

2 关系数据库基本知识

目录

- 2.1 数据模型
- 2.2 关系模型与关系数据库
- 2.3 关系数据库的基本理论

2.1 数据模型



2.1 数据模型

什么是 **模型** ？

对所研究的实体进行必要的**简化**，**用适当的表现形式或规则**描述它的**主要特征**，所得到的简化系统称之为**模型**。



2.1 数据模型

抽象和模仿鸟飞**原理**建飞机模型，再根据模型制造实物，目的是通过建立模型形成原理方案。





2.1 数据模型

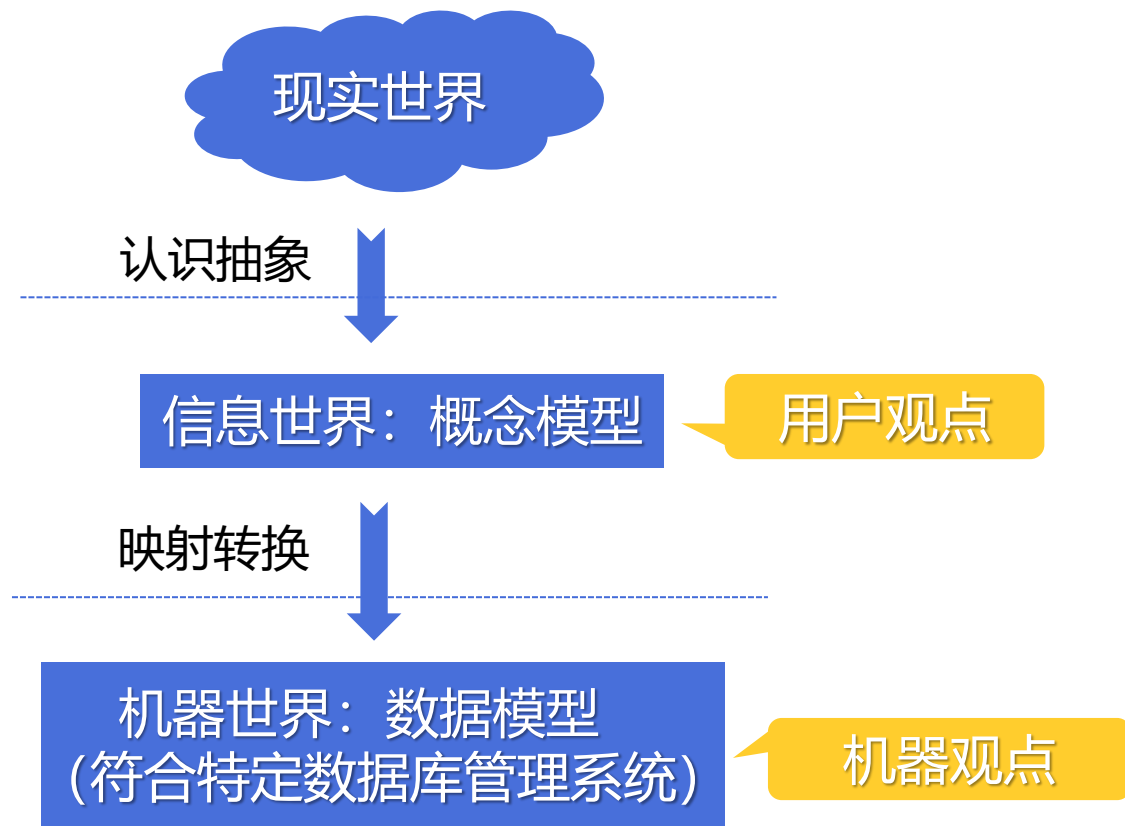
什么是 **数据模型** ？

用来**抽象、表示和处理**现实世界中的信息。

是构造数据**结构**时所遵循的**规则**以及对数据所能进行的**操作**的总称，是实现数据库的原理方案。



2.1 数据模型





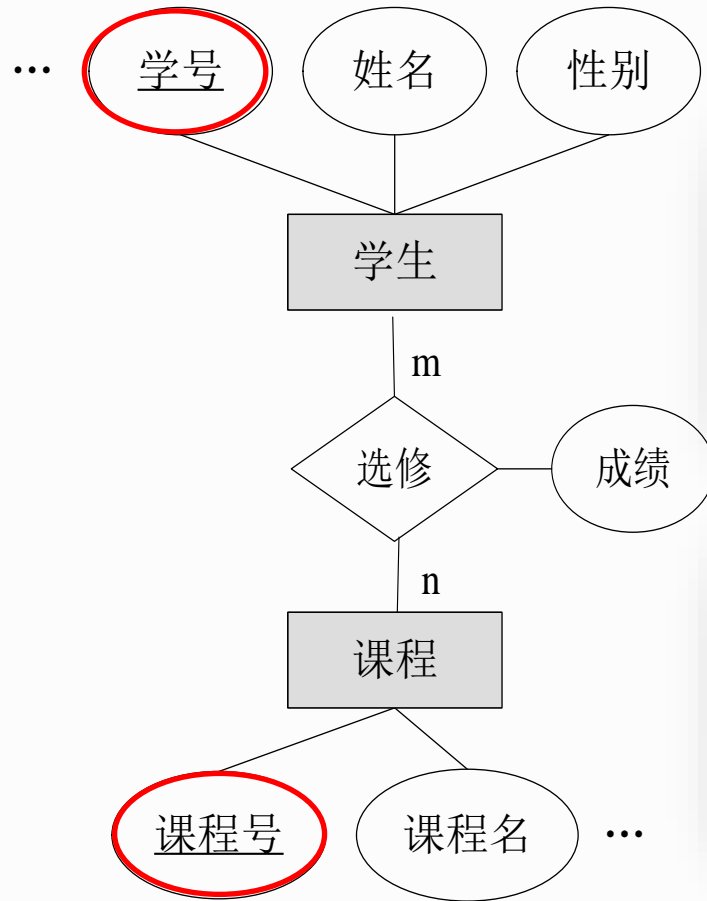
信息世界的概念模型

- ① 筛选出有价值的事物并对其主要**特性**进行**描述**；
- ② 描述这些事物之间的**联系**。



E-R模型 (E-R图)

Entity-Relationship Diagram



实体

现实世界客观存在并可相互区别的事物。用矩形表示

属性

实体中每个成员具有的特性。用椭圆形表示。

关键字

可以唯一标识实体的一组最小的属性集合。用属性加下划线表示。

联系

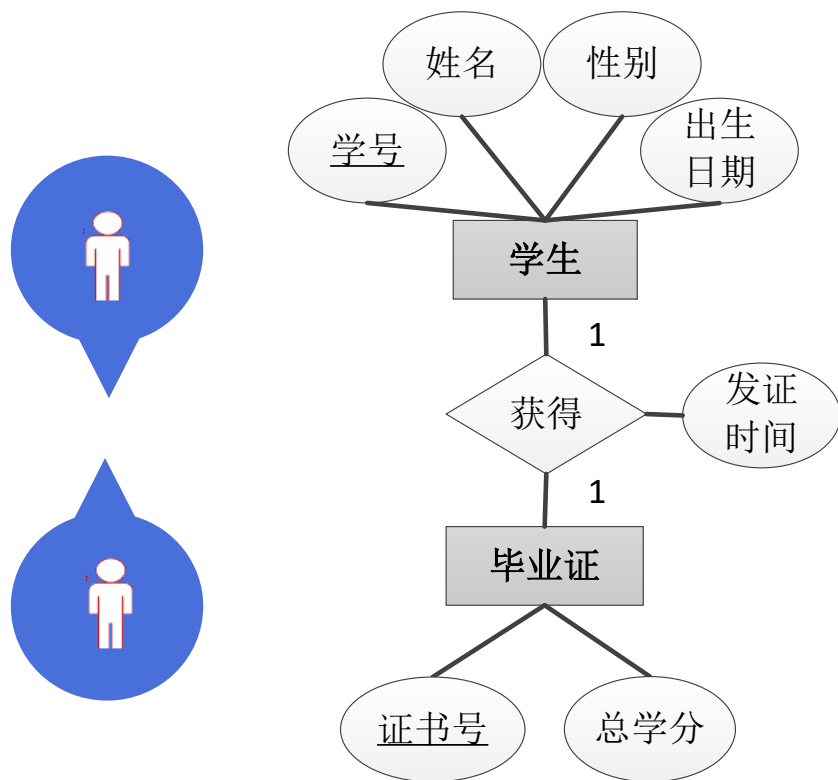
实体之间的一种关联。联系也可以具有属性。用菱形表示。



E-R图中联系的三种类型

一对一联系 1:1

实体集A中的每一个成员最多与实体集B中的一个成员相关联，反之亦然。

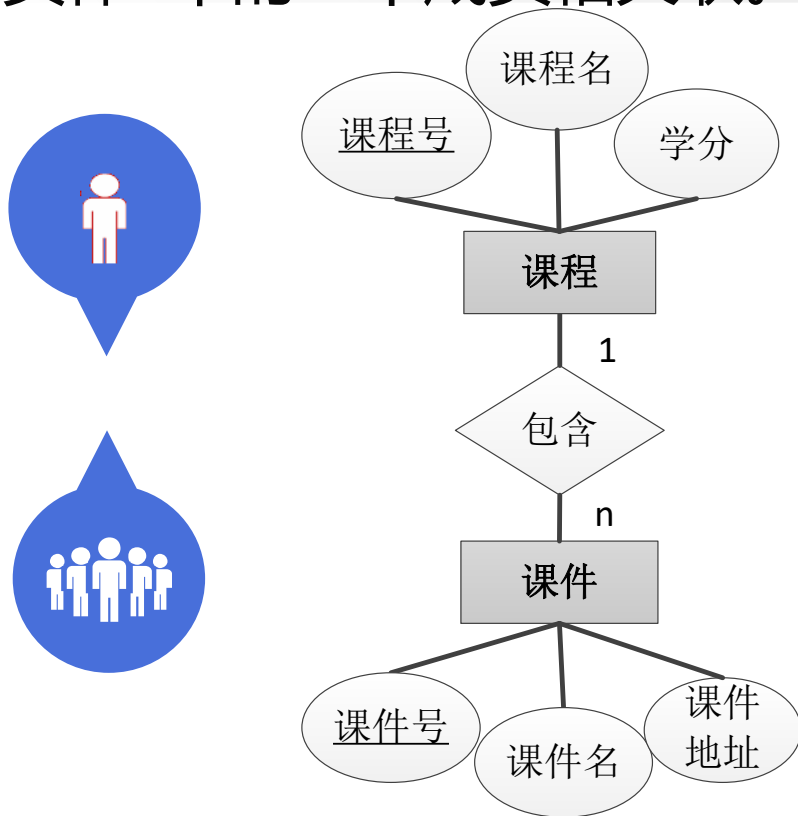




E-R图中联系的三种类型

一对多联系 1:n

实体集A中的每一个成员可以与实体集B中的多个成员相关联，反之，实体B中的每一个成员最多与实体A中的一个成员相关联。

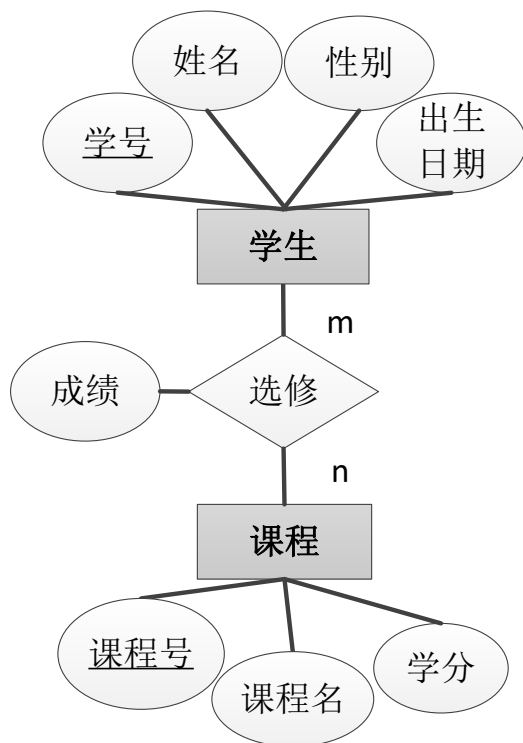
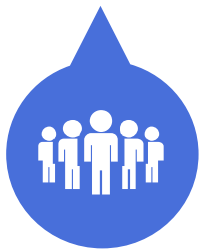




