



国家“十一五”规划教材

数据库原理与应用教程（第4版）



机械工业出版社
China Machine Press



第2章 数据库系统结构



- 2.1 数据和数据模型
- 2.2 概念层数据模型
- 2.3 组织层数据模型
- 2.4 数据库系统的结构

2.1 数据和数据模型



- 现实世界的的数据是散乱无章的，散乱的数据不利于人们对其进行有效的管理和处理。
- 因此，必须把现实世界的的数据按照一定的格式组织起来，以方便对其进行操作和使用。
- 在用数据库技术管理数据时，数据被按照一定的格式组织起来，比如二维表结构，以使数据能够被更高效地管理和处理。

2.1.1 数据与信息



- 描述事物的符号记录称为**数据**。
- 将从数据中获得的有意义的内容称为**信息**
- 数据有一定的格式，这些格式的规定是数据的语法，而数据的含义是数据的语义。
- 数据是信息存在的一种形式，只有通过解释或处理才能成为有用的信息。

示例



- 数据：
(张三, 9912101, 男, 1981, 计算机系, 应用软件)
- 解释：
 - 张三是9912101班的男生, 1981年出生, 计算机系应用软件专业。



数据的静态特征



- 数据的**基本结构**
 - 学生的学号、姓名、性别、出生日期
- 数据间的**联系**
 - 学生选课中的学号与学生基本信息中的学号
- 数据取值范围**约束**
 - 考试成绩在0 ~ 100分之间

动态特征



- 指对数据可以进行的操作以及操作规则。
- 对数据库数据的操作主要有
 - 查询数据
 - 更改数据：插入、删除和更新
- 一般将对数据的静态特征和动态特征的描述称为数据模型三要素

2.1.2 数据模型



- 对于模型，人们并不陌生。

- 建筑模型



- 飞机模型



- 计算机中的模型是对事物、对象、过程等客观系统中感兴趣的内容的模拟和抽象表达，是理解系统的思维工具
- 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（续）



- 数据库管理系统是基于某种数据模型对数据进行组织的，因此，了解数据模型的基本概念是学习数据库知识的基础。
- 在数据库领域中，数据模型用于表达现实世界中的对象，即将现实世界中杂乱的信息用一种规范的、形象化的方式表达出来。

数据模型（续）



- 数据模型即要面向现实世界，又要面向机器世界，因此需满足三个要求：
 - 能够真实地模拟现实世界；
 - 容易被人们理解；
 - 能够方便地在计算机上实现。

数据模型（续）



- 数据模型实际上是模型化数据和信息的工具。根据模型应用的不同目的，可以将模型分为两大类：
- **概念层数据模型**（概念模型），从数据的语义视角来抽取模型，是按用户的观点来对数据和信息进行建模。
- **组织层数据模型**（组织模型）。从数据的组织层次来描述数据。

2.2 概念层数据模型



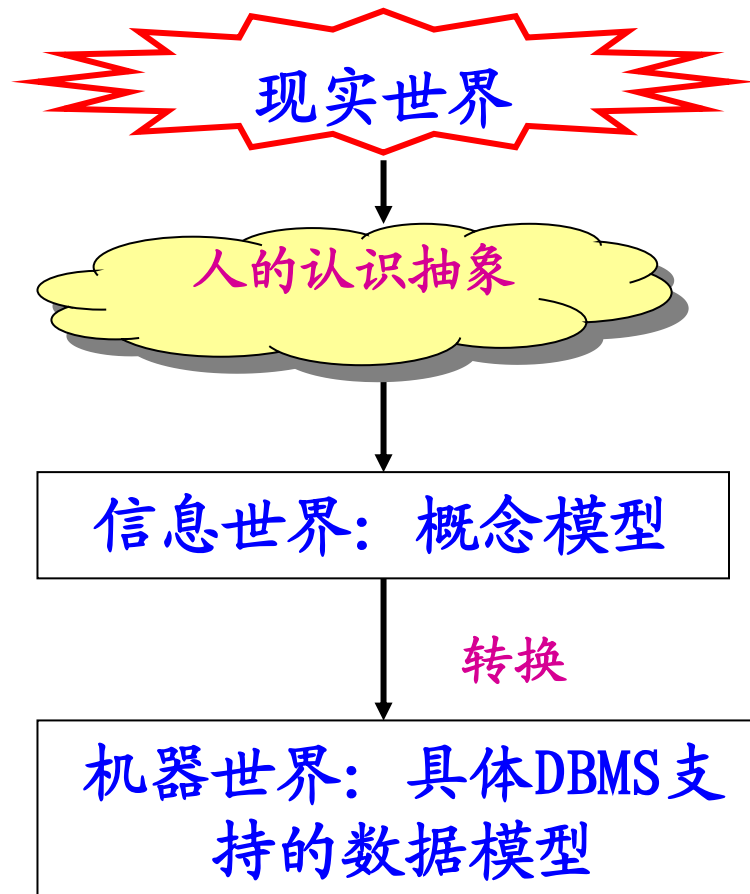
- 从数据的应用语义视角来抽取现实世界中有价值的数据并按用户^{用户}的观点对数据进行建模。
- 主要用在数据库的设计阶段^{设计阶段},
- 与具体的数据库管理系统无关^{与具体的数据库管理系统无关},
- 与具体的实现方式无关。

