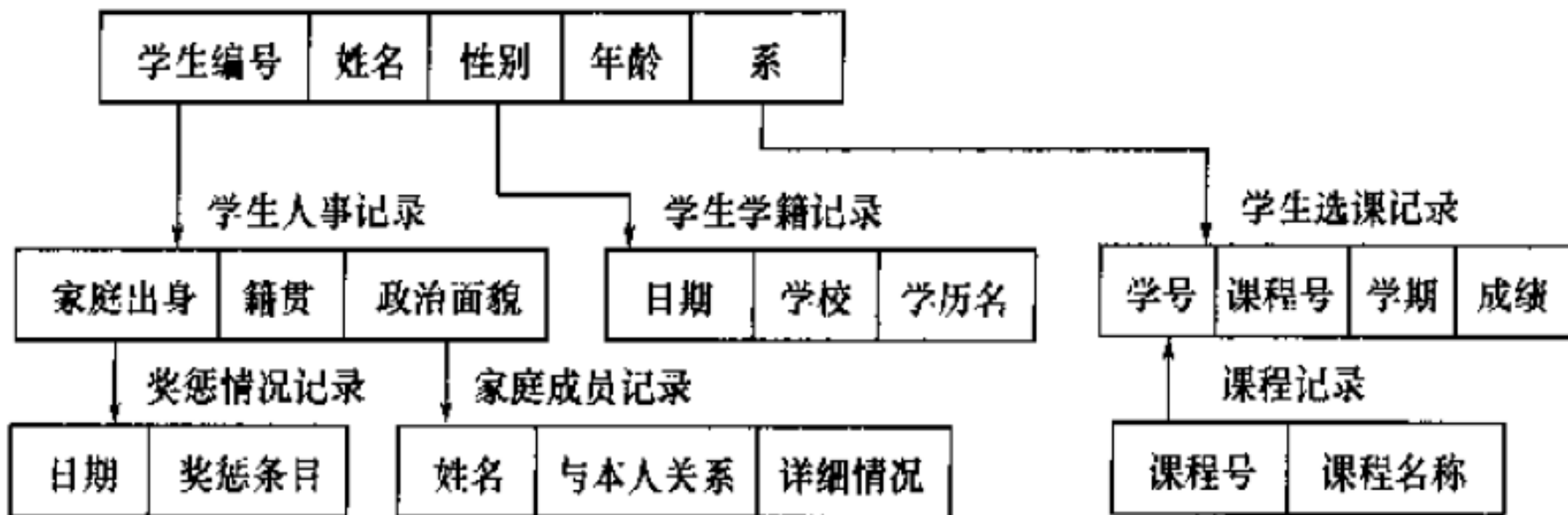
（关系数据库中的参照完整性；
文件系统只能编写应用程序代码实现）

1.1.2 数据管理技术的产生和发展

数据库系统阶段 (20 世纪 60 年代后期—

不仅考虑某个应用的数据结构，还要考虑整个组织的数据结构。因此描述数据时不仅仅描述数据本身，还要描述

学生基本记录





1.1.2 数据管理技术的产生和发展

(2) 数据的共享性高，冗余度低，易扩充

数据可以被多个用户、多个应用共享使用。共享导致数据冗余度低，同时可避免数据之间的不相容性与不一致性。通过重新选择不同的操作子集，易于增加新的应用。

(3) 数据独立性高

物理独立性

指用户的应用程序与存储在磁盘上的数据库中数据是相互独立的。当数据的物理存储改变了，应用程序不用改变。

逻辑独立性

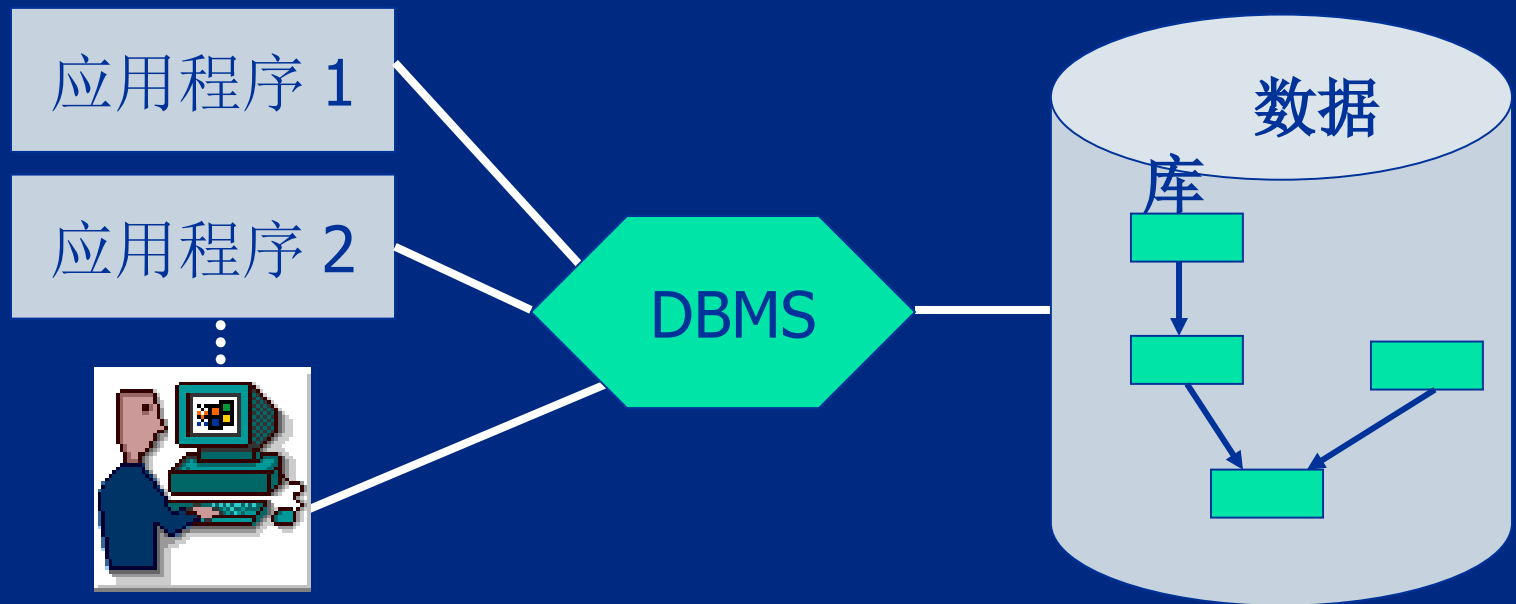
指用户的应用程序与数据库的逻辑结构是相互独立的。数据的逻辑结构改变了，用户程序也可以不变。

1.1.2 数据管理技术的产生和发展

(4) 数据由 DBMS 统一管理和控制

数据的安全性保护 / 数据的完整性检查

并发控制 / 数据库恢复



数据库管理阶段程序与数据的对应关系



1.2 数据模型

在数据库中如何抽象、表示、处理现实世界中的数据 and 信息呢？

数据模型（**Data Model**）这个工具来对现实世界进行抽象的。



1.2 数据模型

模型

是现实世界特征的模拟和抽象。

数据模型

- 也是一种模型，是现实世界数据特征的抽象，表示实体以及实体间的联系。
- 一个用于描述数据、数据间关系、数据语义和数据约束的概念工具的集合。

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。



1.2.1 两类数据模型

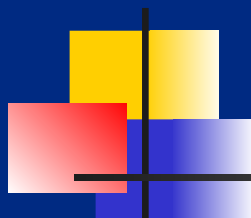
根据应用的不同目的，数据模型划分为两类：

❖ 概念模型（信息模型）

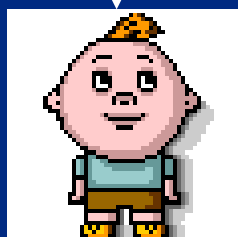
按用户的观点来对数据和信息建模；主要用于数据库设计。

❖ 数据模型（逻辑模型和物理模型）

按计算机系统的观点对数据建模；**逻辑模型**主要包括网状模型、层次模型、关系模型等，用于 DBMS 的实现。**物理模型**是描述数据在系统内部的表示方式和存取方法，由 DBMS 来具体实现。



现实世界



认识
抽象

信息世界
概念模型

转换

机器世界
DBMS 支持的数据模型

现实世界 → 概念模型:

数据库设计人员

概念模型 → 逻辑模型

数据库设计人员
& 数据库设计工具

逻辑模型 → 物理模型

DBMS

客观对象的抽象过程



1.2.2 数据模型的组成要素

由数据结构、数据操作和完整性约束三部分组成

□ 数据结构

是所研究的对象类型的集合。

- 与数据类型、内容、性质有关的对象（例如网状模型中的数据项、记录，关系模型中的域、属性、关系等）。
- 与数据之间联系有关的对象（例如网状模型中的系型）。

数据结构是对系统静态特性的描述。



1.2.2 数据模型的组成要素

□ 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，包括操作及有关的操作规则。

- 数据库主要有检索和更新（包括插入、删除、修改）两大类操作。数据模型必须定义这些操作的含义、符号、规则（如优先级）以及实现操作的语言。

数据操作是对系统动态特性的描述

□ 数据的约束条件

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，用以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（详见第五章）。



1.2 数据模型

1.2.3 概念模型

概念模型是现实世界到机器世界的一个中间层次

- ❖ 概念模型用于信息世界的建模
 - 应该具有较强的语义表达能力
 - 应该简单、清晰、易于用户理解。



1.2.3 概念模型

□ 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。

(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。例如：

(94002268 , 张山, 男, 1976 , 计算机系, 1994)

(3) 码 (Key)

唯一标识实体的属性集称为码 (学号、订单号等)。



1.2.3 概念模型

(4) 域 (Domain)

属性的取值范围称为该属性的域。

例如，学号的域为 8 位整数，姓名的域为字符串集合，性别的域为（男，女）。

(5) 实体型 (EntityType)

用实体名及其属性集合来抽象和刻画同类实体，称为实体型。

学生（学号，姓名，出生年份，系，入学时间）

(6) 实体集 (Entity Set)

同型实体的集合称为实体集。



1.2.3 概念模型

(7) 联系 (Relationship)