E-R图中联系的三种类型

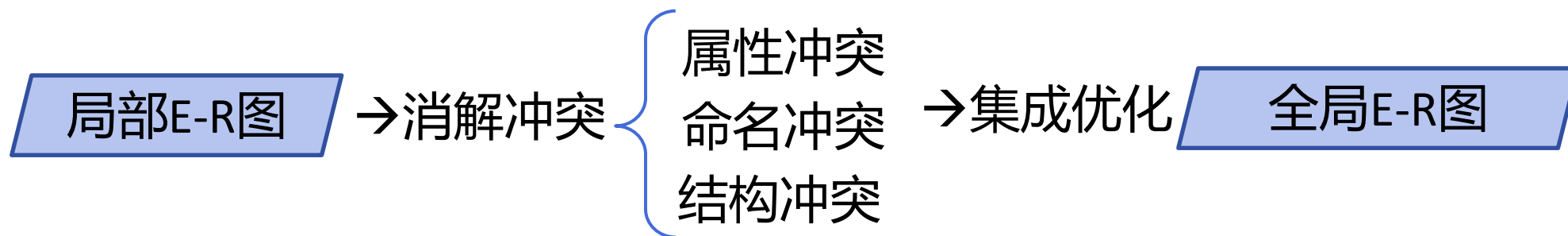
多对多联系 m:n

实体集A中的每一个成员可以与实体集B中的多个成员相关联，反之，实体B中的每一个成员可以与实体A中的多个成员相关联。



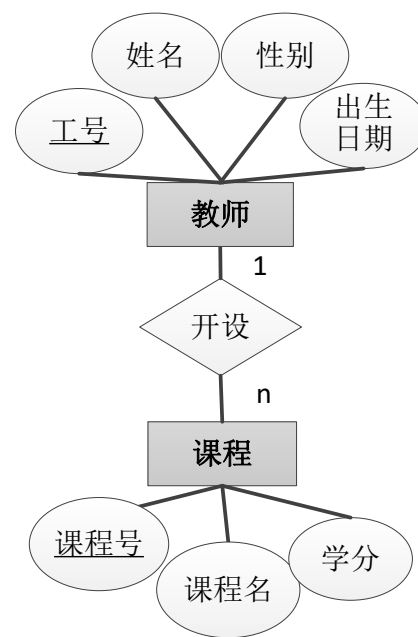
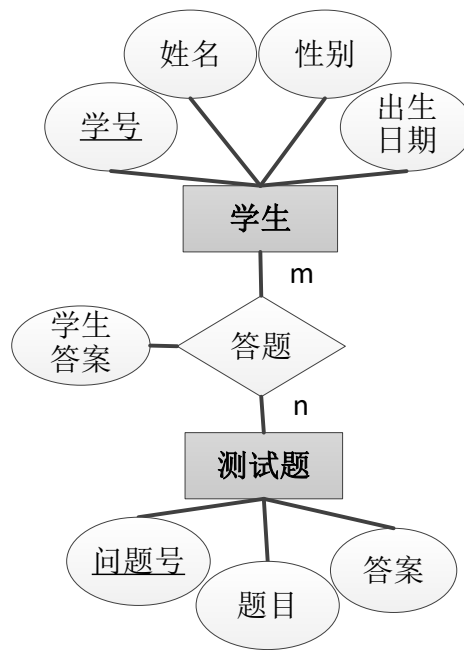
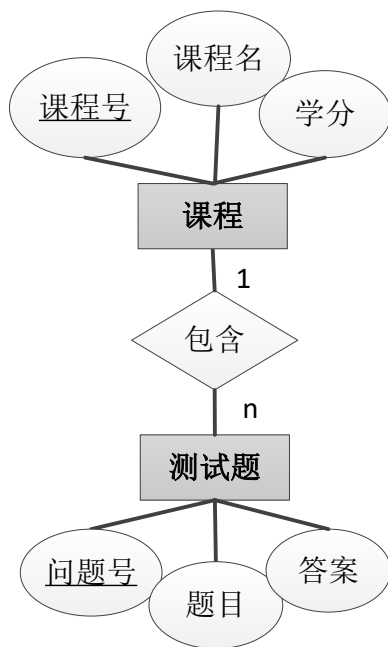
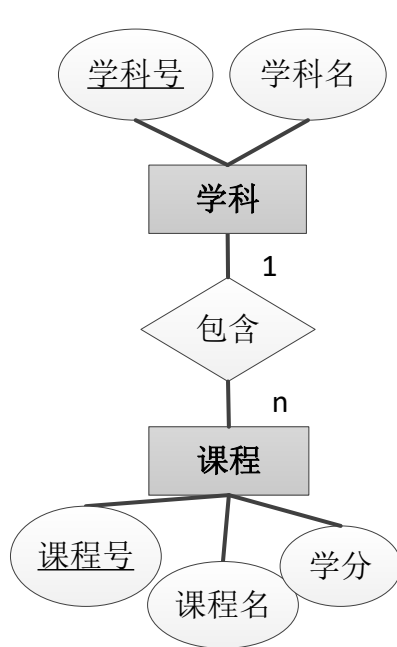
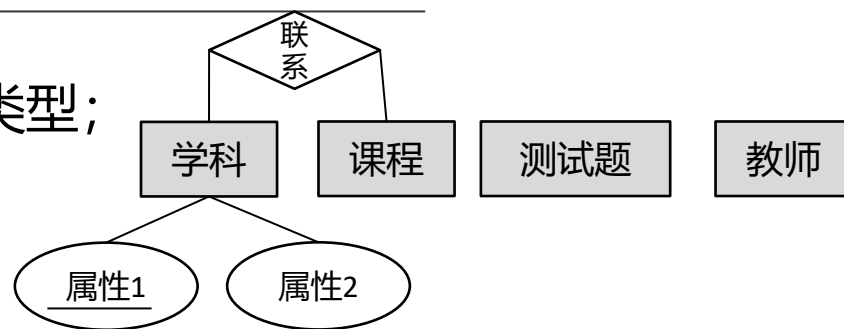


E-R图的设计过程



设计局部E-R图

- 确定实体，分析各实体间存在的联系，并定义联系的类型；
- 确定实体的属性、联系的属性；
- 定义各实体的主关键字。



“e学习”系统的其他局部E-R图

设计全局E-R图

- 将局部E-R图进行合并集成就得到全局E-R图。集成时要消除以下冲突：
 - ✓ **属性冲突**。一是属性域冲突，即属性的类型、取值范围和取值集合不同；二是属性的取值单位冲突。
 - ✓ **命名冲突**。一是同名异义，二是异名同义（通常允许）。
 - ✓ **结构冲突**。一是同一对象有不同的抽象表示，某处为实体，而另一处为属性；二是同一实体在两个图中的属性个数不同。解决这类冲突可以用各图中属性的并集作为该实体的属性。

学生（**学号**【字符串】，姓名，性别，生日）
毕业证（**学号**【整数】，毕业证号，总学分）
选课（**学生号**，课程号，**成绩**）
成绩（数学，外语，计算机）

有冲突的局部E-R图

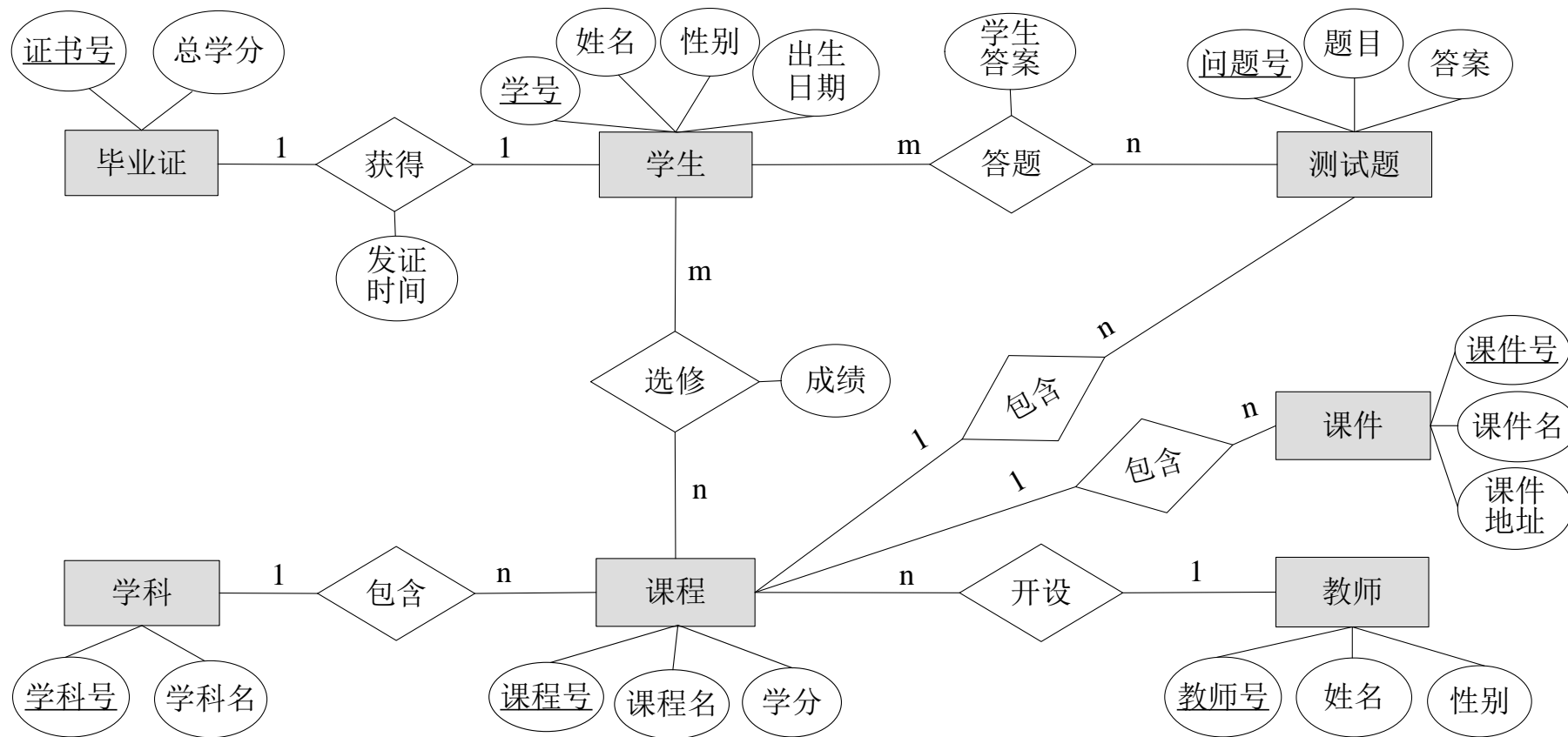
对全局E-R图进行优化

反应
需求

实体少

属性少

联系
无冗余

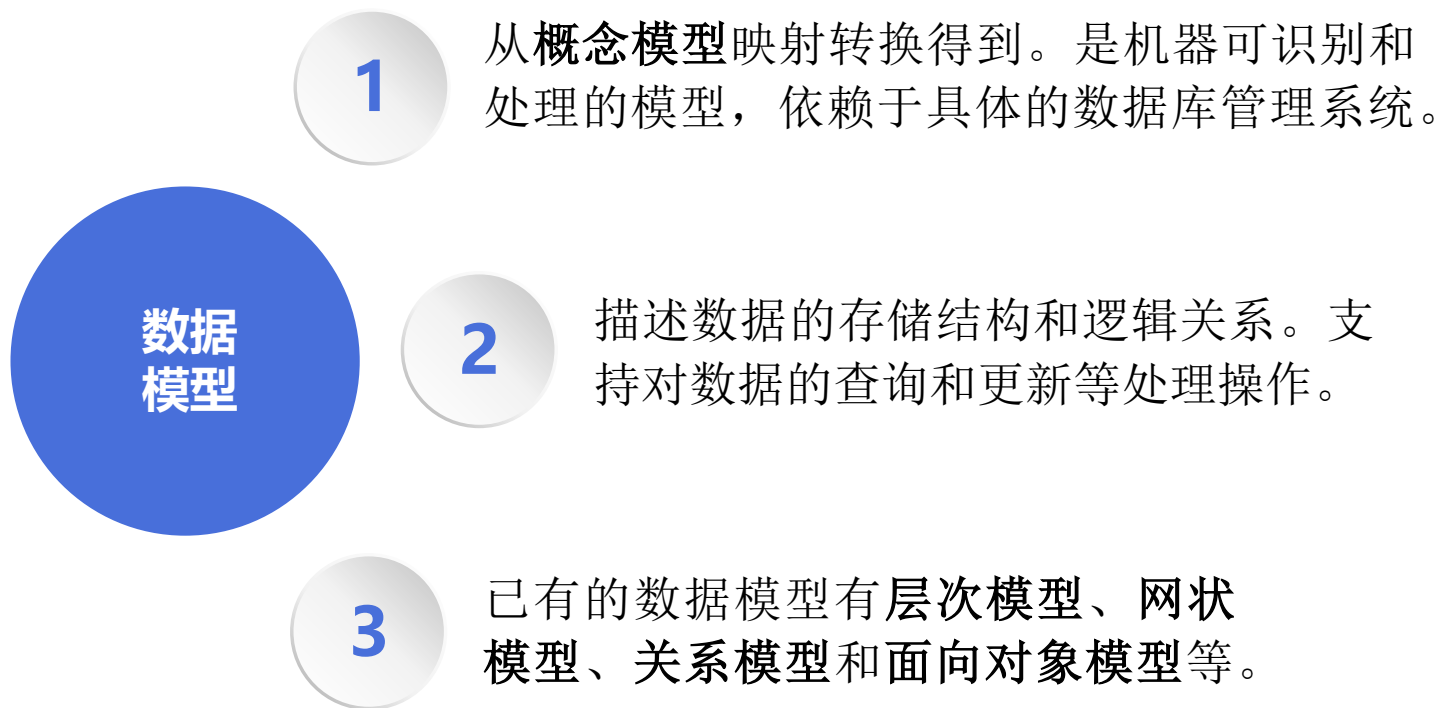


"e学习" 系统的全局E-R图



机器世界的数据库模型

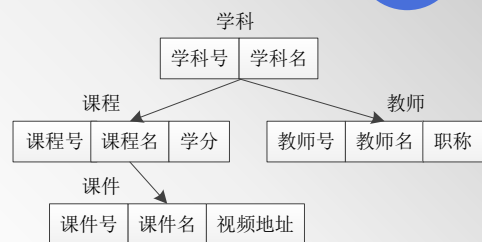
数据模型由数据结构、数据操作和数据完整性约束三部分组成，是严格定义的一组概念的集合，这些概念精确地描述了模型的特性。



