

计算机科学导论

第6章 信息管理

杨老师邮箱: fdteachers@163.com

学习目标

1

了解信息系统概念

2

了解数据库查询语言

3

了解关系数据库设计

4

了解数据库管理

5

了解常用数据库管理系统

6

了解数据库前沿发展技术

7

掌握数据库系统基本概念

8

掌握关系数据库概念

6.1 模型与信息系统

6.1.1 信息(INFORMATION)

- 信息：通常指经过加工了的、有一定意义和价值、且具有特定形式的数据，这些数据能够反映出客观世界事物的内在联系及本质，从而影响信息获取者的行为或决策。
- 信息的特点：时间性、真实性、易于理解、多样性、传播性、不确定性。

6.1.2 信息模型

- 信息模型（Information Model）是一种用来定义信息常规表示方式的方法。通过使用信息模型，可以使用不同的应用程序对所管理的数据进行重用、变更及分享。
- 信息模型的基本构件包括对象、对象视图和对象关系。

6.1.3 信息系统

- 信息系统(information system)是一个由人员、活动、数据、网络和技术等要素组成的集合，主要目的是对组织的业务数据进行采集、存储、处理和交换，以支持和改善组织的日常业务运作，满足管理人员解决问题和制定决策以及对信息的各种需求。

6.2 数据库系统(DATABASE SYSTEM)

6.2.1 数据库

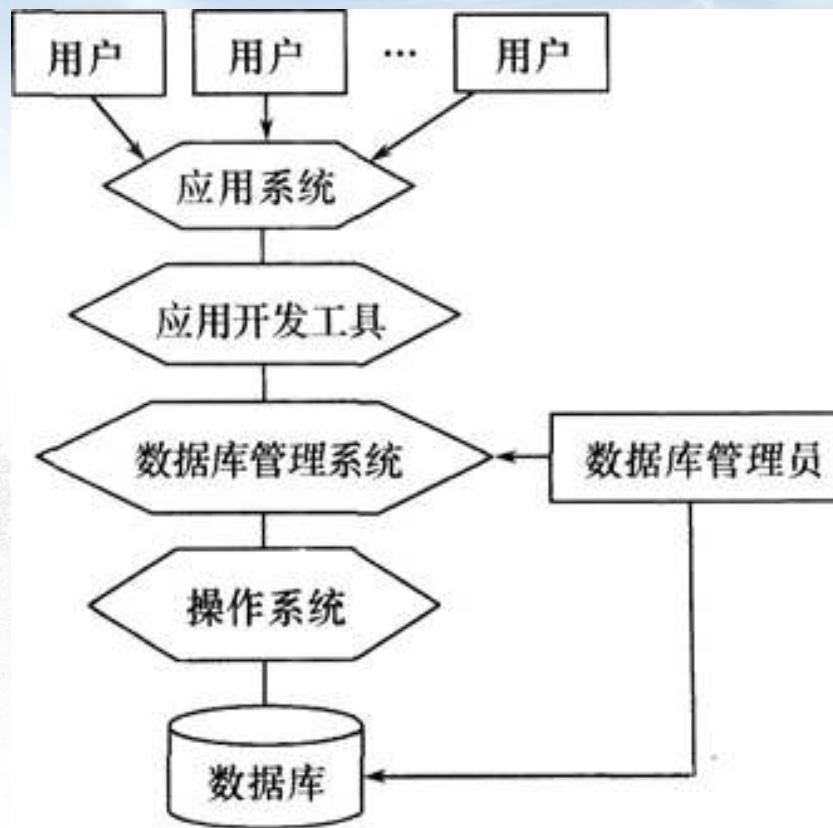
- **数据(Data)**是用来描述事物的符号记录，是数据库中存储的基本对象。数据的种类很多，包括文字、数字、声音、图形及图像等。
- **数据库(DataBase , DB)**是指以一定的组织方式存储的相互关联的数据的集合。这些数据能够长期存储、统一管理和控制，且能够被不同用户所共享，具有数据独立性及最小冗余度。

