



2025年春季学期

数据库系统概论

An Introduction to Database Systems

第六章 关系数据理论

中国科学技术大学

人工智能与数据科学学院

黄振亚, huangzhy@ustc.edu.cn



6.2 规范化

42

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

6.2.8 4NF

6.2.9 规范化小结



6.2.3 范式

43

- ❑ 范式是符合某一种级别的关系模式的集合
- ❑ 关系数据库中的关系必须满足一定的要求。满足不同程度要求的为不同范式
- ❑ 范式的种类：





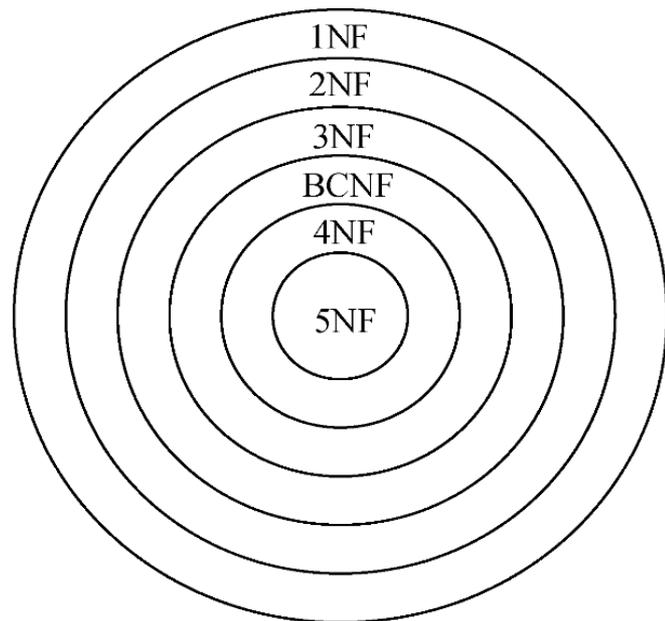
6.2.3 范式

44

- 各种范式之间存在联系：

$$1NF \supset 2NF \supset 3NF \supset BCNF \supset 4NF \supset 5NF$$

- 某一关系模式R为第n范式，可简记为 $R \in nNF$ 。
- 一个低一级范式的关系模式，通过**模式分解**可以转换为若干个高一级范式的关系模式的集合，这种过程就叫**规范化**





6.2 规范化

45

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

6.2.8 4NF

6.2.9 规范化小结



6.2.4 2NF

46

□ 1NF的定义

如果一个关系模式R的所有属性都是**不可分的基本数据项**，
则 $R \in 1NF$

□ 第一范式是对关系模式的最低的要求。**不满足第一范式的数据库模式不能称为关系数据库**

□ 但：满足第一范式的关系模式并不一定是一个好的关系模式



6.2.4 2NF

□ 1NF的例子

如果一个关系模式R的所有属性都是不可分的基本数据项，R是满足最低要求的模式，则 $R \in 1NF$

Sno	School	Mname	Cno	Grade
S1	信息学院	张明	C1	95
S2	信息学院	张明	C1	90
S3	信息学院	张明	C1	88
S4	信息学院	张明	C1	70
S5	信息学院	张明	C1	78
....



2NF—1NF的问题

[例6.4] 关系模式 S-L-C (Sno, School, Sloc, Cno, Grade)

Sloc为学生住处，假设每个学院的学生住在同一个地方

□ 函数依赖包括：

$$(Sno, Cno) \xrightarrow{F} Grade$$

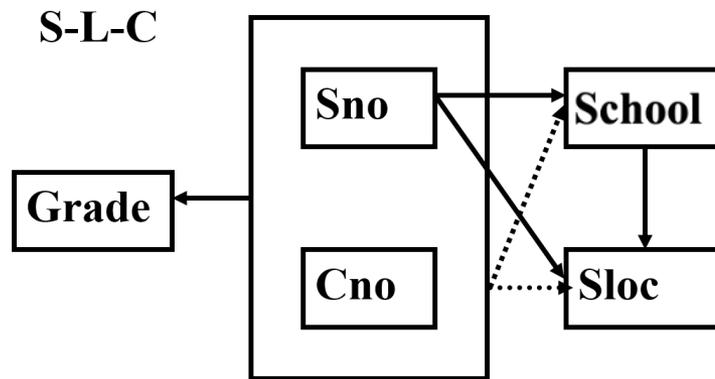
$$Sno \rightarrow School$$

$$(Sno, Cno) \xrightarrow{P} School$$

$$Sno \rightarrow Sloc$$

$$(Sno, Cno) \xrightarrow{P} Sloc$$

$$School \rightarrow Sloc$$



- S-L-C的码为(Sno, Cno)
- S-L-C满足第一范式1NF。
- 非主属性Sdept和Sloc部分函数依赖于码(Sno, Cno)



S-L-C不是一个好的关系模式（续）

关系模式 S-L-C(Sno, School, Sloc, Cno, Grade)

(1) 插入异常

- 插入新学生，未选课（无Cno）的学生无法插入

(2) 删除异常

- 删除只选了一门课的学生的选课信息时，必须把整个元组一起删除，造成学生的School Sloc等信息丢失。

(3) 数据冗余度大

- Sdept, Sloc重复存储

Sno	School	Sloc	Cno	Grade
....

(4) 修改复杂

- 修改学生系别School时需要同时修改Sloc。如果一个学生选的多门课，则School, Sloc被存储了多次。如果该生转系，则需要修改所有相关的School和Sloc，造成修改的复杂化



