

【第 15 周】课程目标及第一章绪论要点

2021 年 12 月 19 日【周日】：201、202 班

2021 年 12 月 20 日【周一】：203、204 班

1、第一章 绪论（要点）

- 1) 数据库的 4 个基本概念
 - 2) 关于数据模型
 - 3) 数据库系统的结构（重点）
 - 4) 数据库系统的组成
- 2、与课程设计的相关性（完善课程设计内容）
- 1) 使用三级模式结构、二级映像，实现逻辑独立性与物理独立性
 - 2) 从数据系统的视角理解掌握各种人员的数据视图

第一章 绪论（要点）

1、数据库的 4 个基本概念

- 1) 数据
 - ✧ 描述事物的符号记录称为数据
 - ✧ 数据的含义称为语义，数据与其语义是不可分的
- 2) 数据库
 - ✧ 数据库是长期存储在计算机内、有组织的、可共享的大量数据的集合。
 - ✧ 数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。
- 3) 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）有以下主要功能
 - ✧ 数据定义功能
 - ✧ 数据组织、存储管理
 - ✧ 数据操纵功能
 - ✧ 数据库的事务管理和运行管理
 - ✧ 数据库建立和维护功能
 - ✧ 等等
- 4) 数据库系统
 - ✧ 数据库系统是由数据库、数据库管理系统（及其应用开发工具）、应用程序和数据库管理员（DBA）组成的存储、管理、处理和维护数据的系统

2、关于数据模型

- 1) 数据库技术的发展是沿着数据模型的主线推进的（P14）（*）
 - ✧ 数据模型（data model）也是一种模型，它是对现实世界数据特征的抽象。数据模型是用来描述数据、组织数据和对数据进行操作的（P14-15）
 - ✧ 通俗地讲，数据模型就是现实世界的模拟（P15）
 - ✧ 数据模型是数据库系统的核心和基础（P15）。各种机器上实现的数据库管理系统软件都是基于某种数据模型或者说支持某种数据模型的（P15）。
 - ✧ 概念模型（conceptual model），也称信息模型，它是按用户观点对数据和信息建模，主要用于数据库设计（P15）（*）
 - ✧ 逻辑模型，主要包括层次模型、网状模型、关系模型（relational model）、面向对象数据模型和对象关系数据模型等。它是按计算机系统的观点对数据建模，主要用于数据库管理系统的实现（P15）

- ✧ 物理模型，是对数据最底层的抽象，它描述数据在系统内部的表示方式和存取方法，或在磁盘或磁带上的存储方式和存取方法，是面向计算机系统的。物理模型的具体实现是数据库管理系统的任务，数据库设计人员要了解和选择物理模型，最终用户则不必考虑物理级的细节（P15）。(*)
- 2) 为了把现实世界的具体事物抽象、组织为某一数据库管理系统支持的数据模型，人们通常
 - ✧ 首先将现实世界抽象为信息世界
 - ✓ 将现实世界的客观对象抽象为某一种信息结构
 - ✓ 这种信息结构不依赖于具体的计算机系统，不是某一个数据库管理系统支持的数据模型，而是概念级数据模型
 - ✧ 然后将信息世界转换为机器世界
 - ✓ 把概念模型转换为计算机上某一数据库管系统支持的数据模型
 - ✧ 完成者 (*)
 - ✓ 从现实世界到概念模型的转换（E-R 图设计），由数据库设计人员完成
 - ✓ 概念模型到逻辑模型的转换，由数据库设计人员完成，也可以用数据库设计工具协助设计人员完成
 - ✓ 逻辑模型到物理模型转换，主要由数据库管理系统完成（P16）
- 3) 概念模型（P16-17）
 - ✧ 信息世界中的基本概念 (*)（需要背一下）
 - ✓ 实体（entity）：客观存在并可相互区分的事物称为实体
 - ✓ 属性（attribute）：实体所具有的某一特性称为属性
 - ✓ 码（key）：唯一标识实体的属性集称为码
 - ✓ 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
 - 问题：关系模式是实体型吗？
 - ✓ 实体集（entity set）：同类实体的集合称为实体集
 - 问题：关系是实体集吗？
 - ✓ 联系（relationship）：
 - 实体内部联系：实体各属性间的联系
 - 实体之间联系：不同实体集之间的联系
 - 实体之间的联系有一对一、一对多、多对多等多种类型
 - ✧ 概念模型中的一种表示方法：实体-联系方法
 - ✓ 概念模型的表示方法很多，其中最为常用的是实体-联系方法
 - ✓ E-R 方法也称为 E-R 模型
- 4) 数据模型的组成要素（P17~18）
 - ✧ 数据模型通常由数据结构、数据操作和数据的完整性约束条件三部分组成
 - ✧ 数据结构描述数据库的组成对象以及对象之间的联系
 - ✧ 数据操作是指对数据库各种对象（型）的实例（值）允许执行的操作集合，包括操作及有关的操作规则（说明：对关系数据库来说，型指关系模式，值指关系）
 - ✧ 数据的完整性约束条件（指一组完整性规则）
 - 在关系模型中，任何关系必须满足实体完整性和参照完整性两个条件
- 5) 数据结构、数据操作和数据完整性约束条件这三个方面的内容完整地描述了一个数据模型，其中数据结构是刻画模型性质的最基本方面（P19）。

