

数据库系统概述

授课教师：陈志泊

联系方式：zhibo@bjfu.edu.cn

数据库技术的 产生与发展



数据库技术的产生与发展

随着计算机硬件和软件的发展，数据管理技术经历了从低级到高级的发展阶段，即：

○ 人工管理阶段

○ 文件系统阶段

○ 数据库管理阶段



人工管理阶段

人工管理阶段 (1946年-50年代中期以前)

1946年，美国宾夕法尼亚大学莫尔学院为美国陆军军械部研制的大型计算机，名为 *ENIAC*（电子数字积分计算机）

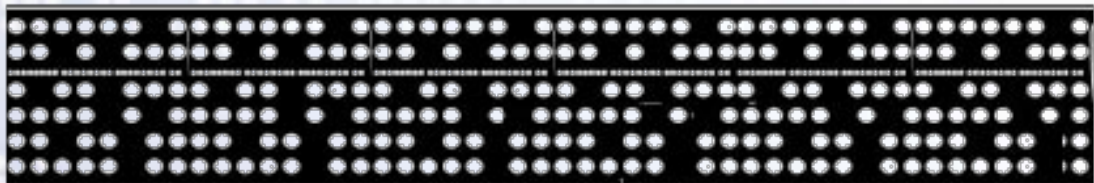


人工管理阶段

计算机软、硬件的特点：

计算机主要用于**科学计算**

硬件中的外存只有卡片、纸带、磁带，没有**磁盘**等直接存取设备





人工管理阶段的特点

(1) 数据不保存

科学计算，数据量少，对于数据保存的需求尚不迫切，只有在计算某一课题时才将原始数据随程序一起输入内存，计算任务完成时，数据和程序一起从内存中被释放。



人工管理阶段的特点

(2) 没有专门管理数据的软件

数据由应用程序自己管理，每个应用程序既要设计数据的逻辑结构，又要设计物理结构以及数据的存取方式，程序员负担重。



人工管理阶段的特点

(3) 数据不能共享，冗余很大

数据是面向程序的，一组数据只能对应一个程序，程序间的数据不能共享，冗余很大。



人工管理阶段应用程序与数据之间的对应关系



人工管理阶段的特点

(4) 数据不具有独立性

数据由应用程序自己管理，数据和程序是不可分割的整体，如果数据的类型、格式或输入/输出方式等逻辑结构或物理结构发生变化，程序员必须对应用程序做出相应的修改。因此，数据无独立性，数据对应用程序的依赖性太强。



文件系统阶段

文件系统阶段（50年代后期至60年代中期）

- ◆ 计算机不仅用于科学计算，还大量用于信息管理。随着数据量的增加，数据的存储、检索和维护成为紧迫的需要
- ◆ 硬件有了磁盘、磁鼓等直接存储设备
- ◆ 软件方面，出现了高级语言和操作系统。操作系统中有了专门管理数据的软件，称为文件系统



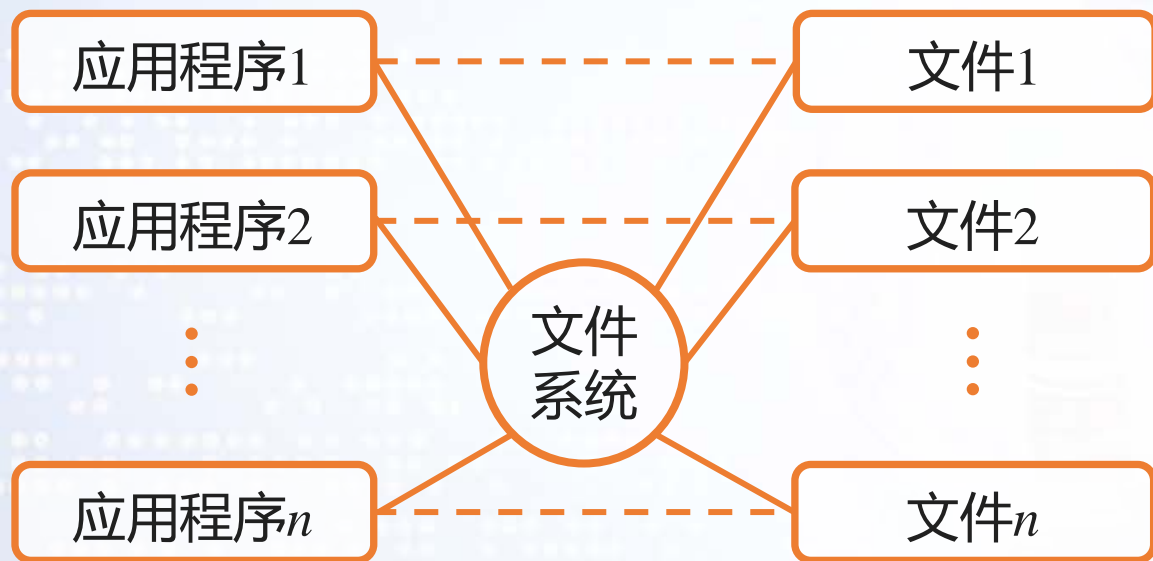
文件系统阶段的特点

数据可以文件形式长期保存在外部存储器的磁盘上，可被多次反复使用，应用程序对文件进行查询、修改和插入操作。

文件系统对数据进行管理。数据组织成具有一定结构的记录，并以文件的形式存储在存储设备上，程序只需用文件名就可与数据打交道，不必关心数据的物理存储（位置、结构等），由文件系统提供存取方法（读/写）。



文件系统阶段的特点

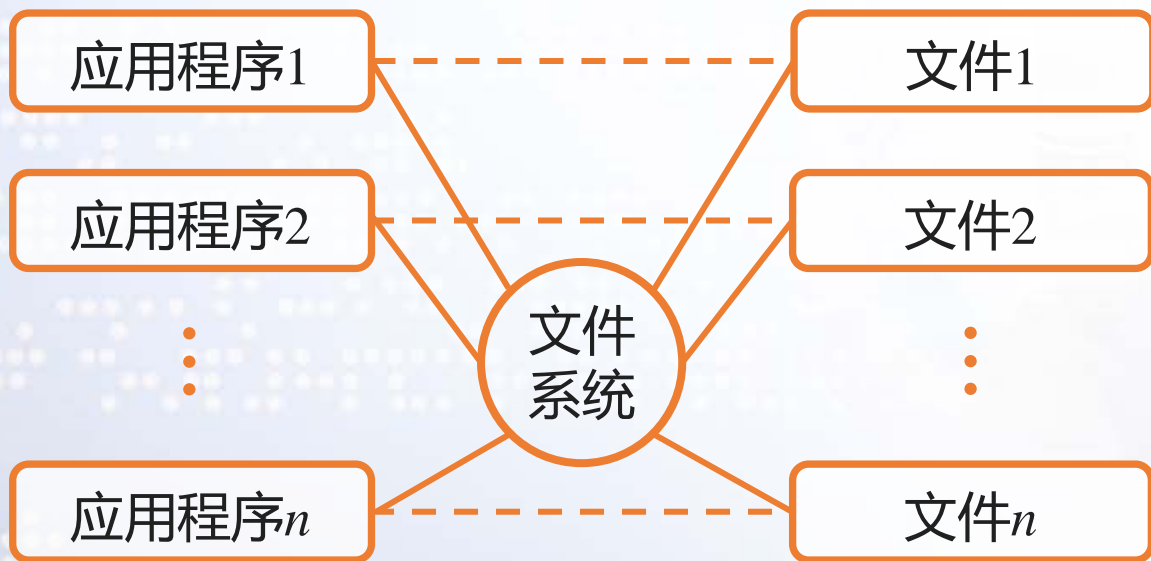


文件系统阶段应用程序与数据之间的对应关系



文件系统阶段的特点

数据和程序有了一定的独立性。 文件系统在程序与数据文件之间的存取转换作用。





文件系统阶段的特点



文件组织形式多样化，便于存储和查找数据，如顺序文件、索引文件等。



数据具有一定的共享性。数据不再属于某个特定的程序，可以重复使用。



文件系统阶段还存在如下一些问题

(1) 数据共享性差、冗余度大。一个文件基本上对应于一个应用程序，即文件仍然是面向应用的，文件间相互独立，缺乏联系。

(2) 数据不一致性。这通常是由数据冗余造成的。



文件系统阶段还存在如下一些问题

(3) 数据独立性差

文件结构的设计仍然基于特定的应用，一旦改变数据的逻辑结构，必须修改相应的应用程序。而应用程序发生变化，如改用另一种程序设计语言来编写程序，也需修改数据结构，程序与数据间的依赖关系并未根本改变。

(4) 数据间的联系弱。 文件与文件之间是独立的，文件间的联系必须通过程序来构造。



数据库系统阶段

数据库系统阶段（60年代中期以后）

- ◆ 硬件方面出现了大容量、存取快速的磁盘，使计算机联机存取大量数据成为可能。
- ◆ 硬件价格下降和软件价格上升，使开发和维护系统软件的成本相对增加。
- ◆ 计算机应用于管理的规模更加庞大，数据量急剧增加，文件系统的管理方法已无法适应各种应用的需要。