数据库领域中应用的数据模型

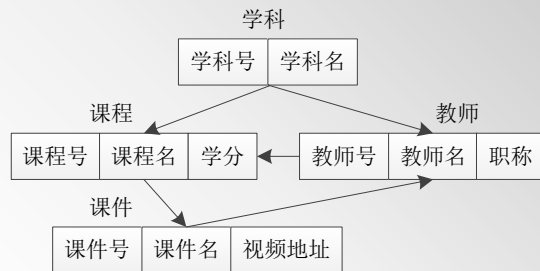
层次模型 (树)

1



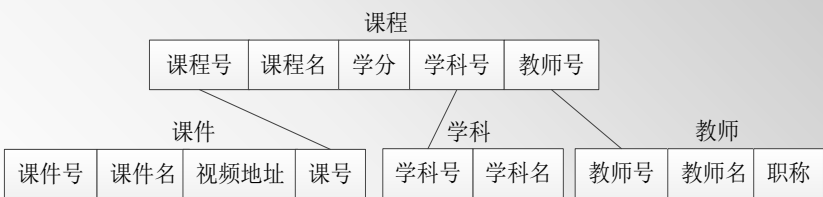
网状模型 (图)

2



关系模型 (表)

3



数据模型

数据库技术发展历程

20世纪60~70年代

1963年GE开发网状数据库IDS
1969年 IBM开发层次数据库IMS

第一代
层次模型
网状模型

第二代
关系模型

20世纪70~80年代

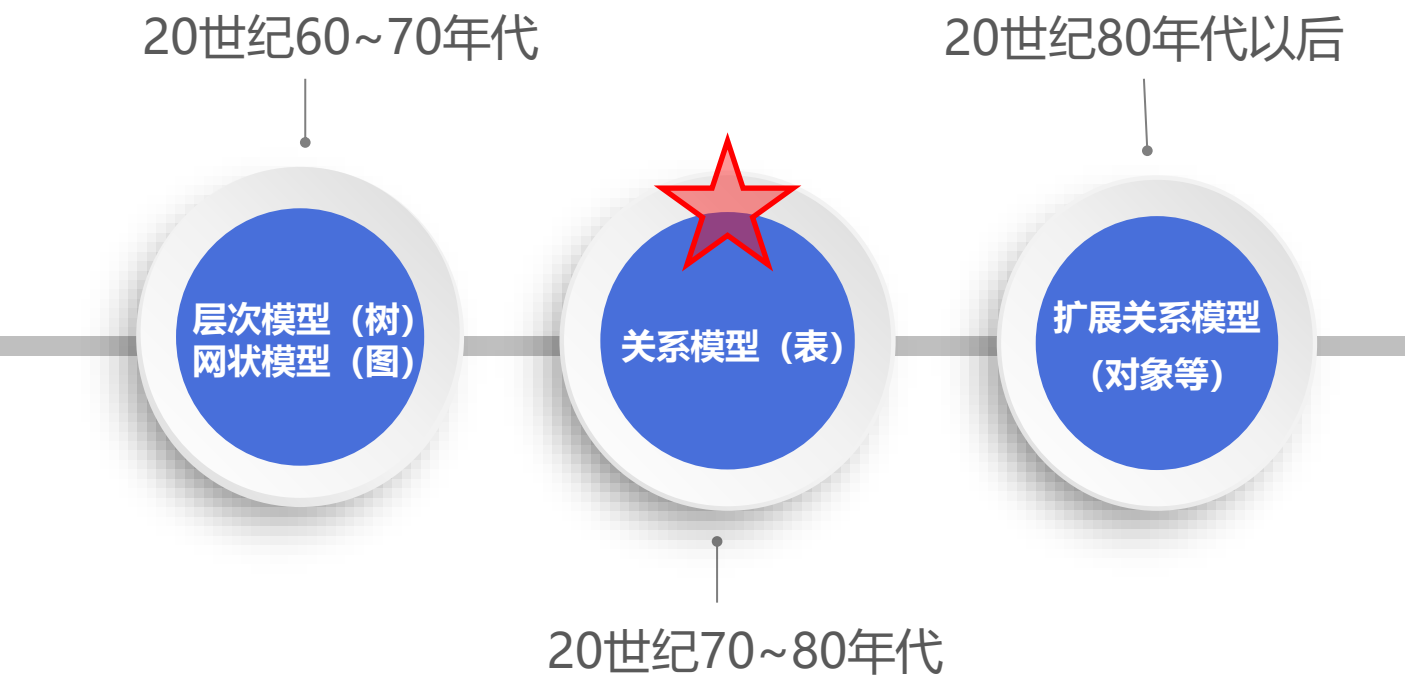
1976年IBM开发关系数据库System R
1982年SQL/DS, 1985年DB2
1979年ORACLE推出Oracle
80年代后走向成熟: 如Oracle、Sybase、
Informix、SQL Server、MySQL等

20世纪80年代以后

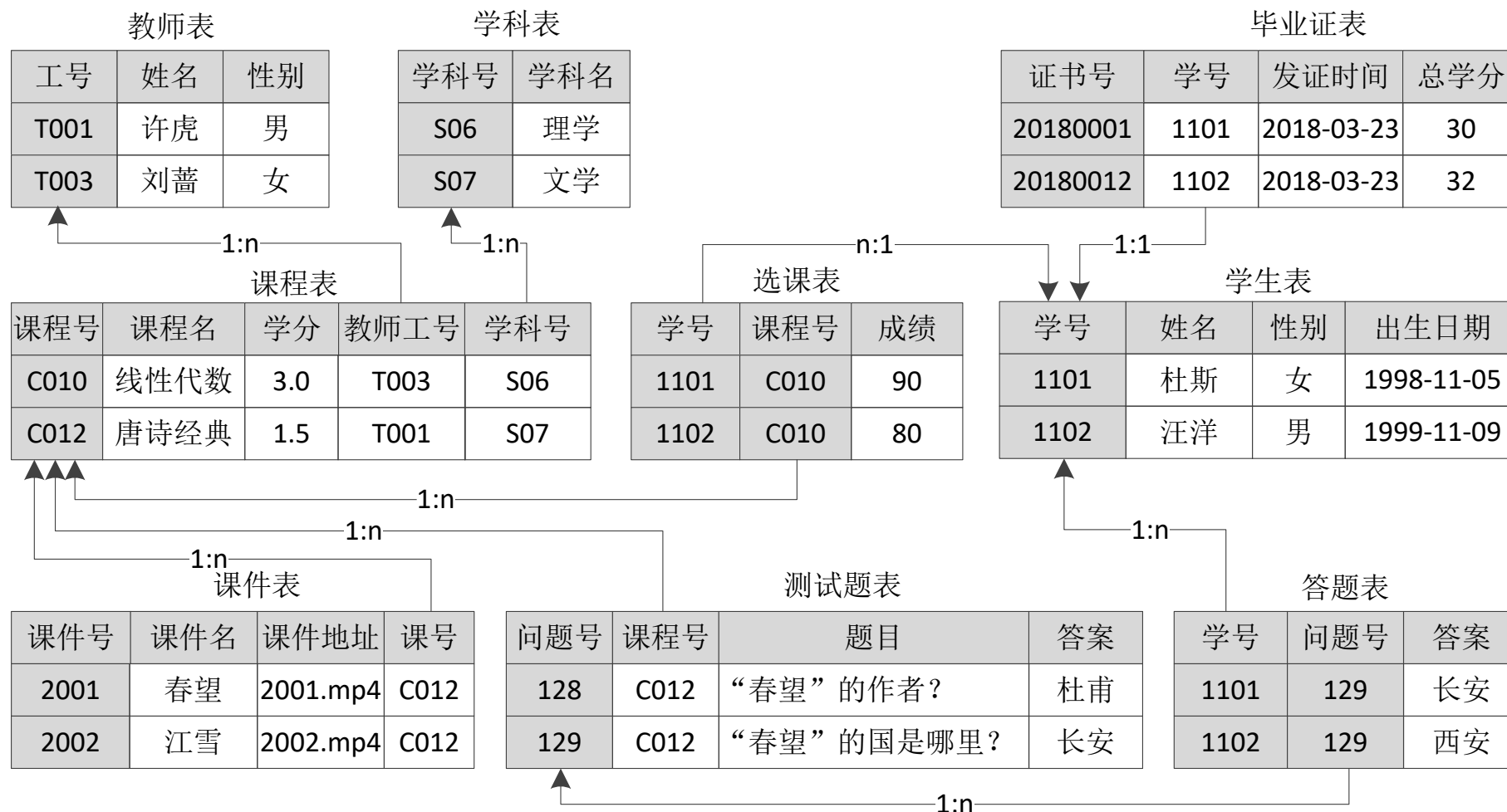
多元化的数据库应用要求 (复杂数据、高性能、
移动、快捷开发等), 发展出各种特色数据库:
面向对象数据库、多媒体数据库、分布式数据
库、移动数据库、NoSQL数据库等

第三代
扩展关系模型

数据模型



“e学习”系统的关系模型



备注：图中灰底字段为表的主关键字

课堂小问答

根据E-R图问答下列问题：

Q1：图中的符号中，哪些符号分别表示实体、属性、关键字、联系？

A1：矩形表示实体，图中实体有学生、课程。

椭圆形表示实体的属性。

带下划线的属性是关键字。

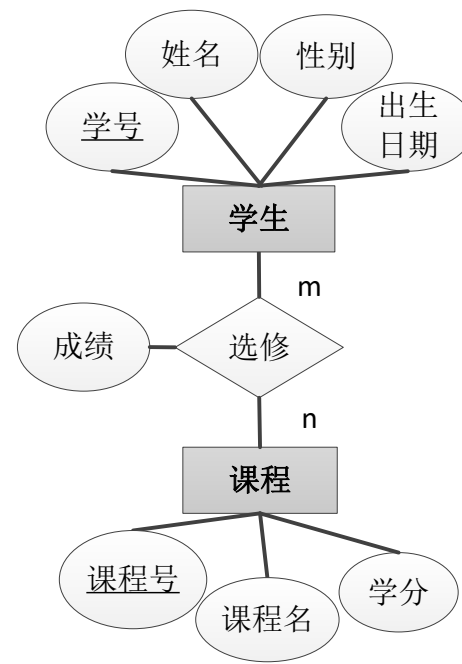
菱形表示联系。

Q2：图中表示的联系是哪种类型的，为什么？

A2：多对多联系

Q3：该E-R图对应上一页的关系模型中的哪些表格？

A3：学生表、课程表、选课表



课堂小故事——关系数据库之父

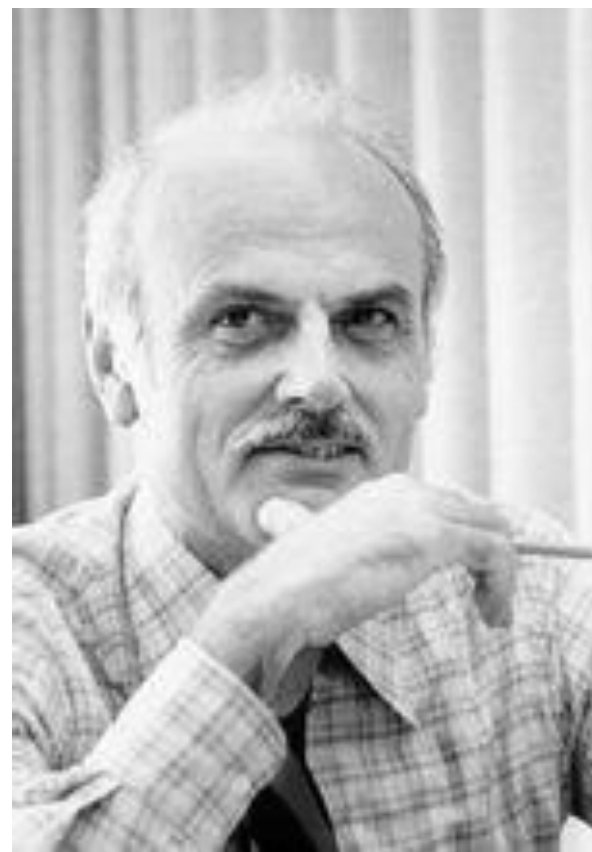
E.F.Codd 是关系数据库的鼻祖。首次提出了数据库系统的关系模型，开创了数据库关系方法和关系数据理论的研究，为数据库技术奠定了理论基础。由于他的杰出贡献，于1981年获得**ACM图灵奖**。图灵奖是计算机界的最高奖项，相当于其他学科的诺贝尔奖。

传奇人生：

1923年出生在英格兰，曾经就读于牛津大学，主修数学和化学专业，第二次世界大战期间曾在皇家空军服役。第二次世界大战后，**Codd**动身前往纽约并成为**IBM**的一名数学程序员。他自觉硬件知识缺乏，于是在60年代初到密歇根大学进修计算机与通信专业(当时他已年近40)，并于1963年获得硕士学位，1965年取得博士学位。

改进数据库：

Codd在1970年发表了具有创新性的技术论文--“**A Relational Model of Data for Large Shared Data Banks**”（大型共享数据库的关系数据模型）。在发表了该论文之后不久，**Codd**又发布了更为详细的指导原则，提出了其指导创建关系数据库的12项原则。



THANK YOU!