6.2.1 数据库

- 数据库管理系统(DataBase Management System , DBMS)是对数据库进行管理的软件系统，是数据库系统的核心。它位于计算机系统中操作系统与用户或应用程序之间，主要功能包括：数据定义、数据操纵、数据组织、存储和管理、数据库的建立和维护、数据通信接口。
- 数据库管理员(Database Administrator,简称DBA)：专门对数据库进行规划、设计、管理、协调和维护的工作人员。

6.2.1 数据库的基本概念

- 数据库系统(DataBase System : DBS)是DB、DBMS、DBA、用户和计算机系统CS(Computer System)的总和。



6.2.2 数据管理技术的发展

- 人工管理阶段：数据不长期保存，应用程序管理数据，数据不能共享，数据不具有独立性。
- 文件系统阶段：数据可以长期保存，文件系统管理数据，数据共享性差，冗余度大，数据独立性差。
- 数据库系统阶段：数据结构化，数据共享性高、冗余度低、易扩充，数据独立性高，有专门的数据库管理系统。

6.2.3 数据模型

◆ 1.概念模型：也称信息模型，是按照用户的观点对数据和信息建模，与计算机系统无关，是概念级的模型，通常用于数据库的设计，它是数据库设计人员和用户之间进行交互的一种形式化描述。

➤ (1) 基本概念

➤ ①实体(entity)

现实世界客观存在并可相互区别的事物称为实体。实体可以是具体的人、事、物，也可以是抽象的概念或联系。

6.2.3 数据模型

➤ ②属性(attribute)

实体所具有的某一特性称为属性。一个实体通常由若干个属性描述。

➤ ③码(key)

唯一标识一个实体的属性或属性组称为码。

➤ ④域(field)

属性的取值范围称为域。

➤ ⑤实体型(Entity Type)

对具有相同属性的若干实体，用实体名及其属性名的集合来抽象和描述同类实体，称为实体型。

6.2.3 数据模型

➤ ⑥实体集(entity set)

同型实体的集合称为实体集。

➤ ⑦联系 (relation)

不同实体集之间的联系，实体间联系通常有一对一、一对多、多对多三种。

一对一联系：记为 “1 : 1”

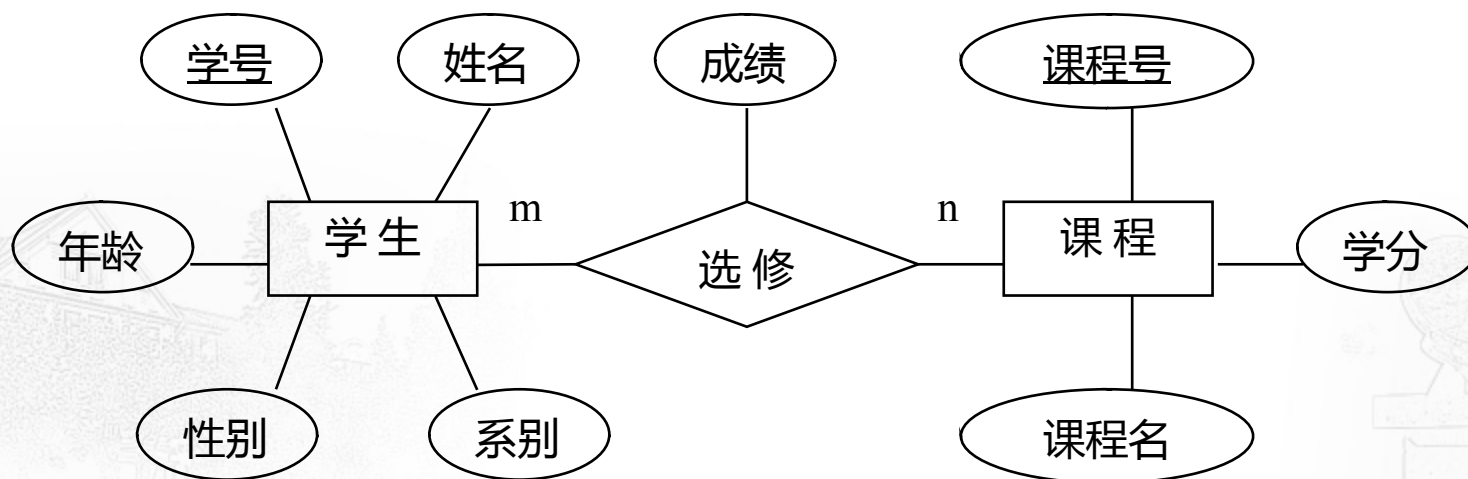
一对多联系：记为 “1 : M”

