



数据库系统概论

An Introduction to Database System

第七章 数据库设计



第七章 数据库设计



7.1 数据库设计概述

7.2 需求分析

7.3 概念结构设计

7.4 逻辑结构设计

7.5 数据库的物理设计

7.6 数据库实施和维护

7.7 小结

数据库设计概述



■ 什么是数据库设计

- 数据库设计是指对于一个给定的应用环境，构造最优的数据库模式，建立数据库及其应用系统，使之能够有效地存储数据，满足各种用户的应用需求（信息管理

要求：在数据库中应该存储和管理哪些数据对象；

数据操作要求：对数据对象需要进行哪些操作，如查询、增、删、改、统计等操作）

- 在数据库领域内，常常把使用数据库的各类系统统称为数据库应用系统。

- 特点：数据量大、保存时间长，数据关联复杂，用户要求多样化。

数据库设计概述（续）



■ 数据库设计

➤ 数据库设计的目标是为用户和各种应用系统提供一个信息基础设施和高效率的运行环境。

➤ 高效率的运行环境



- 数据库数据的存取效率高



- 数据库存储空间的利用率高



- 数据库系统运行管理的效率高



数据库设计准则

- 数据库必须正确反映现实世界，能为某个DBMS所接受
- 应用系统有良好的性能，有利于实施和维护
- 数据库能满足当前和今后相当长时期内的数据需求，使数据库有较长的使用寿命
- 当软件和硬件环境发生变化时容易修改和移植
- 能满足安全性要求，当系统发生故障时，容易恢复数据库
- 数据库存取效率、查询效率要高

7.1 数据库设计概述



7.1.1 数据库设计的特点

7.1.2 数据库设计方法

7.1.3 数据库设计的基本步骤

7.1.4 数据库设计过程中的各级模式



7.1.1 数据库设计的特点

■ 数据库建设的基本规律

➤ 三分技术，七分管理，十二分基础数据

■ 管理

➤ 数据库建设项目管理

➤ 企业（即应用部门）的业务管理

■ 基础数据

➤ 收集、入库

➤ 更新新的数据

■ 结构（数据）设计和行为（处理）设计相结合

➤ 将数据库结构设计和数据处理设计密切结合

7.1 数据库设计概述



7.1.1 数据库设计的特点

7.1.2 数据库设计方法

7.1.3 数据库设计的基本步骤

7.1.4 数据库设计过程中的各级模式



7.1.2 数据库设计方法

- 大型数据库设计是涉及多学科的综合性的技术，又是一项庞大的工程项目。
- 它要求多方面的知识和技术。主要包括：
 - 计算机的基础知识
 - 软件工程的原理和方法
 - 程序设计的方法和技巧
 - 数据库的基本知识
 - 数据库设计技术
 - 应用领域的知识

7.1.2 数据库设计方法

- 手工与经验相结合方法

- 设计质量与设计人员的经验和水平有直接关系
- 缺乏科学理论和工程方法的支持，工程的质量难以保证
- 数据库运行一段时间后常常不同程度地发现各种问题，增加了维护代价