S-L-C不是一个好的关系模式（续）

50

□ 原因

- 存在一类：如Grade，完全函数依赖于码（Sno, Cno）
- 存在另一类：如School、Sloc，部分函数依赖于码（Sno）

□ 解决方法：投影分解法

S-L-C分解为两个关系模式，以消除这些部分函数依赖

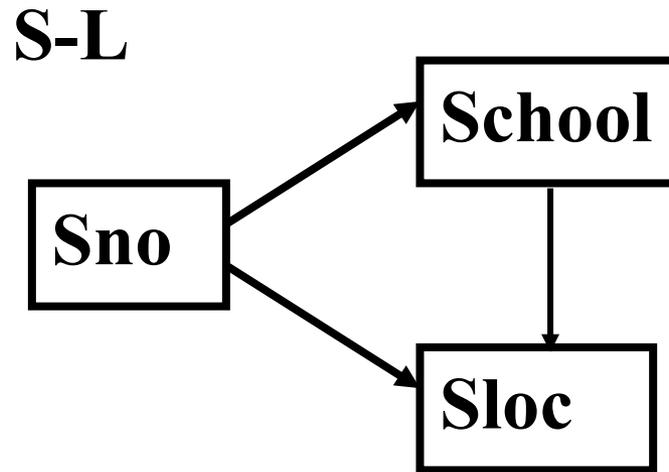
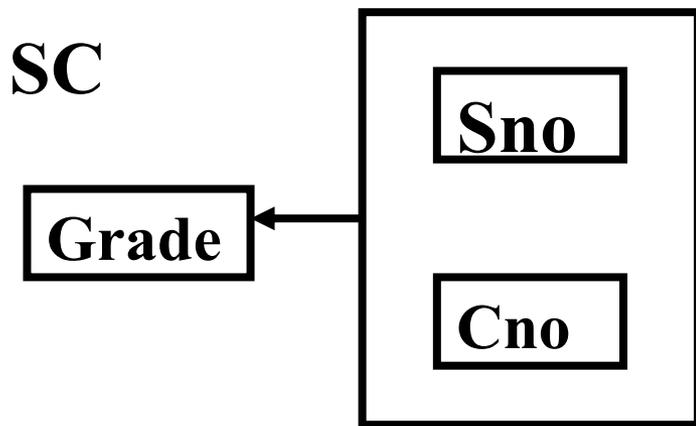
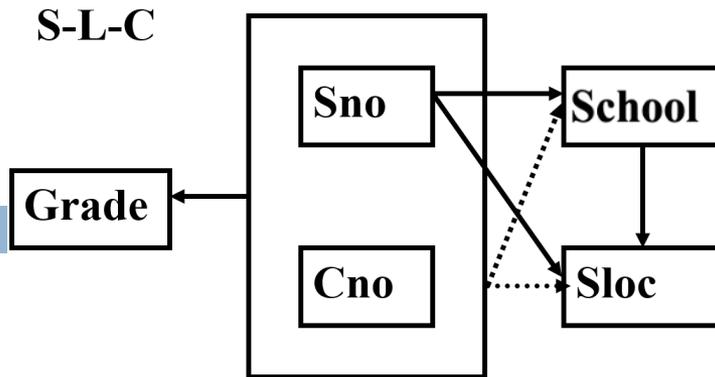
SC (Sno, Cno, Grade)

S-L (Sno, School, Sloc)



2NF (续)

函数依赖图:



- ❖ 关系模式SC的码为 (Sno, Cno)
- ❖ 关系模式S-L的码为Sno
- ❖ 非主属性对码都是完全函数依赖



2NF (续)

52

□ 2NF的定义

定义6.6 若 $R \in 1NF$ ，且每一个非主属性完全函数依赖于码，则 $R \in 2NF$ 。

则：S-L-C (Sno, School, Sloc, Cno, Grade) $\in 1NF$

S-L-C (Sno, School, Sloc, Cno, Grade) $\notin 2NF$

S-C (Sno, Cno, Grade) $\in 2NF$

S-L (Sno, School, Sloc) $\in 2NF$



2NF (续)

SC (Sno, Cno, Grade)
S-L (Sno, School, Sloc)

2NF vs. 1NF (问题解决)

插入异常?

- 插入新学生，未选课（无Cno），可以插入SL关系中

删除异常?

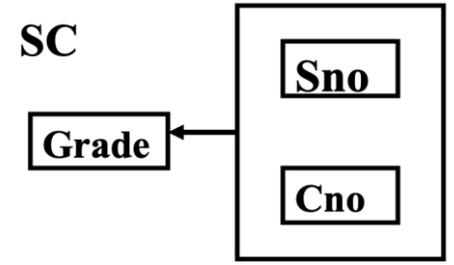
- 删除一个学生所有选课记录，只是SC关系没有记录。SL关系关于该学生不受影响

数据冗余度?

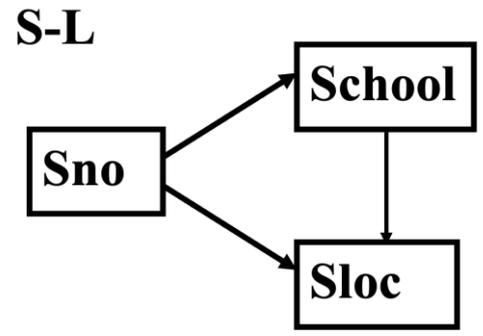
- 不论一个学生选多少课，School和Sloc值都只存储1次（SL关系中）

修改复杂?

- 学生转系：只修改SL关系即可。



Sno	Cno	Grade
...



Sno	School	Sloc
....



2NF (续)

54

- 采用**投影分解法**将一个1NF的关系分解为多个2NF的关系，可以在一定程度上减轻原1NF关系中存在的插入异常、删除异常、数据冗余度大、修改复杂等问题。
- **2NF消除了非主属性对码的部分函数依赖**
 - **判断2NF: 1. 函数依赖 2. 码, 非主属性 3. 部分函数依赖**
- 将一个1NF关系分解为多个2NF的关系, **并不能完全消除关系模式中的各种异常情况和数据冗余**



6.2 规范化

55

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

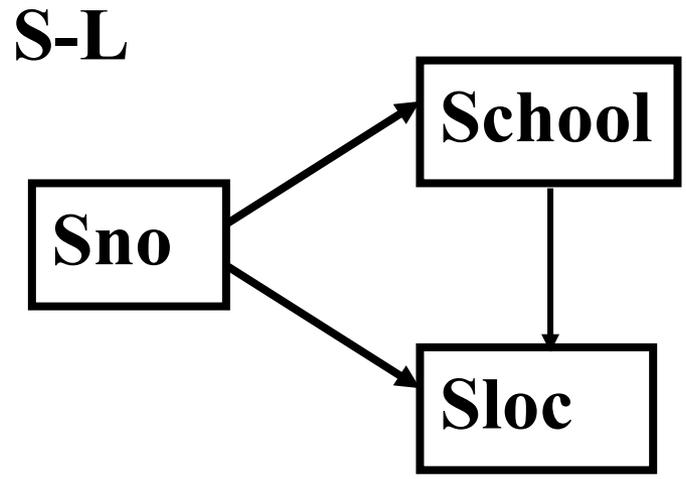
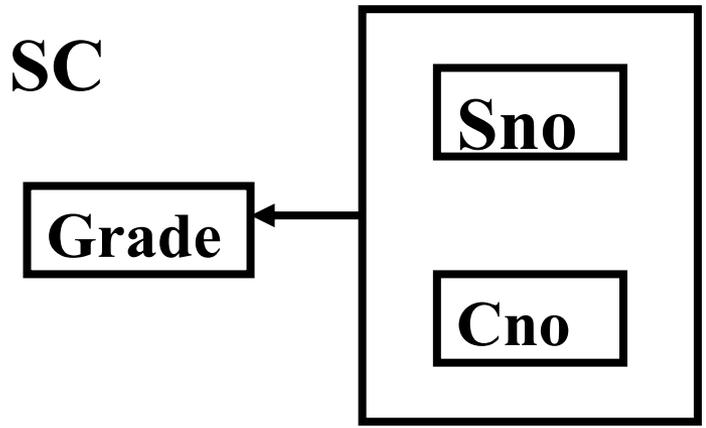
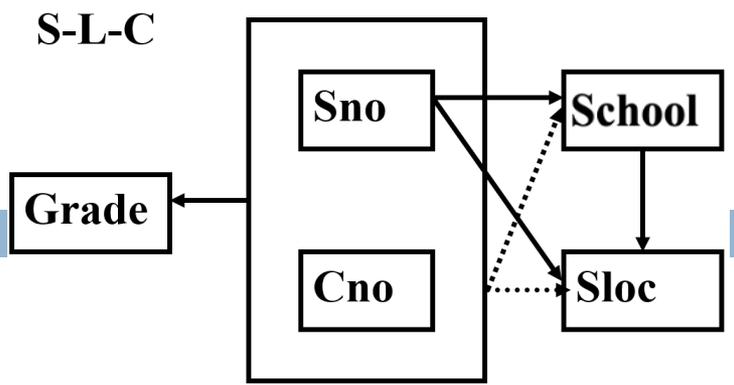
6.2.8 4NF

