



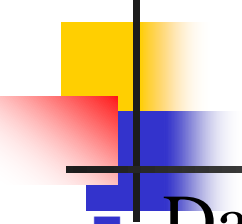
Database Principles

倪巍伟

wni@seu.edu.cn

教材及参考书	<p>以授课讲义为主，推荐如下参考书：</p> <p>✧ 王能斌，数据库系统教程，电子工业出版社，2008</p> <p>✧ Raghu Ramakrishnan, Johannes Gehrke, “Database Management Systems”, 3rd Edition</p> <p>✧ Hector Garcia-Molina, Jeffrey D. Ullman, “Database Systems: the Complete Book”</p>
--------	--

序号	实验项目名称	学时分配（实验学时/总学时）	实验内容提要	实验类型	实验要求	主要软、硬件实验平台和工具软件、人员分组等其他需要说明的情况
1	创建大学数据库	2	创建大学数据库并导入给定样例数据	设计性	必做	单人
2	SQL 查询	2	在大学数据库数据库上完成给定查询	设计性	必做	单人
3	ACCESS 小应用	2	自行构思并开发一个基于 ACCESS 的小应用，只需使用 ACCESS 的工具，不需要编程	设计性	必做	2~3 人合作
4	完整性约束	2	在大学数据库上进行插删改操作，分析并验证各种完整性约束	验证性	必做	单人
5	单机应用开发	2	在实验 3 小应用的设计基础上，完善功能设计，选用合适的程序开发环境开发更加完善的应用	设计性	选做	2~3 人合作
6	网上书店系统	6	自行设计设计开发一个较复杂的 B/S 架构 Web 应用程序——网上书店系统，要求按软件工程规范，分需求分析与概要设计、详细设计、最终结果与可运行程序三次提交作业	综合性	必做	3~4 人合作
			Database Principles, Ni Weiwei,			

- 
-
- Database management has become one of the core components of modern computing platform.
 - Starting from the platform's support for applications, this course discusses **the basic concepts, methods, principles and technologies of DB system** to improve **the design, development and system analysis capabilities**.



Chapter 1 Foreword

1.1 Development of data management

(1) Database technology comes from the needs of data management tasks

Data management refers to how to **classify, organize, code, store, retrieve and maintain data**, which is the central problem of data processing.

(2) Data management has gone through three stages

- . labor management
- . File system
- . Database system

背景	应用背景	人工管理	文件系统	数据库系统
	硬件背景	科学计算	科学计算、管理	大规模管理
	软件背景	无直接存取存储设备	磁盘、磁鼓	大容量磁盘
	处理方式	没有操作系统	有文件系统	有数据库管理系统
特点	数据的管理者	批处理	联机实时处理 批处理	联机实时处理, 分布处理批处理
	数据面向的对象	人	文件系统	数据库管理系统
	数据的共享程度	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



Data intensive application

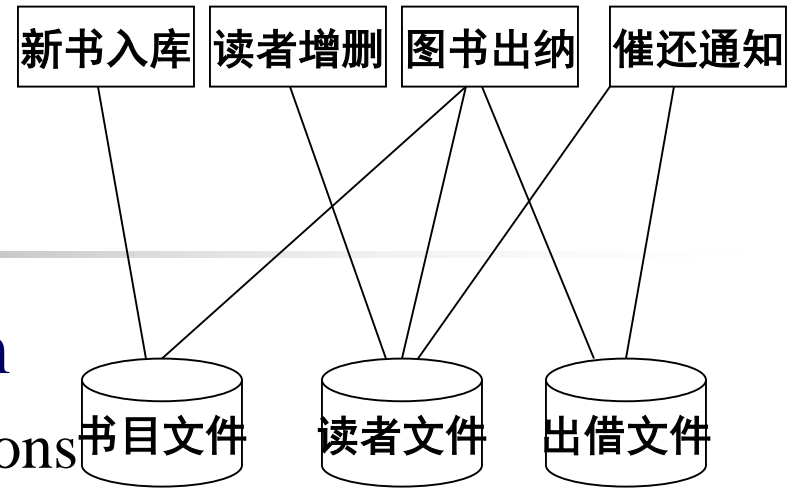
■ Characteristics of data-intensive applications

1. The amount of data is large(TB), which can not be accommodated in general memory, so it needs to be stored in auxiliary memory.
2. Data is persistent.
3. Data is shared by many applications, such as MIS, bank, office system and network data service.

Data management

■ Disadvantages of file system

1. Inconvenient to write applications
2. Data redundancy
3. Every modification of the file structure will lead to the modification of the application program, and the maintenance workload of the application program is heavy
4. Not support concurrent access to files
5. Standardization of data is difficult to achieve.