在现实世界中，事物内部以及事物之间是有联系的，这些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。

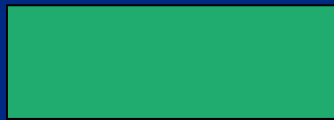
实体之间的联系通常是指不同实体集之间的联系。



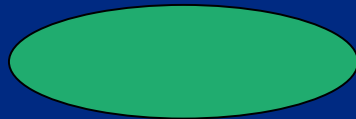
1.2.3 概念模型

□ 实体联系模型（E - R模型）表示方法

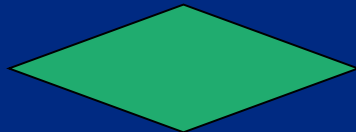
E - R模型的三要素：实体、属性、实体间的联系。



表示实体



表示属性

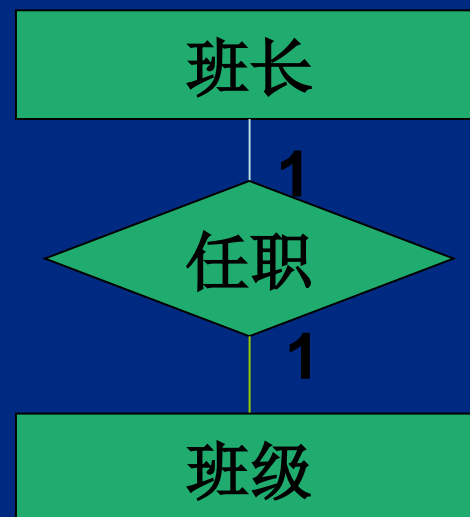
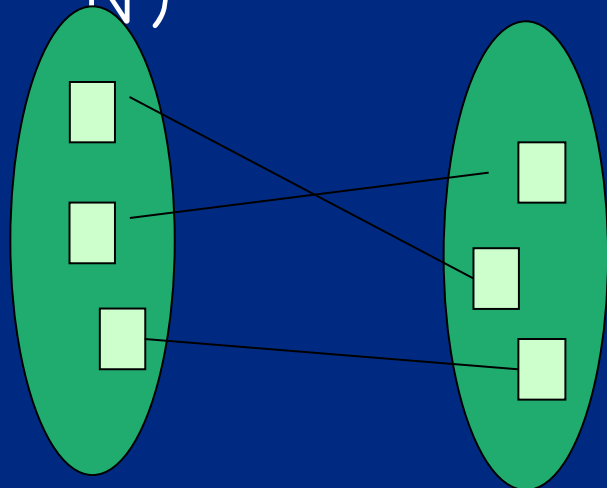


表示实体间联系

1.2.3 概念模型

□ 实体联系模型（E - R模型）表示方法

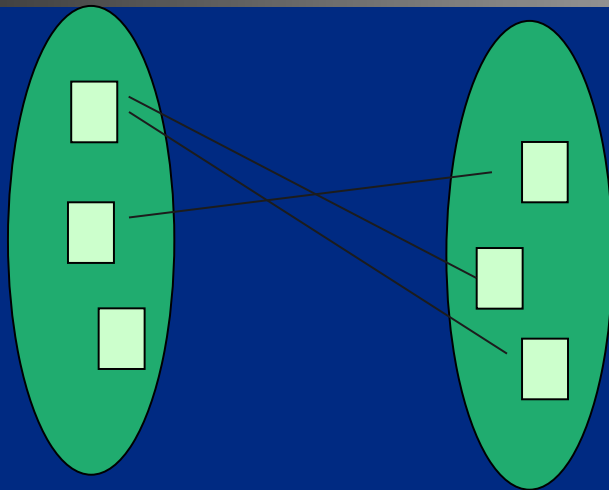
实体型之间的联系：一对一联系（1 : 1）、一对多联系（1 : N）、多对多联系（M : N）



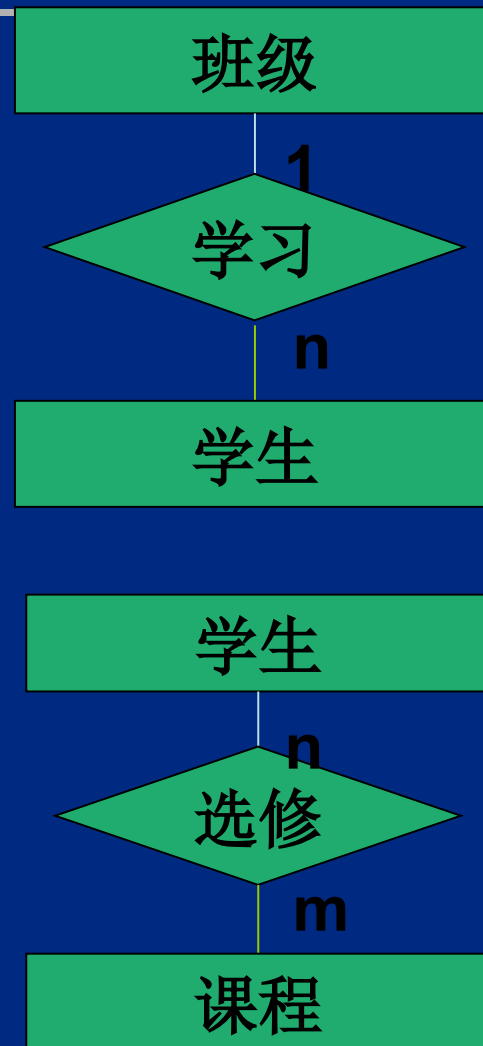
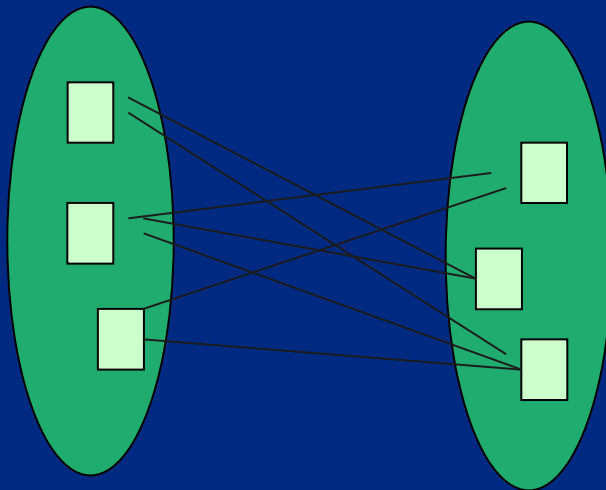
1.2.3 概念模型

实体联系模型（E - R模型）表示方法

1:n

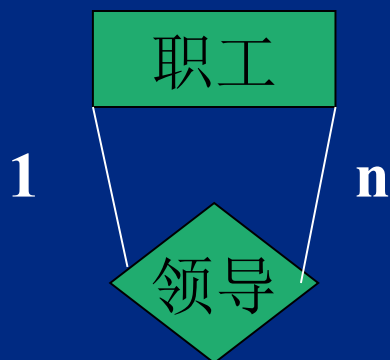


n:m

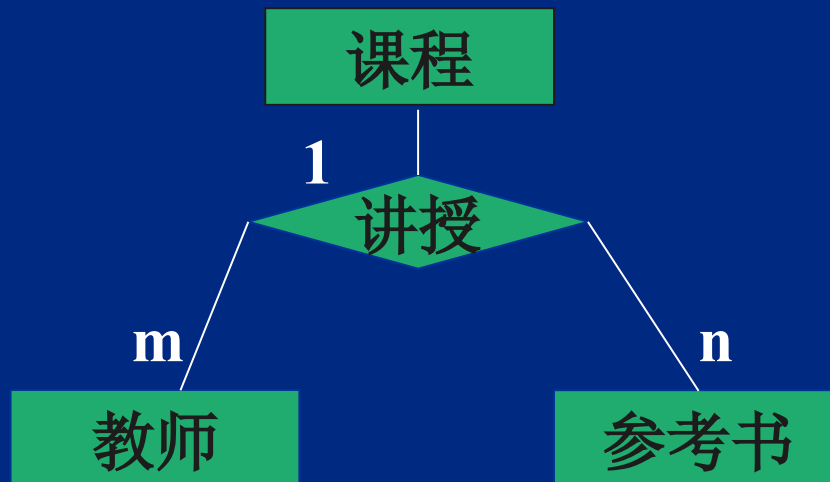


1.2.3 概念模型

实体联系模型（E - R模型）表示方法



单个实体型间的 1:n 联系

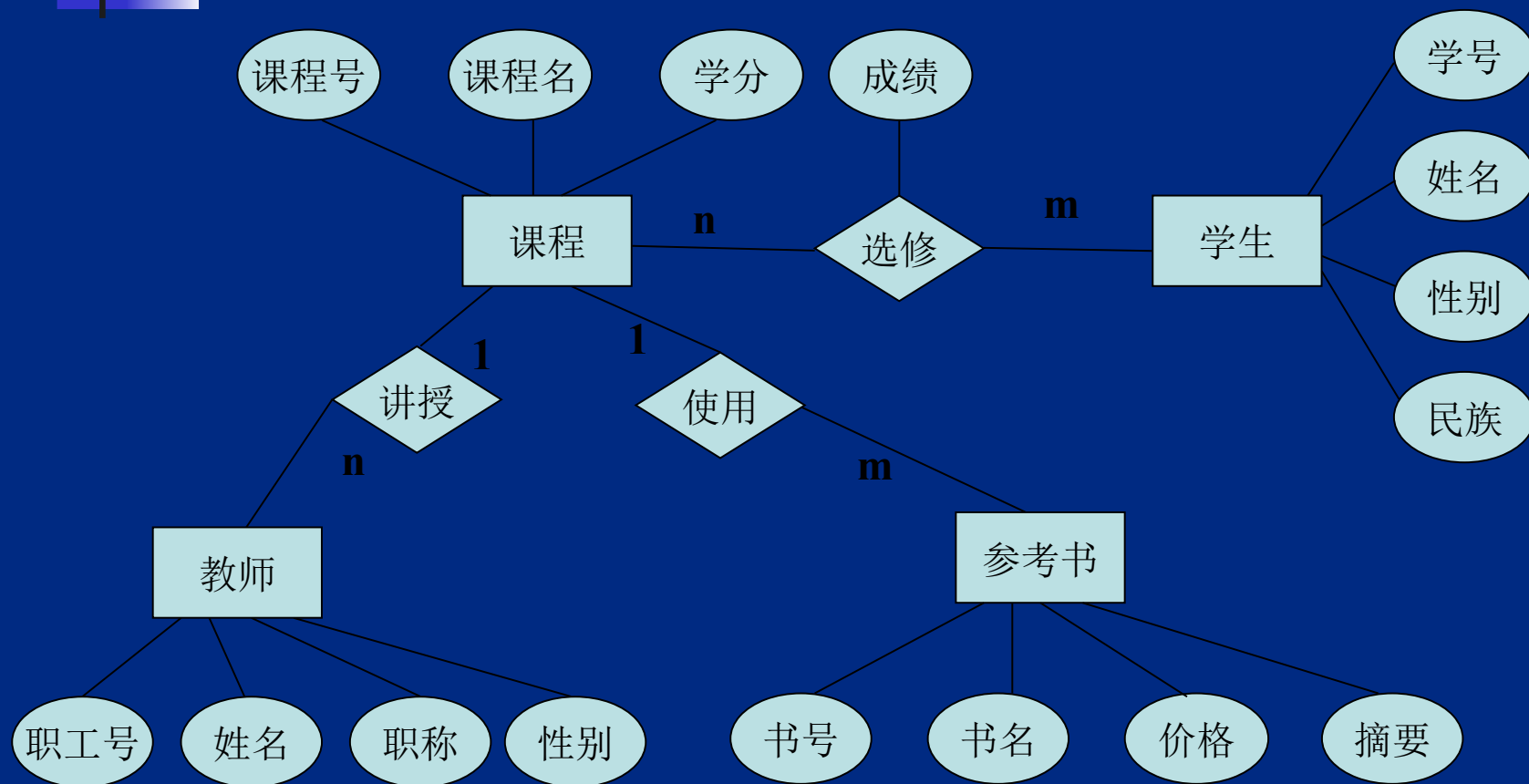


多个实体型间的 1:n 联系

例： 假设一个学生可选多门课程，而一门课程又有多个学生选修，每个学生每选一门课只有一个成绩，一个教师只能讲一门课程，一门课程也可有多个教师讲授，一门课使用多本参考书。画出 E - R 图。