6) 关系模型 (三要素)

✧ 关系模型数据结构 (P25)

- ✓ 每个关系的数据结构是一张规范化的二维表
- ✓ 术语: 关系、元组、属性、码、域、分量、关系模式
- ✓ 关系模型要求关系必须是规范化的 (最基本的规范要求: 关系的每一个分量必须是一个不可分割的数据项)

✧ 关系模型的数据操作

- ✓ 主要包括查询、插入、删除和更新数据

✧ 关系模型的完整性约束条件

- ✓ 操作必须满足完整性约束条件。关系的完整性约束条件包括三大类

3、数据库系统的结构 (重要)

1) 在数据模型中有“型” (type) 与“值” (value) 的概念 (P27)

✧ 型是指对某一类数据的结构和属性的说明

- ✓ 关系模式
- ✓ 关系数据库模式

✧ 值是型的一个赋值

- ✓ 关系
- ✓ 关系数据库

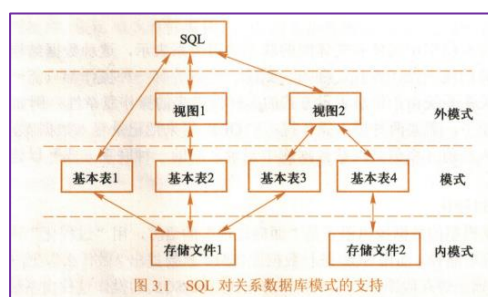
✧ 模式 (schema)

- ✓ 是数据库中全体数据的逻辑结构和特征的描述, 它仅仅涉及型的描述, 不涉及具体的值
- ✓ 是相对稳定的

✧ 实例 (instance) (P28)

- ✓ 模式的一个具体值称为模式的一个实例
- ✓ 是相对变动的

2) 数据库系统的三级模式结构



✧ 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描述, 是所有用户的公共数据视图
- ✓ 是数据库系统模式结构的中间层, 既不涉及数据的物理存储细节和硬件环境, 又与具体的应用程序、所使用的应用开发工具及高级程序设计语言无关
- ✓ 一个数据库只有一个模式。
- ✓ 数据库模式以一种数据模型为基础, 统一综合考虑了所有用户的需求, 并将这些需求有机结合成一个逻辑整体
- ✓ 定义模式时不仅要定义数据的逻辑结构, 例如数据记录由哪些数据项构成, 数据项的名字、类型、取值范围等; 而且要定义数据之间的联系, 定义与

数据有关的**安全性**、完整性要求。(重点)

- ✓ 创建数据库 (**创建模式**)

✧ 外模式

- ✓ 也称子模式或用户模式，它是数据库用户（包括应用程序员和最终用户）能够看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，是与某一应用有关的数据的逻辑表示
- ✓ 外模式通常是模式的子集。一个数据库可以由多个外模式
- ✓ 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，数据库中的**其余数据是不可见的**
(注意：关系数据库不是这样)