Database System

Data is a shared resource of an enterprise, which is managed by the database management system software(DBMS).

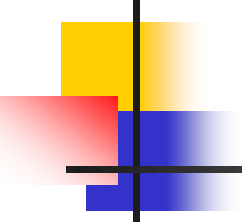
The database system is characterized by **unified management** and **data sharing**.

Data structuring is the **fundamental difference between database and file system**. When describing data, we should not only describe **the data itself**, but also describe **the relationship between the data**.



DBMS Development

- 1964, IDS (Integrated Data Store) network data model, General Electric, Bachman Award (1973)
- 1960s , IMS (Information Management System) , hierarchical data model, IBM
- **1970, E.F Codd, relational data model, Bachman Award(1981)**
- **1977, IBM System R/Berkeley Ingres, relational data model**
- 1980s~now, Relational DBMS becomes the mainstream, DB2, Oracle, Sybase, Informix, MS SQL Server, Access.



Data management will continue to develop, with **application as the driving force, new software and hardware** technology as **the foundation** and database as the main form.

1.2 Database System

Database system: database, DBMS(and other development tools), application systems, DBA and users.

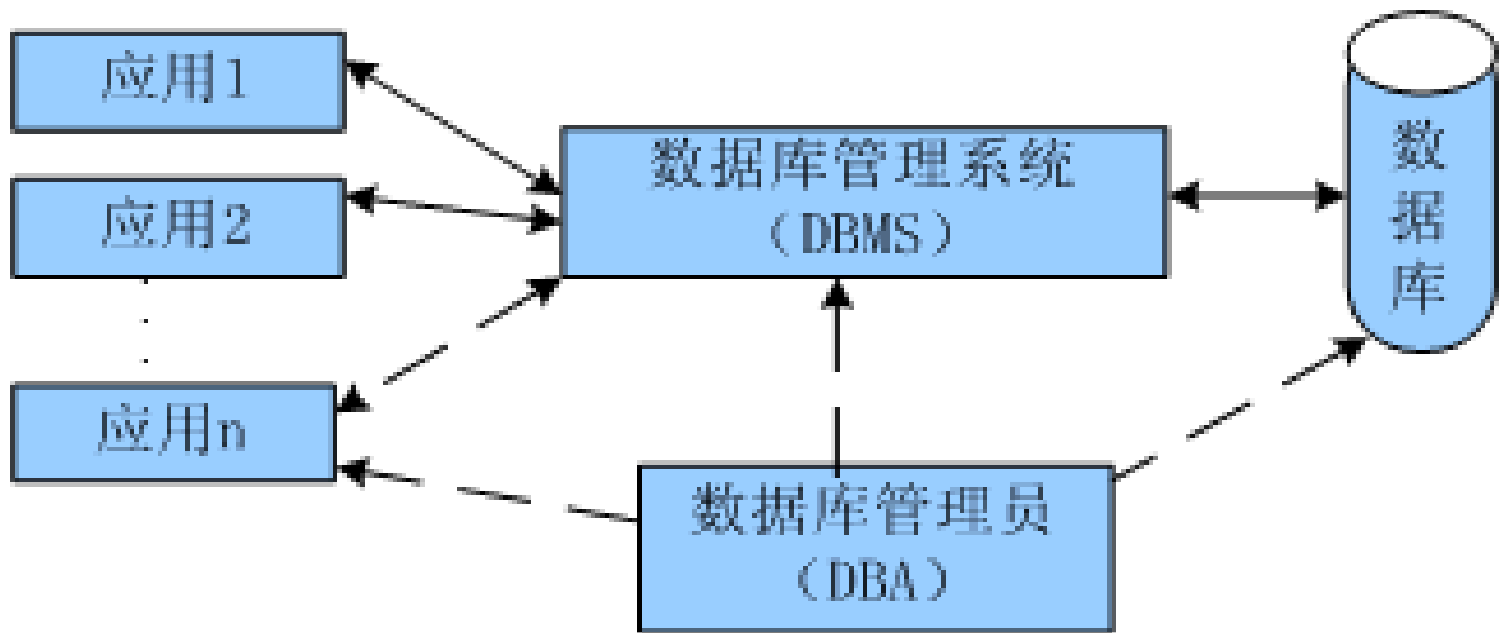


图1-2 数据库系统



Main Functions of DBMS

- Advanced user interface
- Query processing and optimization
- Data directory management
- Concurrency control
- Recovery
- Integrity constraint
- Access control



(1) Hardware& Database System

- 1) The data scale is very large.
- 2) DBMS(software itself) is also very large.