2.2 关系模型与关系数据库

2.2.1 关系模型的数据结构 ——关系（表）



关系模型的数据结构——关系（表）

采用表结构描述实体和实体之间的联系。

课程表

课程号	课程名	学分	教师工号	学科号
C010	线性代数	3.0	T003	S06
C012	唐诗经典	1.5	T001	S07

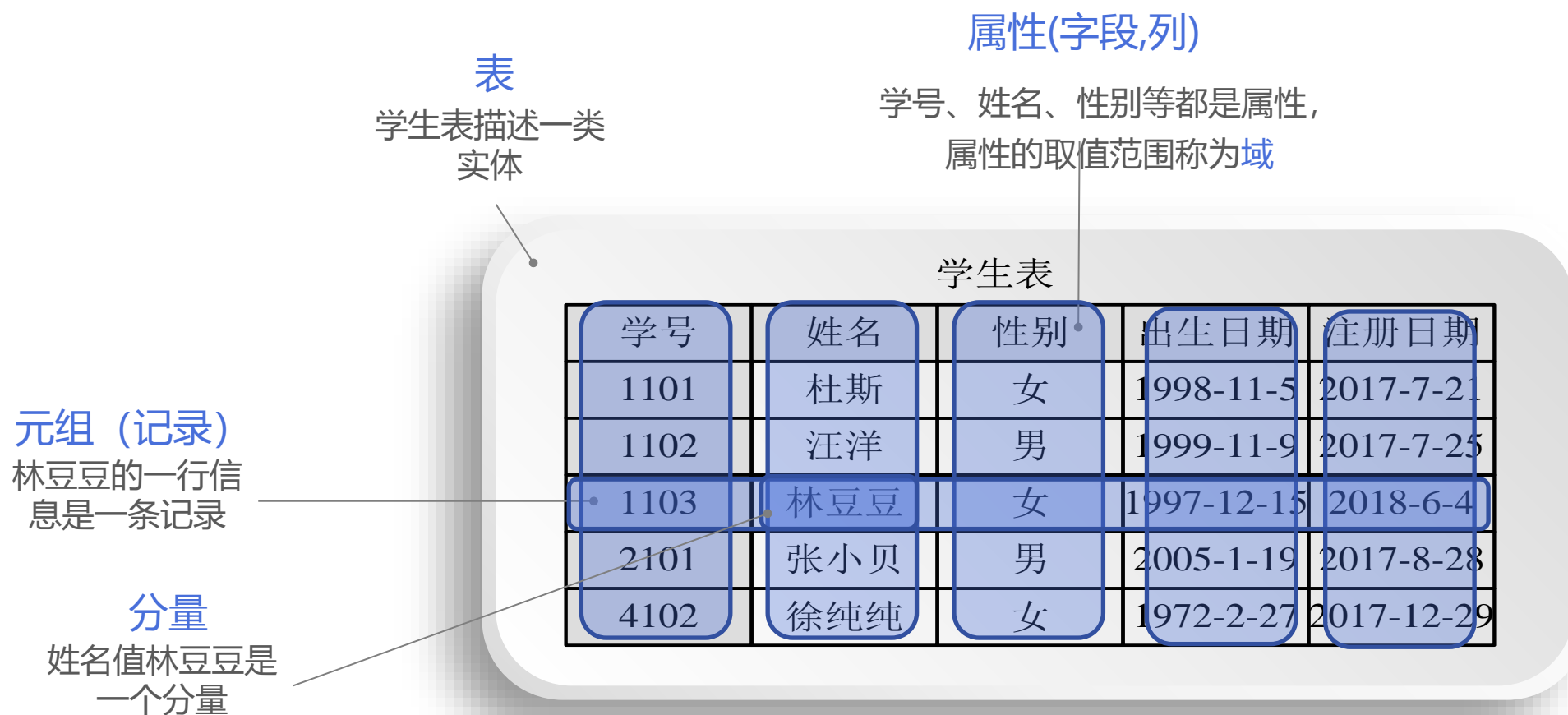
n:1

学科表

学科号	学科名
S06	理学
S07	文学

关系 (表)

一类实体对应一张二维表



关系（表）的一些限制

- 表中的每一个属性必须有唯一的名字，属性在表中的顺序可以任意交换。

学号	姓名	电话	电话
1101	杜斯	13916627891	057167543062
1102	汪洋	13756312276	057154780916
1103	林豆豆	13501702755	051823098588

有不合法的同名属性

通讯录表

学号	手机电话	座机电话	姓名
1101	13916627891	057167543062	杜斯
1102	13756312276	057154780916	汪洋
1103	13501702755	051823098588	林豆豆

- 表中每一个属性的取值是基本数据类型，例如整型、实型、字符型等，数组或结构等不能作为属性的类型。属性必须在所规定的值域范围内取值，不同属性的值域可以相同。

学号	姓名	性别	出生日期	注册日期
1101	杜斯	女	1998-11-05	2017-09-01
1102	汪洋	Male	1999年	2018-05-23
1103	林豆豆	女	1997-12-15	2017-06-16

有不合法的取值

学生表

学号	姓名	性别	出生日期	注册日期
1101	杜斯	女	1998-11-05	2017-09-01
1102	汪洋	男	1999-11-09	2018-05-23
1103	林豆豆	女	1997-12-15	2017-06-16

- 表中任意两条记录不能完全相同，即不允许有重复的行。记录在表中的顺序可以交换。

学号	姓名	性别	出生日期
1101	杜斯	女	1998-11-05
1102	汪洋	男	1999-11-09
1101	杜斯	女	1998-11-05

有不合法的取值

学生表

学号	姓名	性别	出生日期
1101	杜斯	女	1998-11-05
1107	杜斯	女	1998-11-05
1102	汪洋	男	1999-11-09

主关键字 (Primary Key简称主键)

是保证表中记录具有唯一性的一种机制。

- ✓ 候选关键字：按语义能唯一标识元组的最小属性的集合。
- ✓ 主关键字：从候选关键字指定某一个。主关键字的值不能为空、不能重复。

候选关键字1

学号具有唯一性

候选关键字3

如果不允许重名，那么名字也具有唯一性

候选关键字2

身份证号具有唯一性

学生表

学号	姓名	性别	出生日期	身份证号
1101	杜斯	女	1995-11-5	310410199511053273
1102	汪洋	男	1993-11-9	210103199311092180
1103	林豆豆	女	1994-12-15	510234199412156724

主关键字

举例：主关键字

选课表的主关键字？

主关键字可以是一个属性组

主关键字

学号唯一吗？ 课号唯一吗？ 成绩唯一吗？
(学号, 课号) 具有唯一性

选课表

学号	课号	成绩
1101	101	58
1102	101	80
2101	113	95

想一想，如果需要在选课表中记录补考成绩，如何修改表或主关键字？

- ✓ 将（学号、课程号、成绩）做主关键字？

主关键字？

（学号，课号，成绩）具有唯一性？

选课表

学号	课程号	成绩
1101	C010	58
1101	C010	58?
2101	C113	95

- ✓ 增加一个列“补考成绩”？

主关键字

（学号，课号）具有唯一性

选课表

学号	课程号	成绩	补考成绩
1101	C010	58	58
1101	C012	65	
2101	C113	95	

- ✓ 增加一个列“考试时间”，并且把（学号、课程号、考试时间）定义为主关键字！

主关键字

（学号，课程号，考试时间）具有唯一性

选课表

学号	课程号	成绩	考试时间
1101	C010	58	2016-07-11 08:30:00
1101	C010	58	2016-09-05 13:00:00
2101	C113	95	2016-07-08 10:00:00



小结：关系（表）

关系（表）

学生表描述一类
实体

主关键字

“学号” 唯一区
分一个实体

属性(字段,列)

学号、姓名、性别等都是属性，
属性的取值范围称为域

元组（记录）

林豆豆的一行信
息是一条记录

分量

姓名值林豆豆
是一个分量

学生表

学号	姓名	性别	出生日期	注册日期
1101	杜斯	女	1998-11-5	2017-7-21
1102	汪洋	男	1999-11-9	2017-7-25
1103	林豆豆	女	1997-12-15	2018-6-4
2101	张小贝	男	2005-1-19	2017-8-28
4102	徐纯纯	女	1972-2-27	2017-12-29

课堂小问答

Q1: 观察以下关系中存在哪些问题?

关系定义时设定“姓名”字段不允许为空

学号	姓名	性别	联系方式	联系方式
1001	杜斯	男	13916627891	02158785648
1001	汪洋	男	13756312276	01055869548
1003	林豆豆	Gender	13501702548	02755486925
1004		女	13895478561	02154879563

THANK YOU!