- 规范设计法

- 基本思想：过程迭代和逐步求精

数据库设计方法（续）



典型方法

- 新奥尔良（New Orleans）方法

- 将数据库设计分为若干阶段和步骤

- 基于E-R模型的数据库设计方法

- 概念设计阶段广泛采用

- 3NF（第三范式）的设计方法

- 逻辑阶段可采用的有效方法

- ODL（Object Definition Language）方法

- 面向对象的数据库设计方法

- 统一建模语言（UML）方法

7.1 数据库设计概述



7.1.1 数据库设计的特点

7.1.2 数据库设计方法

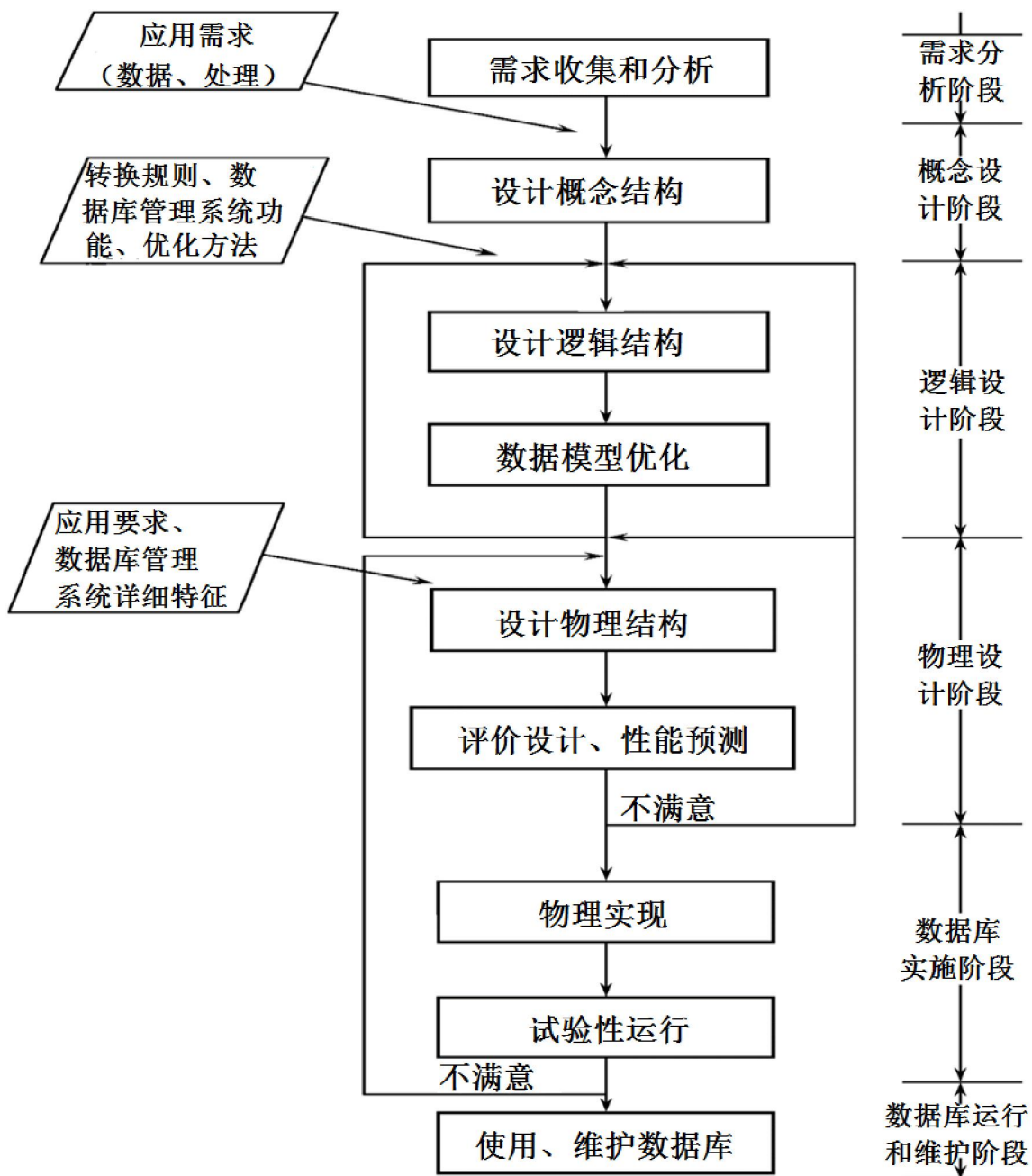
7.1.3 数据库设计的基本步骤

7.1.4 数据库设计过程中的各级模式



7.1.3 数据库设计的基本步骤

- 数据库设计分6个阶段
 - 需求分析
 - 概念结构设计
 - 逻辑结构设计
 - 物理结构设计
 - 数据库实施
 - 数据库运行和维护
- 需求分析和概念设计独立于任何数据库管理系统
- 逻辑设计和物理设计与选用的DBMS密切相关



数据库设计的基本步骤（续）

一、数据库设计的准备工作：

选定参加设计的人

1.系统分析人员、数据库设计人员

- 自始至终参与数据库设计

2.用户和数据库管理员

- 主要参加需求分析和数据库的运行维护

3.应用开发人员（程序员和操作员）

- 在系统实施阶段参与进来，负责编制程序和准备软硬件环境

数据库设计的基本步骤（续）

二、数据库设计的过程(六个阶段)

1.需求分析阶段

➤ 准确了解与分析用户需求（包括数据与处理）

➤ 最困难、最耗费时间的一步

数据库设计的基本步骤（续）

2.概念结构设计阶段

➤ 整个数据库设计的关键

➤ 通过对用户需求进行综合、归纳与抽象，形成一个独立于具体DBMS的概念模型

数据库设计的基本步骤（续）

3.逻辑结构设计阶段

➤ 将概念结构转换为某个DBMS所支持的数据模

型

➤ 对其进行优化

数据库设计的基本步骤（续）

4.数据库物理设计阶段

- 为逻辑数据模型选取一个最适合应用环境的物理结构（包括存储结构和存取方法）

数据库设计的基本步骤（续）

5.数据库实施阶段

➤ 运用**DBMS**提供的数据库语言（如**SQL**）及宿主语言，根据逻辑设计和物理设计的结果

➤ 建立数据库

➤ 编制与调试应用程序

➤ 组织数据入库

➤ 进行试运行

数据库设计的基本步骤（续）

6.数据库运行和维护阶段

➤ 数据库应用系统经过试运行后即可投入正式运

行

➤ 在数据库系统运行过程中必须不断地对其进行

评价、调整与修改

数据库设计的基本步骤（续）

设计一个完善的数据库应用系统往往是上述六个阶段的不断反复

- 这个设计步骤既是数据库设计的过程，也包括了数据库应用系统的设计过程。
- 把数据库设计和对数据库中数据处理的设计紧密结合起来
- 将这两个方面的需求分析、抽象、设计、实现在各个阶段同时进行，相互参照，相互补充，以完善两方面的设计