1.2.3 概念模型

实体联系模型（E - R模型）表示方法



教学 E-R 图 （另： P219 工厂物资管理实例）



1.2.4 最常用的数据模型

数据库领域中最常用的数据模型有四种

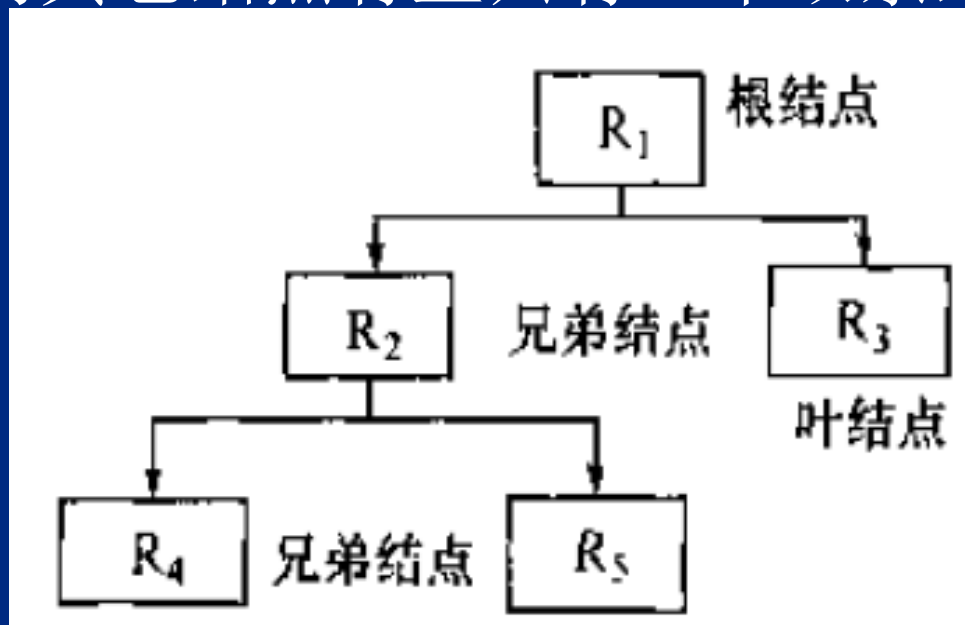
- 层次模型 (Hierarchical Model), 如 IBM 的 IMS 系统
- 网状模型 (Network Model), 如 DBTG 系统
- 关系模型 (Relational Model), 如 Oracle, Sybase
- 面向对象模型 (Object Oriented Model)

1.2.5 层次模型（树型结构）

□ 层次模型定义

满足下面两个条件的基本层次联系的集合为层次模型。

1. 有且只有一个结点没有双亲结点，该结点称为根结点
2. 根以外的其它结点有且只有一个双亲结点。



1.2.5 层次模型

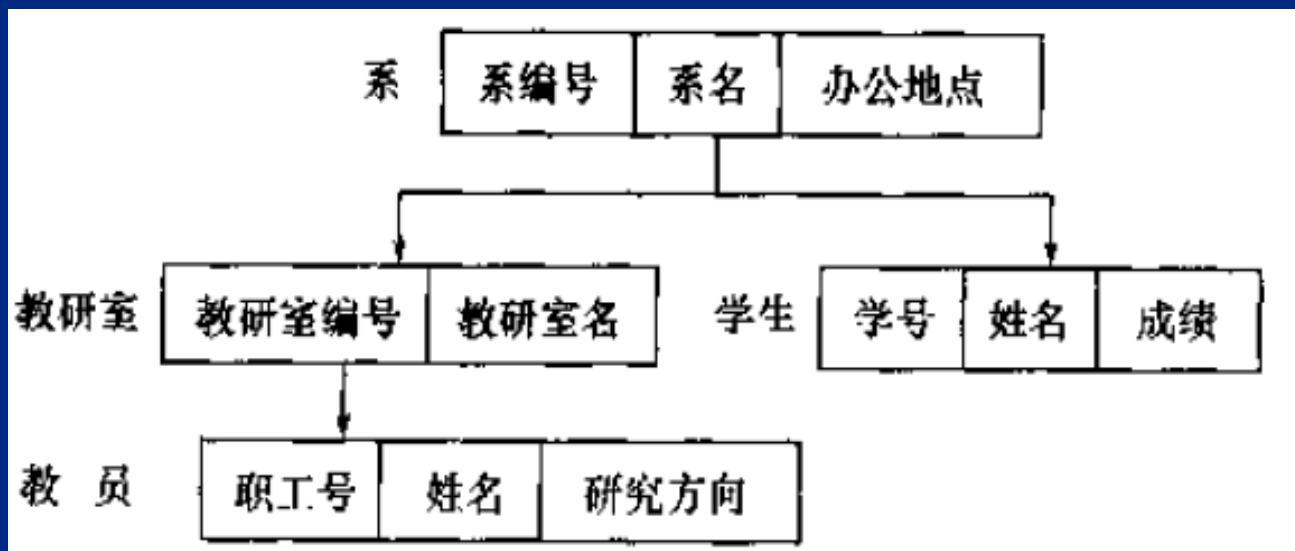
□ 表示方法

实体型：用记录类型描述。

每个结点表示一个记录类型。

属性：用字段描述。每个记录类型可包含若干个字段。

联系：用结点之间的连线表示记录（类型）之间的一对多的联系





1.2.5 层次模型

□ 特点

- 结点的双亲是唯一的
- 只能直接处理一对多的实体联系
- 每个记录类型定义一个排序字段，也称为码字段
- 任何记录值只有按其路径查看时，才能显出它的全部意义
- 没有一个子女记录值能够脱离双亲记录值而独立存在