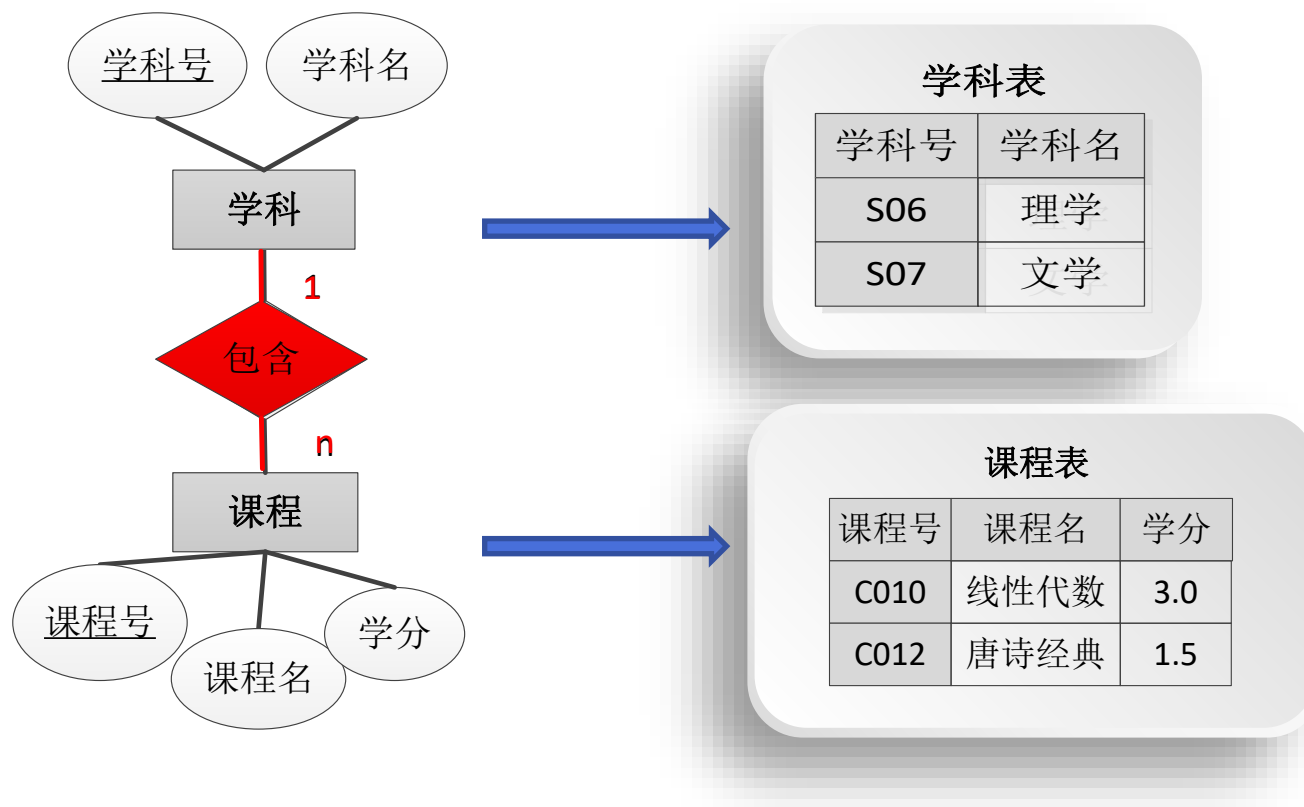
2.2.2 关系模型

——表之间的联系



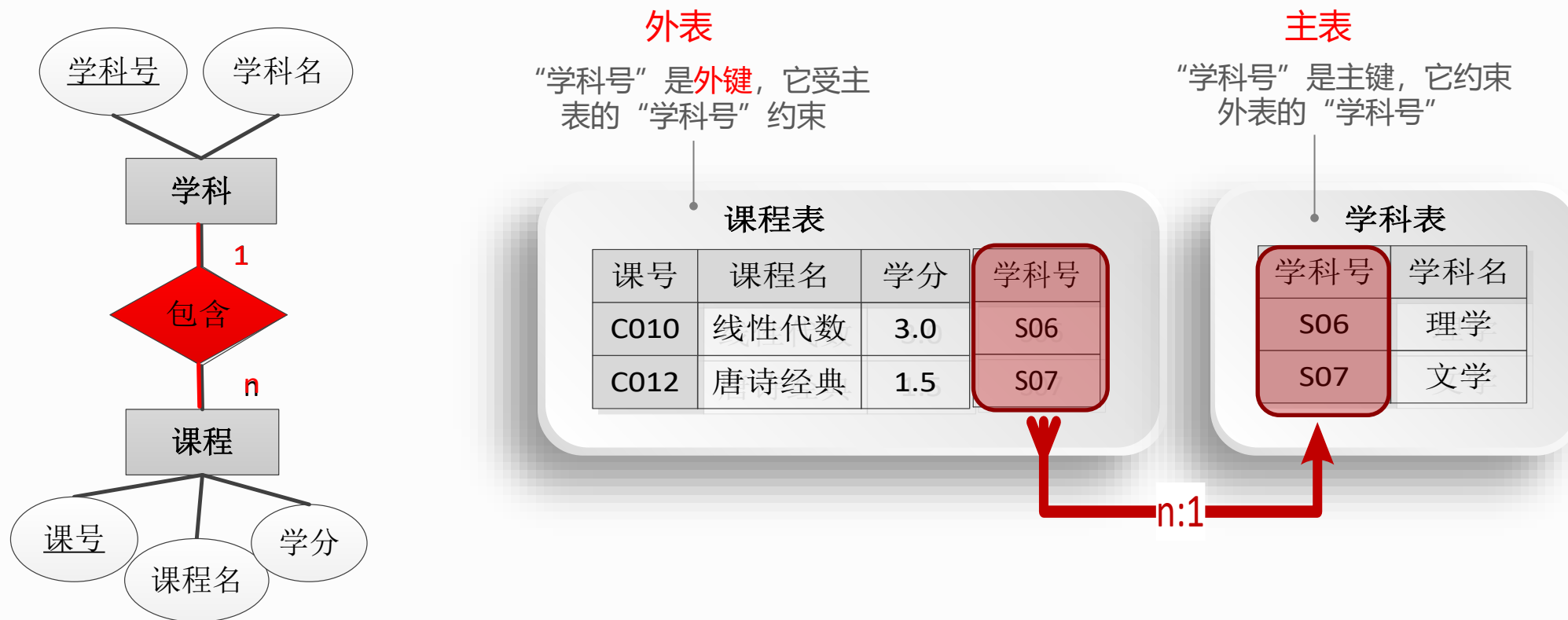
表之间的联系

表之间的联系如何描述？



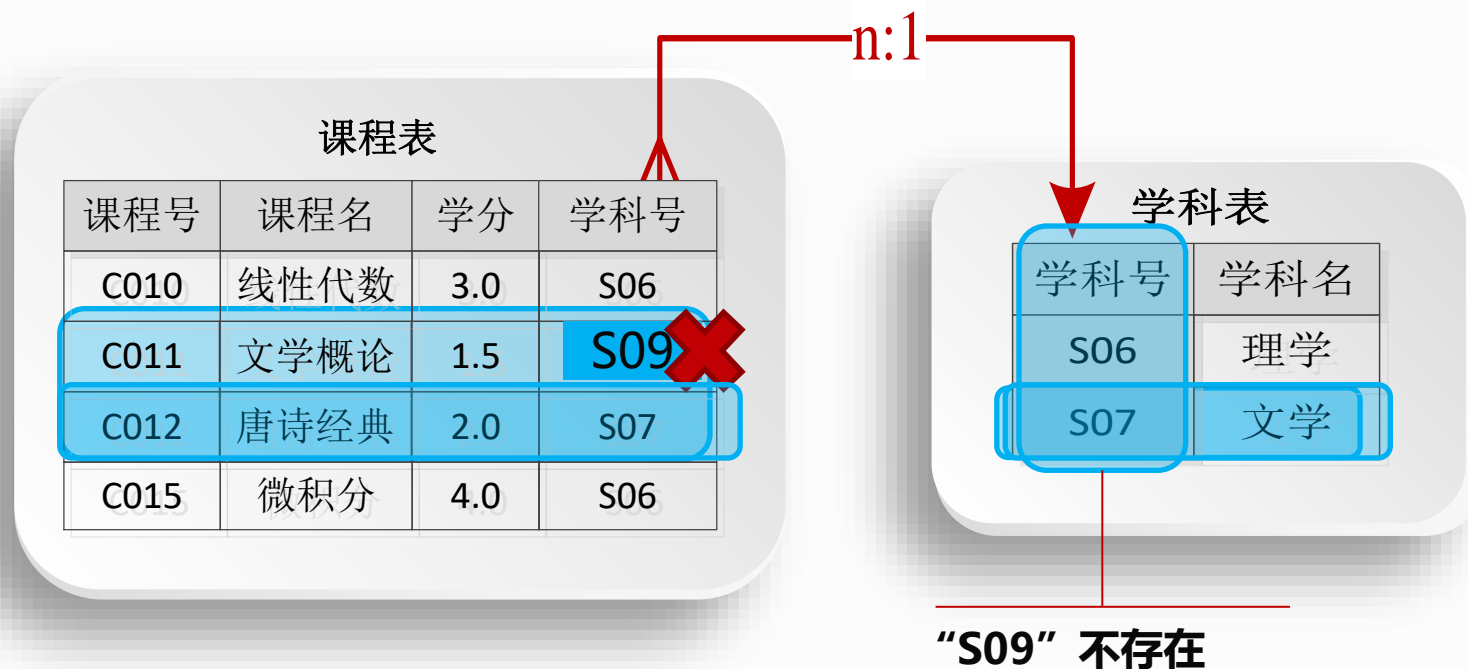
表之间的联系：实现

两个表之间的联系通过它们的**公共属性**（名称可以不同，但数据类型、语义一定相同）来实现，且该公共属性至少在一个表中是主键。



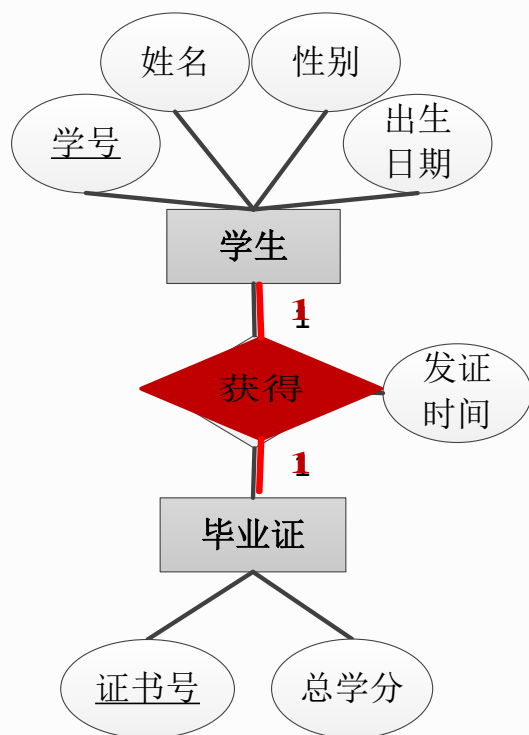
表之间的联系：作用

- ✓ 支持多表联合提取相关信息
- ✓ 表之间互相约束，维护数据的完整性和一致性
- ✓ 减少数据存储冗余



举例：表之间的联系

学生表和毕业证表之间的一对一联系



主表
“学号” 是主键，约束外表的 “学号”

学生表

学号	姓名	性别	出生日期
1101	杜斯	女	1998-11-05
1102	汪洋	男	1999-11-09

毕业证表

证书号	学号	发证时间	总学分
20180001	1101	2018-03-23	30
20180012	1102	2018-03-23	32

“学号” 是**外键**，建立唯一索引，受主表的 “学号” 约束
外表

主表
“学号” 是主键，约束外表的 “学号”

学生表

学号	姓名	性别	出生日期
1101	杜斯	女	1998-11-05
1102	汪洋	男	1999-11-09

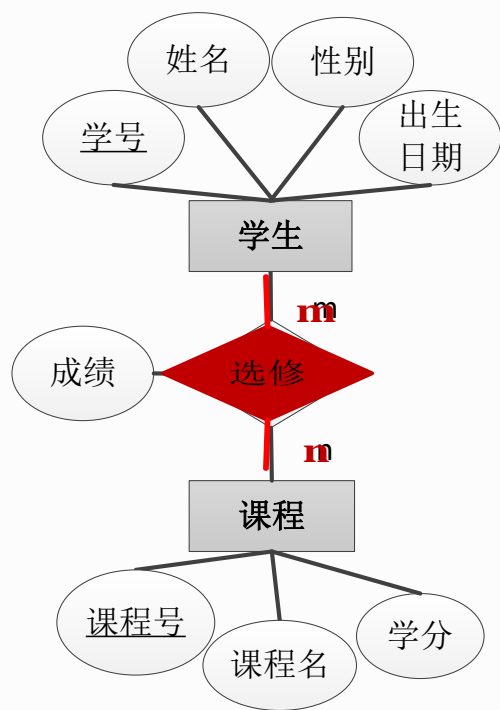
毕业证表

证书号	学号	发证时间	总学分
20180001	1101	2018-03-23	30
20180012	1102	2018-03-23	32

“学号” 即是**主键**也是**外键**，受主表的 “学号” 约束
外表

举例：表之间的联系

学生表和课程表之间的多对多联系



学生表			
学号	姓名	性别	出生日期
1101	杜斯	女	1998-11-05
1102	汪洋	男	1999-11-09
1103	林豆豆	女	1997-12-15

选课表		
学号	课程号	成绩
1101	C010	90
1102	C010	80

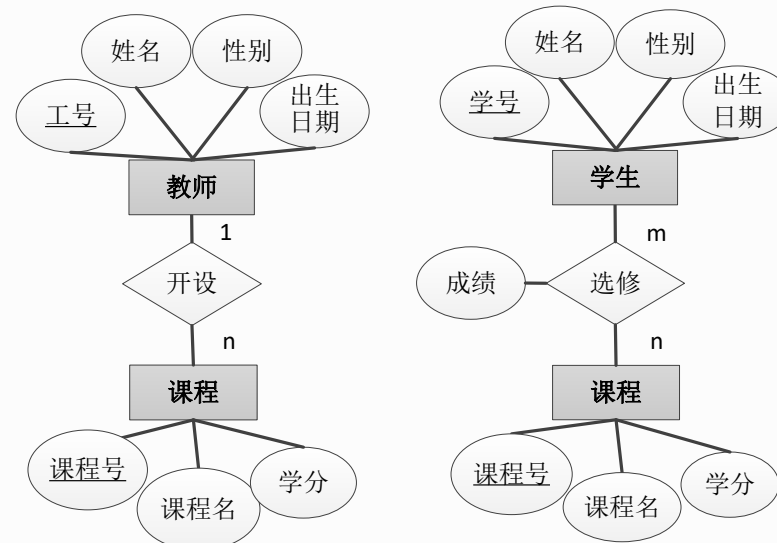
课程表			
课程号	课程名	学分	学科号
C010	线性代数	3.0	S06
C012	唐诗经典	1.5	S07