组织层数据模型



- 从数据的组织方式来描述数据。主要有：
 - 层次模型
 - 网状模型
 - 关系模型
 - 对象-关系模型
- 是从计算机系统的观点对数据进行建模
- 与所使用的数据库管理系统有关。

现实世界客观事物的抽象过程



2.2 概念层数据模型



- 概念层次模型实际上是现实世界到机器世界的一个中间层次。
- **概念层次模型**：抽象现实系统中有应用价值的元素及其关联，反映现实系统中有应用价值的信息结构，不依赖于数据的组织结构。
- 概念模型用于信息世界的建模，是现实世界到信息世界的**第一层抽象**。
- 是数据库设计人员和用户之间进行交流的工具。
- 是面向用户、面向现实世界的数据模型，是**与DBMS无关**。
- 常用的概念模型：实体-联系模型、语义对象模型

实体-联系模型



- 实体
- 属性
- 联系

实体



- 具有公共性质的可相互区分的现实世界对象的集合。
- 可以是具体的事物，也可以是抽象的概念或联系
- 具体的事物： 学生、课程、职工
- 抽象的联系： 学生选课

实体的表示方式



- 在E-R图中用矩形框表示实体，把实体名写在框内，比如学生实体可以表示为：

学生

属性



- 实体所具有的特征或性质。

身高

年龄

性别

体重

.....



联系

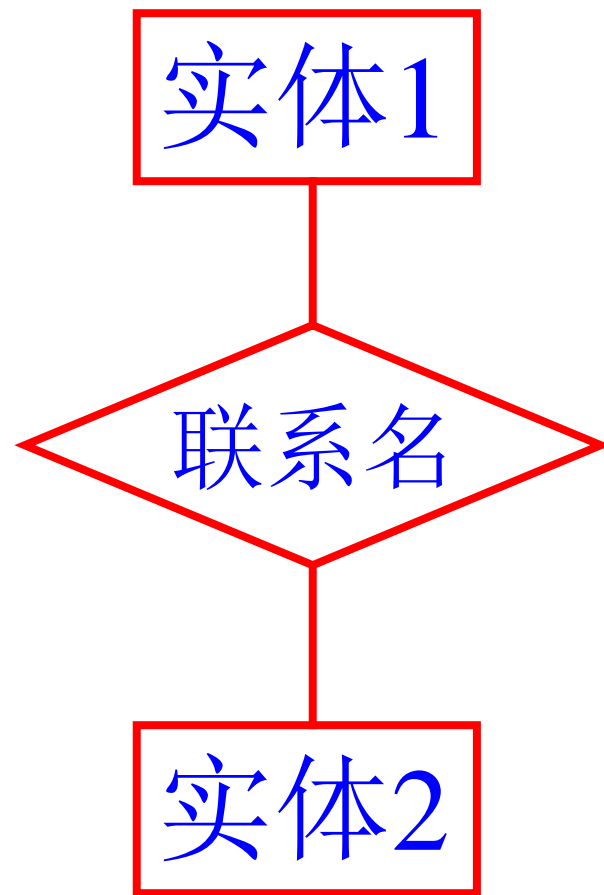


- 联系是数据之间的关联集合，是客观存在的应用语义链。
- 实体内部的联系：一个实体内属性之间的联系。
 - 职工实体内部的职工号和此职工的部门经理号
- 实体之间的联系：不同实体之间的联系。
 - 课程实体和学生实体之间存在选课联系。

联系的表示方式



实体之间的
联系用菱形框表
示，框内写上联
系名，并用连线
与有关的实体相
连。



联系的种类



一对一联系(1:1)

一对多联系(1:n)

多对多联系(m:n)

一对一联系(1:1)