6.2.9 规范化小结



再想想2NF

函数依赖图:



- ❖ 关系模式SC的码为 (Sno, Cno)
- ❖ 关系模式S-L的码为Sno
- ❖ 非主属性对码都是完全函数依赖



再想想2NF

定义6.3 在R(U)中,

▶ 如果 $X \rightarrow Y, Y \not\subseteq X, Y \not\rightarrow X, Y \rightarrow Z, Z \not\subseteq Y$

则称Z对X传递函数依赖。记为: $X \overset{\text{传递}}{\rightarrow} Z$

例: 2NF关系模式S-L(Sno, School, Sloc)中

□ 函数依赖:

$Sno \rightarrow School$

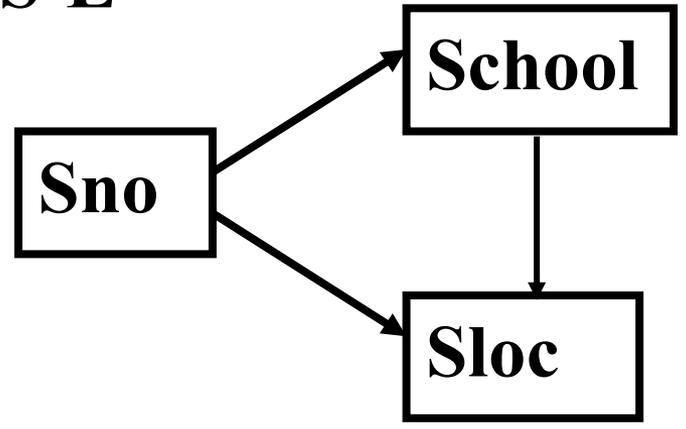
$School \rightarrow Sloc$

$Sno \rightarrow Sloc$

可得:

$Sno \overset{\text{传递}}{\rightarrow} Sloc$, 即S-L中存在非主属性对码的传递函数依赖

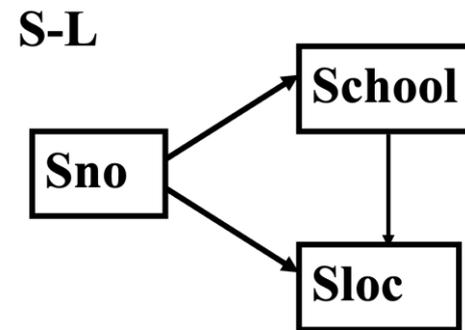
S-L





再想想2NF

Sno	School	Sloc
....



58

□ 在S-L关系中S-L(Sno, School, Sloc)

□ 插入异常

- 如果某个系(DS)刚成立, 没有学生, 则: 无法插入

□ 删除异常

- 如果某个系(DS)的学生全部毕业, 则删除所有学生, 则: 删除了系

□ 数据冗余

- 每个系的学生都住在同一个地方, 该系的sloc信息重复出现

□ 修改复杂

- 如果学校调整学生住处(DS)时, 则需要同时修改所有学生的loc

□ SL不是一个理想的关系模式



6.2.5 3NF

59

□ 3NF的定义

定义6.7 关系模式 $R \langle U, F \rangle$ 中, 若不存在这样的情况: 码 X 、属性组 $Y \subseteq U$ 及非主属性 Z , 且 $Z \not\subseteq Y, Y \twoheadrightarrow X$, 使得 $X \rightarrow Y, Y \rightarrow Z$ 成立, 则称 $R \langle U, F \rangle \in 3NF$ 。

- 若 $R \in 3NF$, 则每一个非主属性既不部分依赖于码也不传递依赖于码?
- 若 $R \in 3NF$, 则必有 $R \in 2NF$, 证明?



3NF (续)

例：2NF关系模式S-L(Sno, School, Sloc)中

□ 函数依赖：

$Sno \rightarrow School, Sno \rightarrow Sloc$

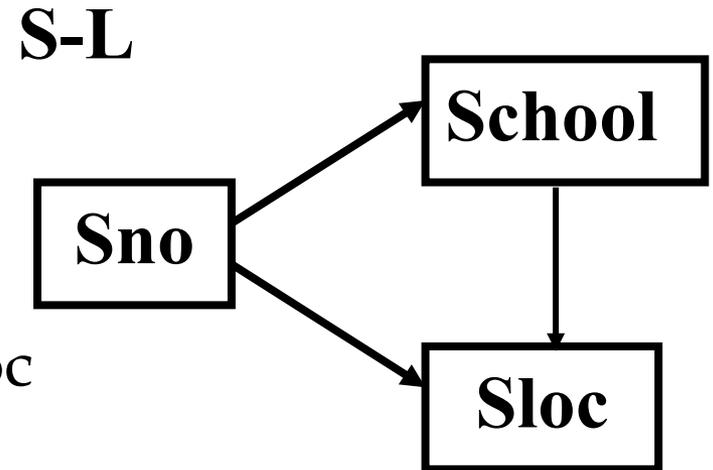
$School \not\rightarrow Sno, School \rightarrow Sloc$

码：Sno，非主属性：School, Sloc

可得：

$Sno \xrightarrow{\text{传递}} Sloc$ ，即S-L中存在非主属性对码的传递函数依赖，

因此， $S-L \notin 3NF$

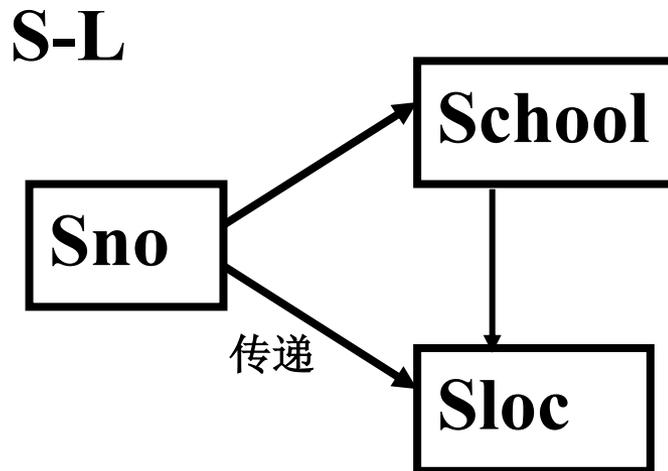




3NF (续)