✧ 内模式

- ✓ 也称为存储模式，一个数据只有内模式。它是数据**物理结构**和**存储方式**的描述，是数据在数据库内部的组织方式
- ✓ **例如**，记录的存储方式是堆存储还是按照某个（些）属性值的升降序存储，或按照属性的聚簇存储；索引依照什么方式组织，是 B+树索引还是 hash 索引；数据是否压缩存储，是否加密；数据的存储记录结构有何规定，如定长结构或变长结构，一个记录不能跨物理页存储等等
(注意：关系数据库的存取方式：主要与索引有关)

3) 数据库的二级映像功能与数据独立性

- ✧ 数据库系统的三级模式是数据的**三个抽象级别**，它把数据的具体组织留给了数据库管理系统，使用户能**逻辑地、抽象地处理数据**，而不关心数据在计算机中地具体表示方式与存储方式。

- ✧ 为了能够在系统内部实现这三个抽象层次的联系和转换，数据库管理系统在这三级模式之间提供了**两层映像**：**外模式/模式映像**和**模式/内模式映像**

- ✧ 两层映像保证了数据库系统中的数据能够具有较高的**逻辑独立性**和**物理独立性**

✧ **外模式/模式映像 (P30)**

- ✓ 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
- ✓ 一个模式可以有任意多个外模式
- ✓ 对于每一个外模式，数据库系统都有一个外模式/模式映像，它定义了该外模式与模式之间的对应关系。这些映像定义通常包含在**各自外模式的描述中**
- ✓ **关系数据模型中，外模式通过视图描述，映像定义在视图中**
- ✓ **数据的逻辑独立性**：当模式改变时（例如增加新的关系、新的属性、改变属性的数据类型等），由数据库管理员对各个外模式/模式的映像做相应改变，可以使外模式保持不变。**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。
- ✓ 注意：关系模型中，应用程序是依据数据的**外模式与模式**编写的

✧ 模式/内模式映像

- ✓ 数据库只有一个模式，也只有一个内模式，所以模式/内模式的映像是唯一的，它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**。
- ✓ 该映像定义通常包含在模式描述中。

- ✓ **数据的物理独立性**：当数据库的存储结构改变时（例如选用了另一种存储结构），由数据库管理员对模式/内模式映像做相应改变，可以使模式保持不变，从而应用程序也不必改变。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
- ✧ 由于数据的存取由 DBMS 管理，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
- 4) 课程设计中的**数据库系统的三级模式结构及数据库的二级映像功能与数据独立性**
 - ✧ 基于**项目视野**看三级模式结构（**关系数据库模式**）
 - ✓ 模式：（基于课程设计请假系统理解模式）
 - ✚ 是数据库中全体数据的**逻辑**结构和特征描述，是所有用户的公共数据视图
 - ✚ 关系数据库模式（P43）：包括若干**域的定义**，以及这些域上定义的**若干关系模式**，即基于需求分析的逻辑设计结果（一组关系模式）
 - ✚ 关系数据库模式也称为关系数据库的**型**，是对关系数据库的**描述**
 - ✚ 关系数据库（P43）：关系数据库的值是这些关系模式在某一时刻对应的关系的集合，通常称为关系数据库
 - ✚ 在一个给定的应用领域（**请假系统**）中，所有关系的集合构成了一个关系数据库
 - ✚ **定义模式**时不仅要定义数据的逻辑结构，例如数据记录由哪些数据项构成，数据项的名字、类型、取值范围等；而且要定义数据之间的联系，定义与数据有关的**安全性**、完整性要求。
 - ✓ 外模式（子模式或用户模式）
 - ✚ 也称子模式或用户模式，它是数据库用户（包括应用程序员和最终用户）能够看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，是与某一应用有关的数据的逻辑表示（P29）
 - ✚ 设计用户子模式（外模式）（P234）
 - 将概念模型转换为全局逻辑模型后，还应该根据局部应用需求，结合具体数据库管理系统的特点设计用户的外模式
 - 目前关系数据库管理系统一般都提供视图概念，可以利用这一功能设计更符合局部用户需求的用户模式
 - ✚ 例：使用视图查找冗余数据，视图中定义了外模式到模式的映像
 - ✓ 内模式（数据库物理结构设计）
 - ✚ 数据库**物理结构**：数据库在物理设备上的存储结构与存取方法称为数据库的物理结构，它依赖于选定的数据库管理系统
 - ✚ 数据库的物理设计：为一个给定的逻辑数据模型选取一个最适合应用要求的物理结构的过程
 - ✚ 数据库的物理设计通常分两步：
 - (1) 确定数据库物理结构，在关系数据库中主要是存取方法和存储结构
 - (2) 对物理结构进行评价，评价的重点是时间和空间效率
 - ✚ 关系模式存取方法选择（常用方法为索引方法和聚簇方法）
 - (1) B+树索引存取方法
 - (2) Hash 索引存取方法