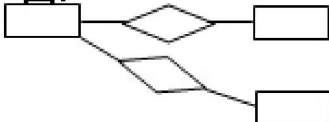
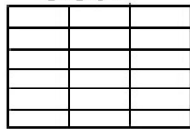
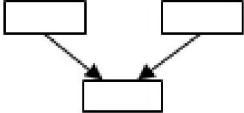
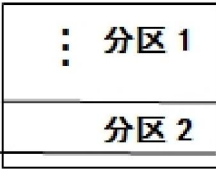
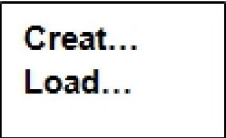
设计阶段	设计描述
需求分析	数字字典、全系统中数据项、数据结构、数据流、数据存储的描述
概念结构设计	概念模型 (E-R 图)  数据字典
逻辑结构设计	某种数据模型 <div> <div>关系</div>  </div> <div> <div>非关系</div>  </div>
物理结构设计	存储安排 存取方法选择 存取路径建立 
数据库实施	创建数据库模式 装入数据 数据库试运行 
数据库运行和维护	性能监测、转储/恢复、数据库重组和重构

图7.3 数据库设计各个阶段的数据设计描述

7.1 数据库设计概述



7.1.1 数据库设计的特点

7.1.2 数据库设计方法