61

函数依赖图:



- SL不是一个好的关系模式



3NF (续)

62

□ 解决方法

采用**投影分解**法，把S-L分解为两个关系模式，以消除传递函数依赖：

S-D (Sno, School)

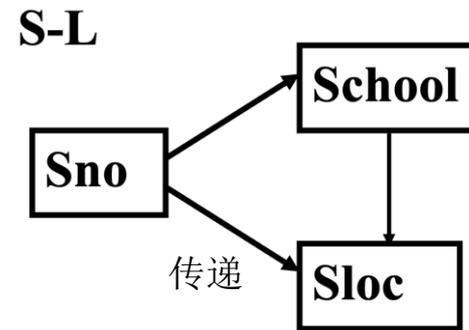
D-L (School, Sloc)

S-D的码为Sno， D-L的码为School。

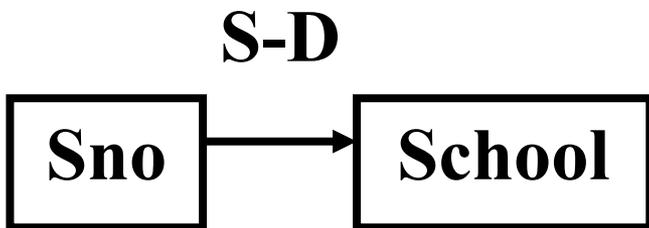
- 分解后的关系模式S-D与D-L中不再存在传递依赖



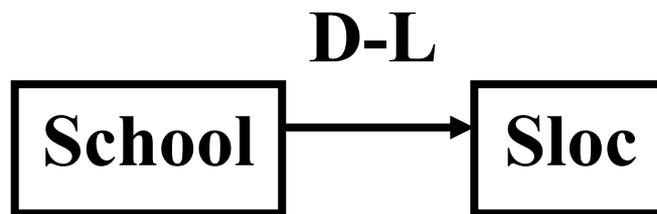
3NF (续)



S-D的码为Sno, D-L的码为Sdept



Sno	School
....	...



School	Sloc
...	...

- ❖ S-L(Sno, School, Sloc) ∈ 2NF
- S-L(Sno, School, Sloc) ∉ 3NF
- S-D(Sno, School) ∈ 3NF
- D-L(School, Sloc) ∈ 3NF

- 解决插入异常: DS刚成立—DL
- 解决删除异常: DS的学生全部毕业—SD
- 解决数据冗余: DS学生住在同一sloc—DL
- 解决修改复杂: 学校调整DS学生sloc—DL



3NF (续)

64

- 采用**投影分解法**将一个2NF的关系分解为多个3NF的关系，可以在一定程度上解决原2NF关系中存在的插入异常、删除异常、数据冗余度大、修改复杂等问题。
- **3NF消除了非主属性对码的传递函数依赖**
 - **判断3NF: 1. 函数依赖 2. 码, 非主属性 3. 传递函数依赖**
- 将一个2NF关系分解为多个3NF的关系后, 仍然不能完全消除关系模式中的各种异常情况和数据冗余。



关系模式规范化的基本步骤

65

□ 关系模式规范化的基本步骤

消除一些非平凡函数依赖，这些非平凡函数依赖的**决定属性集**不是码

1NF

↓ 消除非主属性对码的部分函数依赖

消除**决定属性** **2NF**

集非码的非平 ↓ 消除非主属性对码的传递函数依赖

凡函数依赖 **3NF**

↓ 消除主属性对码的部分和传递函数依赖

BCNF

↓ 消除非平凡且非函数依赖的多值依赖

4NF



6.2 规范化

66

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

6.2.8 4NF

6.2.9 规范化小结



再想想3NF

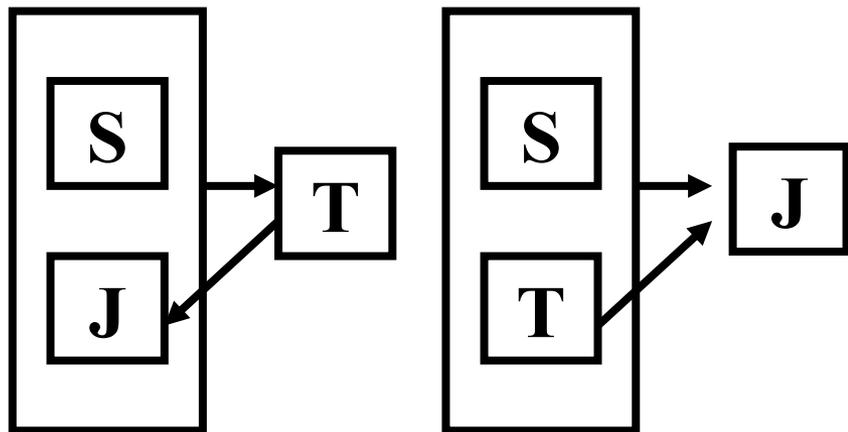
67

[例6.8]在关系模式STJ (S, T, J) 中, S表示学生, T表示教师, J表示课程。有如下函数依赖

- 语义: 每个教师只教一门课, 每门课有若干教师。某一学生选定某门课, 就对应一个固定的教师。某个学生选修某个教师的课就确定了所选课的名称

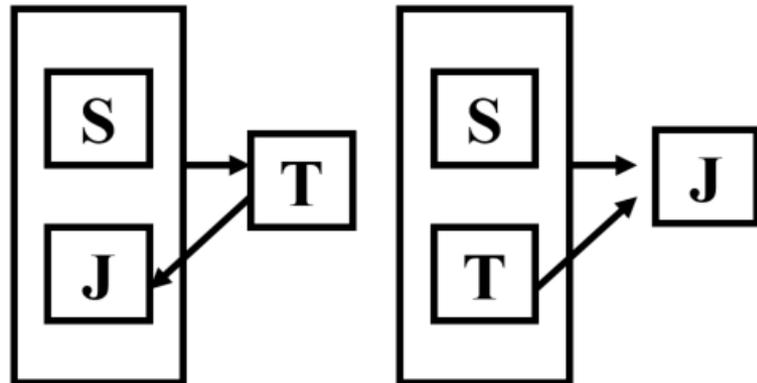
- 函数依赖形式化:

$$T \rightarrow J, (S, J) \rightarrow T, (S, T) \rightarrow J$$





再想想3NF



- 函数依赖:

$T \rightarrow J$, $(S, J) \rightarrow T$, $(S, T) \rightarrow J$

- 码: (S, J) 和 (S, T)

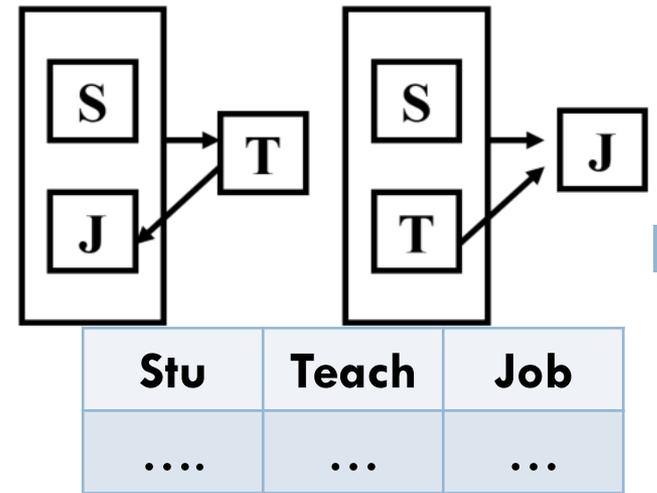
- 主属性: S, T, J都是主属性

- $T \rightarrow J$, 即T是决定因素, 可是T只是主属性, 它既不是候选码, 也不包含候选码

- $STJ \in 3NF$, 不存在非主属性对码的部分和传递函数依赖



再想想3NF



STJ (S, T, J) 是否有问题？

插入异常

- 如果某个教师开设某门课，但尚未有学生选修，则信息无法插入

删除异常

- 如果选修过某门课程的学生全部毕业，删除这些学生元组同时，相应教师开设的该课程也删除了

数据冗余

- 虽然一个教师只教一门课，但是每个选修该教师该门课程的学生元组都要记录这一信息

修改复杂

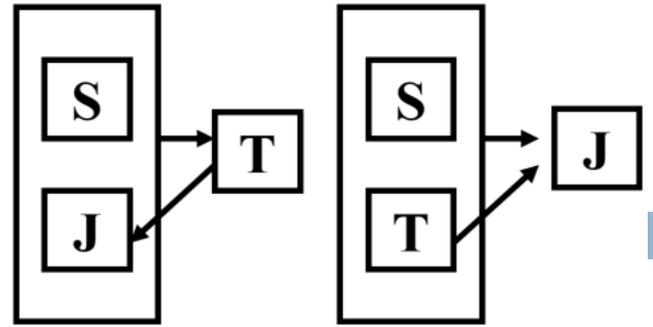
- 某个教师开设课程改名，所有选修该教师该门课学生元组都要进行修改

课后思考：给出数据，验证STJ的异常情况

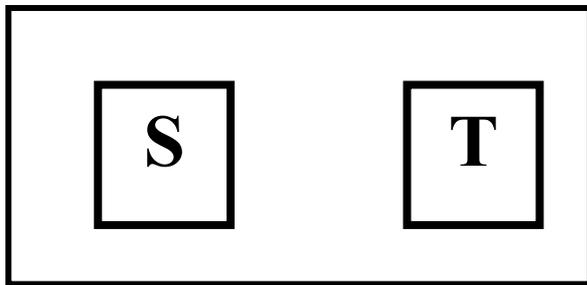
STJ ∈ 3NF，但不是一个理想的关系模式



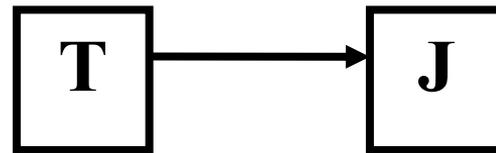
再想想3NF



- STJ (S, T, J) 不是理想的模式
 - 原因：主属性J依赖于T，即主属性J部分依赖于 (S, T)
 - 解决方法：
 - 采用投影分解法，将STJ分解为关系模式：



ST

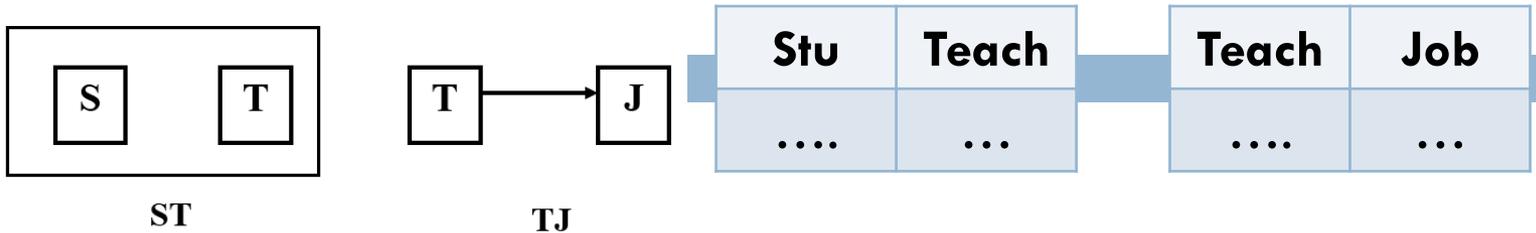


TJ



再想想3NF

S表示学生，T表示教师，J表示课程。



- 分解后的R没有**任何属性对码**的部分函数依赖和传递函数依赖
 - 解决插入异常
 - TJ关系中 can 存储教师开课信息，但无学生选修
 - 解决删除异常
 - 若选修某课程的学生全部毕业，只删除了ST关系的元组，不会影响TJ关系中教师开设课程的信息
 - 解决数据冗余
 - 每个教师开设课程的信息只在TJ关系中存储一次
 - 解决更新异常
 - 某个教师开设的某课程改名，只需要修改TJ关系中的一个元组



6.2.6 BC范式 (BCNF)

72

- **定义6.8** 关系模式 $R\langle U, F \rangle \in 1NF$, 若 $X \rightarrow Y \in F$, 且 $Y \not\subseteq X$ 时, **X必含有码**, 则 $R\langle U, F \rangle \in BCNF$.
 - 1974年, 由Boyce与Codd提出, 比3NF更进了一步。通常认为BCNF是修正的第三范式, 有时也称为扩充的第三范式
- 换言之, 在关系模式 $R\langle U, F \rangle$ 中, 如果每一个**决定因素都包含码 (候选码)**, 则 $R \in BCNF$



BCNF (续)

73

[例6.5] 关系模式Course (Cno, Cname, Ccredit, Cpno)

- $\text{Course} \in 3\text{NF}$: 只有一个码Cno, 没有任何属性对Cno部分依赖或传递依赖
- $\text{Course} \in \text{BCNF}$: 同时Course中Cno是唯一的决定因素, 所以 $\text{Course} \in \text{BCNF}$

[例6.6] 关系模式Student (Sno, Sname, Ssex, Sage, Sdept)