数据库系统阶段

计算机技术的发展、数据管理的需求迫切性，共同促使人们研究一种新的数据管理技术——数据库技术。

数据库技术是把一批相关数据组织成数据库，并对其进行集中、统一的管理，实施很强的安全性和完整性控制的技术。



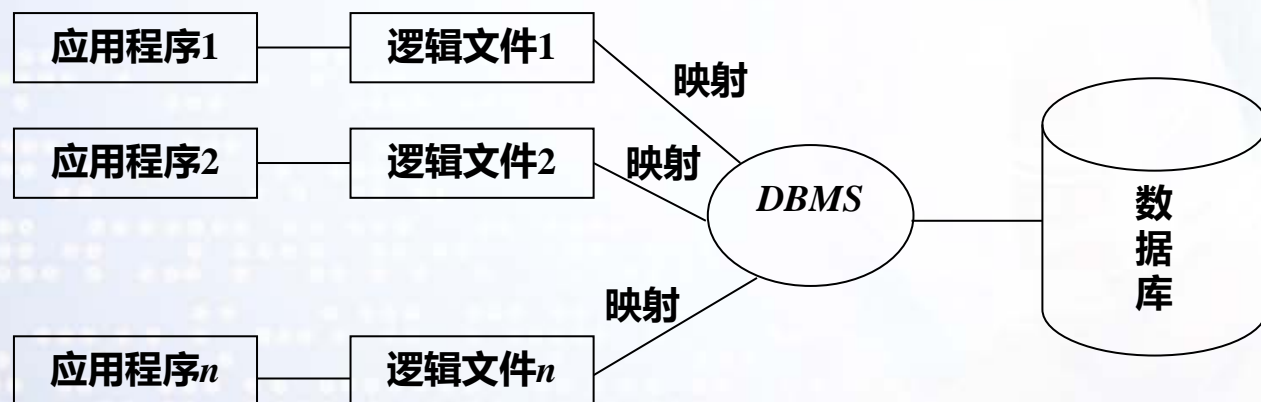
数据库系统阶段的特点

(1) 数据的结构化。 数据及其联系按照数据模型组织到结构化的数据库中，且面向全组织的所有应用。



数据库系统阶段的特点

(2) 数据共享性高、冗余度低。数据库中的一组数据集合可为多个应用和多个用户共同使用。



由数据库管理系统 (DataBase Management System, DBMS) 实现各应用程序对数据库中的数据共享



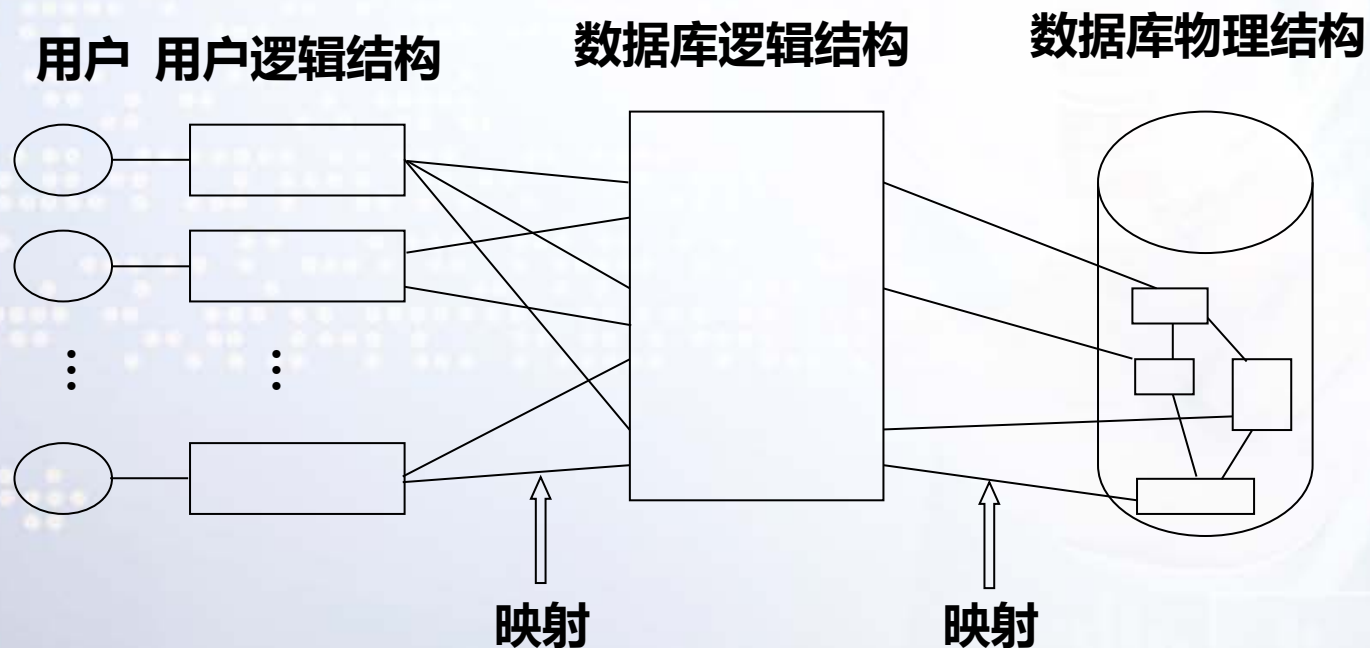
数据库系统阶段的特点

(3) 数据独立性高。即数据库中的数据与应用程序间相互独立，即数据的逻辑结构、存储结构以及存取方式的改变不影响应用程序。



数据库系统阶段的特点

在数据库系统中，整个数据库的结构可分成三级：用户逻辑结构、数据库逻辑结构和物理结构。数据独立性分两级：物理独立性和逻辑独立性。





数据库系统阶段的特点

数据的物理独立性是指当数据库物理结构（如存储结构、存取方式、外部存储设备等）改变时，通过修改映射，使数据库逻辑结构不受影响，进而用户逻辑结构以及应用程序不用改变。

数据的逻辑独立性是指当数据库逻辑结构（如修改数据定义、增加新的数据类型、改变数据间的关系等）发生改变时，通过修改映射，使用户逻辑结构以及应用程序不用改变。



数据库系统阶段的特点



(4) 数据由DBMS统一管理和控制，有统一的数据管理和控制功能。

数据控制功能包括：数据的安全性控制、完整性控制、并发控制、数据恢复。

- ◆ 安全性控制：防止不合法使用数据库造成数据的泄露和破坏。合法用户只能操作有权限的数据，不合法的用户禁止访问。



数据库系统阶段的特点

- ◆ 完整性控制：通过设置一些完整性规则等约束条件，确保数据的正确性、有效性和相容性。
- ◆ 并发控制：多个用户同时存取或修改数据库时，系统可防止由于相互干扰而提供给用户不正确的数据，并防止数据库受到破坏。