1:n

1:n

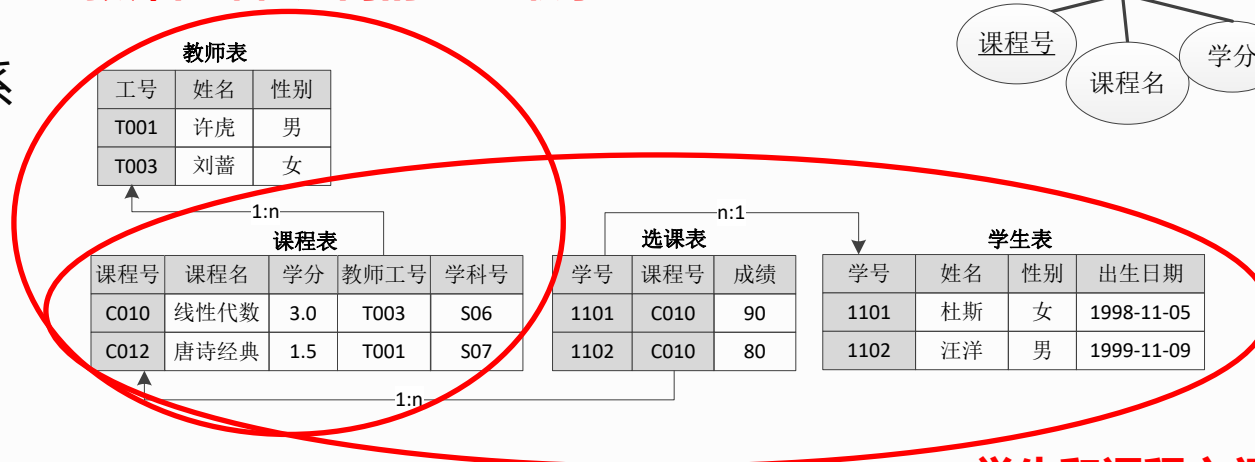
举例：表之间的联系

学生表和教师表之间的联系？



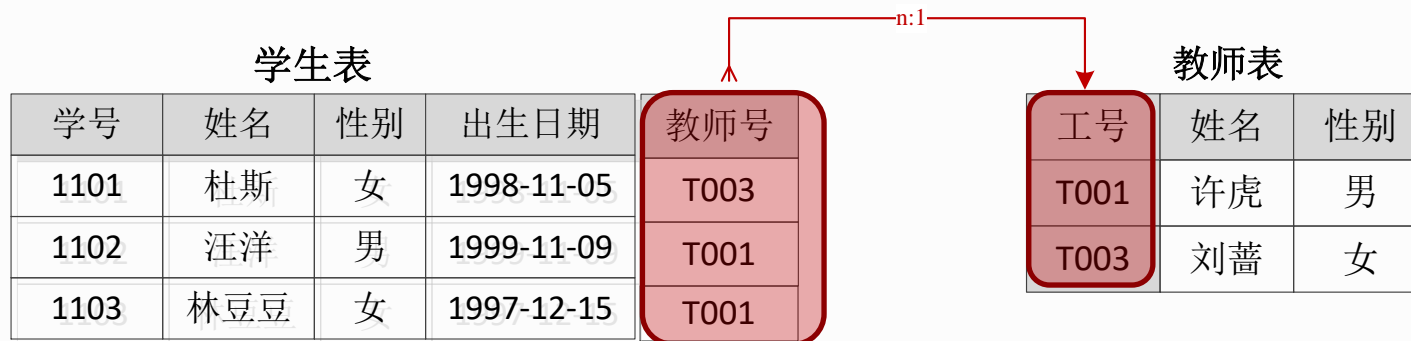
教师 and 课程之间的1: n联系

✓ 教师 and 学生的任课联系



✓ 教师 and 学生的指导联系

学生 and 课程之间的m: n联系



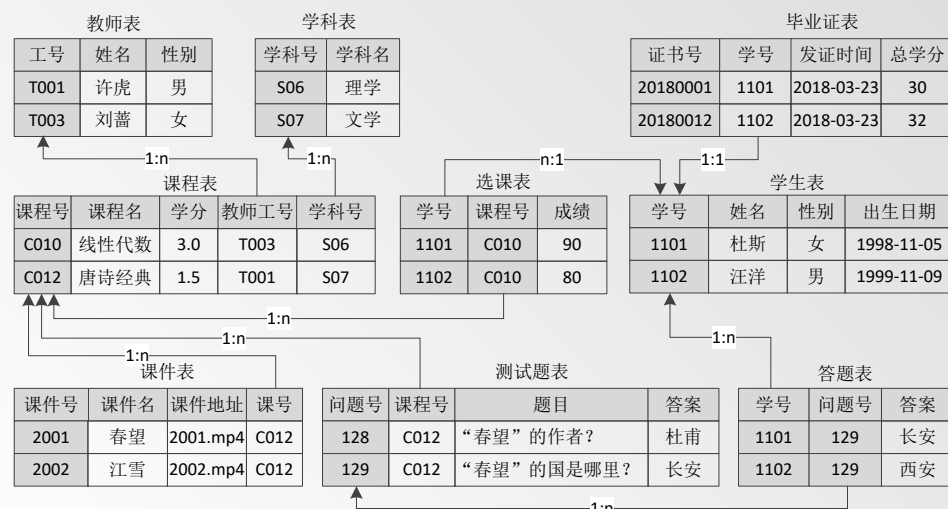
关系模式

关系模式是对关系的描述，包括关系名、属性名以及属性值域，还常标出主键和外键。
关系模式可简记为：关系名(属性名1, 属性名2, ..., 属性名n)

数据库关系模式

- ✓ 学科 (学科号, 学科名)
- ✓ 教师 (工号, 姓名, 性别)
- ✓ 课程 (课程号, 课程名, 学分, 教师工号*, 学科号*)
- ✓ 课件 (课件号, 课件名, 课件地址, 课程号*)
- ✓ 测试题 (问题号, 课号*, 题目, 答案)
- ✓ 学生 (学号, 姓名, 性别, 出生日期)
- ✓ 选课 (学号*, 课程号*, 成绩)
- ✓ 答题 (学号*, 问题号*, 答案)
- ✓ 毕业证 (证书号, 学号*, 发证时间)

“e学习”系统数据库关系图



备注：图中灰底字段为表的主关键字

备注：下划线表示主键字段，加“*”表示外键字段

关系模式

在数据库物理设计阶段，还可以用表格来直观地详细描述关系模式。包括关系名、属性名，以及属性值域的类型和约束、关系模式的主键和外键。

选课（学号*, 课程号*, 成绩）

选课表的关系模式

属性名称	属性说明	类型定义	属性限定	关系（外键）
学号	学生代号	Char(4)	Primary Key	学生表：学号
课程号	课程代号	Char(3)	Primary Key	课程表：课号
成绩	课程成绩	float(5,1)		



小结：表之间的联系

外表

“学科号”是**外键**，它受主表的“学科号”约束

课程表			
课程号	课程名	学分	学科号
C010	线性代数	3.0	S06
C012	唐诗经典	1.5	S07

主表

“学科号”是**主键**，它约束外表的“学科号”

学科表	
学科号	学科名
S06	理学
S07	文学

n:1

表之间建立关系注意：

- 关系是两两之间的
- 关系中有主表、外表角色之分
- 不要随便“拉关系”

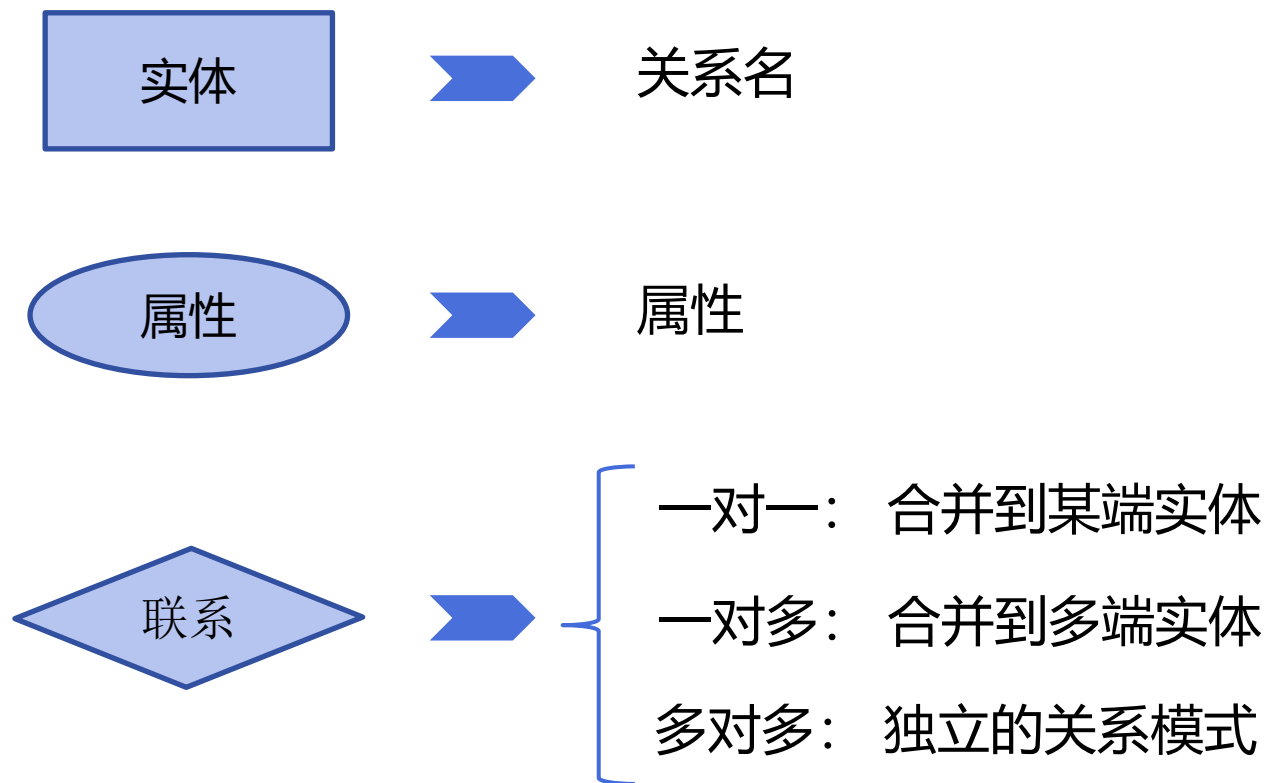
THANK YOU!

2.2.3 关系模型

——E-R图向关系模式转化

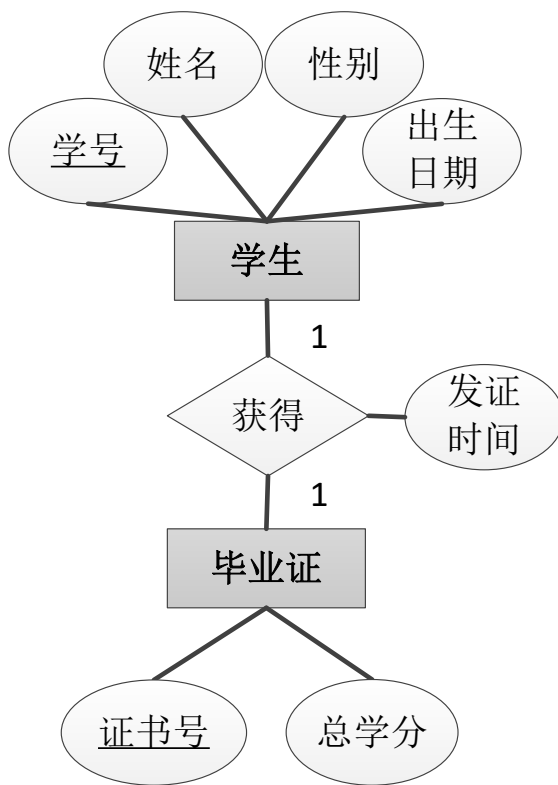


E-R模型转换关系模式



一对一联系的转换规则

一对一联系：将联系的属性和任意“1端”的主键一起合并到另一端的模式模式中。



方案1

学生（学号，姓名，性别，出生日期，**证书号***，**发证时间**）

毕业证（证书号，总学分）

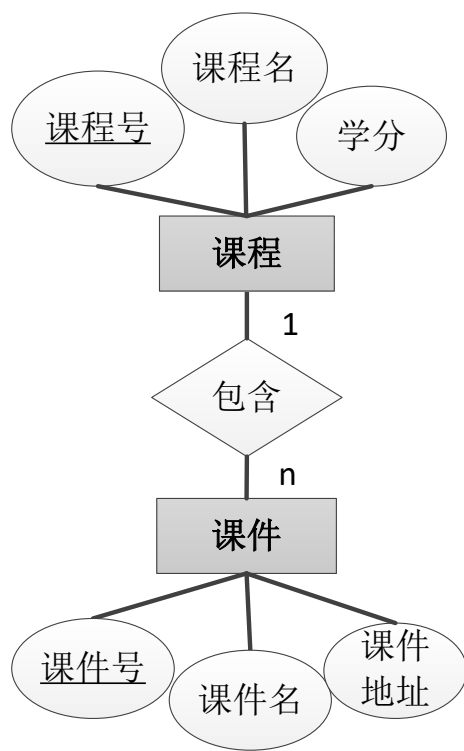
方案2

学生（学号，姓名，性别，出生日期）

毕业证（证书号，总学分，**学号***，**发证时间**）

一对多联系的转换规则

一对多联系：将联系的属性和“1端”的主键合并到“多端”的关系模式中。

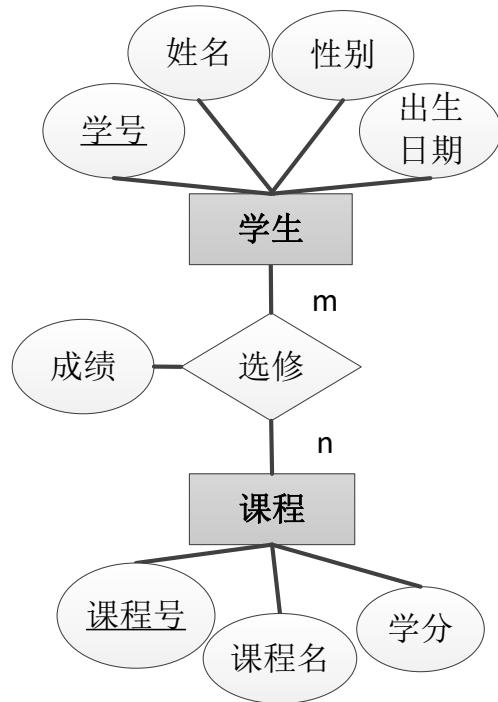


课程（课程号，课程名，学分）

课件（课件号，课件名，课件地址，**课程号***）

多对多联系的转换规则

多对多联系：将联系的属性及两端的主键合并到一个新的关系模式中。



学生（学号，姓名，性别，出生日期）

课程（课程号，课程名，学分）

选课（学号*, 课程号*, 成绩）

其他转换规则

- ✓ 三个或三个以上实体间的一个多元联系可以转化为一个关系模式。与该多元联系相连的各个实体的主键以及联系本身的属性均转换为此关系模式的属性，而此关系模式的主键包含各个实体的主键。
- ✓ 具有相同主键的关系模式可以合并，但是否合并，取决于实际设计需求。