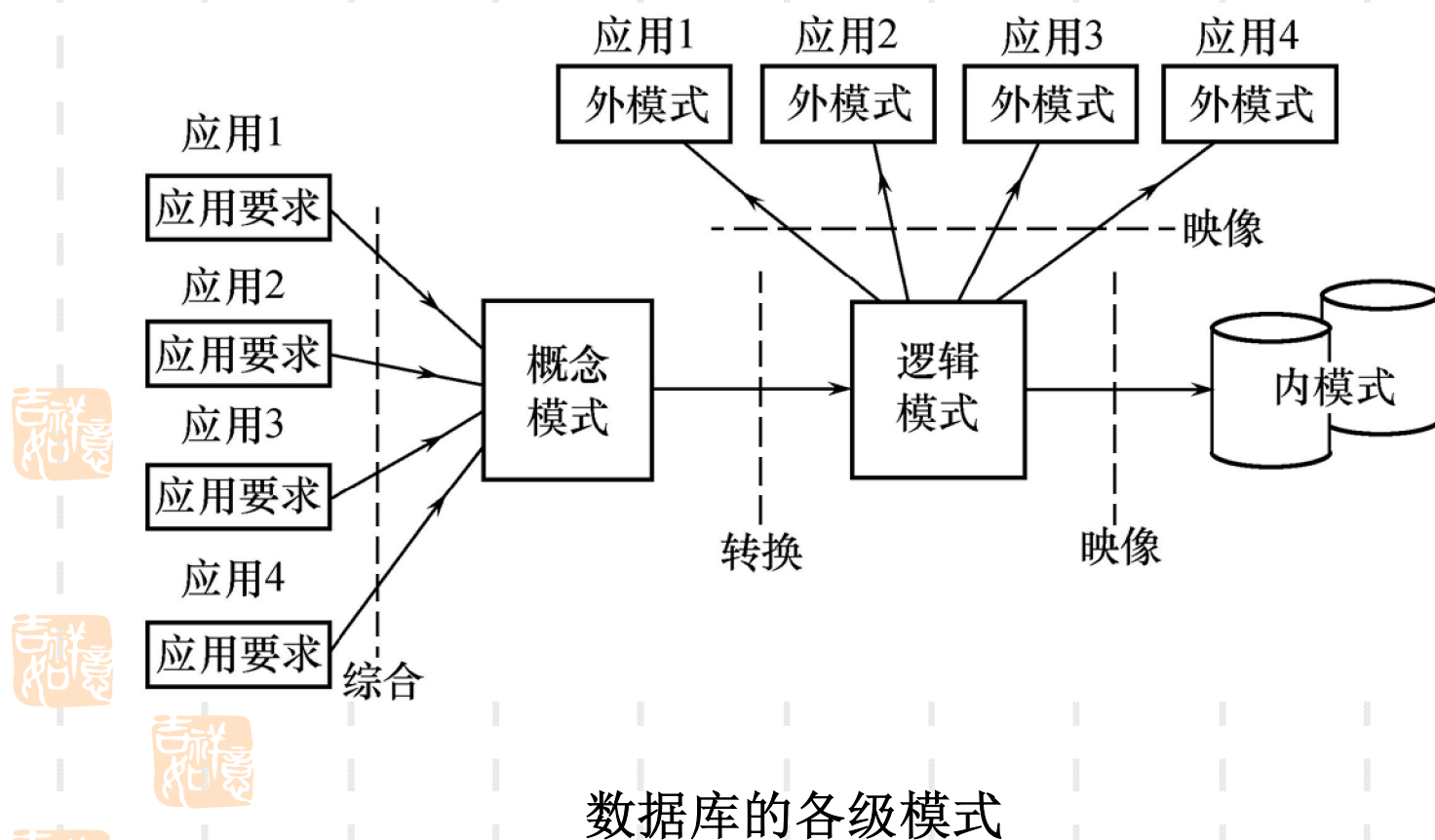
7.1.3 数据库设计的基本步骤

7.1.4 数据库设计过程中的各级模式



7.1.4 数据库设计过程中的各级模式

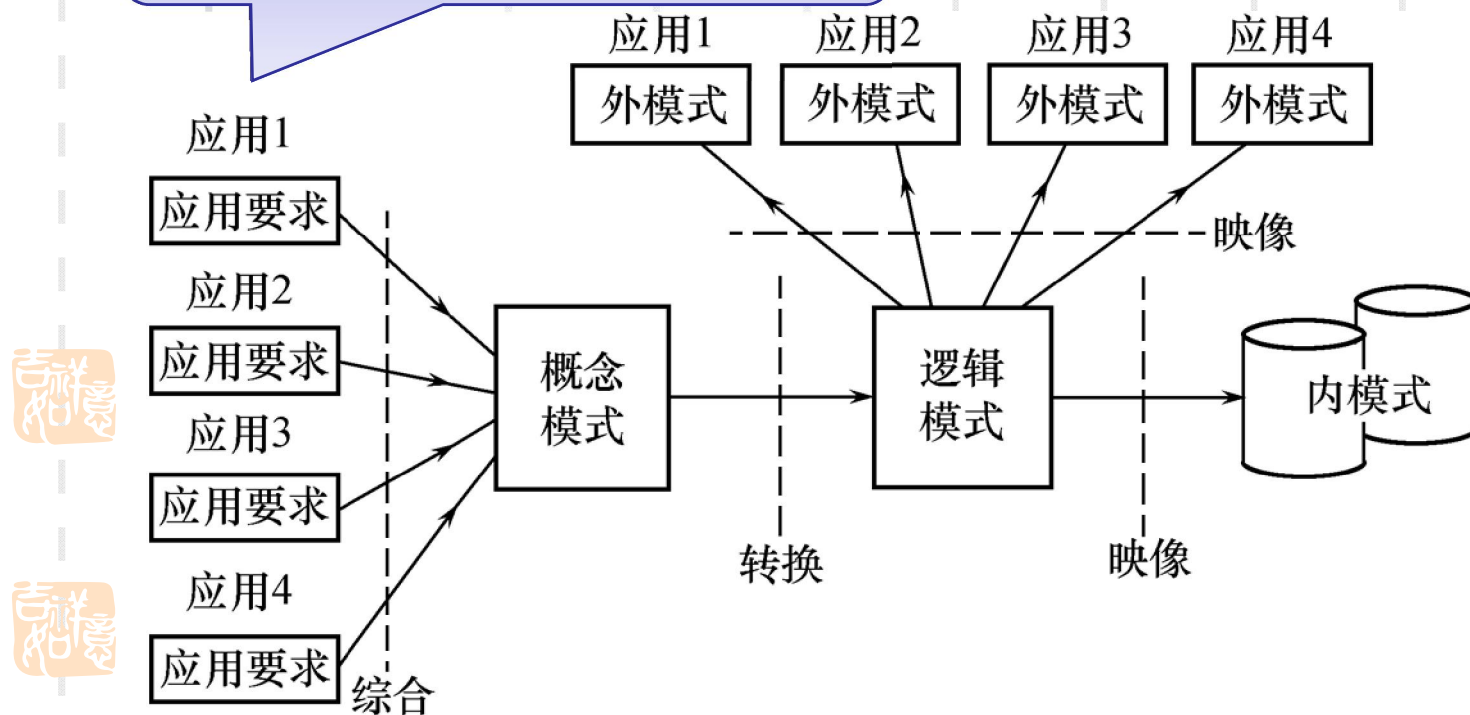
- 数据库设计不同阶段形成的数据库各级模式



数据库设计过程中的各级模式（续）

■ 数据库设计不同阶段形成的数据库各级模

需求分析阶段：
综合各个用户的应用需求



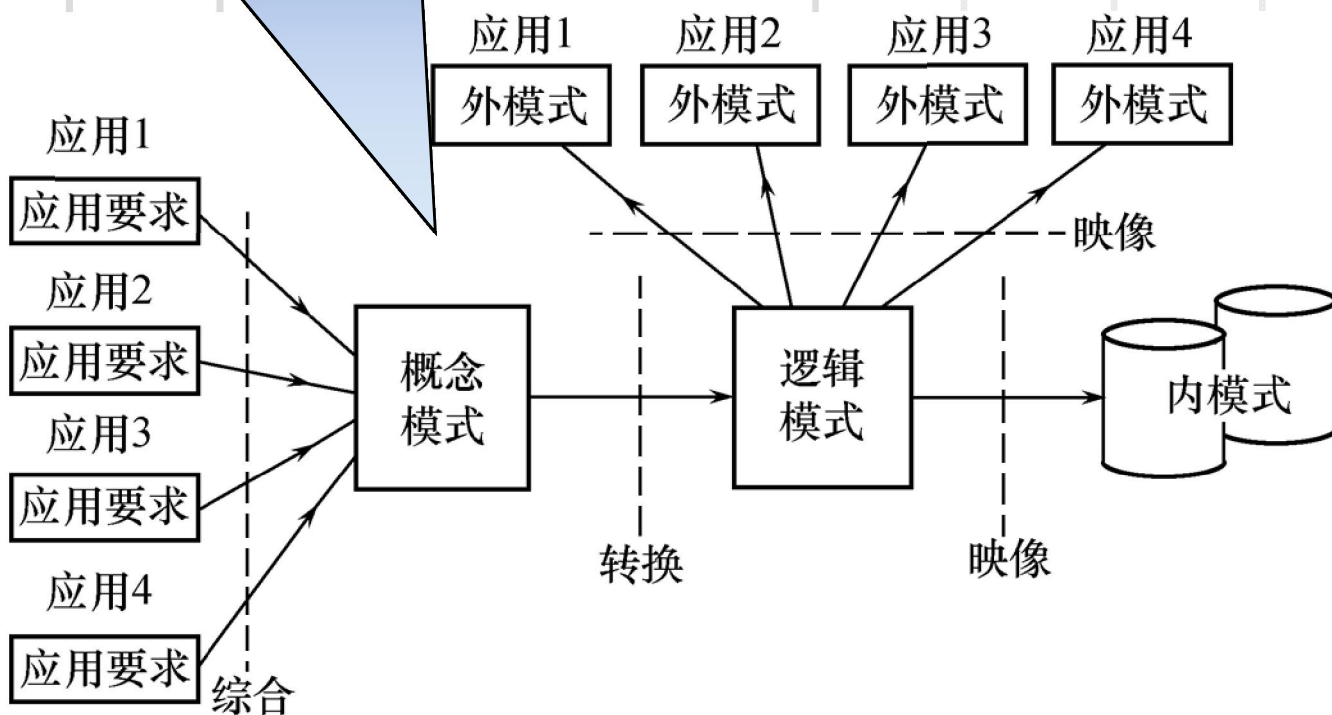
数据库的各级模式

数据库设计过程中的多级模式（续）

概念设计阶段：
形成独立于机器特点，独立于各个数据库管理系统产品的**概念模式（E-R图）**

■ 数
式

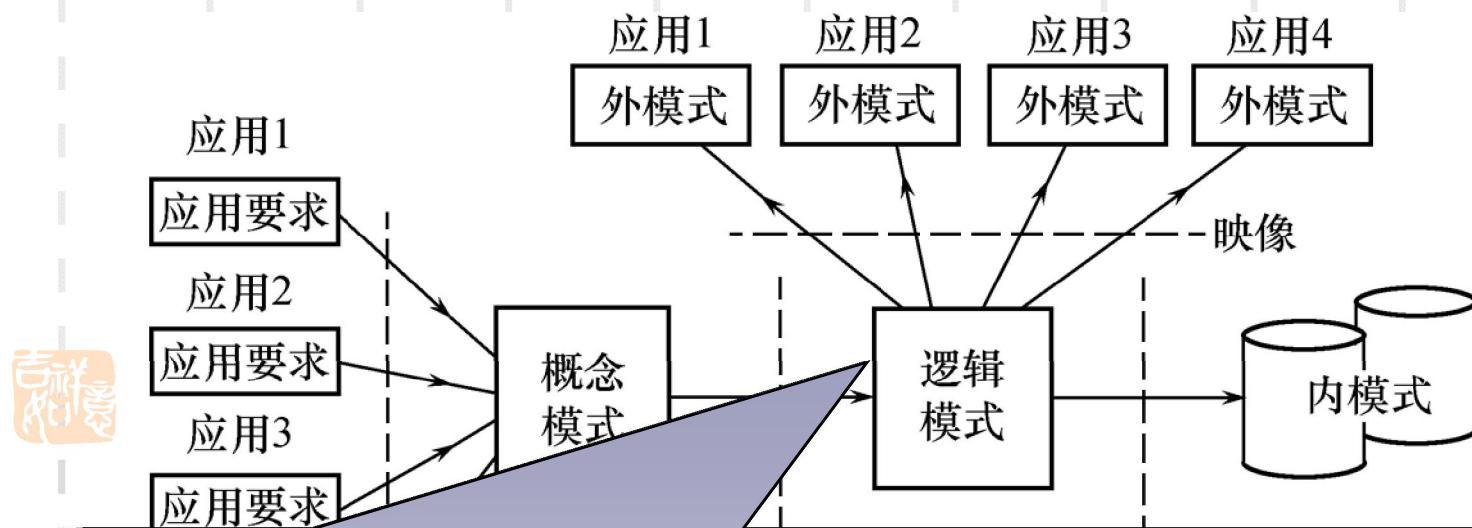
的数据库各级模



数据库的各级模式

数据库设计过程中的各级模式（续）

■ 数据库设计不同阶段形成的数据库各级模式



逻辑设计阶段：

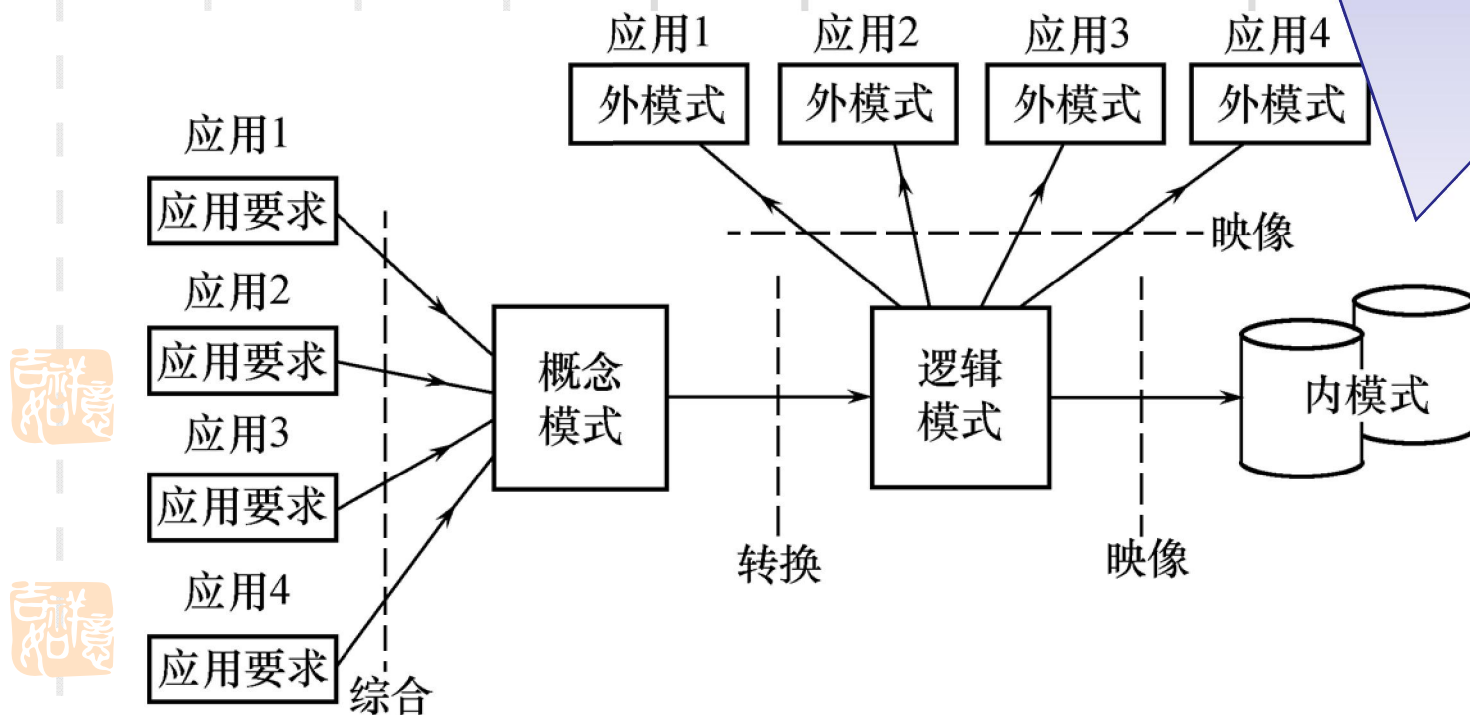