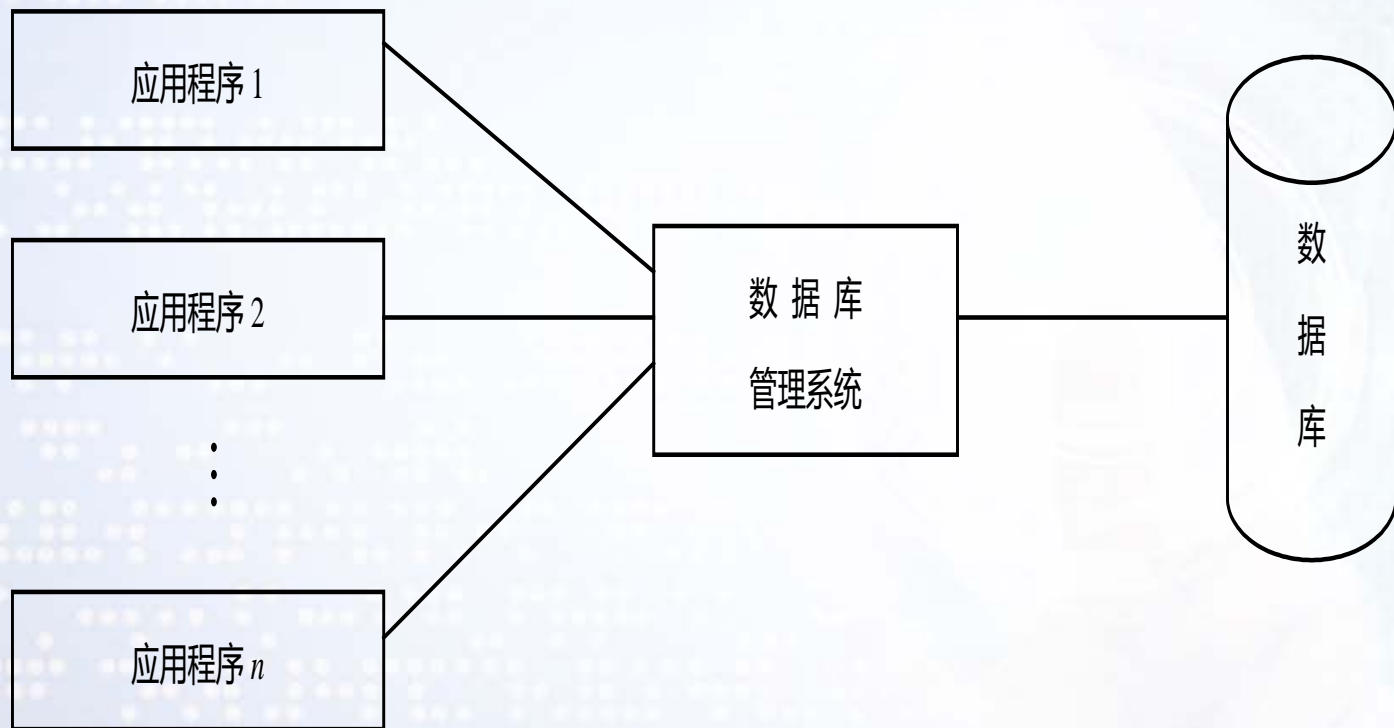


数据库系统阶段的特点

数据恢复：由于计算机系统的软硬件故障、操作员的误操作及其他故意的破坏等原因，造成数据库中的数据不正确或数据丢失时，系统有能力将数据库从错误状态恢复到最近某一时刻的正确状态。



数据库系统阶段的特点



数据库系统阶段程序与数据之间的关系



数据库系统阶段的特点

	人工管理阶段	文件系统阶段	数据库系统阶段
数据的管理者	人	文件系统	数据库管理系统
数据面向的对象	某一应用程序	某一应用程序	整个应用系统
数据的共享程度	无共享,冗余度极大	共享性差,冗余度大	共享性高,冗余度小
数据的独立性	不独立,完全依赖于程序	独立性差	具有高度的物理独立性和逻辑独立性
数据的结构化	无结构	记录内有结构,整体无结构	整体结构化,用数据模型描述
数据控制能力	应用程序自己控制	应用程序自己控制	由数据库管理系统提供数据安全性、完整性、并发控制和恢复能力

北京林业大学

数据库原理与应用

授课教师：陈志泊

联系方式：zhibo@bjfu.edu.cn

数据库系统的组成



数据库系统的组成

数据库系统 (DataBase System,简称DBS)

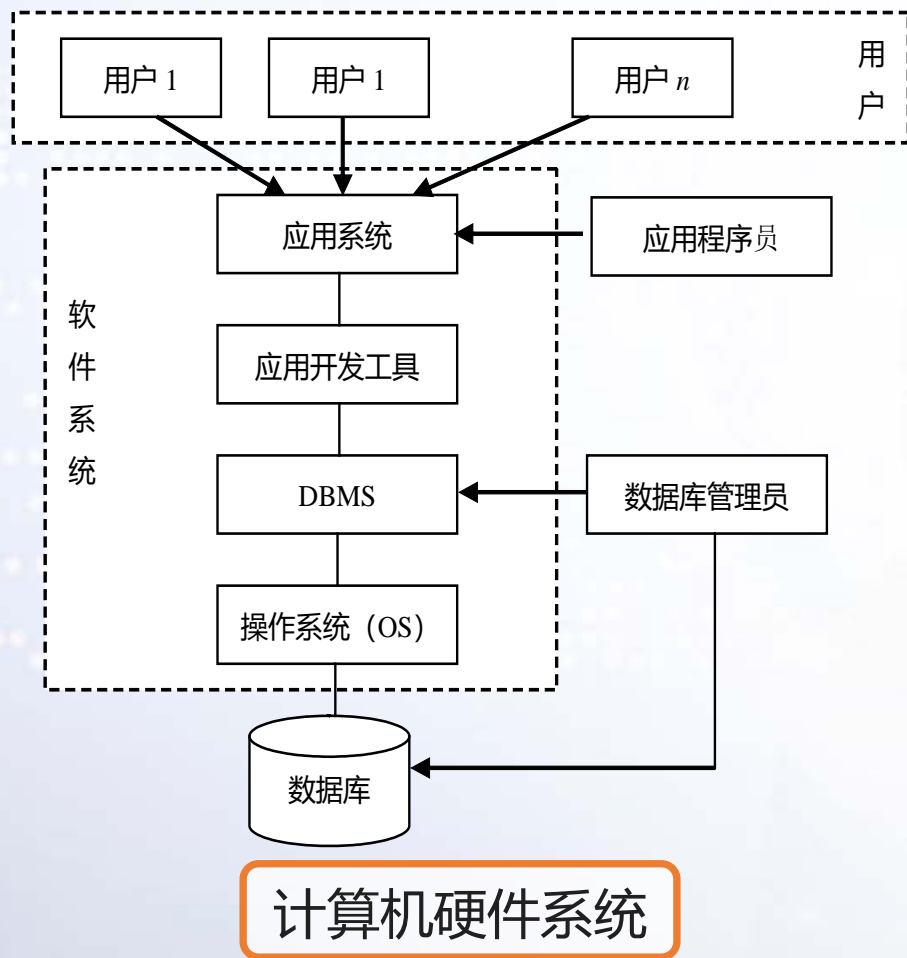
是以计算机软硬件为工具，把数据组织成数据库形式并对其进行存储、管理、处理和维护数据的高效能的信息处理系统。

由计算机硬件系统、数据库、软件系统（含操作系统、应用程序开发工具、数据库应用系统）、数据库管理系统（DBMS）、数据库用户组成。



数据库系统的组成

数据库系统 (DBS)





数据库系统的组成_1、硬件系统

1、硬件 (Hardware) 系统

指存储和运行数据库系统的硬件设备。包括CPU、内存、大容量的存储设备、输入/输出设备和外部设备等。



数据库系统的组成_2、数据库

2、数据库 (DataBase, DB)

是存储在计算机内、有组织的、可共享的数据和数据对象的集合，这种集合按一定的数据模型（或结构）组织、描述并长期存储，同时能以安全和可靠的方法进行数据的检索和存储。



数据库系统的组成_2、数据库

数据库的两个特点

(1) 集成性

将某特定应用环境中的各种应用相关的数据及其数据之间的联系全部集中地并按照一定的结构形式进行存储。

(2) 共享性

数据库中的数据可为多个不同的用户所共享，可同时存取数据库，甚至同时存取数据库中的同一数据。





数据库系统的组成_3、数据库用户

3、数据库用户

即使用数据库的人，对数据库进行存储、维护和检索等操作。



第一类用户：
最终用户 (End User)

非计算机专业人员

利用已编写好的应用程序接口使用数据库



数据库系统的组成_3、数据库用户



第二类用户：
应用程序员 (Application Programmer)

为最终用户设计和编制应用程序
并进行调试和安装



第三类用户：
数据库管理员 (DataBase Administrator, DBA)

负责设计、建立、管理和维护数据库
以及协调用户对数据库要求的个人或
工作团队



数据库系统的组成_3、数据库用户

DBA的主要职责如下：

- 01 参与数据库设计的全过程，决定整个数据库的结构和内容。
- 02 决定数据库的存储结构和存取策略，以获得较高的存取效率和存储空间利用率。
- 03 帮助应用程序员使用数据库系统，如培训、解答应用程序员日常使用数据库系统时遇到的问题等。



数据库系统的组成_3、数据库用户

04

定义数据的安全性和完整性，负责分配各个应用程序对数据库的存取权限。

05

监控数据库的使用和运行，DBA负责定义和实施适当的数据库备份和恢复策略；当数据库的结构需要改变时，完成对数据结构的修改。

06

改进和重构数据库，DBA负责监视数据库系统运行期间的空间利用率、处理效率等性能指标。



数据库系统的组成_4、软件系统

4、软件 (Software) 系统:

软件系统主要包括操作系统 (Operating System , OS) 、 数据库管理系统 (DBMS) 、 应用程序开发工具和数据库应用系统等。

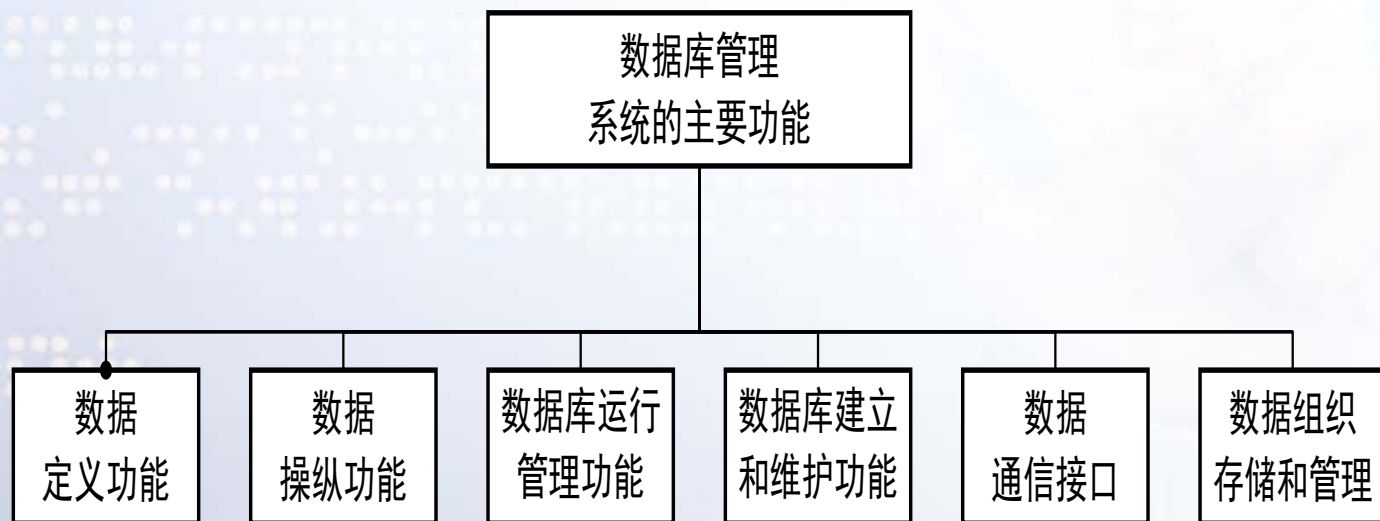


数据库系统的组成_4、软件系统

数据库管理系统:

(DataBase Management System, DBMS)

DBMS是数据库系统的核心软件。





数据库系统的组成_数据库管理系统的功能

① (1) 数据定义功能

DBMS提供数据定义语言 (Data Define Language, DDL)

定义数据的模式、外模式和内模式三级模式结构;

定义模式/内模式和外模式/模式二级映像;

定义有关的约束条件。

如：定义用户口令和存取权限、定义完整性规则；DBMS提供的结构化查询语言 (SQL) 提供Create、Drop、Alter等语句可分别用来建立、删除和修改数据库。



数据库系统的组成_数据库管理系统的功能

◎ (2) 数据操纵功能

DBMS提供数据操纵语言 (Data Manipulation Language, DML) 实现对数据库的基本操作, 包括检索、更新 (包括插入、修改和删除) 等。

DML有两类:

一类是自主型的或自含型的, 可单独使用

另一类是宿主型的, 需要嵌入在高级语言中, 不能单独使用

例如

DBMS提供的结构化查询语言SQL提供INSERT、DELETE、UPDATE、SELECT可分别实现对数据库中数据的增、删、改、查等操作。



数据库系统的组成_数据库管理系统的功能

○ (3) 数据库运行管理功能

是DBMS的核心部分。

DBMS通过对数据库的控制以确保数据正确、有效和数据库系统的正常运行。

DBMS对数据库的控制主要有4个方面：

数据的安全性控制

数据的完整性控制

多用户环境下的数据
并发性控制

数据库的恢复



数据库系统的组成_数据库管理系统的功能



(4) 数据库的建立和维护功能

由DBMS的各个实用程序完成相关功能：

- ◆ 数据库的建立包括数据库的初始数据的装入与数据转换等。
- ◆ 数据库的维护包括数据库的转储、恢复、重组与重构造、系统性能监视与分析等。



数据库系统的组成_数据库管理系统的功能



(5) 数据组织、存储和管理

DBMS负责对需要存放的各种数据的组织、存储和管理工作，确定以何种文件结构和存取方式物理地组织这些数据，以提高存储空间利用率和对数据库进行增、删、查、改的效率。



数据库系统的组成_数据库管理系统的功能

◎ (6) 数据通信接口

DBMS提供与其他软件系统进行通信的功能。

DBMS提供了与其他DBMS或文件系统的接口，实现用户程序与DBMS、DBMS与DBMS、DBMS与文件系统之间的通信与数据交换。通常这些功能要与操作系统协调完成。



数据库系统的组成

DBMS在操作系统支持下工作，应用程序在DBMS支持下才能使用数据库。



数据库管理系统在整个计算机系统中的地位



数据库系统的组成

教学管理系统

主要涉及学生、教师、课程、选课、授课等信息的管理。

主要的数据库项如下：

- ◆ 学生信息：学号、姓名、性别、年龄、系等
- ◆ 教师信息：教师号、姓名、性别、年龄、职称、工资、津贴、系等
- ◆ 课程信息：课程号、课程名称、学时等
- ◆ 选课信息：学号、课程号、成绩等
- ◆ 授课信息：教师号、课程号

典型的操作包括学生成绩统计、学生选课信息统计、教师授课信息统计等。

北京林业大学

数据库原理与应用

授课教师：陈志泊

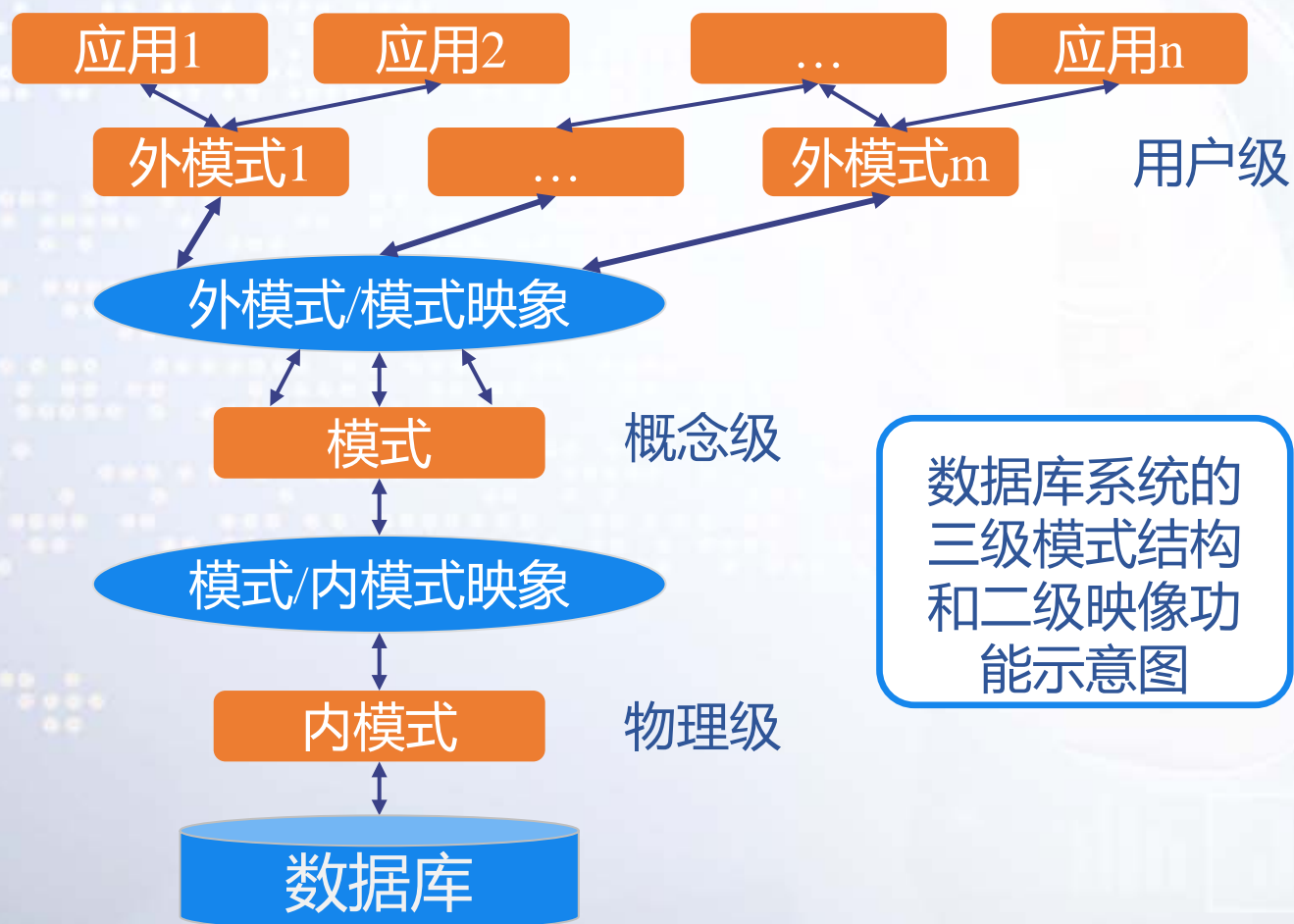
联系方式：zhibo@bjfu.edu.cn

数据库系统的内部 体系结构



数据库系统的内部体系结构

从DBMS角度来看，数据库系统通常采用三级模式结构，这是数据库系统的内部体系结构。





数据库系统的内部体系结构

1、模式 (Schema)