1.2.5 层次模型

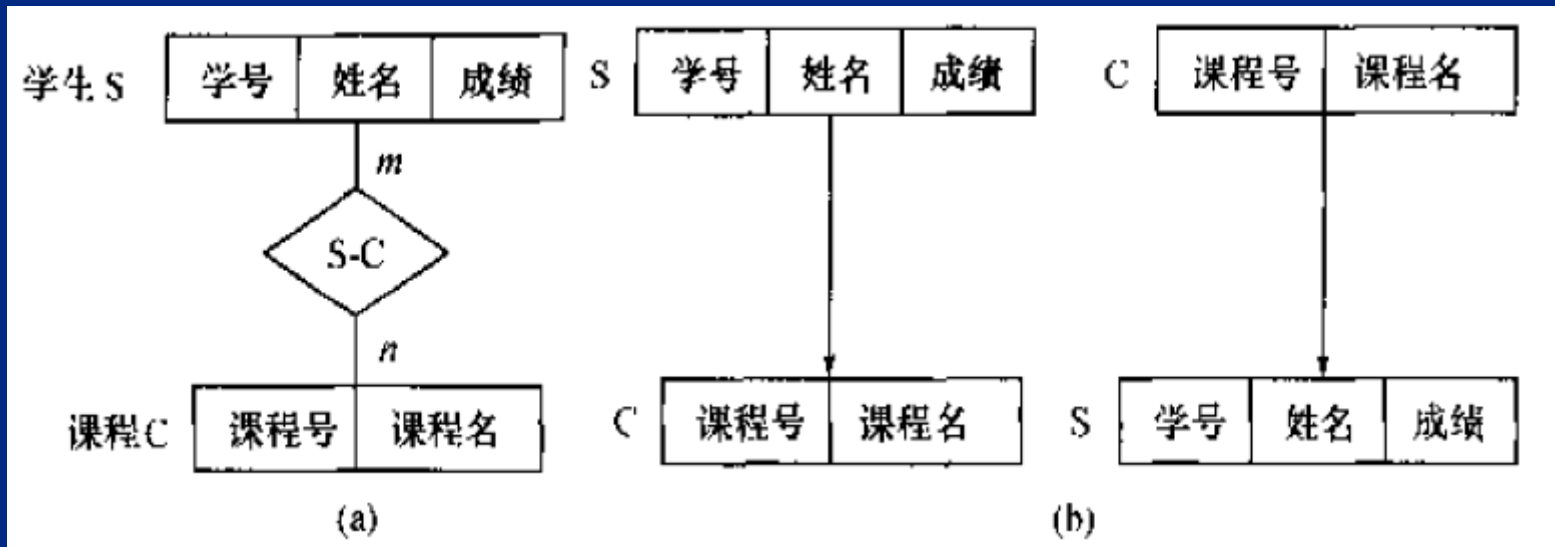
□ 多对多联系在层次模型中的表示

用层次模型间接表示多对多联系，将多对多联系分解成一对多联系。

分解方法：

- 冗余结点法
- 虚拟结点法

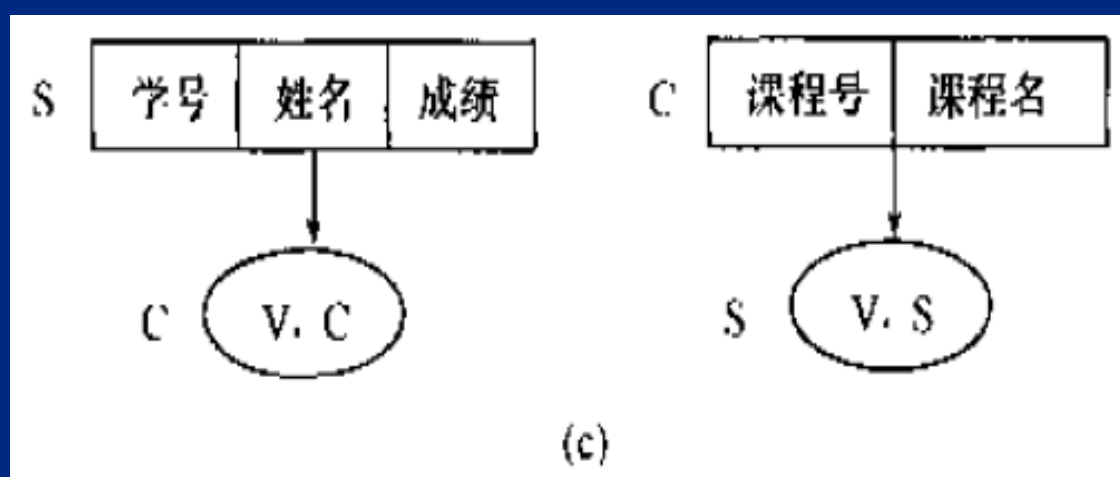
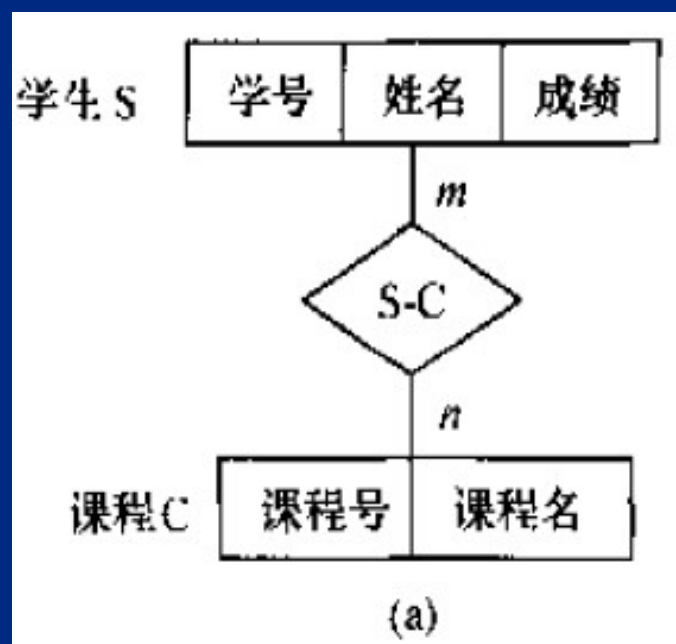
冗余结点法：增设冗余结点，将多对多联系转换为两个一对多联系



优点：结构清晰，允许结点改变存储位置

缺点：需要额外占用存储空间，有潜在的不一致性

虚拟结点法： 将冗余结点换成虚拟结点，
虚拟结点为一个指针，指向替代的节点。



优点：减少对存储空间的浪费，避免产生潜在的不一致性
缺点：结点改变存储位置可能引起虚拟结点中指针的修改

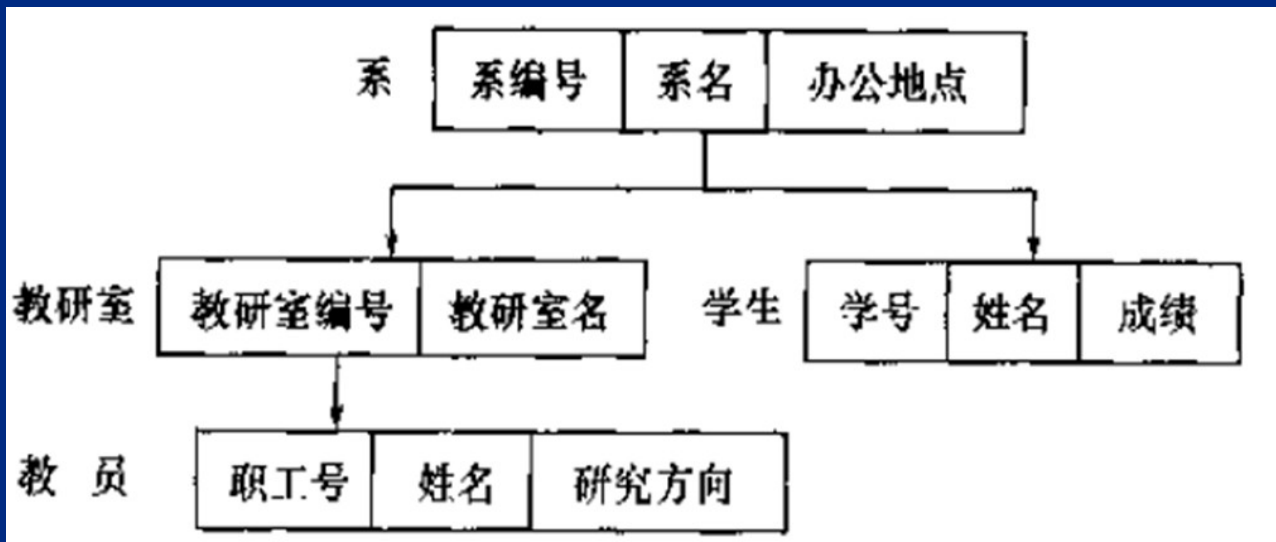
1.2.5 层次模型

□ 层次模型的数据操纵与完整性约束

数据操纵：查询、插入、删除、更新

完整性约束：

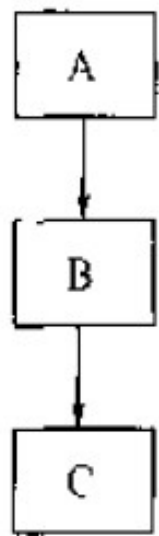
- 无相应的双亲结点值就不能插入子女结点值
- 如果删除双亲结点值，则相应的子女结点值也被同时删除



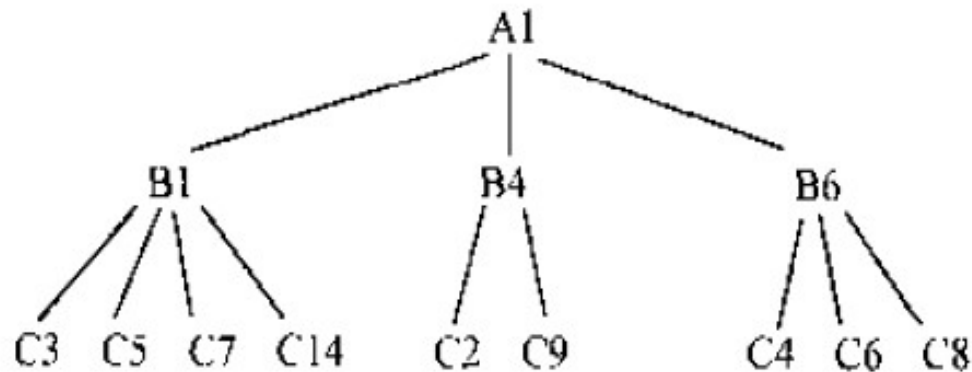
1.2.5 层次模型

层次数据模型的存储结构（邻接法和链接法）

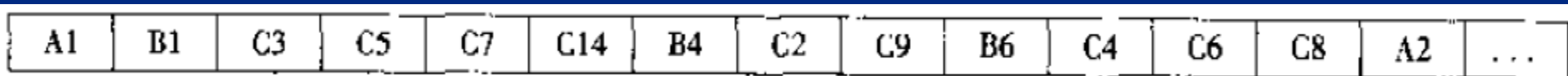
邻接法：按照层次树前序遍历的顺序把所有记录值依次邻接存放，即通过物理空间的位置相邻来实现层次顺序



(a)



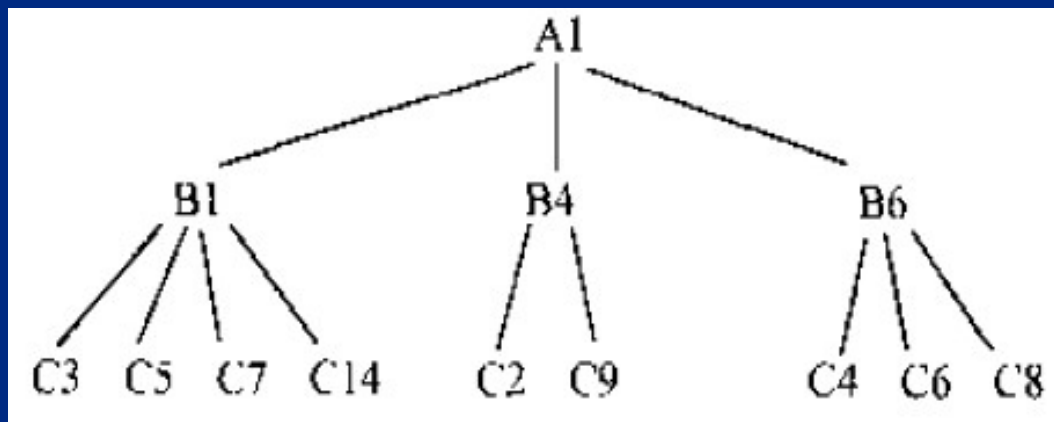
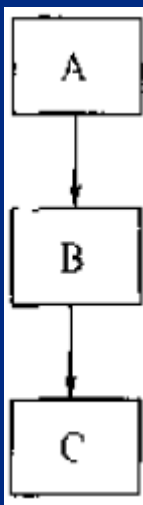
(b)