1. 首先将**E-R**图转换成具体的数据库产品支持的数据模型，如关系模型，形成数据库**逻辑模式**
2. 然后根据用户处理的要求、安全性的考虑，在基本表的基础上再建立必要的视图（**View**），形成数据的**外模式**

数据库设计过程中的各级模式 (续)

■ 数据库设计不式

物理设计阶段:

根据数据库管理系统特点和处理的需要, 进行物理存储安排, 建立索引, 形成数据库内模式



数据库的各级模式

第七章 数据库设计



7.1 数据库设计概述

7.2 需求分析

7.3 概念结构设计

7.4 逻辑结构设计

7.5 数据库的物理设计

7.6 数据库实施和维护

7.7 小结

需求分析



- 需求分析就是确定所要开发的应用系统的目标，收集和分析用户对数据库的要求，了解用户需要什么样的数据库，做什么样的数据库。对用户需求分析的描述是数据库概念设计的基础。
- 需求分析主要是考虑“做什么”的问题，而不是考虑“怎么做”的问题。
- 需求分析的结果是产生用户和设计者都能接受的需求说明书。需求分析简单地说就是分析用户的要求。需求分析是设计数据库的起点，需求分析的结果是否准确的反映了用户的实际要求，将直接影响到后面各个阶段的设计，并影响到设计结果是否合理和实用。



7.2.1 需求分析的任务

- 详细调查现实世界要处理的对象（组织、部门、企业等）
- 充分了解原系统（手工系统或计算机系统）工作概况
- 明确用户的各种需求
- 在此基础上确定新系统的功能
- 新系统必须充分考虑今后可能的扩充和改变

需求分析的任务（续）

- 调查的重点是“数据”和“处理”，获得用户对数据库的要求

（1）信息要求

- 用户需要从数据库中获得信息的内容与性质
- 由信息要求可以导出数据要求，即在数据库中需要存储哪些数据

（2）处理要求

- 用户要完成的处理功能
- 对处理性能的要求

（3）安全性与完整性要求

需求分析的难点

- 用户缺少计算机知识，开始时无法确定计算机究竟能为自己做什么，不能做什么，因此无法一下子准确地表达自己的需求，他们所提出的需求往往不断地变化。
- 设计人员缺少用户的专业知识，不易理解用户的真正需求，甚至误解用户的需求。
- 新的硬件、软件技术的出现也会使用户需求发生变化。