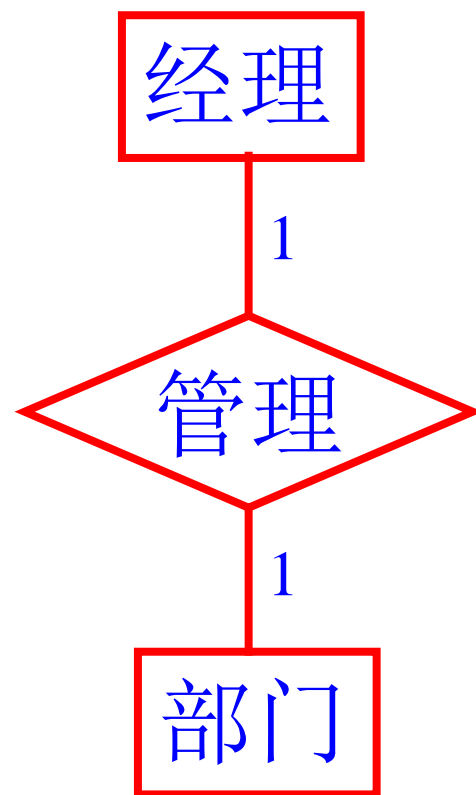


如果实体A中的每个实例在实体B中至多有一个（也可以没有）实例与之关联，反之亦然，则称实体A与实体B具有一对一联系，记作：1 : 1。

一对一联系的例子



部门和正经理（假设一个部门只有一个正经理，一个人只当一个部门的经理）、系和正系主任（假设一个系只有一个正主任，一个人只当一个系的主任）都是一对一联系。



一对多的联系

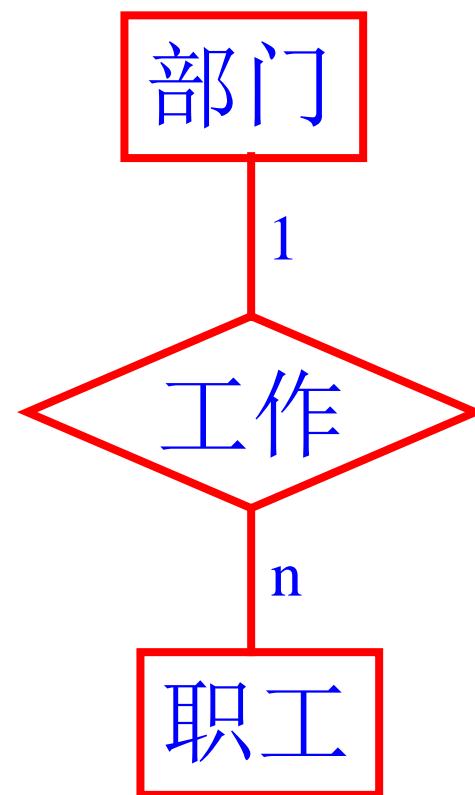


如果实体A与实体B之间存在联系，并且对于实体A中的一个实例，实体B中有多个实例与之对应；而对实体B中的任意一个实例，在实体A中都只有一个实例与之对应，则称实体A到实体B的联系是一对多的，记为 $1:n$ 。

一对多联系的例子



有部门和职工两个实体，
并且有语义：一个部门可以有
多名职工，但是一个职工只在
一个部门工作。则部门和职工
之间的联系是一对多的，我们
把这种联系命名为工作。



多对多的联系

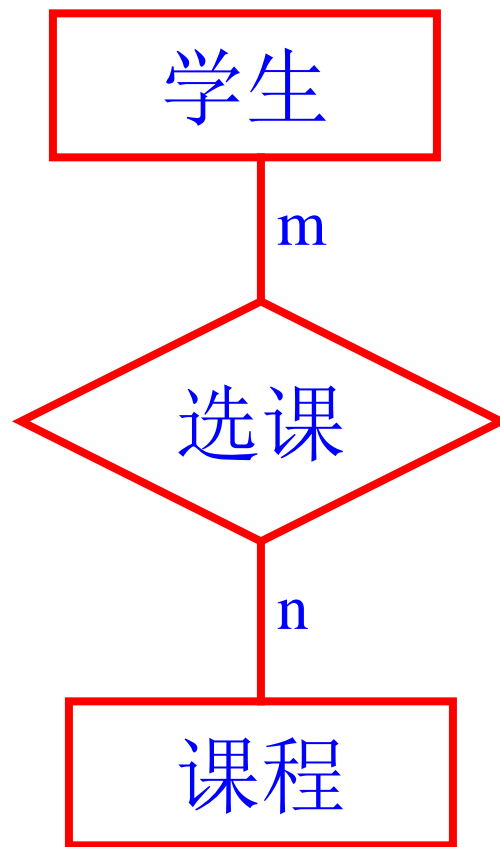


如果实体A与实体B之间存在联系，并且对于实体A中的一个实例，实体B中有多个实例与之对应；而对实体B中的一个实例，在实体A中也有多个实例与之对应，则称实体A到实体B的联系是多对多的，记为 $m:n$ 。