多对多联系：记为 “M : N”

6.2.3 数据模型

➤ (2) 概念模型的表示

用矩形表示实体型；用椭圆表示属性；用菱形表示联系，并标示出联系的类型



6.2.3 数据模型

◆ 2.数据模型

➤ (1) 层次模型

有且只有一个节点没有双亲，该节点为根节点；根以外的其他节点有且只有一个双亲节点。

➤ (2) 网状模型

在层次模型的基础上，允许节点无父节点，或者有多个父节点。

➤ (3) 关系模型

以关系代数为基础，实体间通过公共属性实现联系，与数据的物理结构无关。

6.2.3 数据模型

◆ 3.数据模型的三要素

➤ (1) 数据结构

数据结构是指所研究的数据库组成成分的类型集合，用于描述系统的静态特征。

➤ (2) 数据操作

指对数据库中数据允许执行的操作的集合，用于描述系统的动态特征。数据模型中必须定义操作(如查询、更新)的确切含义、操作符号、操作规则以及实现操作的规则。

➤ (3) 数据的约束条件

是数据完整性规则的集合，指对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。

6.2.3 数据模型

◆ 4. 数据库系统的结构

➤ (1) 模式

模式也称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，是面向全体用户的基本数据视图。

➤ (2) 外模式

外模式也称子模式或用户模式，是数据库用户可见和使用的局部数据的逻辑结构和特征的描述，是数据库用户的数据视图。

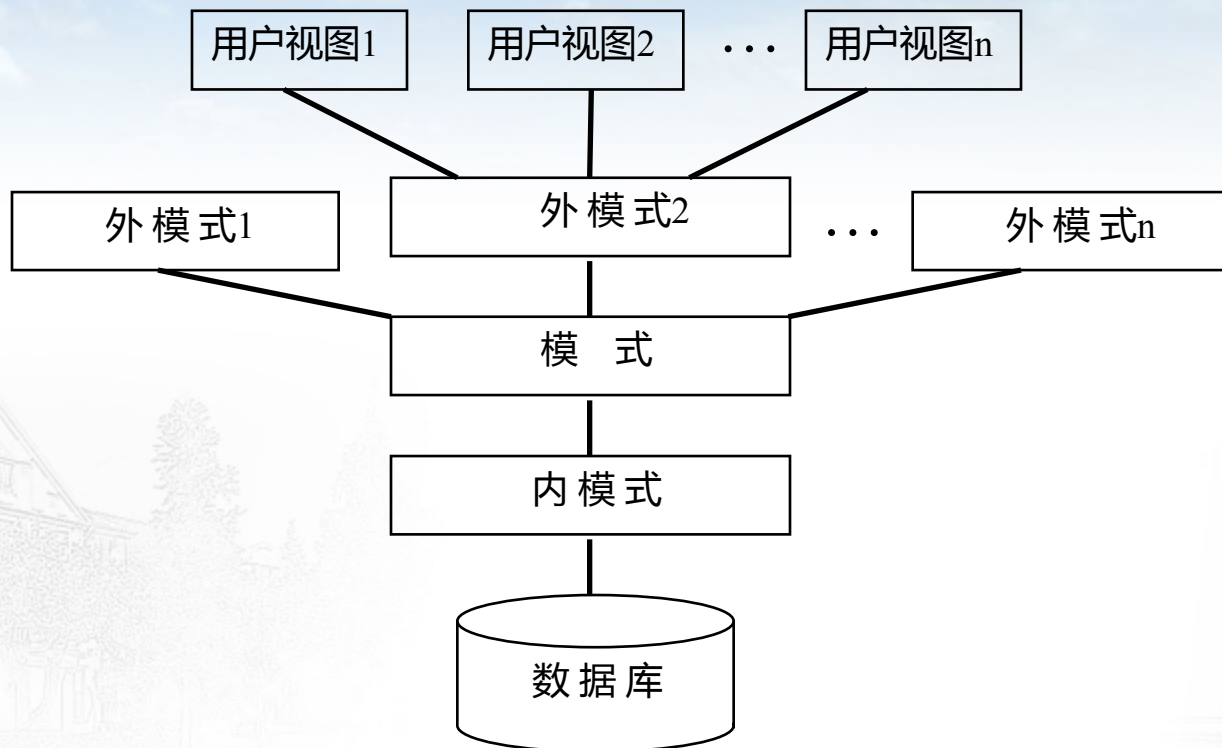
。

➤ (3) 内模式

内模式是数据物理结构和存储结构的描述，是数据在数据库内部的表示。

6.2.3 数据模型

➤ 数据库的三级模式结构

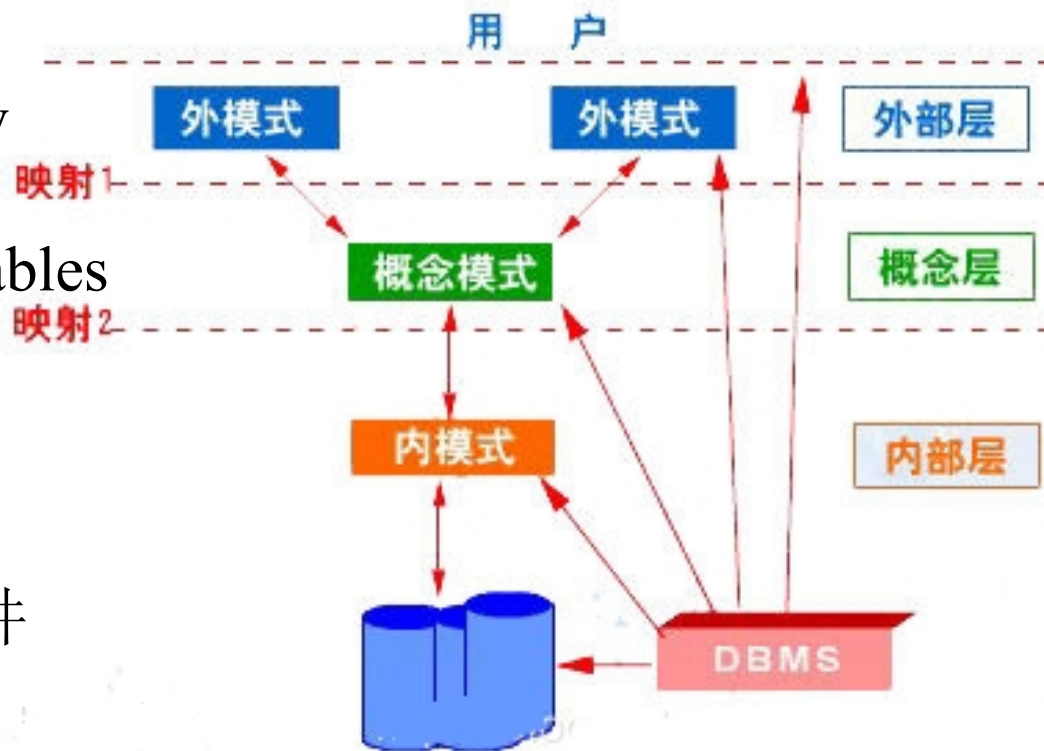




视图view

基本表tables

存储文件



Programmer



DBA

6.2.3 数据模型

◆ 5.数据独立性

➤ (1) 外模式/模式映像

当模式被改变时，可以由DBA改变各个外模式/模式映像，以保证外模式不发生变化，从而使根据外模式编写的应用程序也无需变化，这样就保证了数据与程序的逻辑独立性。

➤ (2) 模式/内模式映像

当数据的物理存储(即内模式)变化了，也可以由DBA来修改模式/内模式的映像，使得模式能够保持不变，外模式也不会变化，当然，应用程序也无需变动，从而保证了数据库中数据与程序的物理独立性。