Database system puts forward higher requirements for hardware resources

- 1) Enough memory: OS, kernel module of DBMS, data cache and application software.
- 2) Enough disk: store data and its backup.
- 3) Higher channel capacity : improve the data transmission rate.



(2) Software

Software related to database system:

- 1) **DBMS**
- 2) OS. Supporting DBMS.
- 3) High-level programming language and its compiling system(with database interface).
- 4) Application software based on database system.



(3) Staff

The personnel involved in the database system include :

1) DBA

2) System analyst and database designer.

3) Application programmer and final users.

Different people have different responsibilities and authorities



1) DataBase Administrator, DBA

Specific duties.

- ① Determine the information content and structure in the database
- ② Determine the storage structure and access strategy of the database
- ③ Define data security requirements and integrity constraints
- ④ Monitor the use and operation of the database.
- ⑤ Reconstruct the database.



2) System analyst & Database designer

System analyst. The **requirements analysis** and specification of the system should be combined with users and DBA to **determine the software and hardware configuration of the system** and participate in the outline design of the database system.


Database designer. Determination of data in the database and the design of various database modes. Database designers must participate in user demand survey and system analysis, and then carry out database design. **In many cases, the database designer is the database administrator.**



3) Application programmer and final users.

Application programmer. Designing, compiling, debugging and installing program modules of the application system.

Final users. Database system and related application softwares' final manipulators.



1.3 Data, data model(数据模型) and data schema(数据模式)

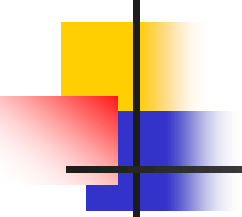
What is data? Symbols describing things .

What is information?

Model , for example, building model, automobile model and so on.

Model is the simulation and abstraction of real world things.

Data Model. It is an abstraction of real-world data characteristics.



In the database, data model is used as a tool to abstract, represent and process data in the real world.

Existing database systems are all based on a certain data model.



The data model should meet three requirements:

1. Truly simulate the real world.(真实地模拟现实世界)
2. Easy to understand.(易为人所理解)
3. Convenient for computer implementation.
(便于计算机实现)

It is difficult to meet these three requirements at the same time. Hence, Usually, different data models are adopted according to different requirements.



Data model is the **core and foundation** of database system.

How to realize?

(1) Firstly, objects in the real world is abstracted into a certain information structure, which does not depend on a specific computer system, as well as some DBMS. It is merely a **conceptual model**(概念模型).

(2) Then, the conceptual model is transformed into a data model supported by a DBMS on the computer (逻辑模型&物理模型) .



Elements of data model(数据模型的组成要素)

A data model is a collection of strictly defined concepts. These concepts accurately describe the following characteristics of the system:

- (1) the static characteristics(静态特性)
- (2) dynamic characteristics(动态特性)
- (3) and integrity constraints(完整性约束)

Therefore, the data model usually consists of three parts: data structure, data operation and integrity constraint.



(1) data structure

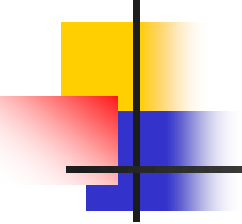
Data structure is a collection of object types studied. These objects are the components of the database, and they include two categories:

One is objects related to data types, contents and properties(数据类型、内容、性质有关的对象)

For example, data items and records in the network model, attributes and relationships in the relational data model, etc.

The other is describing relationships among those data.

For example, set structure in network model, relation in the relational data model, etc.



Data structure is the most important aspect to describe the properties of a data model, and it is a **description of the static characteristics of the system.**

Therefore, in a database system, **the data model is usually named according to the type of its second data structure.**

Hierarchical structure, network structure and relational structure are named as **hierarchical data model(层次数据模型)**, **network data model(网状数据模型)** and **relational data model(关系数据模型)** respectively.



(2) data operation (数据操作)

Data operation——**the set of operations** allowed to be performed on the instances (values) of various objects (types) in the database.

operations and **related operation rules**

——操作及操作规则

There are two main types of database operations: **retrieval and update** (including **insertion**, **deletion** and **modification**).

The data model defines the exact meaning of these operations, **operation symbols**, **operation rules** (such as priority) and **the language in which the operations are implemented**.

数据操作是对系统动态特性的描述。



(3) Data constraint condition(数据约束条件)

A collection of integrity rules——完整性规则的集合

Define the restriction and dependency rules(制约和依存规则) for data and their relations. Limit the database state and the change of state in accordance with the data model(限定符合数据模型的数据库状态以及状态的变化).