(3) 聚簇存取方法

确定数据库的存取结构

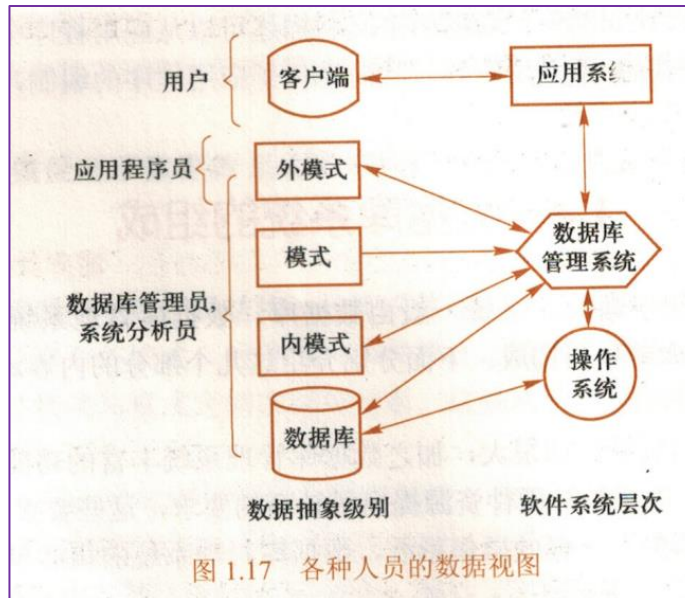
(1) 确定数据的存放位置

(2) 确定系统配置

评价物理结构:

数据库物理设计过程中需要对时间效率、空间效率、维护代价和各种要求进行权衡, 其结果可以产生多种方案

4、数据库系统的组成



5、

课后作业

1、关于数据库系统三级模式结构

1) 判断题

(1) 创建数据库即创建模式 (√)

请验证此题的正确性:

```
drop database if exists schema_test;#删除数据库
create database if not exists schema_test;#建立数据库
改为以下语句是否能正确执行? 是否能实现原有功能?
drop schema if exists schema_test;#删除模式
create schema if not exists schema_test;#建立模式
```

(2) 关系数据库的型是对关系数据库的描述 (√)

(3) 关系数据库的型也称为关系数据库模式 (√)

(4) 关系数据库的型也称为关系数据库 (x)

(5) 关系数据库的值称为关系数据库 (√)

(6) 关系数据库的值是关系数据库模式在某一时刻对应的关系的集合 (√)

2) 掌握:

- (1) **关系数据库模式**包括若干域的定义, 以及在这些域上定义的关系模式 (P43)
- (2) **关系数据库**关系数据库模式在某一时刻对应的关系的集合 (P43)
- (3) 关系数据库模式是关系数据库的型
- (4) 关系数据库是关系数据库的值