■ 解决方法

- 设计人员必须与用户不断深入地进行交流，才能逐步得以确定用户的实际需求

7.2.2 需求分析的方法

- 调查清楚用户的实际需求并进行初步分析
- 与用户达成共识
- 分析与表达这些需求



调查用户需求的步骤



- (1) 调查组织机构情况
- (2) 调查各部门的业务活动情况
- (3) 协助用户明确对新系统的各种要求，包括信息要求、处理要求、完全性与完整性要求
- (4) 确定新系统的边界



常用调查方法

(1) 跟班作业

- 通过亲身参加业务工作了解业务活动的情况

(2) 开调查会

- 通过与用户座谈来了解业务活动情况及用户需求

(3) 请专人介绍

(4) 询问

- 对某些调查中的问题，可以找专人询问

(5) 设计调查表请用户填写

- 调查表设计合理，则很有效

(6) 查阅记录

- 查阅与原系统有关的数据记录

进一步分析和表达用户需求

■ 分析方法

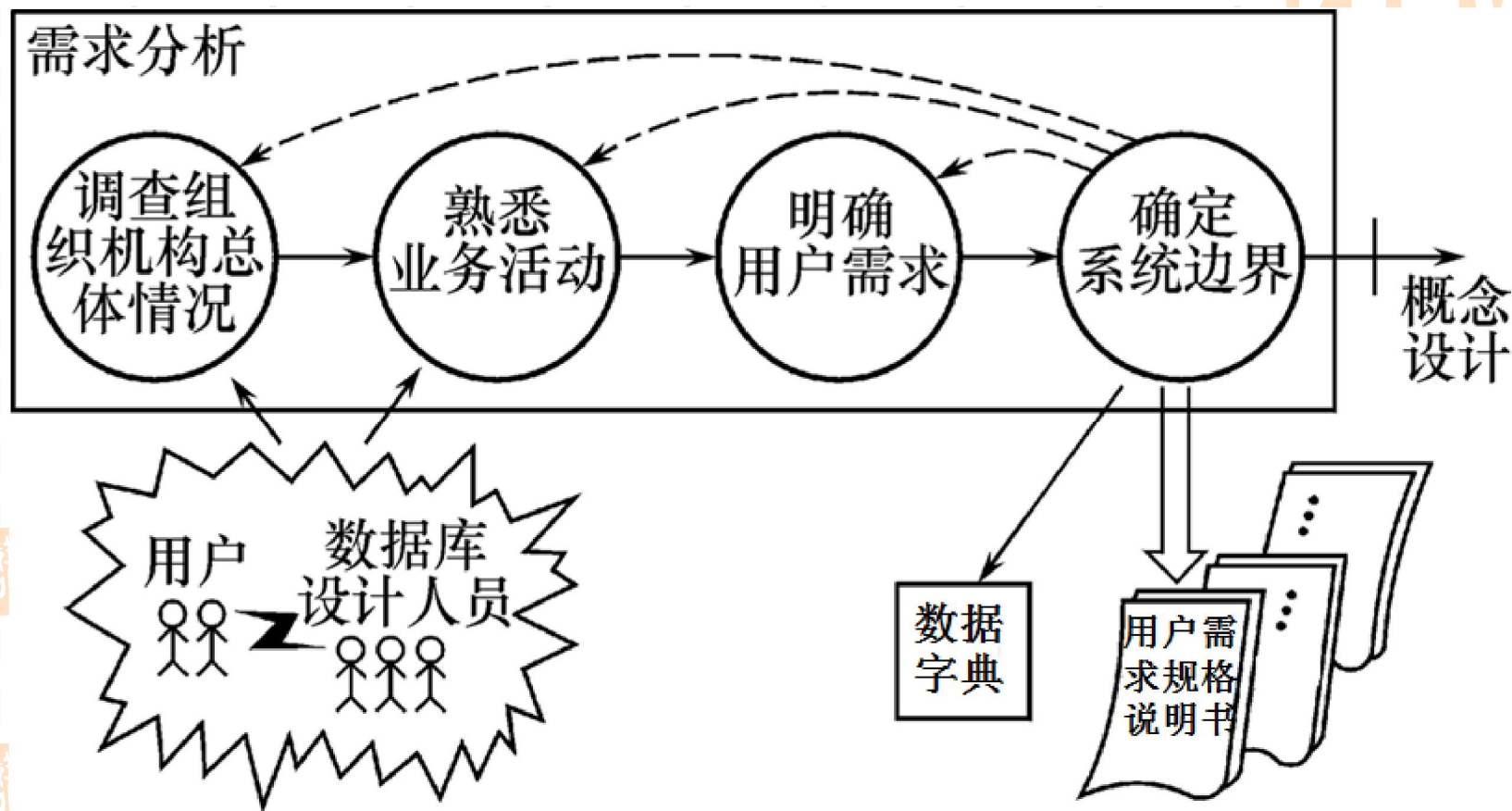
➤ 结构化分析方法（**Structured Analysis**，简称**SA**方法）