多对多联系的例子



有学生和课程两个实体，并且有语义：一个学生可以修多门课程，一门课程可以被多个学生修。那么学生和课程之间的联系就是多对多的，我们把这种联系命名为选课。



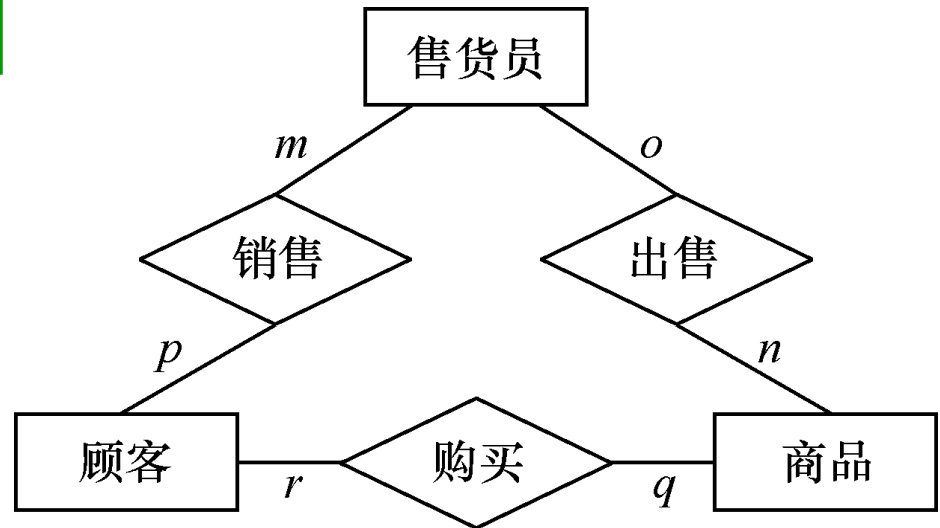
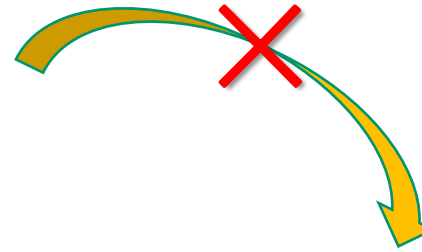
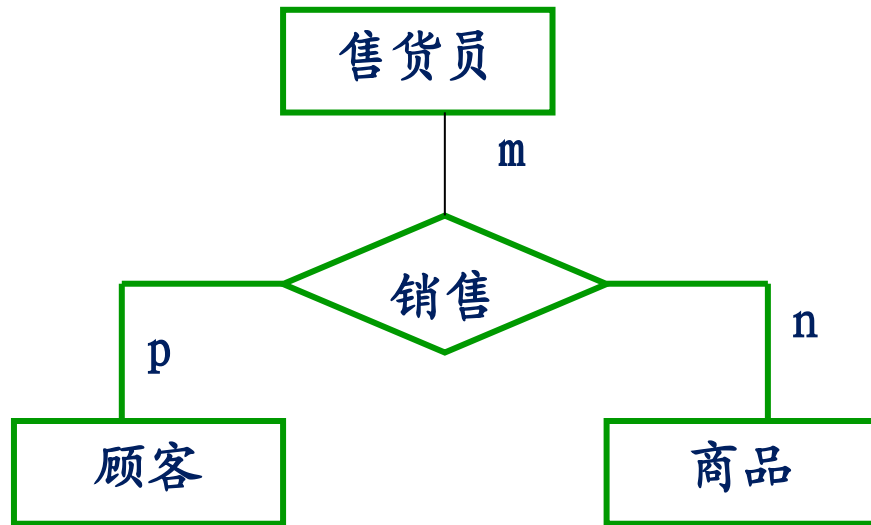
关联多个实体的联系



• 顾客购买商品:

- 每个顾客可 **售货员** 售货员那里购买商品, 并且可以 m 买多种商品;
- 每个售货员可以 **销售** 多名顾客销售商品, 并且可以 p 销售多种商品; n
- 每种商品可由多个售货员 **顾客** 并 **商品** 且可以 n 给多名顾客。

两个图不等价



2.3 组织层数据模型



- 2.3.1 关系数据模型的数据结构
- 2.3.2 关系数据模型的数据操作
- 2.3.3 关系数据模型的数据完整性约束

关系数据模型的数据结构



- 把数据看成是二维表中的元素，而这个二维表就是关系
- 用关系（表格数据）表示实体和实体之间联系的模型称为关系数据模型
- 示例：学生基本信息表

学号	姓名	性别	年龄	所在系
9512101	李勇	男	19	计算机系
9512102	刘晨	男	20	计算机系
9512103	王敏	女	20	计算机系
9521101	张立	男	22	信息系
9521102	吴宾	女	21	信息系

一些基本术语



- **关系**：关系就是二维表。并满足如下性质：
 - 关系表中的每一列都是不可再分的基本属性；
 - 表中的行、列次序并不重要。

不是关
系表

不是基
本属性

学号	姓名	性别	年龄	所在系	出生日期		
					年	月	日
9512101	李勇	男	19	计算机系	1984	4	6
9512102	刘晨	男	20	计算机系	1984	12	15
9512103	王敏	女	20	计算机系	1983	8	21
9521101	张立	男	22	信息系	1983	6	3

一些基本术语（续）



- **元组**：表中的每一行称作是一个元组，它相当于一个记录值。
- **属性**：表中的每一列是一个属性值的集合，列可以命名，称为属性名。
- **主码**：表中用于惟一地确定一个元组的一个属性或最小的属性组。
- **域**：属性的取值范围。如性别的域为：
（‘男’，‘女’）