- 假定Student有两个码Sno, Sname, 假设Sname也是唯一的
- $\text{Student} \in 3\text{NF}$
- $\text{Student} \in \text{BCNF}$



BCNF (续)

74

[例6.7] 关系模式SJP (S, J, P), S是学生, J是课程, P表示名次。
每个学生选修每门课的成绩有一定的名次, 每门课中每一名只有一个学生 (无并列)。

- 函数依赖: $(S, J) \rightarrow P; (J, P) \rightarrow S$
- (S, J) 与 (J, P) 都可以作为候选码, 属性相交
- $SJP \in 3NF,$
- $SJP \in BCNF$



BCNF (续)

[例6.8]在关系模式STJ (S, T, J) 中, S表示学生, T表示教师, J表示课程。每个教师只教一门课, 每门课有若干教师, 某一学生选定某门课, 就对应一个固定的教师。

□ 函数依赖:

$$T \rightarrow J, (S, T) \rightarrow J, (S, J) \rightarrow T$$

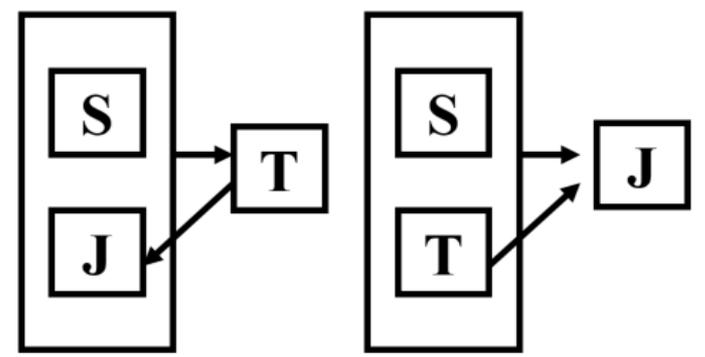
□ (S, J)和(S, T)都是候选码

□ $STJ \in 3NF$

- 没有任何非主属性对码传递依赖或部分依赖
- S,T,J都是主属性

□ $STJ \notin BCNF$ (BCNF每一个决定属性因素都包含码)

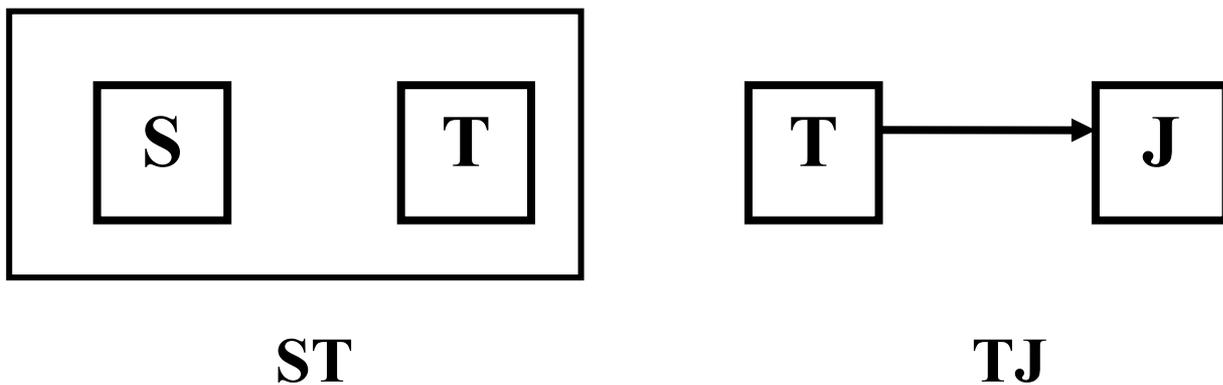
- T是决定因素, T不包含码





BCNF (续)

- 解决方法：将STJ分解为二个关系模式：
 $SJ(S, J) \in BCNF, TJ(T, J) \in BCNF$



没有任何属性（包括主属性）对码的部分函数依赖和传递函数依赖



BCNF (续)

77

- 若 $R \in \text{BCNF}$ ，则含有性质
 - 所有非主属性对每一个码都是完全函数依赖
 - 所有的主属性对每一个不包含它的码，也是完全函数依赖
 - 没有任何属性完全函数依赖于非码的任何一组属性
- 如果关系模式 $R \in \text{BCNF}$ ，则 $R \in 3\text{NF}$
- 如果 $R \in 3\text{NF}$ ，且 R 中只有一个候选码，则 R 必为BCNF
 - Recall: 若 $R \in 3\text{NF}$ ，则每一个非主属性既不部分依赖于码也不传递依赖于码。



3NF与BCNF的关系

78

□ $R \in \text{BCNF} \xrightleftharpoons[\text{不必要}]{\text{充分}} R \in 3\text{NF}$

□ 如果 $R \in 3\text{NF}$ ，且 R 只有一个候选码

$R \in \text{BCNF} \xrightleftharpoons[\text{必要}]{\text{充分}} R \in 3\text{NF}$

□ 3NF和BCNF是在函数依赖的条件下对模式分解所能达到的分离程度的测度

□ 如果一个关系数据库中的所有关系模式都属于BCNF，那么在函数依赖范畴内，它已实现了模式的彻底分解，达到了最高的规范化程度，消除了插入异常和删除异常

□ 3NF的“不彻底”性表现在可能存在主属性对码的部分依赖或传递依赖。



关系模式规范化的基本步骤

□ 关系模式规范化的基本步骤

消除一些非平凡函数依赖，这些非平凡函数依赖的**决定属性集**不是码

1NF

↓ 消除非主属性对码的部分函数依赖

消除**决定属性集**非码的非平凡函数依赖

2NF

↓ 消除非主属性对码的传递函数依赖

3NF

↓ 消除主属性对码的部分和传递函数依赖

BCNF

↓ 消除非平凡且非函数依赖的多值依赖

4NF

范式越高越好？



6.2 规范化

92

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

*6.2.7 多值依赖

*6.2.8 4NF

6.2.9 规范化小结



6.2.7 多值依赖

[例6.9] 学校中某一门课程由多个教师讲授，他们使用相同的一套参考书。每个教师可以讲授多门课程，每种参考书可以供多门课程使用。

- 关系模式Teaching(C,T,B)
 - 课程C
 - 教师T
 - 参考书B



6.2.7 多值依赖

[例6.9] 学校中某一门课程C由多个教师T讲授，他们使用相同的一套参考书B。每个教师可以讲授多门课程，每种参考书可以供多门课程使用。

❖ 表6.3. 非规范化关系

课程 C	教师 T	参考书 B
物理	{ 李勇 } { 王军 }	{ 普通物理学 } { 光学原理 } { 物理习题集 }
数学	{ 李勇 } { 张平 }	{ 数学分析 } { 微分方程 } { 高等代数 }
计算数学	{ 张平 } { 周峰 }	{ 数学分析 } ...
...