2、选择题

- 1) 数据库系统的核心和基础是 (C)
A.物理模型 B.概念模型 **C.数据模型** D.逻辑模型
- 2) 实现将现实世界抽象为信息世界的是 (B)
A.物理模型 **B.概念模型** C.关系模型 D.逻辑模型
- 3) 能够保证数据库系统中的数据具有较高逻辑独立性的是 (A)
A.外模式/模式映像 B.模式 C.模式/内模式映像 D.外模式
- 4) DBMS 是一类系统软件, 它是建立在下列哪些系统之上的? (C)
A.应用系统 B.编译系统 **C.操作系统** D.硬件系统
- 5) 下列说法正确的是 (D)
A.数据库的概念模型与具体的 DBMS 有关
B.三级模式中描述全体数据的逻辑结构和特征的是外模式
C.数据库管理员负责设计和编写应用系统的程序模块
D.从逻辑模型到物理模型的转换一般是由 DBMS 完成的
- 6) 长期存储在计算机内, 有组织、可共享的大量数据的集合是 (B)
A.数据 (data) **B.数据库 (Database)**
C.数据库管理系统 (DBMS) D.数据库系统 (DBS)

3、判断题

- 1) 通常情况下, 外模式是模式的子集 (√)
- 2) 数据库管理系统是指在计算机系统引入数据库后的系统, 一般由 DB、DBS、应用系统和 DBA 组成 (×)
- 3) 在数据模型的组成要素中, 数据结构是刻画一个数据模型性质最重要的方面, 人们通常按数据结构的类型来命名数据模型 (√)
- 4) 数据库系统的三级模式是对数据进行抽象的 3 个级别, 把数据的具体组织留给 DBMS 管理 (√)

4、填空题

- 1) 数据库系统的逻辑模型按照计算机的观点对数据建模, 主要包括层次模型、网状模型、关系模型、面向对象模型、对象关系模型和半结构化数据模型等。
- 2) 最经常使用的概念模型是 E-R 模型
- 3) 数据独立性是数据库领域的重要概念, 包括数据的物理独立性和数据的逻辑独立性
- 4) 数据库系统的三级模式结构是指数据库系统是由外模式、模式和内模式三级构成
- 5) 两个实体型之间的联系可以分为三种: 一对一联系、一对多联系和多对多联系。
- 6) 数据库管理系统提供的数据控制方面的功能包括数据的安全性保护、数据的完整性检查、并发控制和数据库恢复。
- 7) 数据库的三级模式结构中, 描述局部数据的逻辑结构和特征的是外模式。
- 8) 数据模型的组成要素中描述系统的静态特性和动态特性的分别是数据结构和数据操作。

5、完善课程设计内容 (在基于 MySQL 实现中, 增加以下标题内容)

- 1) 建立模式 (并说明在模式描述中定义的模式/内模式映像内容)

2) 建立外模式（并说明在各自的外模式描述中定义的外模式/模式映像内容）

举例：

1、定义外模式举例：（）

- ✧ 关系数据模型中，外模式通过视图描述，映像定义在视图中
- ✧ 每一个外模式，数据库系统都有一个外模式/模式映像，它定义了该外模式与模式之间的对应关系
- ✧ 数据的逻辑独立性：当模式改变时（例如增加新的关系、新的属性、改变属性的数据类型等），由数据库管理员对各个外模式/模式的映像做相应改变，可以使外模式保持不变。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。

/*定义视图

建立包含导出学生数的系视图

*/

drop view if exists 系_导出学生数；

create view 系_包含导出学生数(系号,系名,学生数,地址)

as

select 系.系号,系名,count(*) 学生数,地址

from 系,学生

where 系.系号=学生.系号

group by 系.系号;

#执行创建视图语句的结果：只是把视图定义存入数据字典，并不执行其中的 select 语句

#查询视图：

select * from 系_包含导出学生数;

2、外模式/模式映像：

- ✧ 映像
- ✧ 程序与数据的逻辑独立性

3、模式/内模式映像

存储方式：不同索引

- ✓ 对于每一个外模式，数据库系统都有一个外模式/模式映像，它定义了该外模式与模式之间的对应关系。这些映像定义通常包含在各自外模式的描述中
- ✓ 关系数据模型中，外模式通过视图描述，映像定义在视图中
- ✓
- ✓ 注意：关系模型中，应用程序是依据数据的外模式与模式编写的

3、：

4、