● SA方法从最上层的系统组织机构入手

● 采用自顶向下、逐层分解的方式分析系统

■ 对用户需求进行分析与表达后，需求分析报告必须提交给用户，征得用户的认可

需求分析过程



需求分析过程

7.2.3 数据字典

- 数据字典是关于数据库中数据的描述，即元数据，不是数据本身
- 数据字典在需求分析阶段建立，在数据库设计过程中不断修改、充实、完善
- 数据字典是进行详细的数据收集和数据分析所获得的主要结果

注意：

和关系数据库管理系统中数据字典的区别和联系

数据字典（续）

- 数据字典的内容

- 数据项
- 数据结构
- 数据流
- 数据存储
- 处理过程

- 数据项是数据的最小组成单位

- 若干个数据项可以组成一个数据结构

- 数据字典通过对数据项和数据结构的定义来描述数据流、数据存储的逻辑内容

1. 数据项

- 数据项是不可再分的数据单位
- 对数据项的描述

数据项描述={数据项名,数据项含义说明,别名,
数据类型,长度,取值范围,取值含义,
与其他数据项的逻辑关系,
数据项之间的联系}

- “取值范围”、“与其他数据项的逻辑关系”定义了数据的完整性约束条件，是设计数据检验功能的依据
- 可以用关系规范化理论为指导，用数据依赖的概念分析和表示数据项之间的联系

2. 数据结构

- 数据结构反映了数据之间的组合关系。
- 一个数据结构可以由若干个数据项组成，也可以由若干个数据结构组成，或由若干个数据项和数据结构混合组成。

■ 对数据结构的描述

数据结构描述=

{数据结构名, 含义说明, 组成:{数据项或数据结构}}

3. 数据流

- 数据流是数据结构在系统内传输的路径。
- 对数据流的描述

数据流描述={数据流名,说明,数据流来源,
数据流去向,组成:{数据结构},
平均流量,高峰期流量}

- 数据流来源：说明该数据流来自哪个过程
- 数据流去向：说明该数据流将到哪个过程去
- 平均流量：在单位时间（每天、每周、每月等）里的传输次数
- 高峰期流量：在高峰时期的数据流量

4. 数据存储

- 数据存储是数据结构停留或保存的地方，也是数据流的来源和去向之一。
- 对数据存储的描述

数据存储描述={数据存储名,说明,编号,输入的数据流,输出的数据流,组成:{数据结构},数据量,存取频度,存取方式}

- 存取频度：每小时、每天或每周存取次数，每次存取的数据量等信息
- 存取方法：批处理 / 联机处理；检索 / 更新；顺序检索 / 随机检索
- 输入的数据流：数据来源
- 输出的数据流：数据去向

5. 处理过程

- 处理过程的具体处理逻辑一般用判定表或判定树来描述。数据字典中只需要描述处理过程的说明性信息
- 处理过程说明性信息的描述

处理过程描述={处理过程名,说明,输入:{数据流},
输出:{数据流},处理:{简要说明}}

➤ 简要说明：说明该处理过程的功能及处理要求

- 功能：该处理过程用来做什么

- 处理要求：处理频度要求，如单位时间里处理多少事务，多少数据量、响应时间要求等

- 处理要求是后面物理设计的输入及性能评价的标准

数据字典举例

例：学生学籍管理子系统的数据字典。

数据项，以“学号”为例：

数据项：学号

含义说明：唯一标识每个学生

别名：学生编号

类型：字符型

长度：8

取值范围：00000000至99999999

取值含义：前两位标别该学生所在年级，
后六位按顺序编号

与其他数据项的逻辑关系：



数据字典举例（续）



数据结构，以“学生”为例

“学生”是该系统中的一个核心数据结构：

数据结构：学生

含义说明：是学籍管理子系统的主体数据结构，定义了一个学生的有关信息

组成：学号，姓名，性别，年龄，所在系，年级



数据字典举例（续）



数据流，“体检结果”可如下描述：

数据流： 体检结果

说明： 学生参加体格检查的最终结果

数据流来源：体检

数据流去向：批准

组成：

平均流量：

高峰期流量：



数据字典举例（续）



数据存储，“学生登记表”可如下描述：

数据存储： 学生登记表

说明： 记录学生的基本情况

流入数据流：

流出数据流：

组成：

数据量： 每年3000张

存取方式： 随机存取



数据字典举例（续）



处理过程 “分配宿舍” 可如下描述：

处理过程：分配宿舍

说明：为所有新生分配学生宿舍

输入：学生，宿舍

输出：宿舍安排

处理：在新生报到后，为所有新生分配学生宿舍。

要求同一间宿舍只能安排同一性别的学生，

同一个学生只能安排在一个宿舍中。

每个学生的居住面积不小于3平方米。

安排新生宿舍其处理时间应不超过15分钟。



需求分析小结

- 把需求收集和分析作为数据库设计的第一阶段是十分重要的。
- 第一阶段收集的基础数据（用数据字典来表达）是下一步进行概念设计的基础。

■ 强调两点

(1) 设计人员应充分考虑到可能的扩充和改变，使设计易于更改，系统易于扩充

(2) 必须强调用户的参与