多值依赖 (续)

❖ 表6.4 用二维表表示Teaching

课程C	教师T	参考书B
物理	李勇	普通物理学
物理	李勇	光学原理
物理	李勇	物理习题集
物理	王军	普通物理学
物理	王军	光学原理
物理	王军	物理习题集
数学	李勇	数学分析
数学	李勇	微分方程
数学	李勇	高等代数
数学	张平	数学分析
数学	张平	微分方程
数学	张平	高等代数
...



多值依赖（续）

96

[例6.9] 学校中某一门课程由多个教师讲授，他们使用相同的一套参考书。每个教师可以讲授多门课程，每种参考书可以供多门课程使用。

- Teaching具有**唯一候选码**(C, T, B)，即全码all-key
- Teaching \in BCNF



多值依赖 (续)



Teaching模式中存在的问题

(1) 数据冗余度大

存储教师与参考书?

(2) 插入操作复杂

课程增加一个教师, 参考书?

(3) 删除操作复杂

课程删除一本参考书, 教师?

(4) 修改操作复杂

课程修改一本参考书, 教师?

原因: 参考书B和教师T之间的关系是笛卡尔积, 则彼此独立, 毫无关系, 都只取决于课程C

课程C	教师T	参考书B
物理	李勇	普通物理学
物理	李勇	光学原理
物理	李勇	物理习题集
物理	王军	普通物理学
物理	王军	光学原理
物理	王军	物理习题集
数学	李勇	数学分析
数学	李勇	微分方程
数学	李勇	高等代数
数学	张平	数学分析
数学	张平	微分方程
数学	张平	高等代数
...



多值依赖（续）

98

□ 定义6.9 设 $R(U, F)$ 是一个属性集 U 上的一个关系模式，

X 、 Y 和 Z 是 U 的子集，并且 $Z = U - X - Y$ 。关系模式 $R(U)$ 中多值依赖

$X \twoheadrightarrow Y$ 成立，当且仅当对 $R(U)$ 的任一关系 r ，给定的一对 (x, z) 值，有一组 Y 的值，这组值仅仅决定于 x 值而与 z 值无关

□ 例 Teaching (C, T, B)

□ 对于 C 的每一个值， T 有一组值与之对应，而不论 B 取何值。因此， T 多值依赖于 C ，即 $C \twoheadrightarrow T$

■ 分析：对于一个 $(x_{\text{物理}}, z_{\text{光学原理}})$ 有一组 T {李勇, 王军}，这组 T 仅取决于 C 的值。即，对于另一个 $(\text{物理}, \text{普通物理学})$ ，它对应的一组 T 仍然是{李勇, 王军}，(尽管此时参考书 B 已经改变)。所以， T 多值依赖于 C ，即 $C \twoheadrightarrow T$

■ 同理：对于一个 $(x_{\text{物理}}, z_{\text{李勇}})$ 也是如此，即 $C \twoheadrightarrow B$ 也成立



多值依赖（续）

99

- 另一个等价的形式化的定义：

在 $R(U)$ 的任一关系 r 中，如果存在元组 t, s 使得

$t[X]=s[X]$ ，那么就必然存在元组 $w, v \in r$ ，(w, v 可以与 s, t 相同)，使得(1) $w[X]=v[X]=t[X]=s[X]$ ，而(2) $w[Y]=t[Y]$ ， $w[Z]=s[Z]$ ，(3) $v[Y]=s[Y]$ ， $v[Z]=t[Z]$ 均成立（即交换 s, t 元组的 Y 值所得的两个新元组必在 r 中），

则 Y 多值依赖于 X ，记为 $X \twoheadrightarrow Y$ 。

这里， X, Y 是 U 的子集， $Z=U-X-Y$ 。



多值依赖 (续)

在 $R(U, F)$ 的任一关系 r 中,
 如果存在元组 t, s , 使得 $t[X]=s[X]$,
 那么就必然存在元组 $w, v \in r$,
 (w, v 可以与 s, t 相同), 使得

- (1) $w[X]=v[X]=t[X]=s[X]$,
- (2) $w[Y]=t[Y], w[Z]=s[Z]$,
- (3) $v[Y]=s[Y], v[Z]=t[Z]$

(即交换 s, t 元组的 Y 值所得的两个新元组必在 r 中)均成立。

则 Y 多值依赖于 X , 记为 $X \twoheadrightarrow Y$ 。

- 这里, X, Y 是 U 的子集,
 $Z=U-X-Y$ 。

X	Y	Z
课程C	教师T	参考书B
t 物 理	李 勇	普通物理学
w 物 理	李 勇	光学原理
物 理	李 勇	物理习题集
v 物 理	王 军	普通物理学
s 物 理	王 军	光学原理
物 理	王 军	物理习题集
数 学	李 勇	数学分析
数 学	李 勇	微分方程
数 学	李 勇	高等代数
数 学	张 平	数学分析
数 学	张 平	微分方程
数 学	张 平	高等代数
...



多值依赖（续）

101

- 平凡多值依赖和非平凡的多值依赖
 - 若 $X \twoheadrightarrow Y$ ，而 $Z = \emptyset$ （空），则称 $X \twoheadrightarrow Y$ 为平凡的多值依赖
 - 否则称 $X \twoheadrightarrow Y$ 为非平凡的多值依赖



多值依赖 (续)

[例6.10] 关系模式WSC (W, S, C)