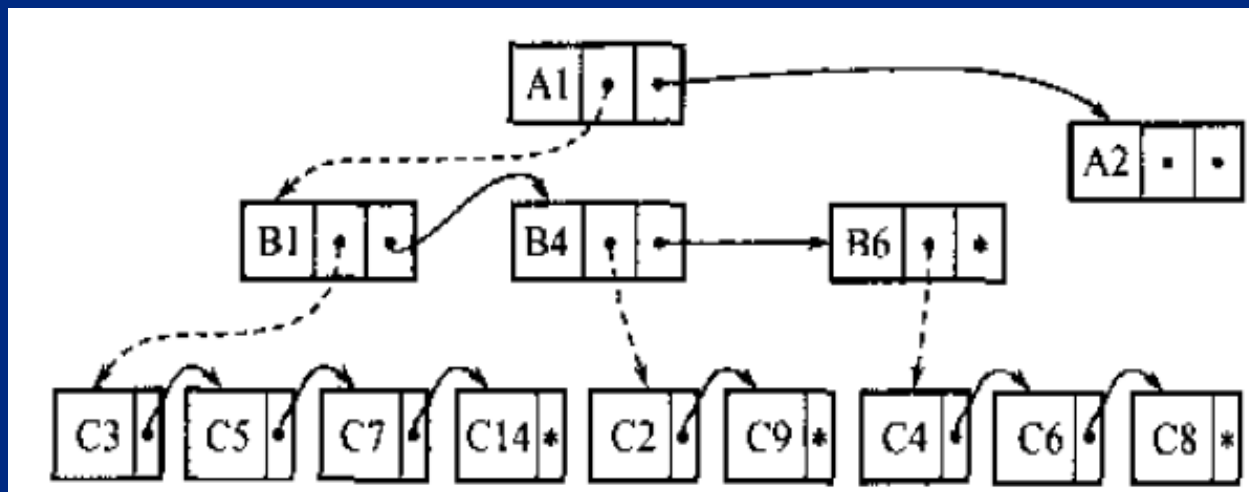
1.2.5 层次模型

层次数据模型的存储结构（邻接法和链接法）

链接法：用指针反映数据之间的层次关系。

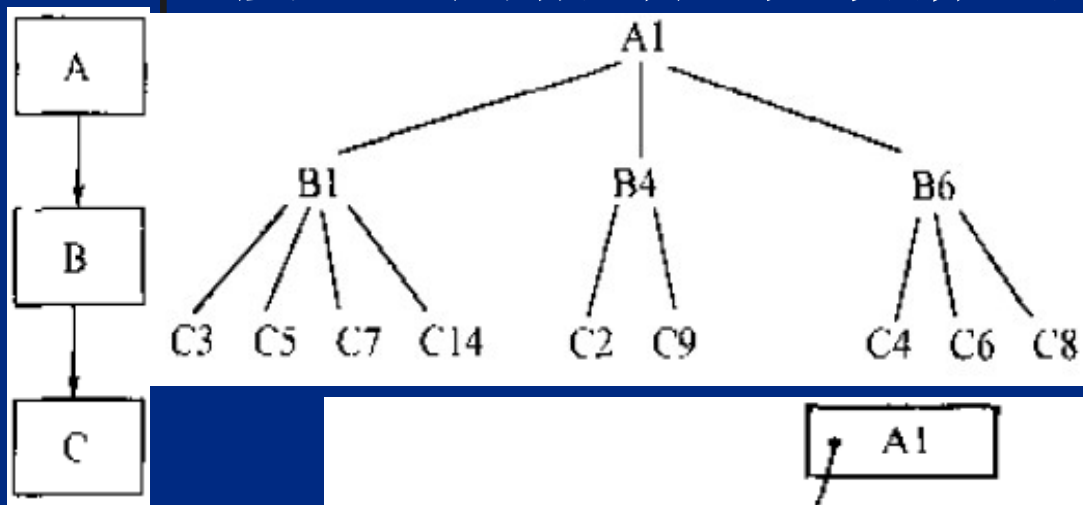


子女-兄弟链接法：
每个记录设两个指针，
分别指向最左边的子女
和最近的兄弟



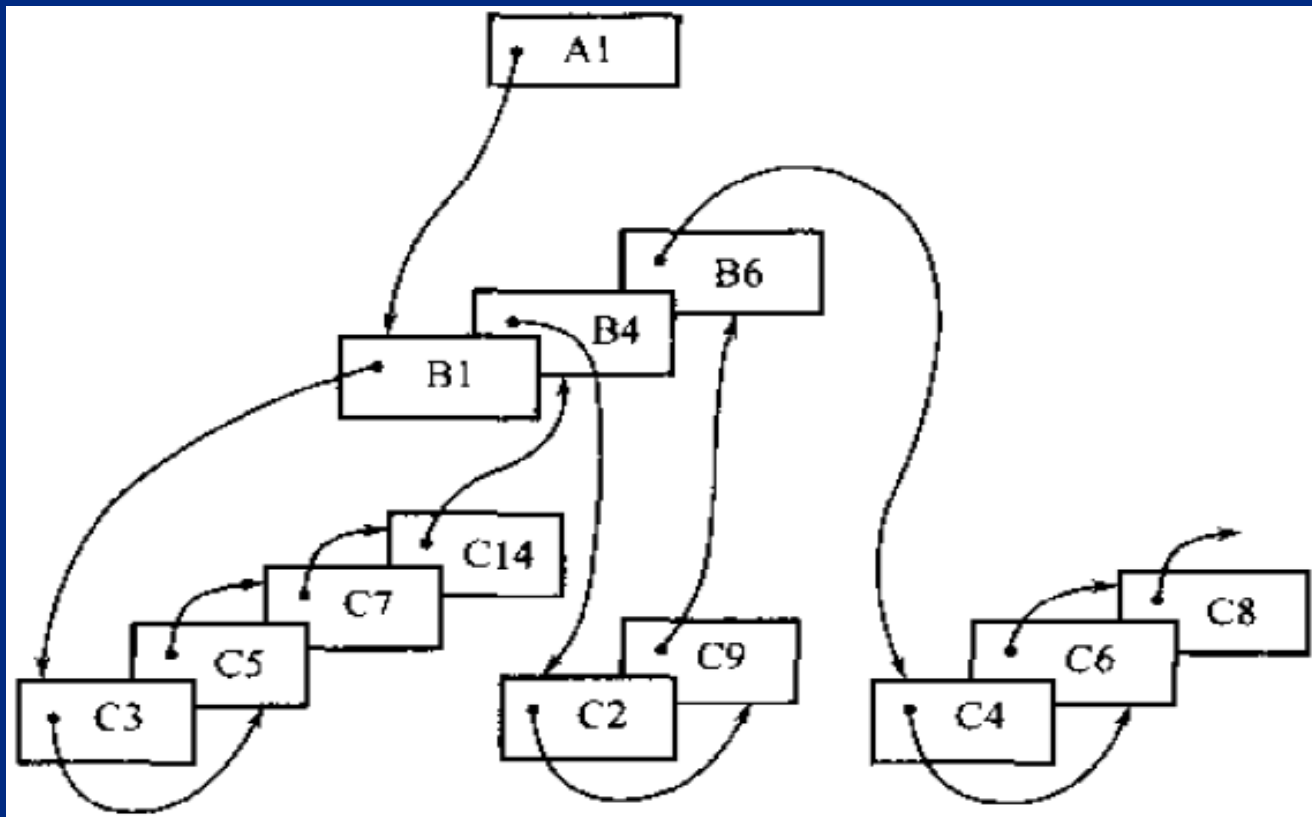
□ 层次数据模型的存储结构（邻接法和链接法）

链接法：用指针反映数据之间的层次关系。



层次序列链接法：

按树的前序遍历顺序链接各记录值





1.2.5 层次模型

□ 层次模型的优缺点

优点:

- 数据结构简单明晰
- 数据库查询效率高
- 良好的完整性支持

缺点:

- 多对多联系表示不自然
- 对插入和删除操作的限制多
- 查询子女结点必须通过双亲结点
- 层次命令趋于程序化



1.2.5 层次模型

□ 典型的层次数据库系统

IMS 数据库管理系统

- 第一个大型商用 DBMS
- 1968 年推出
- IBM 公司研制



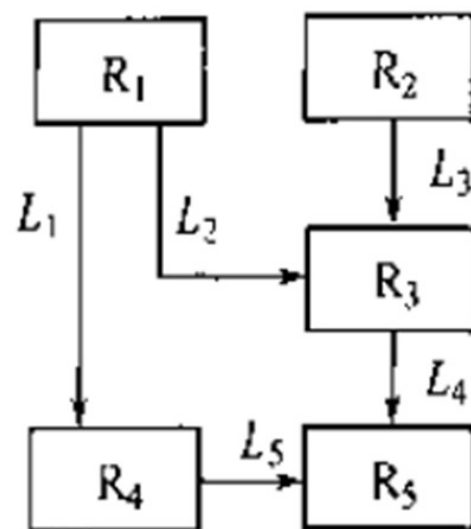
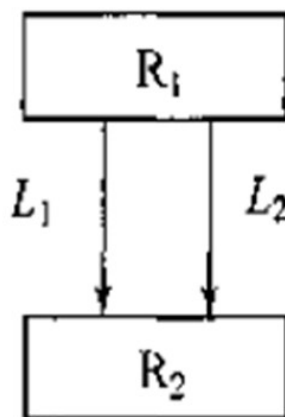
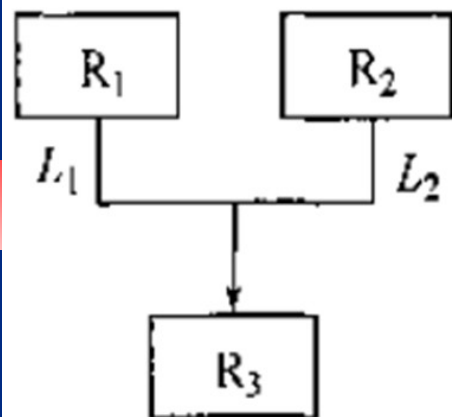
1.2.6 网状模型

□ 定义

满足下面两个条件的基本层次联系的集合为网状模型。

1. 允许一个以上的结点无双亲；
2. 一个结点可以有多个的双亲。

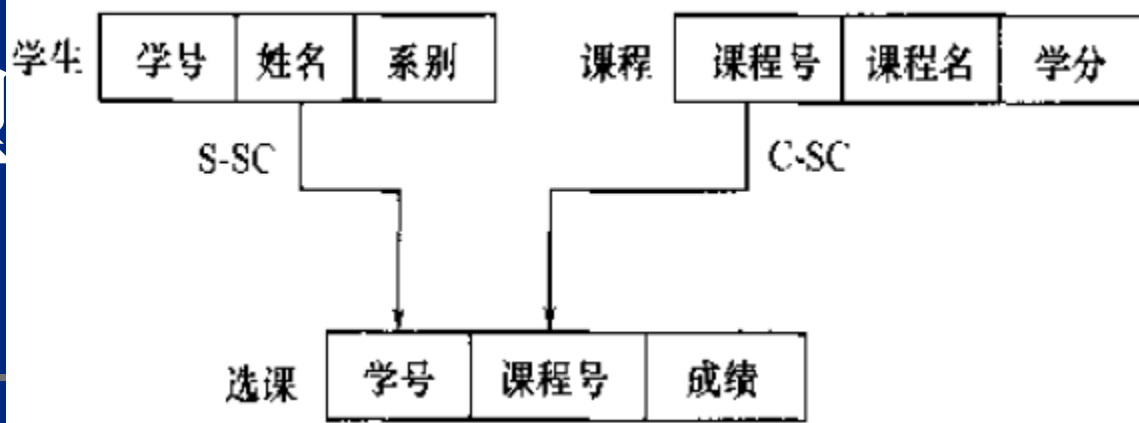
层次模型中子女结点和双亲结点的联系是唯一的，而网状模型中这种联系则不唯一。因此要为每个联系命名，指出与该联系有关的双亲记录和子女记录。



□ 网状模型与层次模型的区别

- 网状模型允许多个结点没有双亲结点
- 网状模型允许结点有多个双亲结点
- 网状模型允许两个结点之间有多种联系（复合联系）
- 网状模型可以更直接地去描述现实世界
- 层次模型实际上是网状模型的一个特例