又称概念模式，处于中间层，是对数据库中全体数据的逻辑结构和特征的描述，是数据库的整体逻辑，即概念视图、概念级数据库。



数据库系统的内部体系结构

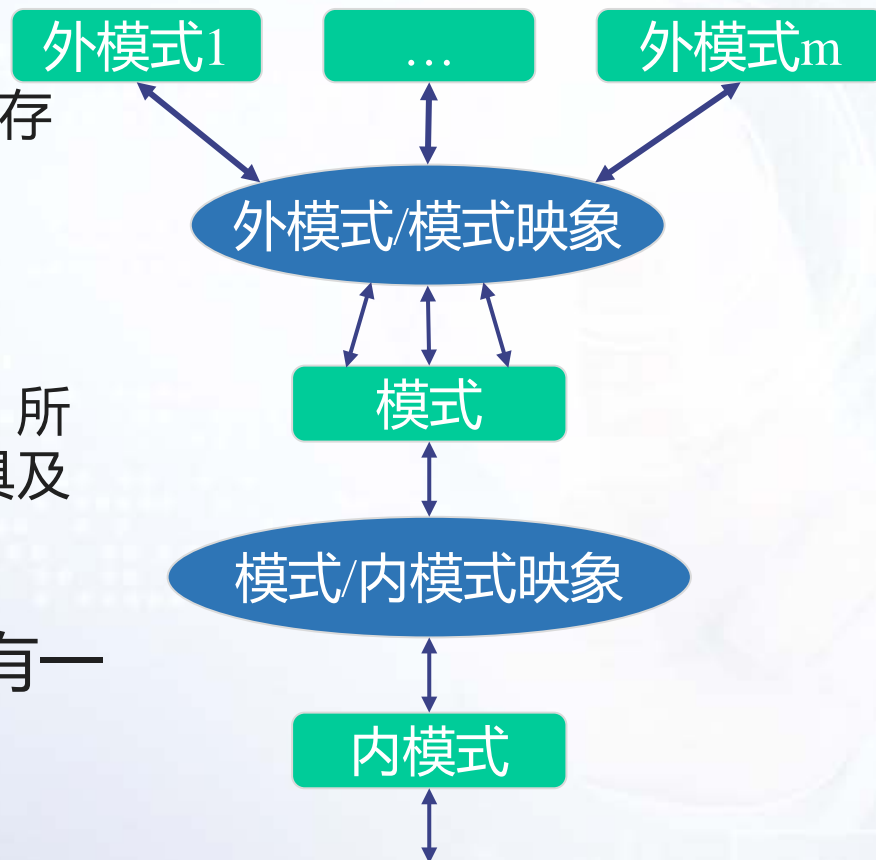
对下层：

不涉及数据的物理存储细节和硬件环境；

对上层：

与具体的应用程序、所使用的应用开发工具及程序设计语言无关。

◆ 一个数据库只有一个模式。





数据库系统的内部体系结构

2、外模式 (External Schema)

又称子模式或用户模式，处于最外层，是对数据库用户能看到并允许使用的那部分局部数据的逻辑结构和特征的描述，是与某一应用有关的数据的逻辑表示，即用户视图、用户数据库。

◆ 外模式是模式的子集，可有多，同一个外模式可以为多个应用程序使用。

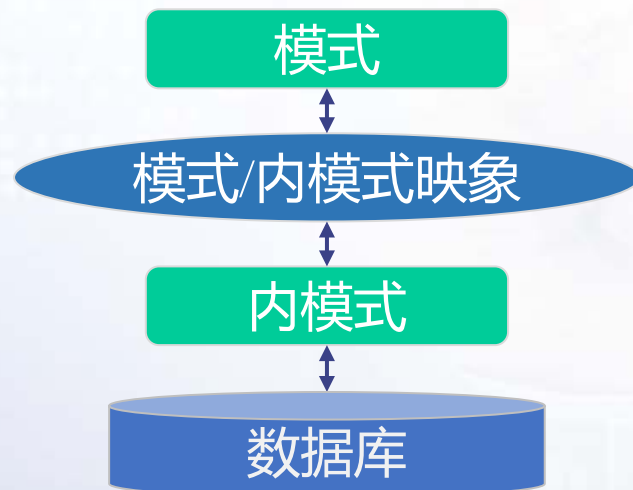


数据库系统的内部体系结构

3、内模式 (Internal Schema)

又称为存储模式或物理模式，处于最内层，也是靠近物理存储的一层，是对整个数据库存储结构的描述，是数据在数据库内部的表示方式，又叫物理级数据库、物理视图。

◆ 一个数据库只有一个内模式。





数据库系统的内部体系结构

小结:

内模式:

物理视图，物理级数据库，是整个数据库实际存储的表示。

模 式:

概念视图，概念级数据库，是整个数据库实际存储的抽象表示。

外模式:

用户视图，用户级数据库，是概念模式的某一部分的抽象表示。



数据库系统的内部体系结构

小结:

- ◆ 一个数据库系统，实际存在的只是物理级数据库，即内模式，是数据访问的基础。
- ◆ 概念数据库只不过是物理级数据库的一种抽象描述
- 用户级数据库是用户与数据库的接口。