Ensure **the correctness**(正确), **validity**(有效) and **compatibility**(相容) of the data.

According to different use objects and application purposes, a multi-level data model is adopted, which can be generally divided into three levels:

(采用多级数据模型，一般分为三级)

(1) Conceptual data model(概念数据模型,与DBMS无关)

Mainly used to describe the conceptual structure of an enterprise.

The most representative model is **E-R data model**.

(2) Logical data model(逻辑数据模型,与DBMS有关)

Logical data model is the intermediate state from conceptual data model to computer realization

——逻辑模型既要面向用户，也要面向实现。



(3) Physical data model

(物理数据模型，与DBMS、OS硬件有关)

It reflects the storage structure of data, such as physical blocks, pointers, indexes and so on.

What is data schema (数据模式) ?

A description of the type and structure of an enterprise and their relationship with a certain data model.

数据模式有型与值之分： $\left\{ \begin{array}{l} \text{型} \text{——} \text{框架} \\ \text{值} \text{——} \text{实例} \end{array} \right.$

for example, student record

型:

姓名	学号	性别	出生年份	籍贯	系别	入学时间
----	----	----	------	----	----	------

(a)

一个值:

王彤	9098135	女	1980	江苏	计算机	1998
----	---------	---	------	----	-----	------

(b)

图1-3 数据模式及其实例



▲ Difference between data model and data schema

Data model is a means and tool to describe data in real world data.

(数据模型是描述现实世界数据的手段、工具)

Data schema——Using this tool to describe specific objects and their relationships.

(利用这个工具对具体对象及相互间关系进行描述得到的结果)

For example, the difference between C language and programs written in C language.



▲ Data schema are usually divided into three levels

(1) Conceptual schema(概念模式)

A description of a unit's data with a logical data model. Is a global schema of a unit and the logical structure of the whole database.

用逻辑数据模型对一个单位的数据的描述，是一个单位的全局模式，是整个数据库的逻辑结构。

(2) External schema (外模式)

Use logical data model to describe the data used by users. It is part of or derived from the conceptual model.