6.3 关系数据库

6.3.1 关系数据库

◆ 1 . 关系(Relation)

- 关系可以看作是一个行与列交叉的二维表，每一个交叉点都必须都是单值的，每一列的所有数据都是同一类型的，每一列都有唯一的列名，行和列在表中的顺序无关紧要，表中任意两行不能相同。

◆ 2 . 属性(Attribute)

- 关系中的每一列称为属性，属性的总数称为关系的度。

◆ 3 . 元组(Tuple)

- 关系中的行称为元组，元组包含了一组属性，元组的总数称为关系的基数。

6.3.1 关系数据库

- ◆ 4 . 候选码(Candidate key)
 - 候选码是关系中能够唯一标识一个元组的某个属性或属性组，一个关系可以有多个候选码。
- ◆ 5 . 主码(key)
 - 一个关系中选定一个候选码作为该关系的主码。
- ◆ 6 . 主属性(Prime attribute)
 - 主码的各个属性称为主属性。
- ◆ 7 . 外码(foreign key)
 - 在关系数据库中，为了实现表与表之间的联系，将一个表的主码作为公共属性放到另一个关系中，在另一个关系中起连接作用的属性称为外码。

6.3.2 关系模型与关系模式

◆ 1.关系模型

➤ (1) 关系数据结构

关系模型中，关系数据结构就是表。

➤ (2) 关系操作集合

选择、投影、连接、除、并、交、差、查询、插入、删除、修改等。

➤ (3) 关系的数据约束

三种完整性约束条件：实体完整性、参照完整性、用户自定义的完整性。

6.3.2 关系模型与关系模式

◆ 2.关系模式

- 关系的描述称为关系模式，一个完整的关系模式通常需要描述出一个关系的属性的集合、属性所取值的域的集合、属性向域的映像的集合、属性间数据依赖关系的集合。通常，使用关系模式的简单描述：

$$R(A_1, A_2, \dots, A_n)$$

- 其中，R为关系名， A_1, A_2, \dots, A_n 为属性名。域名和属性向域的映像直接说明为属性的类型和长度。

6.3.2 关系模型与关系模式

◆ 3.关系运算

➤ (1) 传统的关系运算

A	B	C
a1	b1	c1
a2	b2	c2
a2	b2	c1

(a)

A	B	C
a1	b2	c2
a1	b3	c2
a2	b2	c1

(b)

A	B	C
a1	b1	c1
a2	b2	c2
a2	b2	c1
a1	b3	c2
a1	b2	c2

(c)

A	B	C
a2	b2	c1

(d)

A	B	C
a1	b1	c1
a2	b2	c2

(e)

(a)关系R ; (b)关系S ;

(c)关系的并运算(d)关系的交运算(e)关系的差运算

6.3.2 关系模型与关系模式

➤ (2) 专门的关系运算

学生表

学号	姓名	性别	年龄	所在系
01001	王立	男	19	电信系
02013	李强	男	20	计算机
03050	陈然	女	19	管理系
02010	唐杰	男	20	电信系

课程表

课程号	课程名	学分
001	数据库	3
002	英语	4
101	C语言	3
102	物理	4

成绩表

学号	课程号	成绩
01001	001	90
01001	002	85
03050	002	90
03050	102	95

6.3.2 关系模型与关系模式

➤ ①选择运算(Selection)

从给出的关系中选出满足给定条件的那些元组，
是对关系的行进行的运算。

[例] 查询电信系所有学生的情况。

学号	姓名	性别	年龄	所在系
01001	王立	男	19	电信系
02010	唐杰	男	20	电信系

Select * from student where dept='电信系'

6.3.2 关系模型与关系模式

➤ ②投影运算 (Projection)