用户根据外模式进行的操作，通过外模式到模式的映射与概念级数据库联系起来，又通过模式到内模式的映射与物理级数据库联系起来。



数据库系统的内部体系结构

数据库系统的二级映象与数据独立性

◆ DBMS在三级模式之间提供了二级映象功能，保证了数据库系统中的数据能够具有较高的**逻辑独立性**与**物理独立性**。

外模式/
模式映象



保证了数据与程序间的逻辑独立性

模式/内
模式映象



确保了数据的物理独立性



数据库系统的内部体系结构

数据库系统的三级模式与二级映像的优点



数据库系统的外部 体系结构



数据库系统的外部体系结构

从最终用户角度来看，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

浏览器/服务器结构

这是数据库系统的外部体系结构。



数据库系统的外部体系结构

1、单用户结构的数据库系统

又称桌面型数据库系统

将应用程序、DBMS和数据库都装在一台计算机上，由一个用户独占使用。适合未联网用户、个人用户等。

DBMS提供较弱的数据库管理和较强的应用程序和界面开发工具，既是数据库管理工具，同时又是数据库应用程序和界面的前端工具。

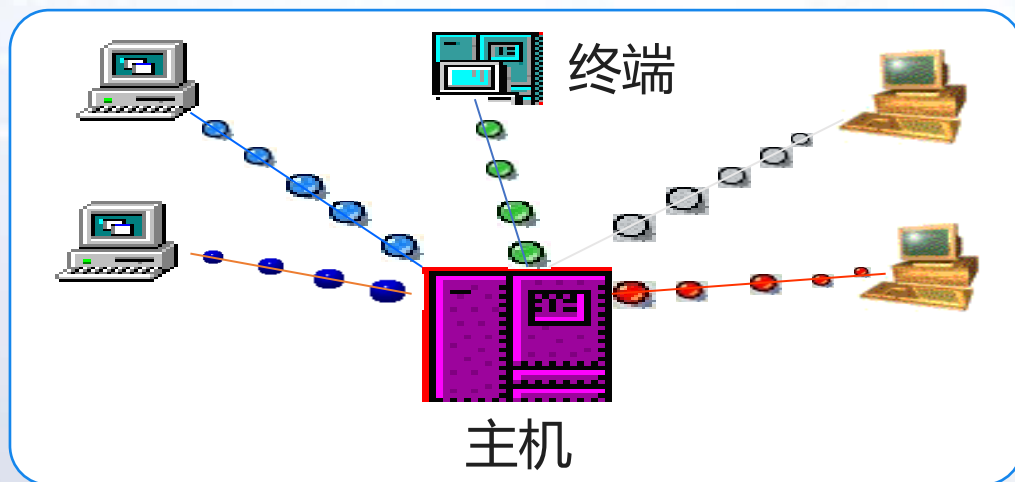
如：Microsoft Access, Visual Foxpro等



数据库系统的外部体系结构

2、主从式结构的数据库系统

是大型主机带多终端的多用户结构的系统，又称主机/终端模式





数据库系统的外部体系结构

优点

- ◆ 结构简单，易于管理、控制与维护。

缺点

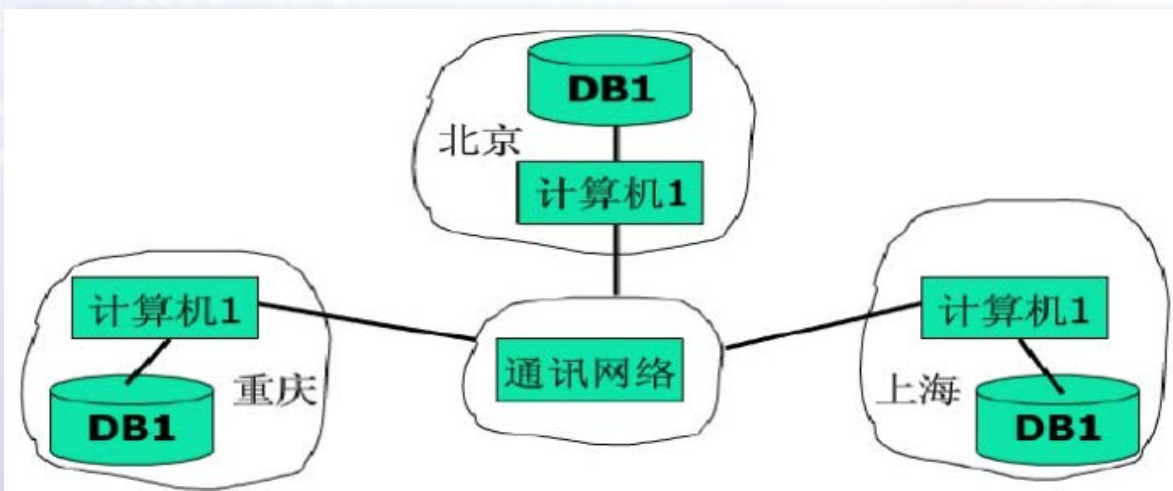
- ◆ 当终端数目太多时，主机的任务会过分繁重，成为系统瓶颈。
- ◆ 系统的可靠性依赖主机，当主机出现故障时，整个系统都不能使用。



数据库系统的外部体系结构

3、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。数据库分布存储在计算机网络的
不同结点上。





数据库系统的外部体系结构

- (1) 数据在物理上是分布的
- (2) 所有数据在逻辑上是一个整体
- (3) 结点上分布存储的数据相对独立

优点 多台服务器并发的处理数据，提高效率

缺点 数据的分布式存储给数据处理任务协调与维护带来困难。



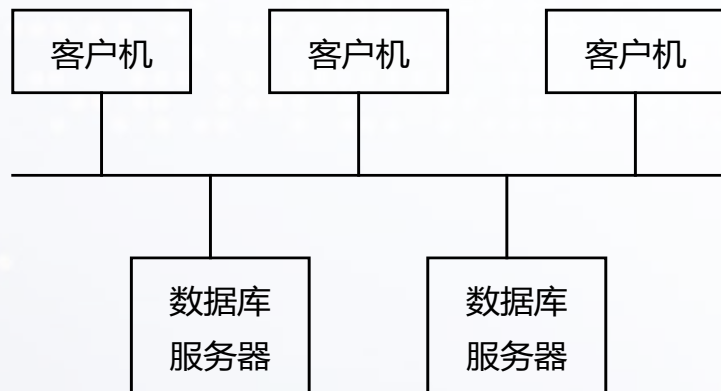
数据库系统的外部体系结构

4、客户/服务器结构的数据库系统

(Client/Server, C/S结构)

把DBMS的功能与应用程序分开。分为：

- 数据库服务器（简称服务器，Server）
- 客户机（Client）



DBMS 应用开发工具、应用程序等

网络环境

DBMS、数据库等



数据库系统的外部体系结构



胖客户机结构，两层结构。

优点

网络运行效率大大提高

缺点

维护升级很不方便



数据库系统的外部体系结构

5、浏览器/服务器结构的数据库系统 (Browser/Server, B/S结构)

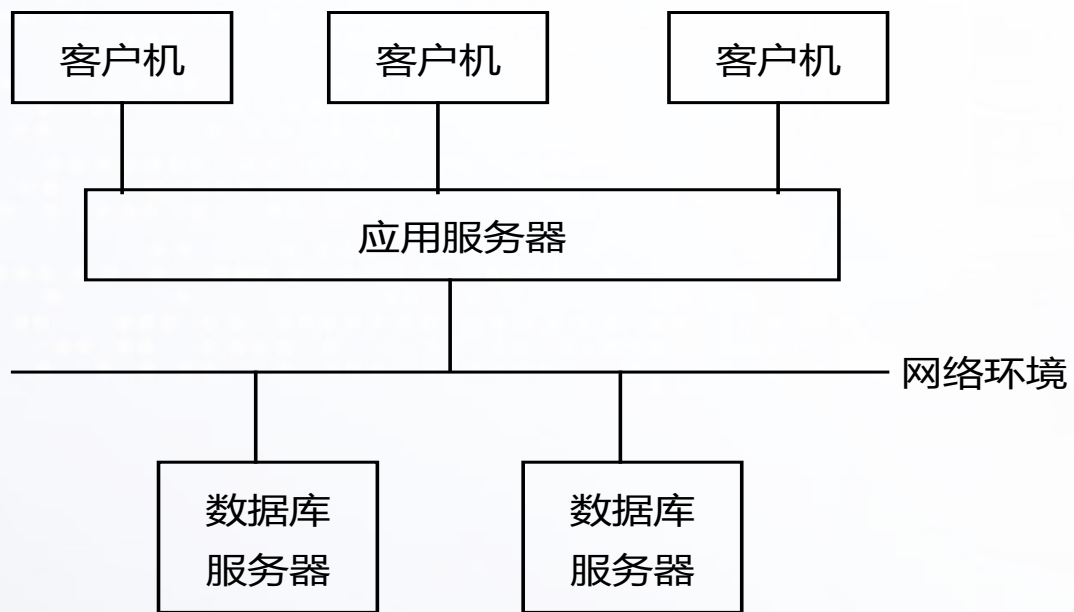
- 针对客户机/服务器结构的不足而提出的。
- 客户机仅安装通用的浏览器软件，实现输入/输出
- 应用程序安装在应用服务器上，充当了中介



数据库系统的外部体系结构



瘦客户机结构，三层结构。



北京林业大学

数据库原理与应用

授课教师：陈志泊

联系方式：zhibo@bjfu.edu.cn

三个世界及其 有关概念



三个世界及其有关概念



数据库管理的对象(数据)存在于现实世界中,即现实世界中的事物及其各种联系。从现实世界的事物到存储到计算机的数据库中的数据,要数经历**现实世界**、**信息世界**和**计算机世界**三个不同的世界,经历两级抽象和转换完成。