主码示例



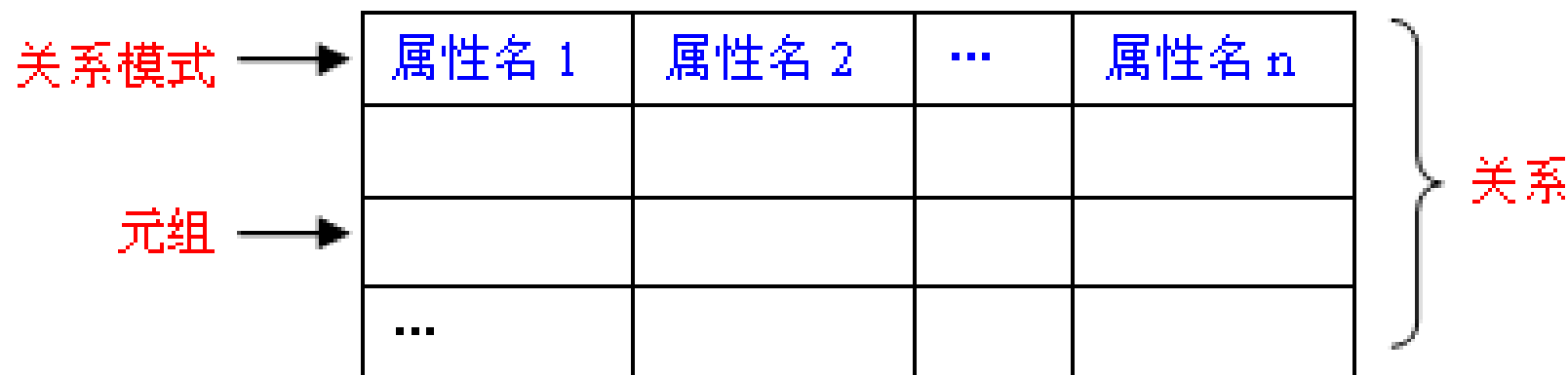
- 学生基本信息表：
(学号, 姓名, 年龄, 性别, 所在系)
- 学生修课信息表：
(学号, 课程号, 成绩)

一些基本术语（续）



- **关系模式**：二维表的结构称为关系模式，即，关系模式是二维表的表框架或表头结构。
- 关系模式一般表示为：
关系名（属性1，属性2，...，属性n）
- 例如，学生关系模式为：
学生（学号，姓名，性别，年龄，所在系）

各概念之间的关系



关系模型的数据操作



- 关系模型的操作对象是集合（也就是关系）。
- 非关系型数据库系统中典型的操作是一次一行或一次一个记录。
- 因此，集合处理能力是关系系统区别于其他系统的一个重要特征。
- 关系数据模型的数据操作主要包括四种：查询、插入、删除和修改数据。

关系模型的数据完整性约束



- 实体完整性
- 参照完整性
- 用户定义的完整性

实体完整性



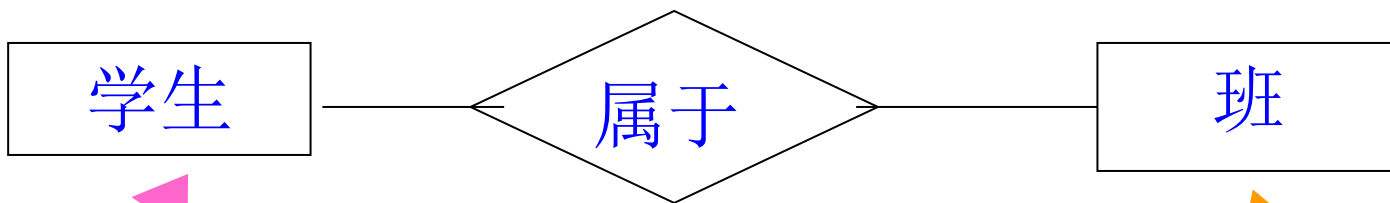
- 实体完整性是保证关系中的每个元组都是可识别的和惟一的。
- 而且表中不允许存在如下的记录：
 - 无主码值的记录
 - 主码值相同的记录

参照完整性



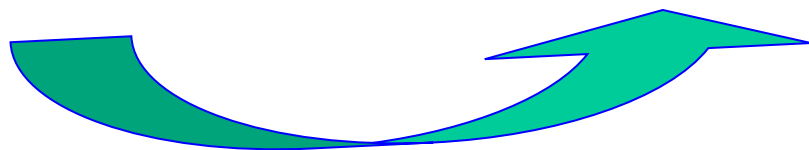
- 用于描述实体之间的联系。
- 参照完整性一般是指多个实体（表）之间的关联关系。
- 一般用外码实现。
 - 外码：**取作本表（子表）属性之一的外表（父表，主表）主码。
- 主码值先在主表中生成，后在子表中引用