用逻辑数据模型，对用户所用到的那部分数据的描述。它是概念模式的一部分或由前者推导而来的。

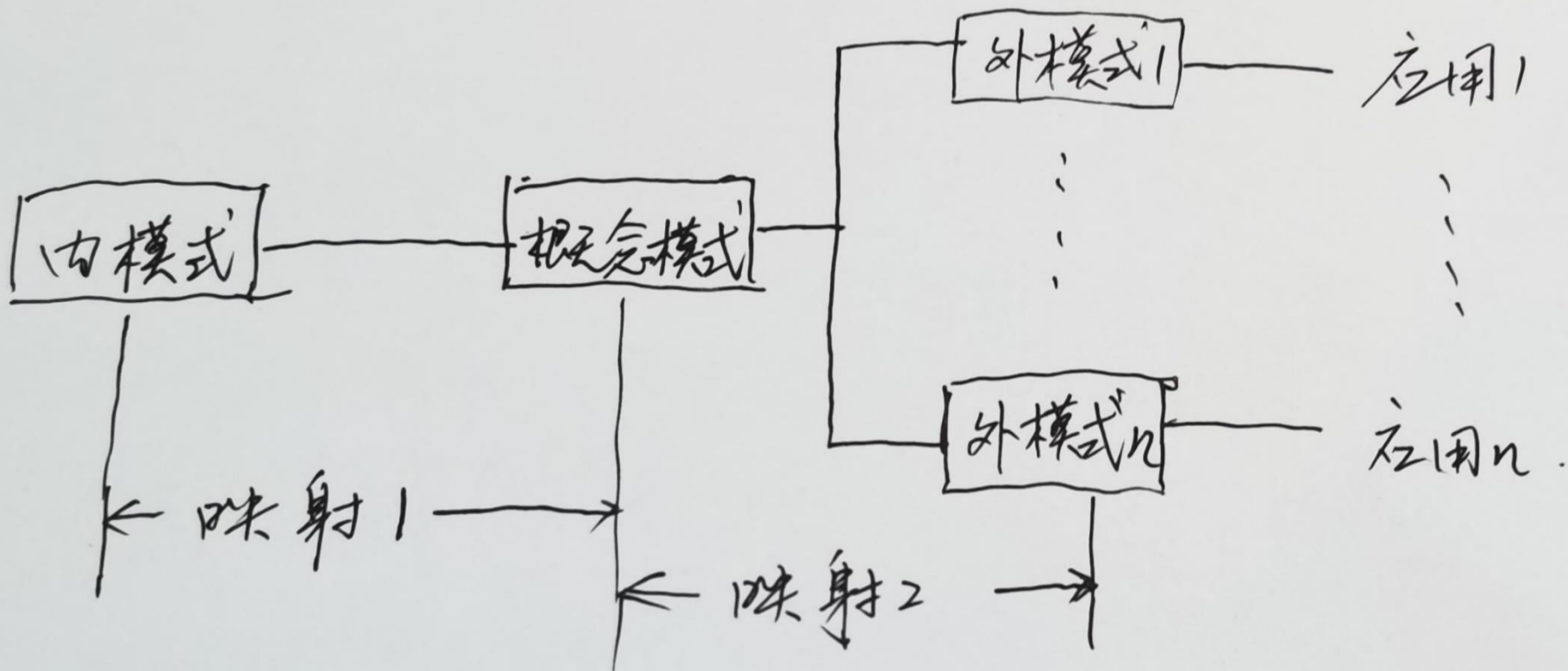


(3) Inner schema(内模式)

It is **the description of data with physical data model**. Represents the database storage structure, that is, the physical structure.

Conceptual schema, external schema and inner schema all exist in the data directory, which is the most basic content of the data directory.

概念模式、外模式和内模式都存于数据目录中,是数据目录的最基本内容。DBMS通过数据目录,管理和访问数据模式。





1.4 Database application

Choose the appropriate DBMS, design, establish, maintain and manage the database system, and serve users.

(1). Planning of database system(规划)

(2). Database design (设计)

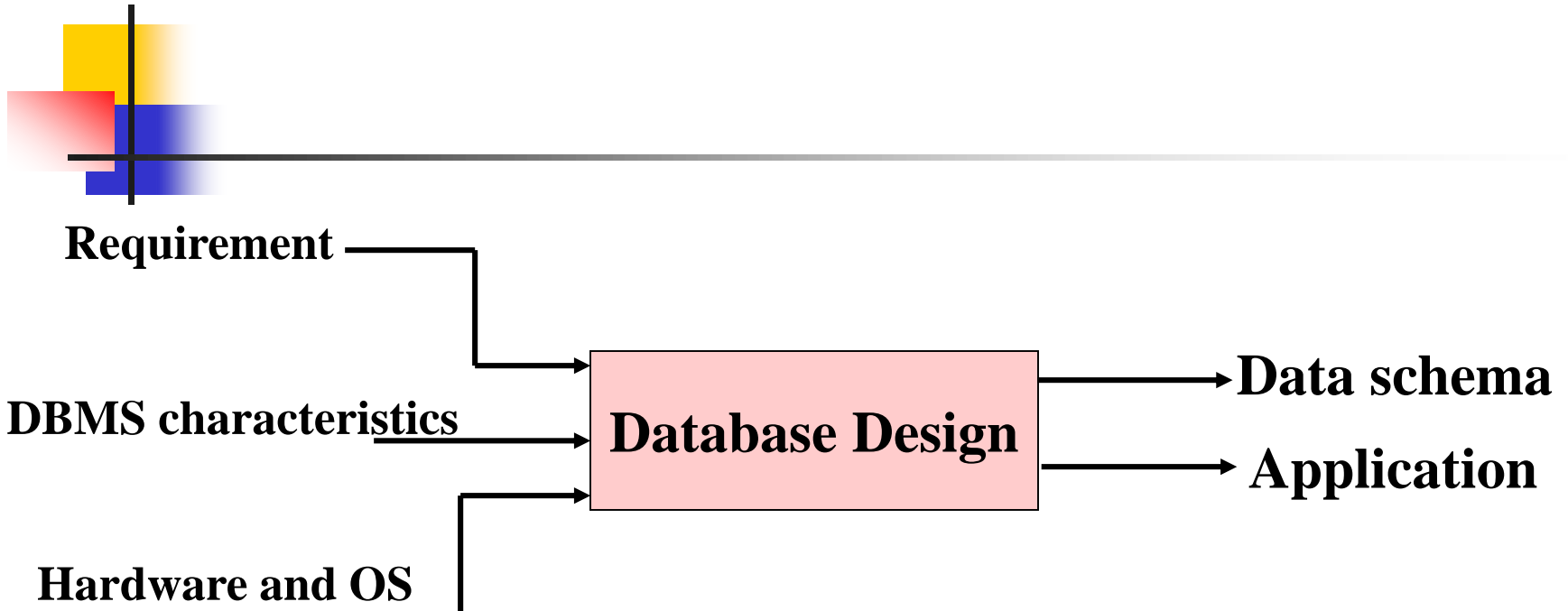


Fig 1-5 Basic tasks of database design



(3). Database establishment

(4). Operation, management and maintenance of database