三个世界及其有关概念



数据处理的抽象和转换过程



三个世界及其有关概念

1、现实世界

即客观存在的世界，由客观存在的事物及其联系所组成。

人们总是选用感兴趣的最能表征一个事物的若干特征来描述该事物，例如，选用学号、姓名、性别、年龄、系等来描述学生，有了这些特征，就能区分不同的学生。

客观世界中，事物之间是相互联系的，但人们只选择那些感兴趣的联系。如可以选择“学生选修课程”这一联系表示学生和课程之间的关系。



三个世界及其有关概念

2、信息世界（概念世界）

是现实世界在人们头脑中的反映，经过人脑的分析、归纳和抽象，形成信息，人们把这些信息进行记录、整理、归类 and 格式化后，就构成了信息世界。

信息世界是对客观事物及其联系的一种抽象描述。
如：学生信息、教师信息等



三个世界及其有关概念

信息世界（概念世界）

从现实世界到概念世界是通过概念模型来表达的。

如：E-R模型

概念模型：又叫信息模型，是按用户的观点对数据和信息建模，不依赖于具体的计算机系统，只是用来描述某个特定组织所关心的信息结构。

客观事物在信息世界中的抽象表示，如：学生、教师



三个世界及其有关概念

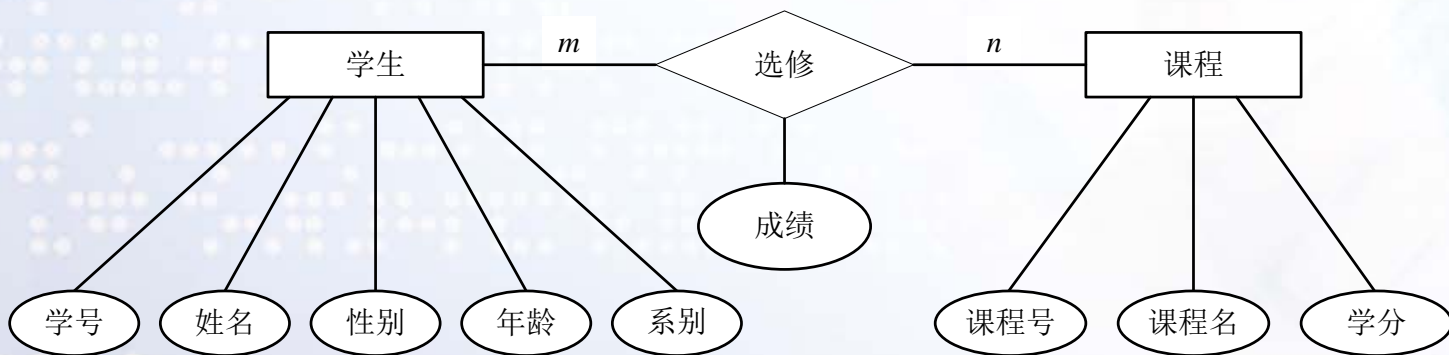




三个世界及其有关概念

信息世界（概念世界）

客观事物之间的联系在信息世界中的抽象表示。
如，学生和课程间的联系





三个世界及其有关概念



信息世界（概念世界）

在现实世界中事物、事物的属性以及事物之间的联系，抽象到信息世界对应的就是实体、实体的属性以及实体之间的联系。



三个世界及其有关概念

信息世界及其有关基本概念

实体

客观存在并且可以相互区别的“事物”称为实体

实体可以是具体的人、事、物，也可以是抽象的事件

属性

实体所具有的某一特性称为属性

型

——> 属性名

值

——> 具体值



三个世界及其有关概念



如学生实体由学号、姓名、性别、年龄、系等方面的属性组成。

(990001、张立、20、男、计算机)，这些属性值的集合表示了一个学生实体。



三个世界及其有关概念

实体型

具有相同属性的实体必然具有共同的特征

用实体名及其属性名集合来抽象和描述同类实体，称为实体型，如：学生（学号，姓名，年龄，性别，系）

实体集

同型实体的集合称为实体集

同型实体的集合称为实体集。如所有的学生、所有的课程等。



三个世界及其有关概念



码

能唯一标识一个实体的属性或属性集称为实体的码

如：学生的学号就是学生实体的码。



域

某一属性的取值范围称为该属性的域

如：性别的域为男或女



三个世界及其有关概念

联系

客观事物内部以及事物之间是联系的，它们分别被抽象为：

- 单个实体型内部的联系，是指组成实体的各属性之间的联系
- 实体型之间的联系，是指不同实体集之间的联系



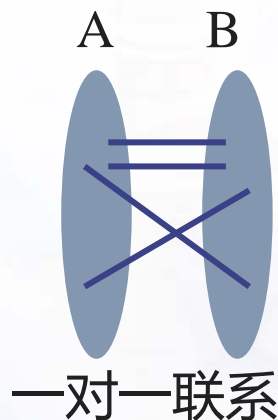
三个世界及其有关概念

两个实体型间的联系

两个实体型之间的联系是指两个不同的实体集间的联系，有三种类型：一对一联系、一对多联系、多对多联系。

一对一联系 (1:1)

实体集A中的一个实体至多与实体集B中的一个实体相对应，反之，实体集B中的一个实体至多与实体集A中的一个实体相对应，则称实体集A与实体集B为一对一的联系，记作1:1。如，班级与班长。

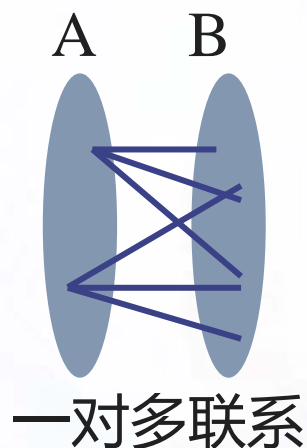




三个世界及其有关概念

一对多联系 (1:n)

实体集A中的一个实体与实体集B中的 n ($n \geq 0$) 个实体相联系，反之，实体集B中的一个实体至多与实体集A中的一个实体相联系，记作1:n。例如，班级与学生、公司与职员、省与市之间的联系。

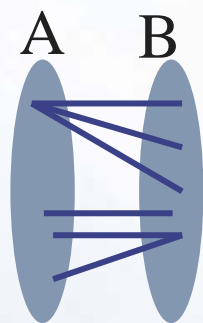




三个世界及其有关概念

多对多联系 (m:n)

实体集A中的一个实体与实体集B中的n ($n \geq 0$) 个实体相联系, 反之, 实体集B中的一个实体与实体集A中的m ($m \geq 0$) 个实体相联系, 记作m:n。如, 教师与学生、学生与课程、工厂与产品之间的联系。



多对多联系



三个世界及其有关概念

3、计算机世界

又叫数据世界，是对现实世界的第二层抽象，即对信息世界中信息的数据化，将信息用字符和数值等数据表示，**使用计算机存储并管理概念世界中描述的实体集、实体、属性和联系的数据**

信息世界到数据世界，使用数据模型来描述，数据库中存放数据的结构是由数据模型决定的



三个世界及其有关概念

计算机世界的概念

字段 Field

标记实体属性的命名单位称为字段，字段名往往和属性名相同。如学生有学号、姓名、年龄、性别和系等字段。

记录 Record

一个记录描述一个实体，字段的有序集合称为记录。如，一个学生（990001，张立，20，男，计算机）为一个记录。



三个世界及其有关概念

计算机世界的概念

文件

File

文件是用来描述实体集的。同一类记录的集合称为文件。如，所有学生的记录组成了一个学生文件。

关键字

Key

能唯一标识文件中每个记录的字段或字段集，称为记录的关键字。



三个世界及其有关概念

现实世界	→	信息世界	→	计算机世界
事物总体	→	实体集	→	文件
事物个体	→	实体	→	记录
特征	→	属性	→	字段
事物间联系	→	实体模型	→	数据模型

三个世界各术语的对应关系

北京林业大学

数据库原理及应用

授课教师：陈志泊

联系方式：zhibo@bjfu.edu.cn

数据模型



数据模型

- ✓ **现实世界中的事物及其联系，经过两级抽象和转换后形成了计算机世界中的数据及其联系，而数据模型就是用来描述数据及其联系的。**

数据库中存放数据的结构是由数据模型决定的，数据模型是数据库的框架，是数据库系统的核心和基础



数据模型

数据模型的概念

数据模型是描述数据、数据联系、数据的语义和完整性约束的概念集合，由**数据结构**、**数据操作**和**完整性约束**三要素组成。



数据模型

数据模型的组成要素

(1) 数据结构

即数据组织的结构，用于描述系统的静态特征，描述数据库的组成对象以及对象间的联系。

即：

一是描述数据对象的类型、内容、性质等；

二是描述数据对象间的联系。



数据模型

数据模型的组成要素

常用的数据结构有：

层次结构----层次模型----层次数据库

网状结构----网状模型----网状数据库

关系结构----关系模型----关系数据库



数据模型

数据模型的组成要素

(2) 数据操作

对数据库中的数据允许执行的操作的集合，包括操作及相应的操作规则（优先级）等，描述了数据库的动态特性。



一类是查询操作



一类是更新操作
(含插入、删除和修改)



数据模型

数据模型的组成要素

(3) 数据的完整性约束

一组完整性规则的集合。

完整性规则是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。



数据模型



常用的数据模型

(1)
层次模型

(2)
网状模型

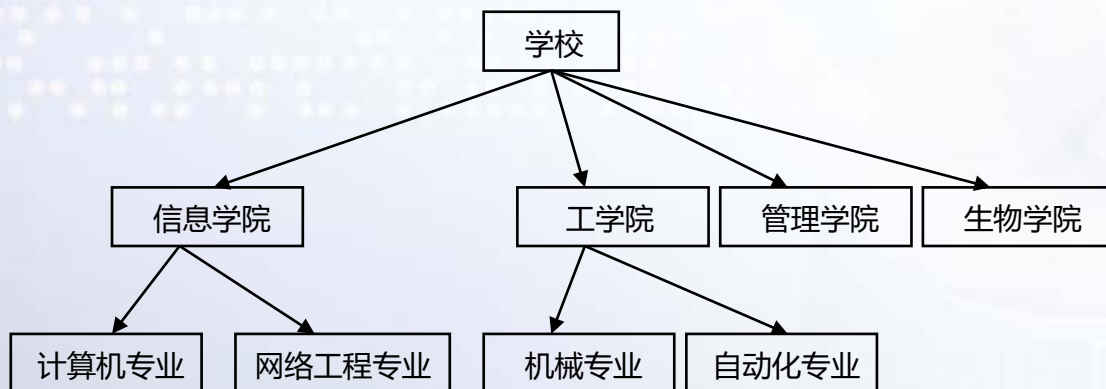
(3)
关系模型



数据模型

(1) 层次模型

- ◆ 层次模型是采用**树型结构（有根树）**来表示**实体及其实体间的联系**的模型。树形结构中的结点表示实体型，实体型间的联系用指针表示。
- ◆ 采用层次模型的数据库的典型代表是IBM公司1968年推出的IMS数据库管理系统