1.2.6 网状模型



□ 多对多联系在网状模型中的表示

用网状模型间接表示多对多联系

方法：将多对多联系直接分解成一对多联系

□ 网状数据模型的操纵和完整性约束

操纵：查询、插入、删除、更新

完整性约束：

- 支持记录码的概念（唯一标识、不许重复）
- 保证一个记录中双亲记录与子女记录之间是一对多联系
- 支持双亲记录和子女记录之间某些约束条件，比如允许插入尚未确定双亲结点值的子女结点值；允许只删除双亲结点值等

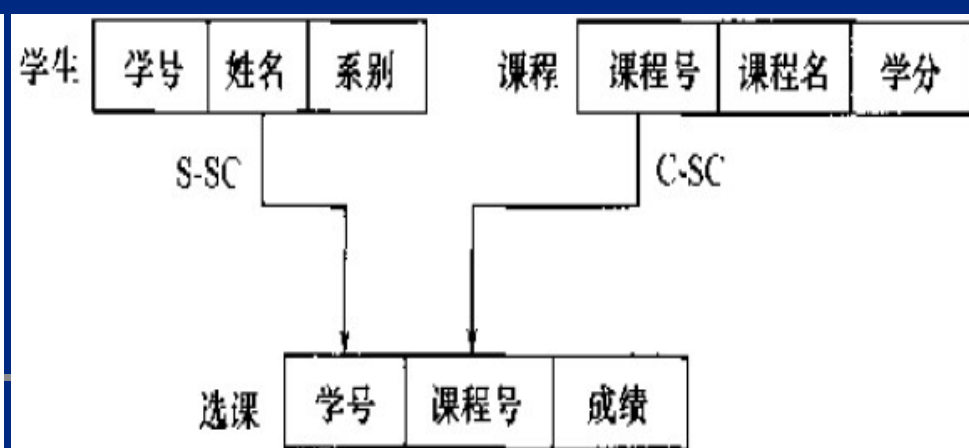
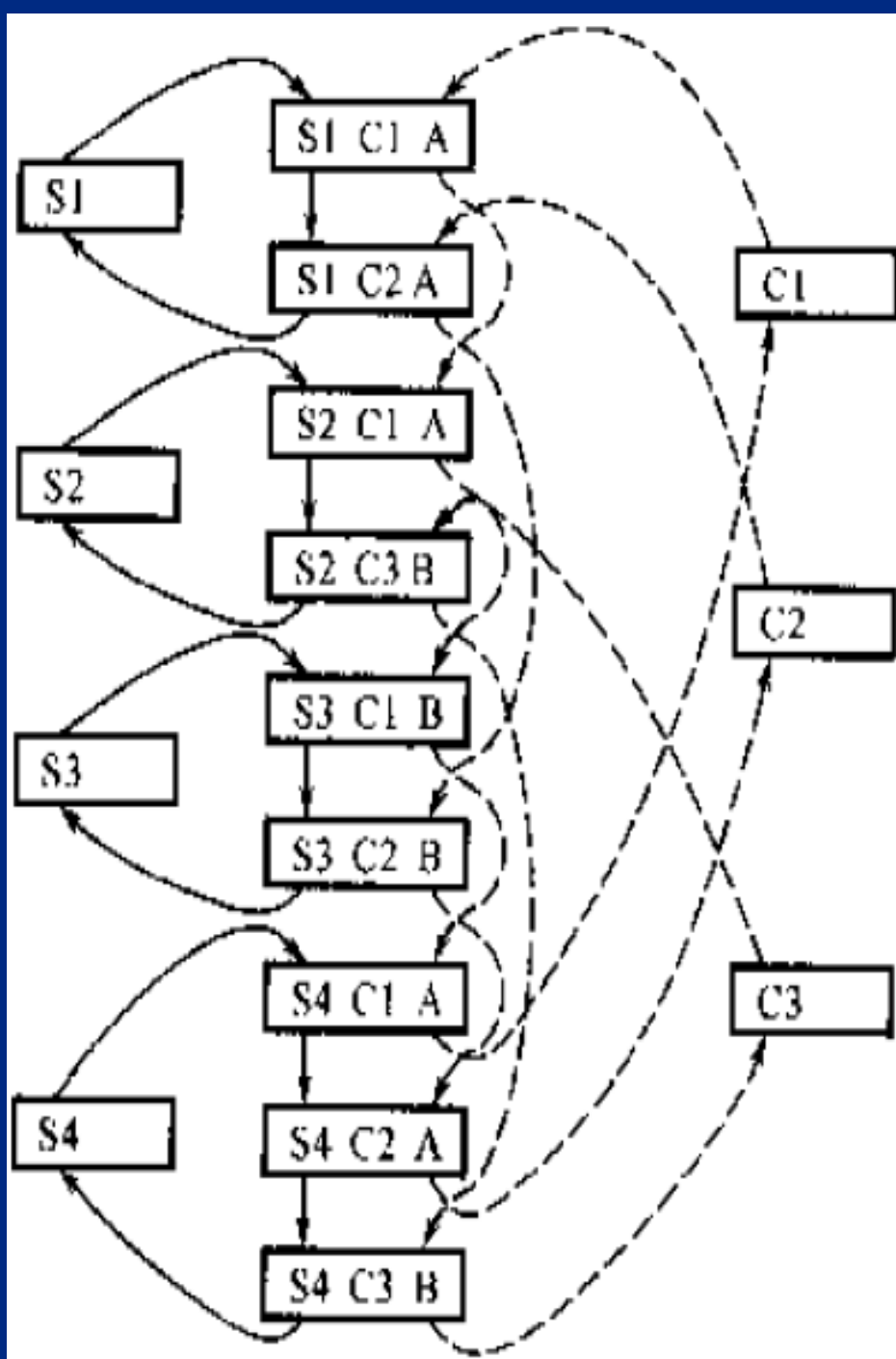


1.2.6 网状模型

- 网状数据模型的存储结构
 关键：实现记录之间的联系

常用方法

- 单向链接
- 双向链接
- 环状链接
- 向首链接



单向链接法



1.2.6 网状模型

□ 网状模型的优缺点

优点：

- 能够更为直接地描述现实世界，如一个结点可以有多个双亲。
- 具有良好的性能，存取效率较高。

缺点

- 结构比较复杂，而且随着应用环境的扩大，数据库的结构就变得越来越复杂，不利于最终用户掌握。
- 数据定义语言（**DDL**）、数据操作语言（**DML**）复杂，用户不容易使用。



1.2.6 网状模型

□ 典型的网状数据库系统

DBTG 系统，亦称 CODASYL 系统。由 DBTG 提出的一个系统方案，奠定了数据库系统的基本概念、方法和技术。70 年代推出。

实际系统

- Cullinet Software Inc. 公司的 IDMS
- Univac 公司的 DMS1100
- Honeywell 公司的 IDS/2
- HP 公司的 IMAGE



1.2.7 关系模型

- 最重要的一种数据模型。也是目前主要采用的数据模型
- **1970** 年由美国 **IBM** 公司 **San Jose** 研究室的研究员 **E.F.Codd** 提出
- 本课程的重点



1.2.7 关系模型

□ 关系数据模型的数据结构（I）

在用户观点下，关系模型中数据的逻辑结构是一张二维表，它由行和列组成。

学生登记表

学 号	姓 名	年 令	性 别	系 名	年 级
95004	王小明	19	女	社会学	95
95006	黄大鹏	20	男	商品学	95
95008	张文斌	18	女	法律学	95
...

1.2.7 关系模型

学生登记表

学 号	姓 名	年 令	性 别	系 名	年 级
95004	王小明	19	女	社会学	95
95006	黄大鹏	20	男	商品学	95
95008	张文斌	18	女	法律学	95
...

□ 关系数据模型的数据结构（ II ）

- 关系（ **Relation** ）

一个关系对应通常说的一张表。

- 元组（ **Tuple** ）

表中的一行即为一个元组。

- 属性（ **Attribute** ）

表中的一列即为一个属性，给每一个属性起一个名称即属性名。

1.2.7 关系模型

学生登记表

学 号	姓 名	年 令	性 别	系 名	年 级
95004	王小明	19	女	社会学	95
95006	黄大鹏	20	男	商品学	95
95008	张文斌	18	女	法律学	95
...

□ 关系数据模型的数据结构（**III**）

- 主码（**Key**）

表中的某个属性组，它可以唯一确定一个元组。

- 域（**Domain**）

属性的取值范围。

- 分量

元组中的一个属性值。

学生登记表

学 号	姓 名	年 令	性 别	系 名	年 级
95004	王小明	19	女	社会学	95
95006	黄大鹏	20	男	商品学	95
95008	张文斌	18	女	法律学	95
...

□ 关系数据模型的数据结构（ IV ）

● 关系模式

对关系的描述

关系名（属性 1，属性 2，…，属性 n）

学生（学号，姓名，年龄，性别，系别，年级）



1.2.7 关系模型

□ 实体及实体间联系的表示方法

- 实体型：直接用关系（表）表示。
- 属性：用属性名表示。
- 一对一联系：隐含在实体对应的关系中。
- 一对多联系：隐含在实体对应的关系中。
- 多对多联系：直接用关系表示。



1.2.7 关系模型

□ 实体及实体间联系的表示方法（例）

例 1：学生、系、系与学生之间的一对多联系：

学生（学号，姓名，年龄，性别，系号，年级）

系（系号，系名，办公地点）

例 2：系、系主任、系与系主任间的一对一联系：

系（系号，系名，系主任姓名）

例 3：学生、课程、学生与课程之间的多对多联系：

学生（学号，姓名，年龄，性别，系号，年级）

课程（课程号，课程名，学分）

选修（学号，课程号，成绩）



1.2.7 关系模型

□ 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。

职工号	姓名	职 称	工 资			扣 除		实 发
			基 本	津 贴	职 务	房 租	水 电	
86051	陈 平	讲 师	1305	1200	50	160	112	2283
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮



1.2.7 关系模型

□ 关系模型的数据操纵

- ✓ 查询、插入、删除、更新
- ✓ 数据操作是集合操作，操作对象和操作结果都是关系，即若干元组的集合
- ✓ 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”

□ 关系模型的完整性约束

实体完整性 / 参照完整性 / 用户定义的完整性（第 5 章）



1.2.7 关系模型

□ 关系数据模型的存储结构

实体及实体间的联系都用表来表示。在关系数据库的物理组织中，有的 DBMS 一个表对应一个操作系统文件，有的 DBMS 从操作系统获得若干大的文件，自己设计表、索引等存储结构。



1.2.7 关系模型

□ 关系数据模型的优缺点

优点：

- 建立在严格的数学概念的基础上
- 概念单一，数据结构简单、清晰，用户易懂易用。

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透明

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



1.2.7 关系模型

□ 关系数据模型的优缺点

缺点：

- 存取路径对用户透明导致查询效率往往不如非关系数据模型。

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。

（第 9 章）



1.2.7 关系模型

□ 典型的关系数据库系统

- ORACLE
- SYBASE
- INFORMIX
- DB/2
- COBASE
- PBASE
- EasyBase
- DM/2
- OpenBase