从给出的关系中选出关系的一个或多个属性列，是对关系的列进行的运算。

[例]查询学生的姓名和所在的系。

姓 名	所在系
王立	电信系
李强	计算机
陈然	管理系
唐杰	电信系

6.3.2 关系模型与关系模式

➤ ③连接运算 (Join)

将两个关系模式通过共有的属性名拼接成一个更多属性列的关系。

[例]学生关系与成绩关系的自然连接。

学 号	姓 名	性 别	年 龄	所在系	课程号	成绩
01001	王立	男	19	电信系	001	90
01001	王立	男	19	电信系	002	85
03050	陈然	女	19	管理系	002	90
03050	陈然	女	19	管理系	102	95

6.4 数据库查询语言SQL

6.4.1 SQL的特点

- 功能统一
- 非过程性语言
- 面向集合的操作方式
- SQL提供两种灵活的使用方式
- SQL简单、易学

6.4.2 SQL的功能

◆ 1 . 数据定义(DDL)

- 供用户简便地建立数据库和表的结构，包括定义、删除、修改表，建立和删除索引等。create, drop, alter;

◆ 2 . 数据操纵(DML)

- 供用户实现表中数据的插入、删除和修改等操作。insert,update,delete

◆ 3 . 数据查询(DQL)

- 供用户实现对表中数据内容的各种查询。Select;

◆ 4 . 数据控制(DCL)

- 提供数据库系统并发控制、数据库恢复、数据库安全性和完整性等功能。

grant, revoke

6.4.3 SQL基本概念

- 基本表：一个关系即对应一个基本表。
- 视图：视图是由一个或几个基本表导出的表。
- 存储过程(procedure)：存储过程是用SQL语言编写成的SQL语句序列，它存储于SQL服务器上供用户调用执行。
- 触发器(trigger)：触发器是一种特殊的存储过程，用来维护不同表中的相关数据的一致性。通过事件触发。

6.5 关系数据库设计

6.5.1 关系数据库的设计原则

按照DBMS提供的功能和描述工具，设计出规模适当、能够正确反映数据关系、数据冗余少、存取效率高、能满足多种查询要求的数据模型。

6.5.2 关系数据库的设计步骤

- 需求分析
- 概念结构设计
- 逻辑结构设计
- 数据库物理设计
- 数据库实施
- 运行与维护

6.6 数据库管理

6.6.1 DBA和数据字典

- 职责：向终端用户提供数据和信息；制定数据库管理的政策、标准和规程；确保数据库的安全性；对数据库进行备份和恢复；培训终端用户。
- 管理工具：DBA进行数据库管理的重要工具是数据字典。数据字典通常是DBMS的一部分，它用来存储数据的特性和关系的定义。