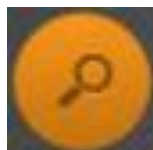
2.2.4 关系模型

——关系（表）的数据操作



关系模型的数据操作

以关系代数和关系演算为理论基础，支持对关系的数据操作。



查询



插入



修改



删除

数据操作

关系演算

集合的单目或二目运算，结果仍然是一个集合。



数据操作

在关系演算支持下，对关系进行数据操作，结果仍然是表。

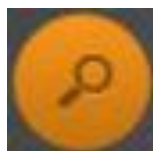
- ✓ 传统集合运算：
并、交、差、广义笛卡儿积等
对应了数据表针对行的操作
- ✓ 专门定义运算：
选择、投影、连接等
对应了数据表针对列的操作

- ✓ 插入：向表中插入新记录
- ✓ 删除：从表中删除满足条件的记录
- ✓ 修改：修改表中某些记录属性的值
- ✓ 查询：从若干个表中提取满足条件的数据、生成计算列或汇总数据。

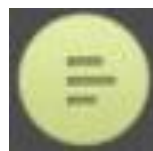


小结：关系模型的数据操作

以关系运算为理论基础，支持对关系表的数据操作。



查询



插入



修改



删除

SQL语言实现了这些操作。

THANK YOU!

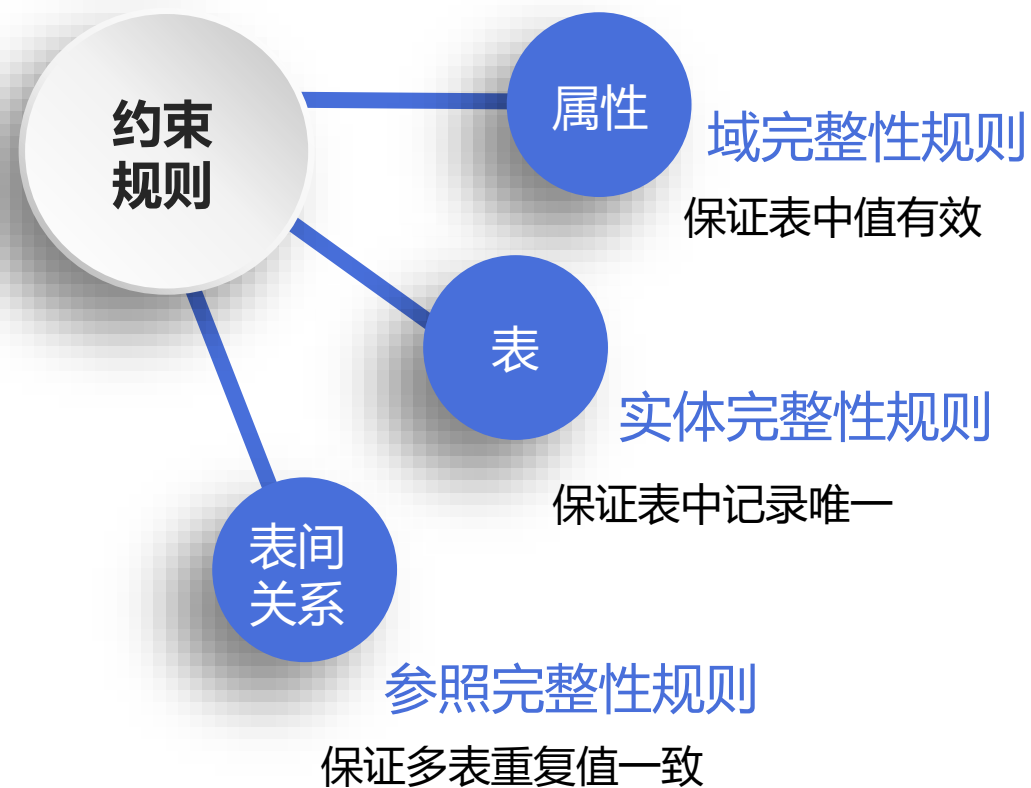
2.2.5 关系模型

——关系（表）的数据完整性约束



关系模型的数据完整性约束

通过约束规则保证数据状态转换后的数据完整性和一致性。



域完整性

属性的值必须是属性值域中的值：保证表中值有效

学生表

学号	姓名	性别	出生日期	班号
1101	杜斯	女	1998-11-5	11
1102	汪洋	男	1999-11-9	11
1103	林豆豆	女	32.5	11
2101	张小贝		1995-1-19	21
4102	徐纯纯	女	1992-2-27	41

学生表关系模式

属性名	类型定义	属性限定
学号	int, 自增	Primary Key
姓名	varchar(16)	Not Null
性别	char(1)	Not Null, 默认值 “男”
出生日期	datetime	>=1990-01-01

实体完整性

主关键字不能有空值且值不能相同，它保证表中记录唯一。

学生表

学号	姓名	性别	出生日期	班号
1101	杜斯	女	1998-11-5	11
1102	汪洋	男	1999-11-9	11
1101	林豆豆	女	1997-12-15	11
2101	张小贝	男	1995-1-19	21
4102	徐纯纯	女	1992-2-27	41
	王秋水	女	1996-3-4	21

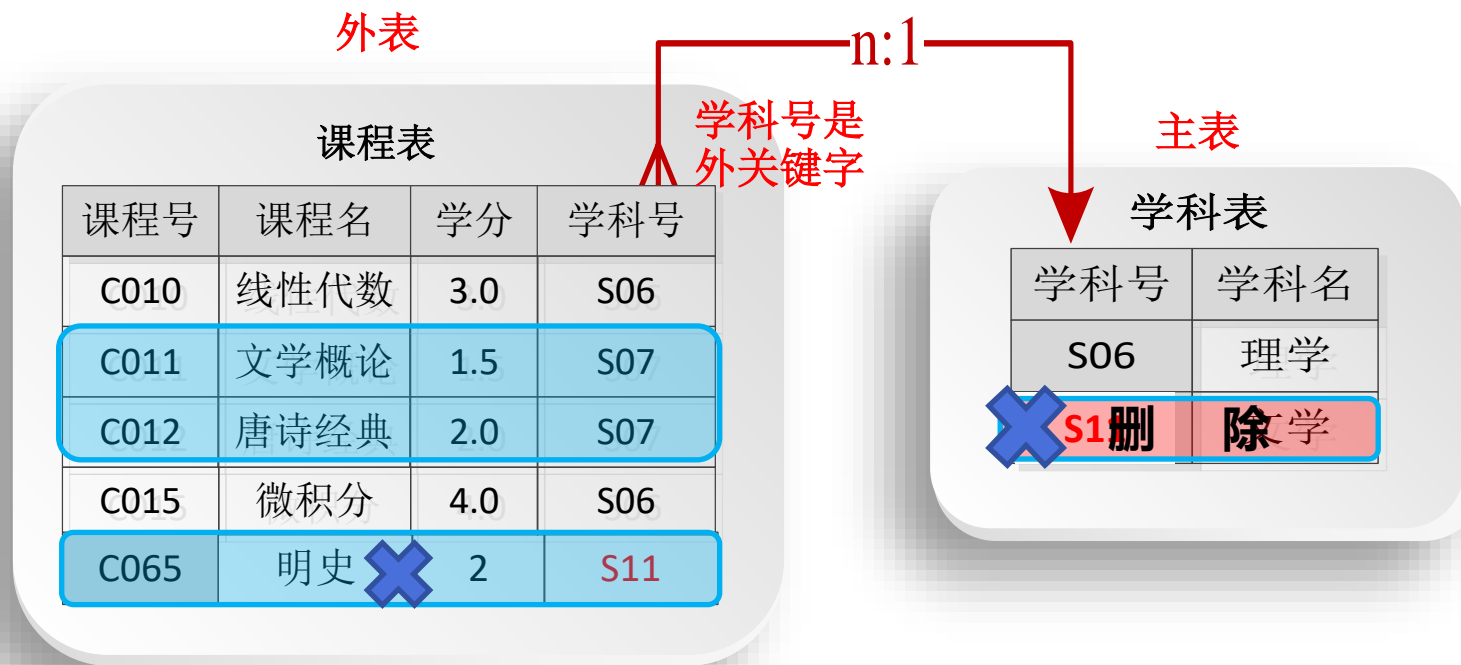
参照完整性

表的外关键字的取值必须是其主表主关键字的存在值或空值。它保证多表重复值一致

插入约束：当在外表中插入记录时，保证外关键字值在主表中出现过。

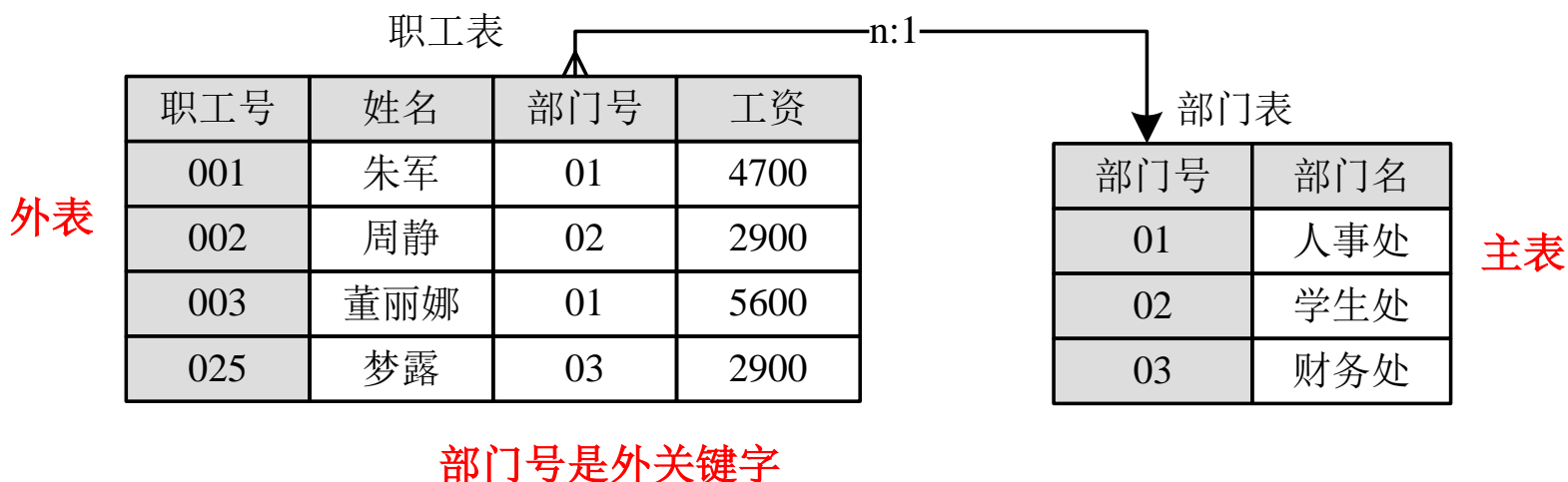
删除约束：当要从主表中删除一条记录时，必须考虑外表数据的完整性。

更新约束：修改主表中的主关键字值，必须考虑外表数据的一致性。



思考题

职工表和部门表通过“部门号”建立关系并实施了参照完整性约束。职工表主键是职工号，外键是部门号；部门表主键是部门号。



在如下数据库操作中，不能成功执行的是_____。

- A. 从职工表中删除行（'025'，'梦露'，'03'，2900）
- B. 将行（'005'，'乔兴'，'04'，750）插入到职工表中
- C. 将职工号为“001”的工资改为5700
- D. 将职工号为“003”的部门号改为'03'

答案：B

不满足参照完整性的插入约束，
部门号‘04’在部门表中不存在



小结：关系模型的数据完整性约束

通过约束规则保证数据状态转换后的数据完整性和一致性。



THANK YOU!