1.3 数据库系统结构

- 数据库系统模式的概念
- 数据库系统的三级模式结构
- 数据库的二级映象功能与数据独立性

从**数据库管理系统**角度看：数据库系统通常采用三级模式结构；从**数据库最终用户**角度看：数据库系统的结构分为单用户结构、主从式结构、分布式结构、客户 / 服务器、浏览器 / 应用服务器 / 数据库服务器多层结构等。



1.3.1 数据库系统模式的概念

□ 数据库系统 “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

记录型：（学号，姓名，性别，系别，籍贯）

记录值：（900201，李明，男，计算机，江苏）

1.3.1 数据库系统模式的概念

“型” 和 “值” 的区别

模式（ Schema ）

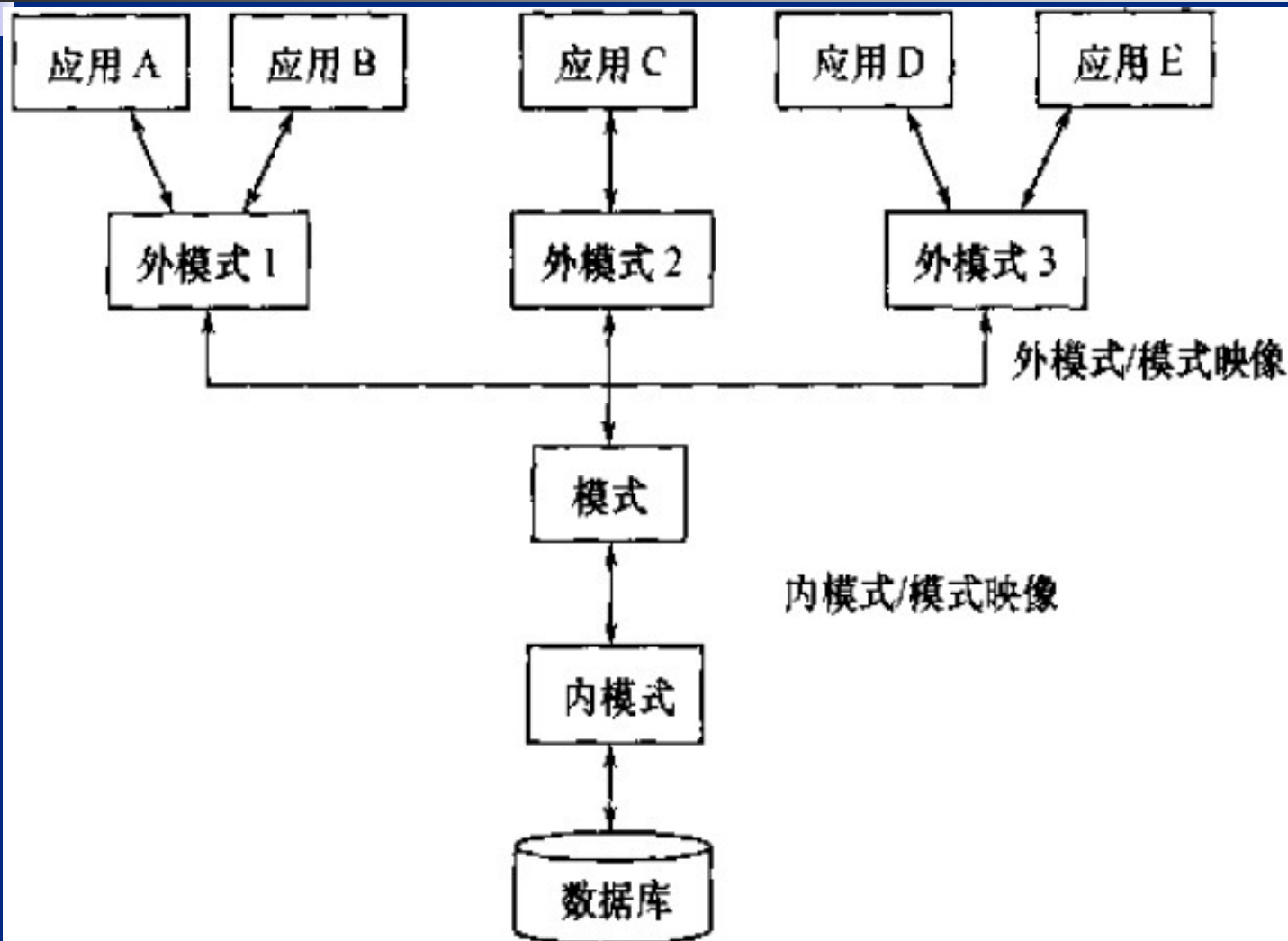
- 数据库逻辑结构和特征的描述
- 是型的描述
- 反映的是数据的结构及其联系
- 模式是相对稳定的

模式的一个实例（ 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更新而变动

1.3.2 数据库系统的三级模式结构

数据库系统由外模式、模式和内模式三级构成





1.3.2 数据库系统的三级模式结构

□ 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，综合了所有用户的需求

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



1.3.2 数据库系统的三级模式结构

□ 外模式（External Schema）

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

模式与外模式的关系：一对多

通常是模式的子集；一个数据库可以有多个外模式

外模式与应用的关系：一对多

同一外模式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



1.3.2 数据库系统的三级模式结构

□ 外模式（External Schema）

外模式是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数据

□ 内模式（Internal Schema）

内模式（也称存储模式）

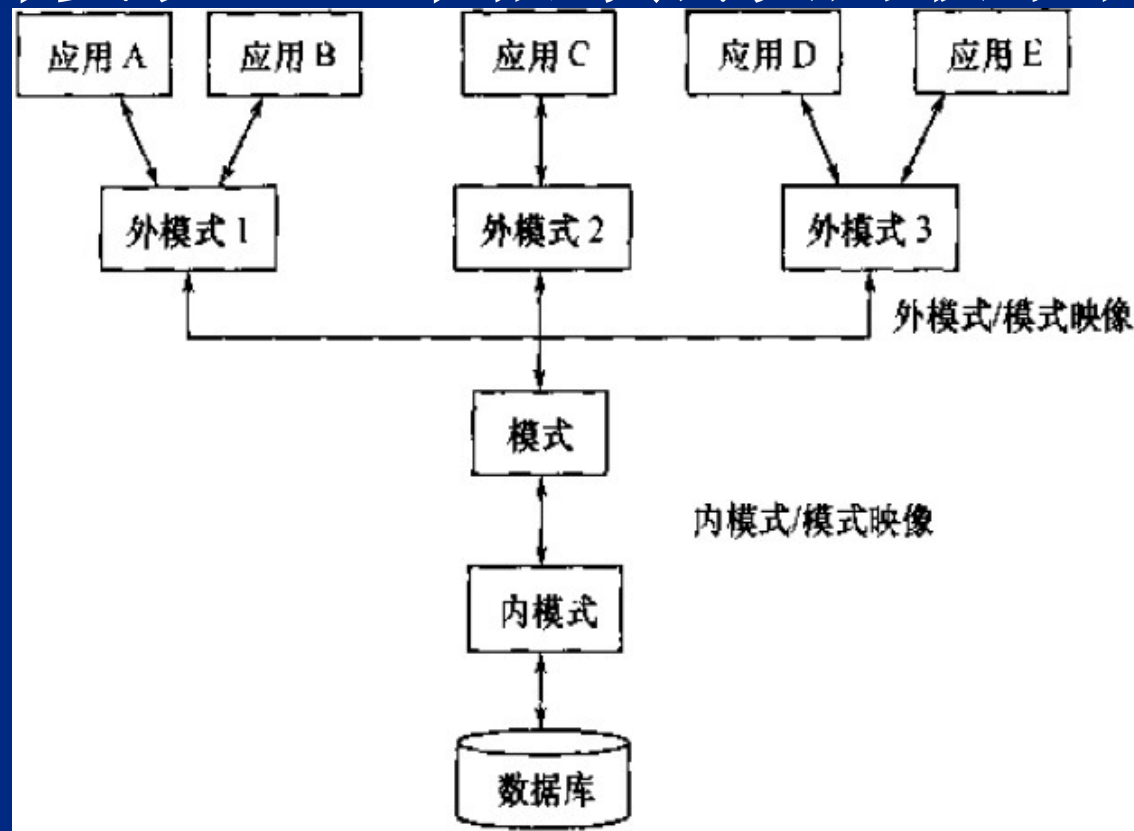
- 是数据物理结构和存储方式的描述
- 是数据在数据库内部的表示方式

记录的存储方式（顺序存储，按照 B 树结构存储，按 hash 方法存储） / 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

1.3.3 数据库的二级映像功能与数据独立性

三级模式是对数据的三个抽象级别，二级映像是在 DBMS 内部实现这三个抽象层次的联系和转换。



1.3.3 数据库的二级映像功能与数据独立性

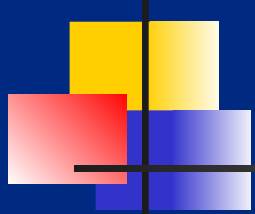
□ 外模式 / 模式映像

- ✓ 一个模式可以定义任意多个外模式
- ✓ 对每一个外模式，数据库系统都有一个外模式 / 模式映射，定义该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

保证数据的逻辑独立性

- 当模式改变时，数据库管理员修改有关的外模式 / 模式映像，使外模式保持不变
- 应用程序是依据数据的外模式编写的，从而应用程序不必修改，保证了数据与程序的逻辑独立性，简称数据的逻辑独立性。

1.3.3 数据库的二级映像功能与数据独立性



□ 模式 / 内模式映像

- ✓ 定义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 / 内模式映象是唯一的
- ✓ 该映象定义通常包含在模式描述中

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 / 内模式映象，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。