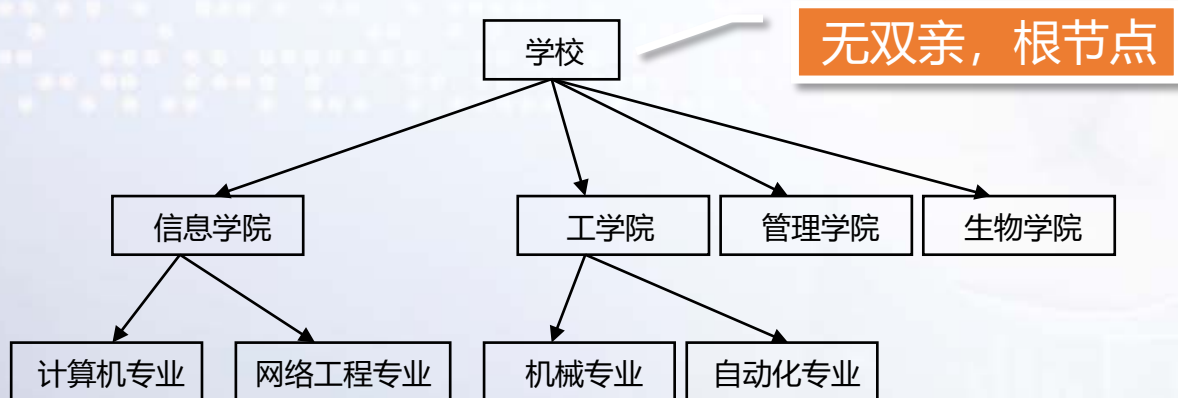
数据模型

(1) 层次模型

特点

- ◆ 有且仅有一个结点没有双亲，根结点
- ◆ 根以外的其他结点有且仅有一个双亲结点
- ◆ 父子结点之间的联系是一对多 (1:n) 的联系

任何一个给定的记录值只有按其路径查看时，才能显出它的全部意义





数据模型

(1) 层次模型

层次模型的数据操纵与数据完整性约束

- ◆ 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，如果删除双亲结点值，则相应的子女结点值也被同时删除。
- ◆ 修改操作时，应修改所有相应的记录，以保证数据的一致性。



数据模型

(1) 层次模型

优点

- ◆ 结构简单，层次分明
- ◆ 查询效率高，从根结点到树中任一结点均存在一条唯一的层次路径
- ◆ 提供良好的数据完整性支持

缺点

- ◆ 不能直接表示多对多联系
- ◆ 插入和删除数据限制太多
- ◆ 查询子女结点必须通过双亲结点



数据模型

(2) 网状模型

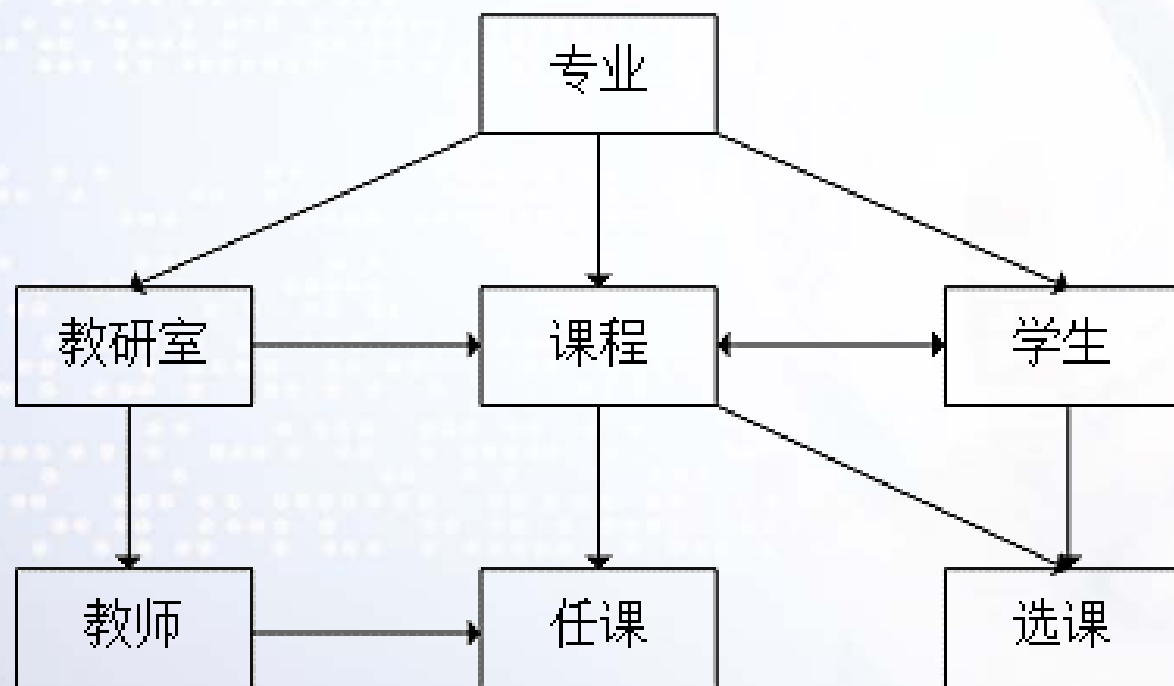
网状模型的数据结构

- ◆ 网状模型是采用有向图结构表示实体以及实体之间联系的数据模型。
- ◆ 每个结点表示一个实体型。
- ◆ 结点间的带箭头的连线（或有向边）表示记录型间的1：n的父子联系。
- ◆ 比较有代表性的系统是，20世纪70年代，数据系统语言研究会CODASYL提出的DBTG系统



数据模型

(2) 网状模型



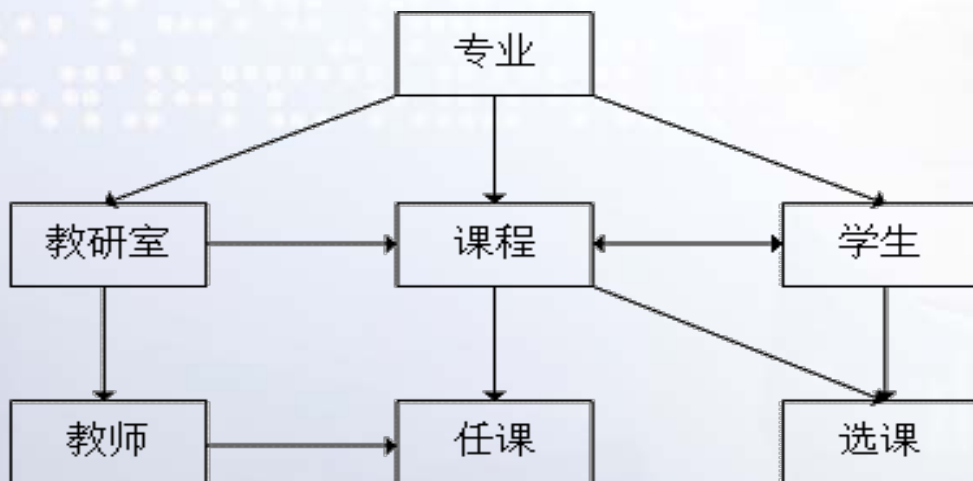


数据模型

(2) 网状模型

网状模型的特点：

- ◆ 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（复合联系）





数据模型

(2) 网状模型

网状模型的数据操纵与完整性约束

插入数据时，允许插入尚未确定双亲结点值的子女结点值。

删除数据时，允许只删除双亲结点值。

修改数据时，只需更新指定记录即可。



数据模型

(2) 网状模型

网状模型的优缺点

优点

- ◆ 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
- ◆ 数据定义语言、数据操纵语言复杂
- ◆ 用户需要了解网状模型的实现细节



数据模型

(3) 关系模型

关系模型的数据结构

- ◆ 以二维表（关系）的形式表示实体和实体之间联系的数据模型。
- ◆ 1970年美国IBM公司的研究员E.F.Codd提出，1977年IBM公司研制的关系数据库的代表System R。

小型数据库系统：
Foxpro、Access

大型数据库系统：
Oracle、SQL Server、
Informix、Sybase、MySQL



(3) 关系模型

关系模型的数据结构是一张规范化的二维表，它由表名、表头和表体三部分构成

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

属性

关系 ↔ 二维表

元组

域：
属性的取值范围
(男，女)



(3) 关系模型

分量

每一行对应的列的属性值，即为元组中的一个属性值。

候选码

可惟一标识一个元组的属性或属性集。
如S表中学号可以惟一确定一个学生，为学生关系的主码。



(3) 关系模型

关系模式

关系模式是对关系的描述，是关系模型的“型”。

一般表示为：关系名（属性1，属性2，……属性n）

如：学生（学号，姓名，性别，年龄，系别）



(3) 关系模型

关系模型的数据操纵与完整性约束

- 关系模型的数据操纵主要包括查询、插入、删除和修改
- 关系模型中的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。
- 关系模型把对数据的存取路径隐蔽起来，用户只要指出“干什么”，而不必详细说明“怎么干”，从而大大地提高了数据的独立性，提高了用户操作效率。



(3) 关系模型

关系模型的优缺点

优点:

- ◆ 有严格的数学理论根据
- ◆ 数据结构简单、清晰，用关系描述实体及其联系
- ◆ 具有更高的数据独立性、更好的安全保密性

缺点:

查询效率不如非关系模型