外码引用例1

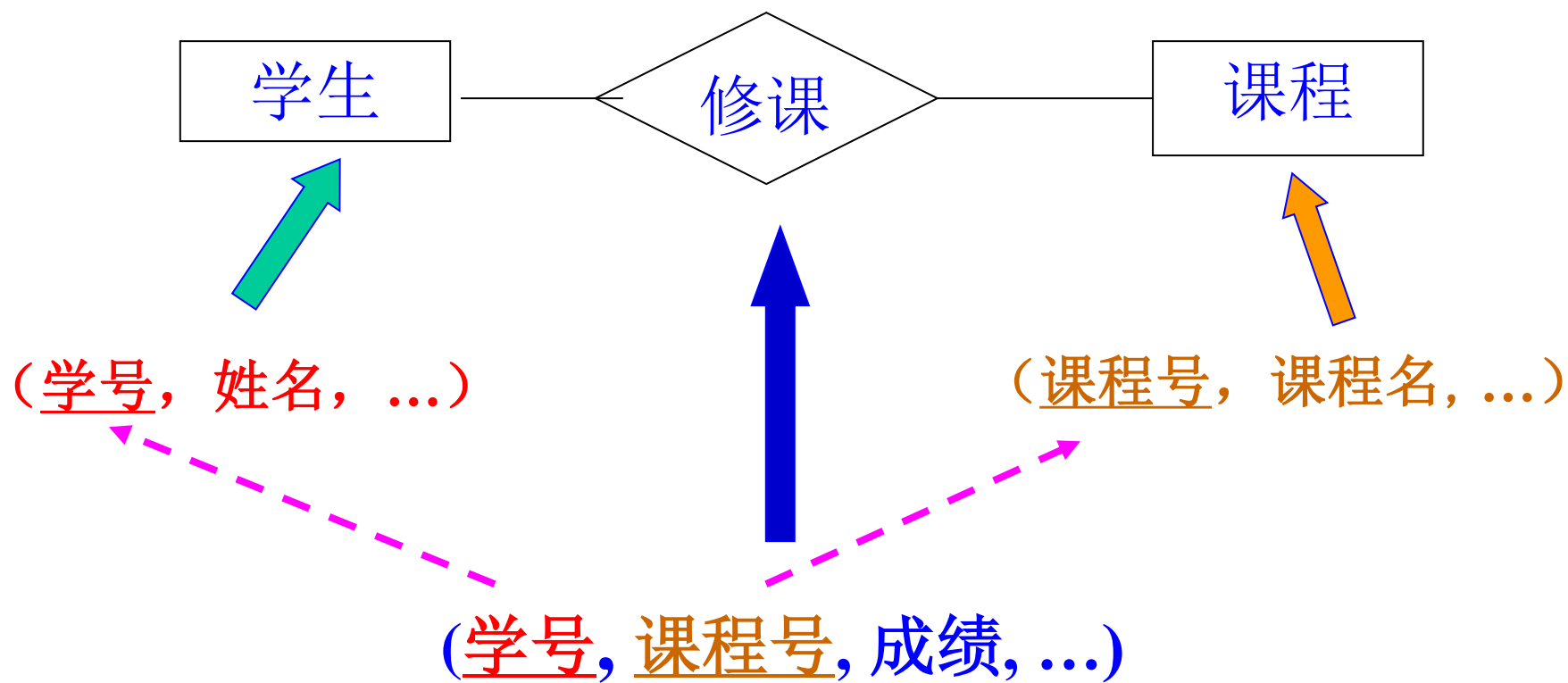


(学号, 姓名, 班号, ...)

(班号, 专业, 人数, ...)



外码引用例2



参照完整性规则



- 参照完整性规则就是定义外码与主码之间的引用规则。
- 对于外码，一般应符合如下要求：
 - 或者值为空；
 - 或者等于其所应用的关系中的某个元组的主码值。

外码性质



- 外码可以有重复值。
- 例：如下两个关系模式表示：

外码

职工（职工号，职工名，部门号，工资级别）

部门（部门号，部门名）
- 如果某新来职工还没有被分配到具体的部门，则其“部门号”就为空值；
- 如果职工已经被分配到了某个部门，则其部门号就有了确定的值（非空值）。

用户定义的完整性



- 也称为域完整性或语义完整性。
- 按应用语义，属性数据有：
 - 类型与长度限制：方便计算机操作
 - 取值范围限制：防止属性值与应用语义矛盾



语义许可取值范围约束例:

成绩取

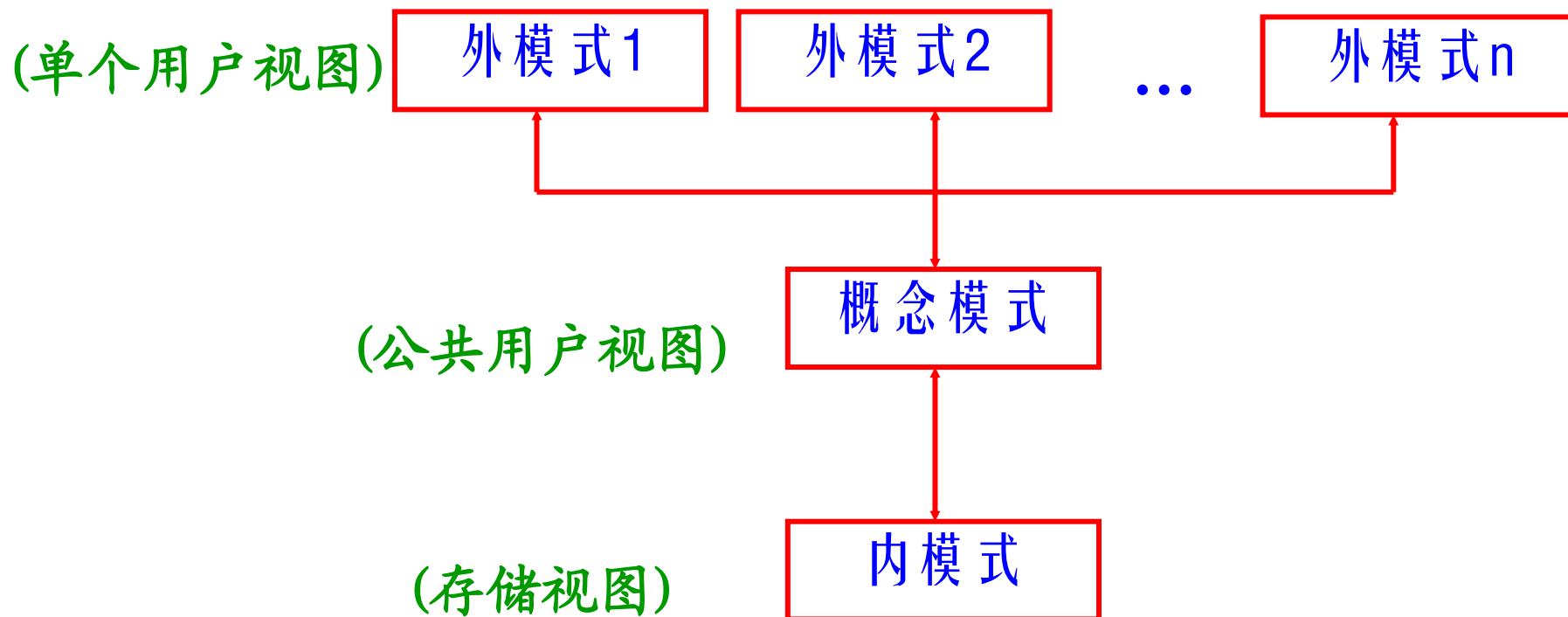
- $[0 .. 100]$,
- {优、良、中、及格、不及格}

2.4 数据库系统的结构



- 2.4.1 三级模式结构
- 2.4.2 模式映象与数据独立性

三级模式结构





- 概念模式是关系的。
- 外模式也是关系的或接近关系的，其内容来自概念模式，它最接近用户。
- 内模式不是关系的，它是数据的物理存储方式。

外模式



- 也称为用户模式、子模式。
- 用户对现实系统中**感兴趣**整体的局部数据结构的描述。
- 是DB整体数据结构的子集或局部重构。
- 用外模式定义语言表达。例，对RDB，定义视图结构。
- 是保证数据库安全的一个措施。

示例1——子集



学生

学 号	姓 名	年 龄	性 别	所 在 系
0611101	李勇	21	男	计算机系
0611102	刘晨	20	男	计算机系
0611103	王敏	20	女	计算机系
0621101	张立	20	男	信息管理系
0621102	吴宾	19	女	信息管理系

模式

学 号	姓 名	性 别
0611101	李勇	男
0611102	刘晨	男
0611103	王敏	女
0621101	张立	男
0621102	吴宾	女

外模式

示例2——重构



- 学生（学号，**姓名**，性别，年龄，所在系）
- 课程（课程号，**课程名**，学分）
- 选课（学号，课程号，**成绩**）

学生（姓名，课程名，成绩）

示例3——安全性



职工表（职工号，姓名，所在部门，基本工资，职务工资，~~奖励工资~~）



职工信息（职工号，姓名，所在部门，基本工资，职务工资）

概念模式



- 也称为逻辑模式或模式，是数据库中**全体数据**的逻辑结构和特征的描述，是所有用户的公共数据描述。
- 是数据库系统结构的**中间层**。
- 用模式定义语言表达。例，对关系数据库，定义：表头结构、属性取值范围、…。



- 模式实际上是数据库数据在逻辑级上的视图。
- 一个数据库只有一种模式。
- 不涉及存储记录对列、索引、指针或其他存储的访问细节。

内模式



- 也称为存储模式。
- 数据的内部存储结构描述（数据的组织与存储）。
- 用内模式定义语言描述。例，对RDB，定义数据文件的位置、索引的依据等。

2.4.2 模式映像与数据独立性



- 数据库系统的三级模式是对数据的三个抽象级别，它把数据的具体组织留给DBMS管理，使用户能逻辑地、抽象地处理数据，而不必关心数据在计算机中的具体表示方式与存储方式。
- 为了能够在内部实现这三个抽象层次的联系和转换，数据库管理系统在三级模式之间提供了两层映像：
 - 外模式/模式映像
 - 模式/内模式映像



模式间的映象 (DBMS基本功能之一)