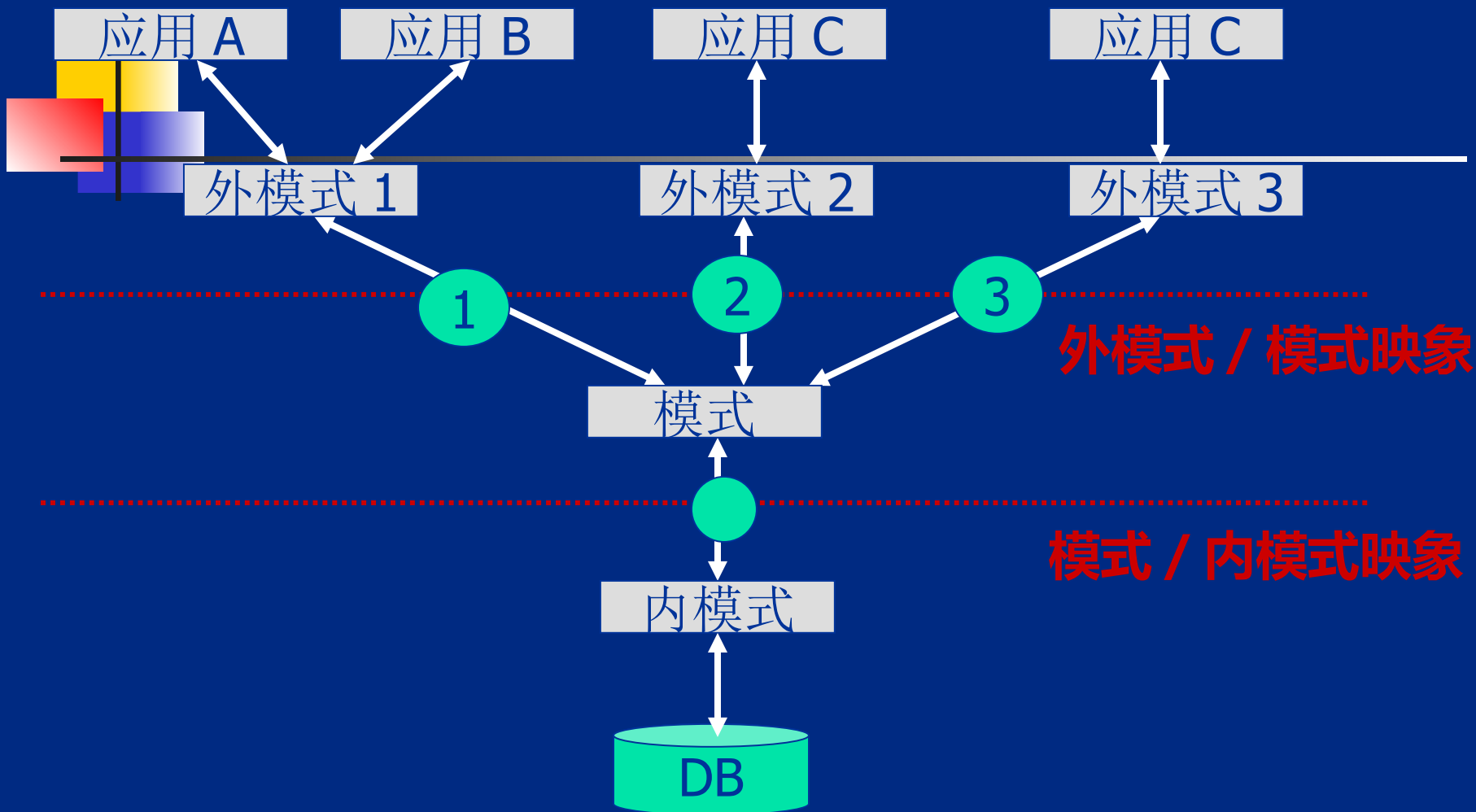
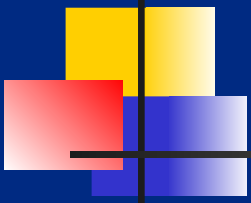


图 1-24 数据库系统的模式结构

1.3.3 数据库的二级映像功能与数据独立性



优点:

- ❖ 数据库的二级映象保证了数据库外模式的稳定性，从而从底层保证了应用程序的稳定性。
- ❖ 数据和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。另外，由于数据的存取由 **DBMS** 管理，用户不必考虑存取路径等细节，从而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。



1.4 数据库系统的组成

数据库系统一般由数据库、数据库管理系统（及其开发工具）、应用系统和数据库管理员组成。

数据库系统对硬件资源的要求

1. 足够大的内存，存放操作系统、**DBMS** 的核心模块、数据缓冲区和应用程序。
2. 足够大的外存，存放数据库以及做数据备份
3. 较高的通道能力，提高数据传送率



1.4 数据库系统的组成

软件

主要包括：

- DBMS
- 操作系统
- 与数据库接口的高级语言及其编译系统
- 以 DBMS 为核心的应用开发工具
- 为特定应用环境开发的数据库应用系统

1.4 数据库系统的组成

人员

□数据库管理员：

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

周期性转储数据库：数据文件 / 日志文件

系统故障恢复

介质故障恢复

监视审计文件

- ✓数据库的改进和重组重构

性能监控和调优

数据重组



1.4 数据库系统的组成

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 DBA 协商，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

1.4 数据库系统的组成



人员

□ 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

□ 用户

- ✓ 偶然用户：企业或组织机构的高中级管理人员
- ✓ 简单用户：银行的职员、机票预定人员、旅馆总台服务员
- ✓ 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API 编制自己的应用程序



小结


- 数据库系统概述
 - 数据库的基本概念（掌握）
 - 数据管理的发展过程（了解）
 - 数据库系统的研究领域（了解）
- 数据模型
 - 数据模型的三要素（掌握）
 - 概念模型， **E-R** 模型（掌握）
 - 三种主要数据模型（了解）



小结

- 数据库系统的结构（掌握）
 - 数据库系统三级模式结构
 - 数据库系统的体系结构
- 数据库系统的组成（掌握）

下面一章我们将从数据库的基础关系数学给大家阐述数据库的应用原理。

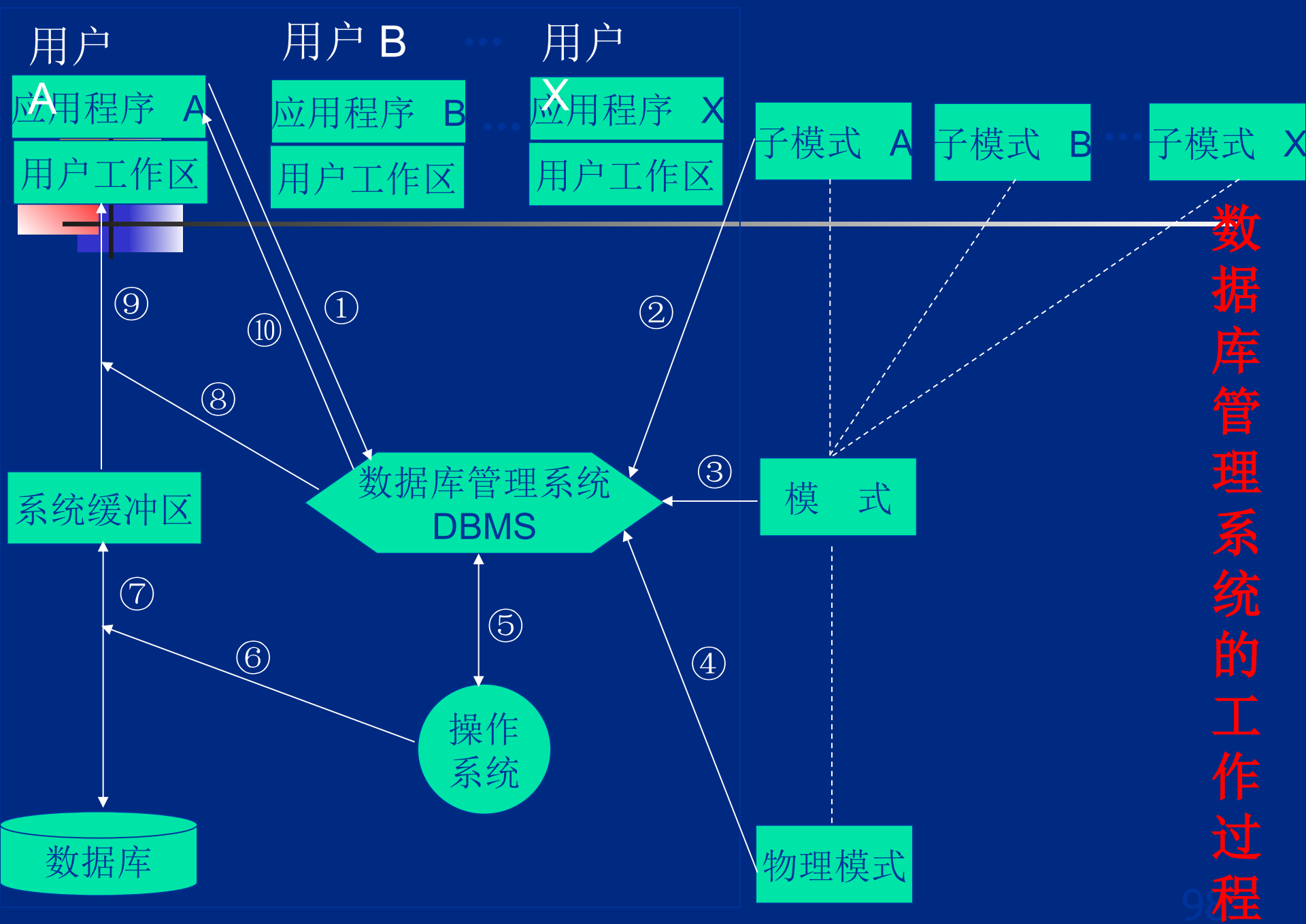


学校中有若干系，每个系有若干班级和教研室，每个教研室有若干教师（不兼职其他部门），每个教师教若干门课程；每个班有若干学生，每个学生选修若干课程。请根据逻辑分析，用 E-R 图画出此学校的概念模型。



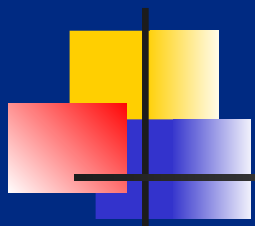
DBMS 层次结构和工作过程

- 应用层;
- 语言处理层;
- 数据存取层;
- 数据存储层;



1. 应用程序 **A** 向 **DBMS** 发出从数据库重读数据记录的命令；
2. **DBMS** 对该命令进行语法检查、语义检查，并调用应用程序 **A** 对应的子模式，检查 **A** 的存取权限，决定是否执行该命令。如果拒绝执行，则向用户返回错误信息；
3. 在决定执行该命令后，**DBMS** 调用模式，依据子模式 / 模式映象的定义，确定应读入模式中的哪些记录；
4. **DBMS** 调用物理模式，依据模式 / 物理模式映象的定义，决定应从哪个文件、用什么存取方式、读入哪个或物理记录；

5. **DBMS** 向操作系统发出执行读取所需物理记录
的命令；
6. 操作系统执行读数据的有关操作；
7. 操作系统将数据从数据库的存储区送至系统缓
冲区；
8. **DBMS** 依据子模式 / 模式映象的定义，导出应
用程序 **A** 所要读取得记录格式；
9. **DBMS** 将数据记录从系统缓冲区传送到应用程
序 **A** 的用户工作区；
10. **DBMS** 向应用程序 **A** 返回命令执行情况的状态
信息；



作业 13

- 作业：

习题 9 （教材 page 36）