6.6.1 DBA和数据字典

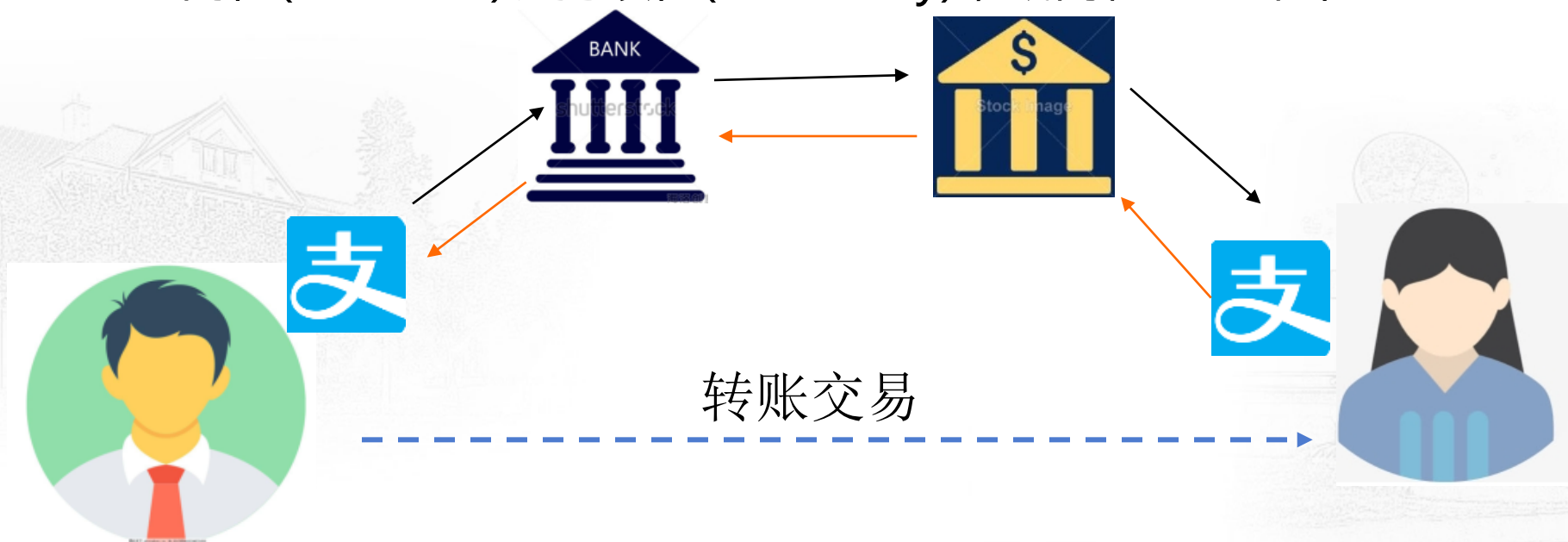
- ◆ 数据字典通常包含以下信息。
 - 每个数据库的创建者、创建日期。
 - 数据库中表的定义。
 - 表中所有列的列名、列的数据类型、显示格式。
 - 索引的定义。
 - 数据库中所有终端用户和管理员的访问权限。
 - 所有表之间的联系，完整性约束定义。

6.6.2 数据库的安全性控制

- 访问权限控制
- 定义视图
- 监控DBMS操作

6.6.3 事务处理技术

- 事务(transaction)的概念：事务是用户定义的一个数据库操作序列，这些操作要么全做，要么全都不做，是一个不可分割的工作单位。
- 事务的特性：原子性(Atomicity)、一致性(Consistency)、隔离性(Isolation)、持续性(Durability)，或简称ACID特性。



6.6.3 事务处理技术

◆ 并发控制技术：封锁

- (1) 共享锁：也称读锁，如果事务T对数据对象X加了共享锁，则其他事务只能对X加共享锁，而不能加排它锁。
- (2) 排它锁：也称写锁，如果事务T对数据对象X加了共享锁，则只允许事务T独占X，其他任何事务不能对X再加任何类型的锁，直到T释放X上的锁。

◆ 数据库的恢复技术：数据库的恢复指使数据库从一种不一致的状态恢复到原来一致性的状态。数据库恢复的主要依据是备份的数据副本和日志文件。

6.6.4 数据库的完整性控制

- 指保证数据库中数据的正确性、有效性、相容性，防止错误的数据进入数据库。

6.7常用数据库管理系统

- ORACLE
- SYBASE
- DB2
- MS SQL SERVER
- MYSQL
- ACCESS
- SQLite

6.8 数据库新发展

- 分布式数据库
- 空间数据库
- 非关系型的数据库NoSQL
- 多媒体数据库
- 数据仓库
- 信息存储与检索
- 超文本和超媒体
- 多媒体信息与多媒体系统
- 数字地球与智慧城市

本章小结

- ▶本章介绍了信息系统中的核心基础——数据库技术，包括数据、数据库、数据库管理系统、数据库管理员的基本概念，关系数据库设计的过程，数据库管理的基本内容，常用的数据库管理系统以及数据库发展的方向。
- ▶通过本章的学习，读者应该了解数据发展的阶段——人工管理、文件系统、数据库系统；熟悉数据库系统的构成，清楚数据、数据库、数据库管理系统、数据库管理员之间的关系和作用；掌握关系数据库系统中经常使用的一些术语，简单运用SQL进行数据库的数据查询、数据操纵、数据定义、数据控制等操作；了解关系数据库设计的原则和步骤、数据库的新发展。

谢谢,再见!