- 维持DB数据与DB应用程序相互独立，从而使DB应用程序不随DB数据的逻辑或存储结构的变动而变动。

模式/内模式映像



- 当存储结构改变时，可（由DBA）用内模式定义语句，调整内模式定义，从而保持模式不变。

数据和程序物理独立（存储独立）

外模式/模式映像



- 当模式改变时，可用外模式/模式定义语句，调整外模式/模式映像定义，从而保持外模式不变。

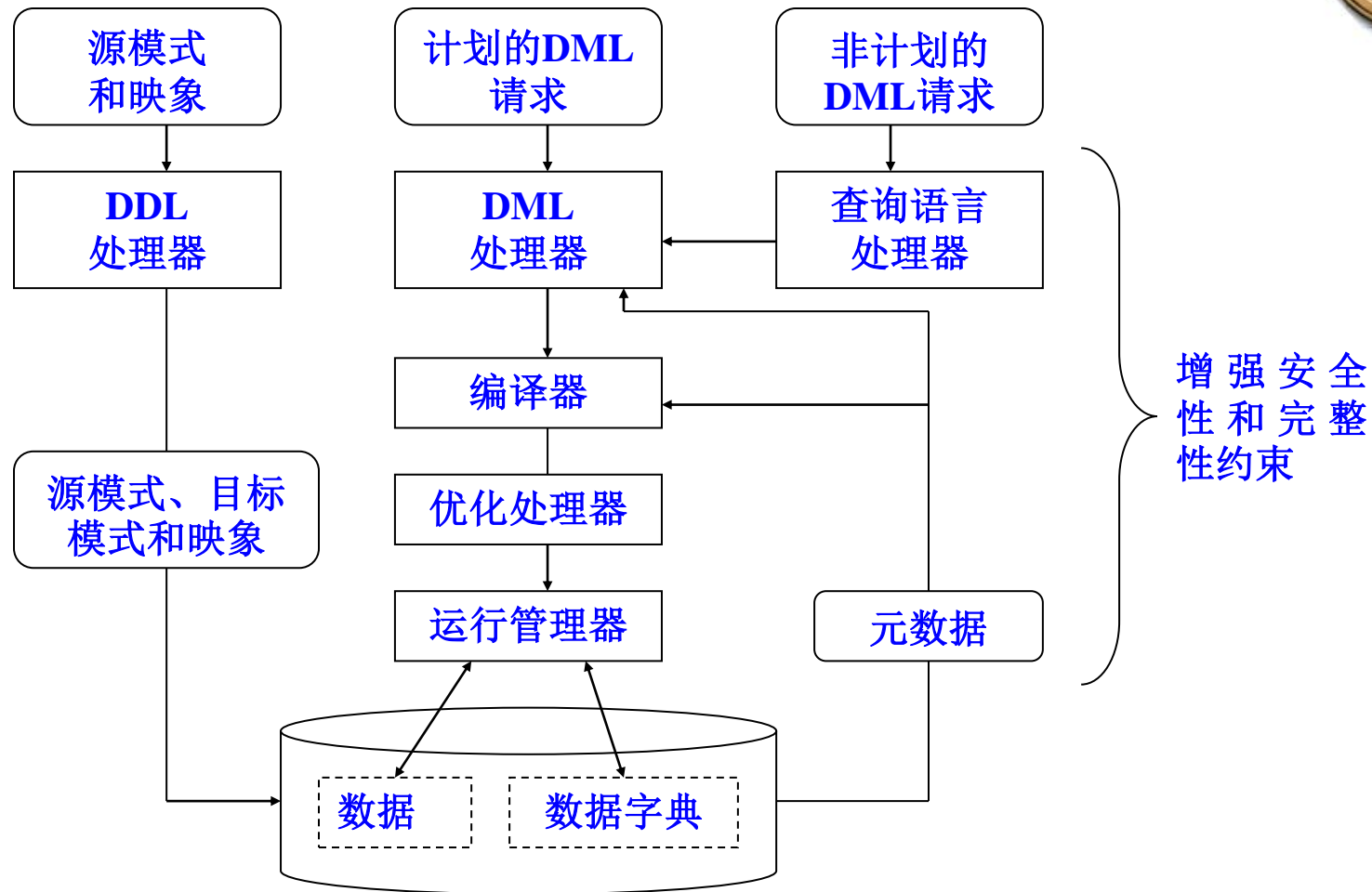
数据和程序逻辑独立（概念独立）

2.5 数据库管理系统



- **DBMS**是处理数据库访问的系统软件。
- 处理过程：
 - 用户使用数据库语言发出一个访问请求；
 - **DBMS**接受请求并分析；
 - **DBMS**检查用户外模式、相应的外模式 / 概念模式间的映象、概念模式、概念模式 / 内模式间的映象和存储结构定义。

DBMS的功能和组成



数据库管理系统的功能



- 数据定义
- 数据操作
- 优化和执行
- 数据安全和完整性
- 数据恢复和并发
- 数据字典
- 性能