2.3 关系数据库的基础理论



2.3.1 关系模型规范化理论

关系模式规范化理论包括一系列**范式**（Normal Forms, 简记为NF），即规定了关系模式需要满足的一些约束条件。

满足高级范式首先要满足低级范式。一般设计要求是满足第三范式。

关系模式的规范化就是将一个低一级范式的关系模式，通过模式分解转换为高一级范式的过程。



关系模型规范化的目标？

- ✓ 减少数据冗余
- ✓ 避免操作异常
- ✓ 避免数据不一致

三个范式

1 第一范式

关系模式中的每一分量都不可分

2 第二范式

关系模式中的所有非主键属性都完全依赖于主关键字。（不允许部分依赖）

3 第三范式

关系模式中的任意非主键属性都不传递依赖于主关键字。（不存在传递依赖）

举例：关系模型规范化——1NF

生源表的正确描述。

主关键字？

姓名	地区	
	省	市
杜斯	浙江	杭州
汪洋		
林豆豆	江苏	连云港

非1NF

规范化为
1NF

设置主关键字 列属性分解

姓名	省	市
杜斯	浙江	杭州
汪洋	浙江	杭州
林豆豆	江苏	连云港

设置主关键字 列属性合并

姓名	地区
杜斯	浙江杭州
汪洋	浙江杭州
林豆豆	江苏连云港

第一范式：每一分量都不可分

关系模型规范化举例——2NF

主关键字是 (学号, 课程名)

R

学号	姓名	课程名	学分	类别	成绩
1001	陈佳迪	高等数学	4	理学	90
1001	陈佳迪	大学物理	3	理学	80
1001	陈佳迪	哲学概论	2	哲学	85
1002	张君悦	英语	4	外语	78
1002	张君悦	哲学概论	2	哲学	86
1003	王霄	高等数学	4	理学	78

第二范式：所有非主键属性都完全依赖于主关键字

非主属性是指不包含在任何候选码中的属性。

转化为2NF的投影分解法

R (学号, 姓名, 课程名, 学分, 类别, 成绩)



用主键属性集的每一个子集作主键构成一个关系模式

R1 (学号,)

R2 (课程名,)

R3 (学号, 课程名,)



将依赖于最小主键的属性放在相应的关系模式中

R1 (学号, 姓名)

R2 (课程名, 学分, 类别)

R3 (学号, 课程名, 成绩)



去掉只由主键的子集构成的关系模式

R1 (学号, 姓名)

R2 (课程名, 学分, 类别)

R3 (学号, 课程名, 成绩)

关系模型规范化举例——2NF

第二范式：所有非主键属性都完全依赖于主关键字

R

学号	姓名	课程名	学分	类别	成绩
1001	陈佳迪	高等数学	4	理学	90
1001	陈佳迪	大学物理	3	理学	80
1001	陈佳迪	哲学概论	2	哲学	85
1002	张君悦	英语	4	外语	78
1002	张君悦	哲学概论	2	哲学	86
1003	王霄	高等数学	4	理学	78

规范化为
2NF

R1

学号	姓名
1001	陈佳迪
1002	张君悦
1003	王霄

R2

课程名	学分	类别
高等数学	4	理学
大学物理	3	理学
英语	4	外语
哲学概论	2	哲学

R3

学号	课程名	成绩
1001	高等数学	90
1001	大学物理	80
1001	哲学概论	85
1002	英语	78
1002	哲学概论	86
1003	高等数学	78

关系模型规范化举例——3NF

主关键字是 (课程名)

R

课程名	学分	教师	职称
高等数学	4	张晓芸	讲师
大学物理	3	马建军	教授
英语	4	李明	副教授
数学分析	3	张晓芸	讲师

第三范式：任意非主键属性
都不传递依赖于主关键字

转化为3NF的投影分解法

R2 (课程名, 学分, 教师, 职称)



删除不依赖于主键的属性, 以每个被依赖属性为主键新建一个关系模式

R21 (课程名, 学分, 教师)

R22 (教师,



在新关系模式中放入依赖它的属性, 并设被依赖者为主键

R21 (课程名, 学分, 教师)

R22 (教师, 职称)

关系模型规范化举例——3NF

R1

课程名	学分	教师
高等数学	4	张晓芸
大学物理	3	马建军
英语	4	李明
数学分析	3	张晓芸

R2

教师	职称
张晓芸	讲师
马建军	教授
李明	副教授

规范化为
3NF

R

课程名	学分	教师	职称
高等数学	4	张晓芸	讲师
大学物理	3	马建军	教授
英语	4	李明	副教授
数学分析	3	张晓芸	讲师

第三范式：任意非主键属性
都不传递依赖于主关键字

举例：“e学习”系统数据库逻辑结构

- ✓ 学科（学科号，学科名）
- ✓ 教师（工号，姓名，性别）
- ✓ 课程（课程号，课程名，学分，教师工号*，学科号*）
- ✓ 课件（课件号，课件名，课件地址，课程号*）
- ✓ 测试题（问题号，课号*，题目，答案）
- ✓ 学生（学号，姓名，性别，出生日期）
- ✓ 选课（学号*, 课程号*, 成绩）
- ✓ 答题（学号*, 问题号*, 答案）
- ✓ 毕业证（证书号，学号*，发证时间）

满足
3NF

2.3.2 关系模型的运算理论简介



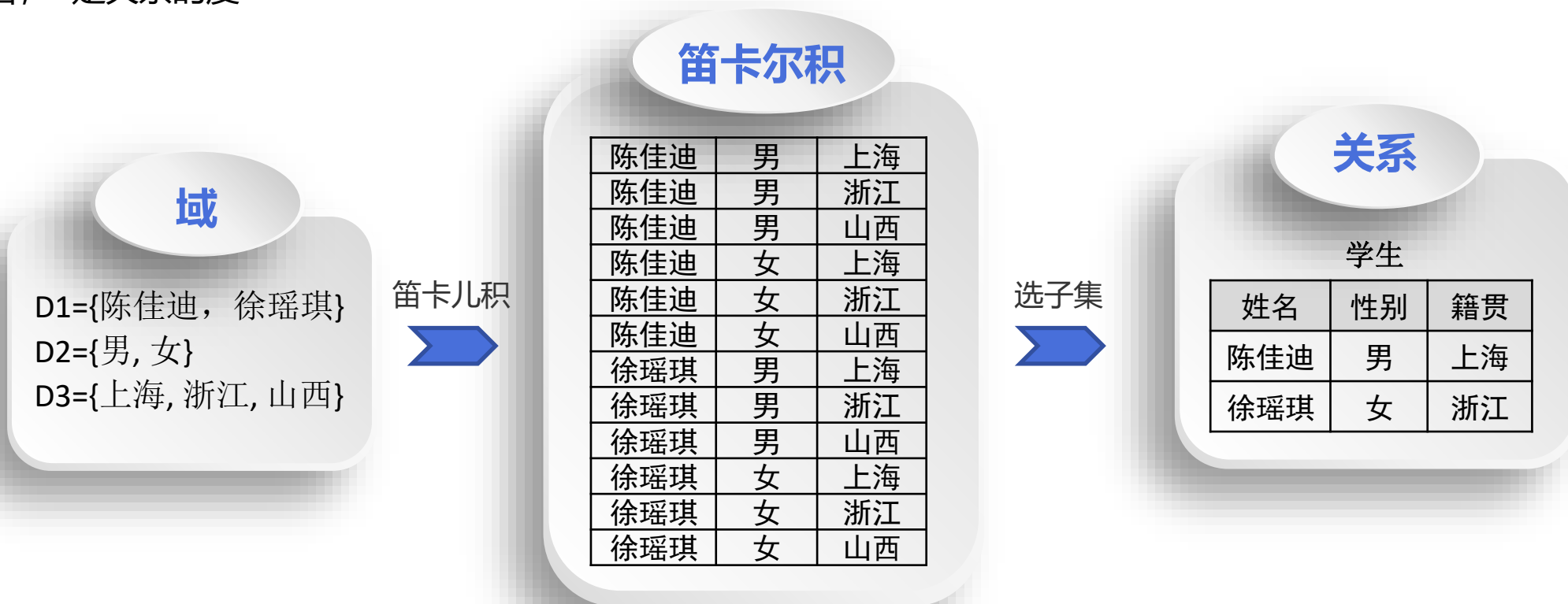
关系定义

域：一组同类型数据值的集合

笛卡尔积：给定一组域D1, D2, ..., Dn, 其笛卡尔积为:

$$D1 \times D2 \times \dots \times Dn = \{(d1, d2, \dots, dn) | di \in Di, i = 1, 2, \dots, n\}$$

关系：笛卡尔积的任意一个子集称为一个n元关系, 通常用R(D1, D2, ..., Dn)表示, 这里R为关系名, n是关系的度





关系运算

- 传统集合运算

并、交、差、广义笛卡尔积等
对应了数据表针对**行**的操作

- 专门定义运算

选择、投影、连接等对应了数
据表针对**列**的操作

传统集合运算

如果两个关系具有相同或相容的关系模式，即可进行并、交、差运算

学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

学号	名字	性别
1101	杜斯	女
1103	林豆豆	女
4102	徐纯纯	女

合唱团表S

- ✓ 参加运算的集合都以元组（记录）为元素
- ✓ 其运算是从“水平”角度进行，即行运算。
- ✓ 都是二元运算，由两个关系产生一个新的关系

传统的集合运算：并

$R \cup S = \{t \mid t \in R \vee t \in S\}$, t 是元组变量

学号	姓名	性别
1101	杜斯	女
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男
4102	徐纯纯	女



学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$R \cup S$

学号	名字	性别
1101	杜斯	女
1103	林豆豆	女
4102	徐纯纯	女

合唱团表S

传统的集合运算：交

$R \cap S = \{t \mid t \in R \wedge t \in S\}$, t 是元组变量

学号	姓名	性别
1103	林豆豆	女



学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$R \cap S$

学号	名字	性别
1101	杜斯	女
1103	林豆豆	女
4102	徐纯纯	女

合唱团表S

传统的集合运算：差

$R - S = \{t \mid t \in R \wedge t \notin S\}$, t 是元组变量

学号	姓名	性别
1102	汪洋	男
2101	张小贝	男



学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$R - S$

学号	名字	性别
1101	杜斯	女
1103	林豆豆	女
4102	徐纯纯	女

合唱团表S

传统的集合运算：广义笛卡尔积

$R \times U = \{t_r t_u \mid t_r \in R \wedge t_u \in U\}, t_r, t_u$ 是元组变量

学号	姓名	性别	学号	课程名	任课教师
1102	汪洋	男	1101	高等数学	黎念青
1102	汪洋	男	2101	武术	刘延琳
1103	林豆豆	女	1101	高等数学	黎念青
1103	林豆豆	女	2101	武术	刘延琳
2101	张小贝	男	1101	高等数学	黎念青
2101	张小贝	男	2101	武术	林燕琳

学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$R \times U$

学号	课程名	任课教师
1101	高等数学	黎念青
2101	武术	刘延琳

课程表U

专门的关系运算：选择

一元运算，它选择满足某些条件的元组。

$\sigma_F(R) = \{t \mid t \in R \wedge F(t) = \text{True}\}$, t 是元组变量

学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$\sigma_{\text{性别}=\text{“男”}}(R)$



学号	姓名	性别
1102	汪洋	男
2101	张小贝	男

文学社男生

- ✓ 结果与原关系模式相同，数据是原数据的子集
- ✓ 它实现了关系的**水平分割**，支持根据条件查询数据记录。

专门的关系运算：投影

一元运算，它在关系中选择某些属性

$$\pi_A(R) = \{t \mid t \in R \wedge t \in A\}, t \text{ 是分量}$$

学号	姓名	性别
1102	汪洋	男
1103	林豆豆	女
2101	张小贝	男

文学社表R

$\pi_{\text{姓名, 性别}}(R)$



姓名	性别
汪洋	男
林豆豆	女
张小贝	男

文学社表R'

- ✓ 结果的关系模式是原关系模式的子集。
- ✓ 它实现了关系的**垂直分割**，支持查询包含部分属性数据记录的集合。

专门的关系运算：连接

二元运算，从两个关系的笛卡尔积中选取满足一定条件的元组



✓ 它实现了针对多表的查询操作



关系运算

- 传统集合运算

并、交、差、广义笛卡尔积等对应了数据表针对行的操作

- 专门定义运算

选择、投影、连接等对应了数据表针对列的操作

THANK YOU!

本章作业

- 1) 完成canvas上的**测验2_数据库基本知识**（允许重做一次，重做重新抽题，取最高分）
- 2) 在自己的电脑上安装Visio或亿图ERD软件
- 3) 完成Canvas上的**实验2_关系数据库基本知识**