- W表示仓库，S表示保管员，C表示商品
- 假设每个仓库有若干个保管员，有若干种商品
- 每个保管员保管所在的仓库的所有商品
- 每种商品被所有保管员保管

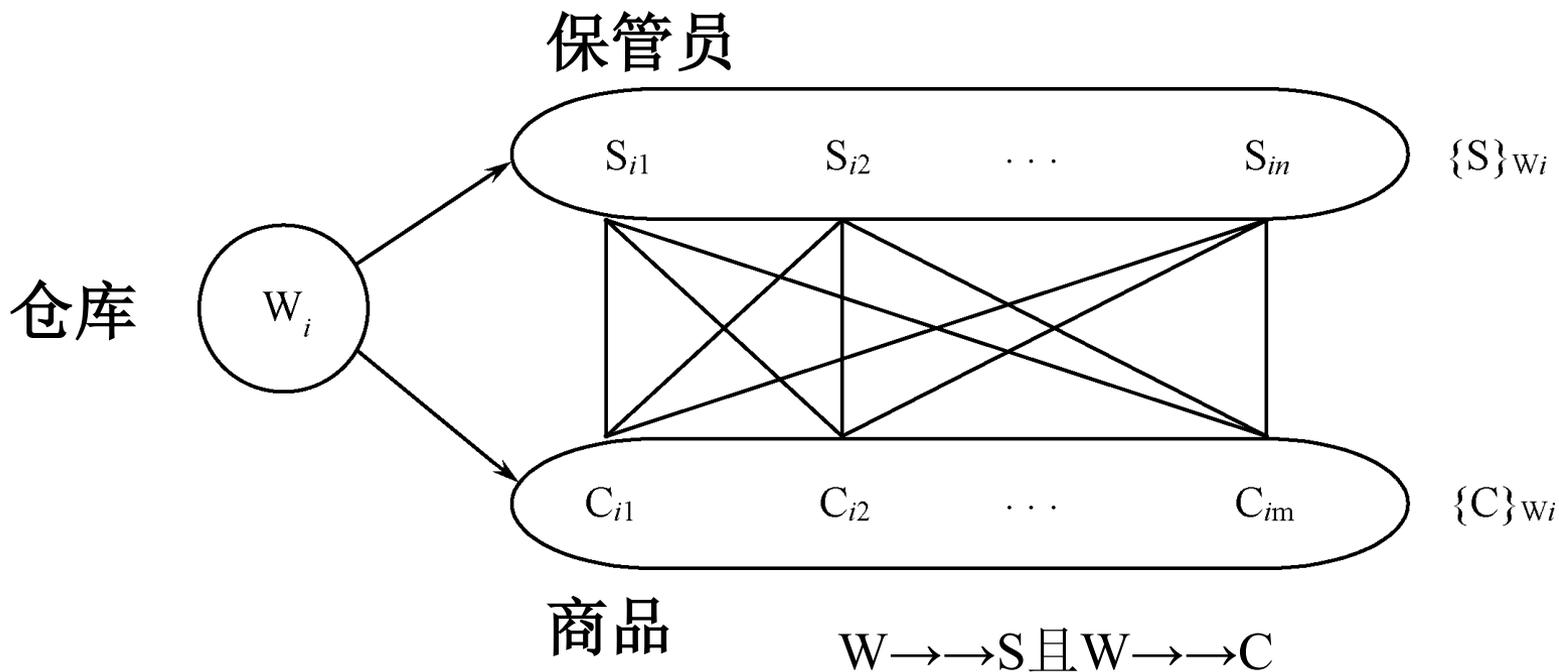
W	S	C
W1	S1	C1
W1	S1	C2
W1	S1	C3
W1	S2	C1
W1	S2	C2
W1	S2	C3
W2	S3	C4
W2	S3	C5
W2	S4	C4
W2	S4	C5



多值依赖 (续)

[例6.10] 关系模式WSC (W, S, C)

- W表示仓库, S表示保管员, C表示商品
- 假设每个仓库有若干个保管员, 有若干种商品
- 每个保管员保管所在的仓库的所有商品
- 每种商品被所有保管员保管





多值依赖的性质

104

(1) 多值依赖具有对称性

若 $X \twoheadrightarrow Y$, 则 $X \twoheadrightarrow Z$, 其中 $Z = U - X - Y$

(2) 多值依赖具有传递性

若 $X \twoheadrightarrow Y$, $Y \twoheadrightarrow Z$, 则 $X \twoheadrightarrow Z - Y$

(3) 函数依赖是多值依赖的特例

若 $X \rightarrow Y$, 则 $X \twoheadrightarrow Y$ 。

(4) 若 $X \twoheadrightarrow Y$, $X \twoheadrightarrow Z$, 则 $X \twoheadrightarrow Y \cup Z$ 。

(5) 若 $X \twoheadrightarrow Y$, $X \twoheadrightarrow Z$, 则 $X \twoheadrightarrow Y \cap Z$ 。

(6) 若 $X \twoheadrightarrow Y$, $X \twoheadrightarrow Z$, 则 $X \twoheadrightarrow Y - Z$, $X \twoheadrightarrow Z - Y$ 。

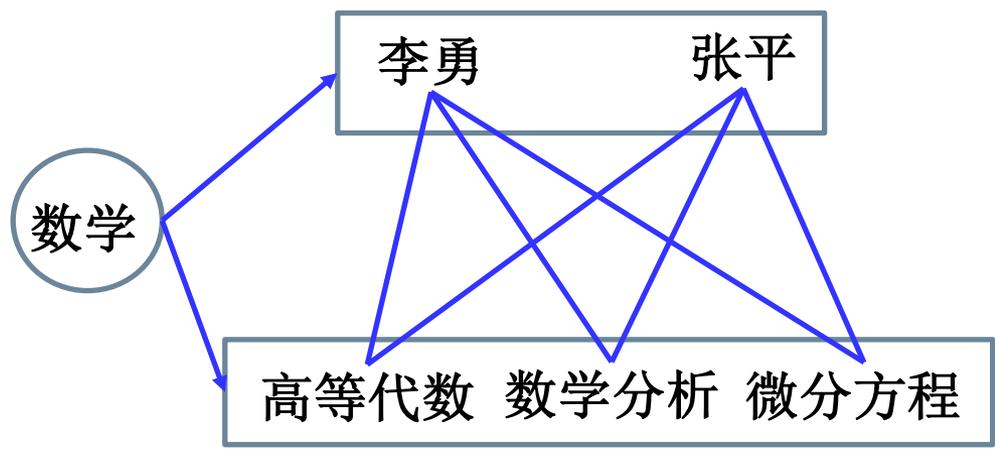
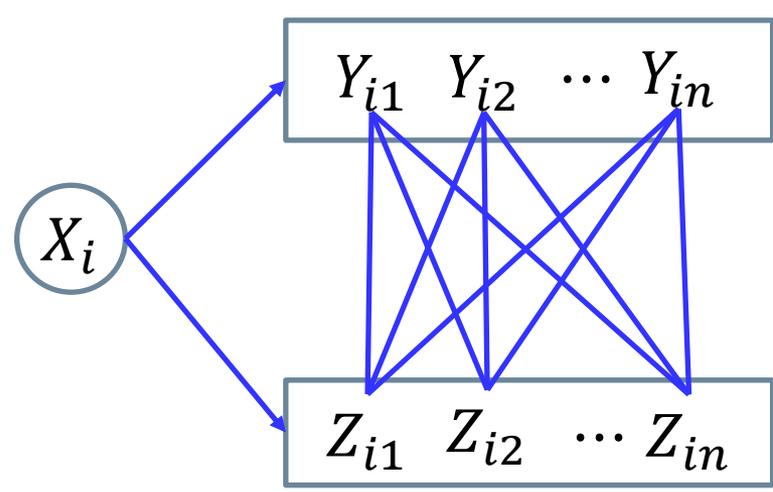
课后思考：举出实际的例子验证或证明上述性质



多值依赖的性质

(1) 多值依赖具有对称性

若 $X \twoheadrightarrow Y$ ，则 $X \twoheadrightarrow Z$ ，其中 $Z = U - X - Y$





多值依赖与函数依赖的区别

106

(1) 多值依赖的有效性与属性集的范围有关

- 若 $X \twoheadrightarrow Y$ 在 U 上成立，则在 W ($XY \subseteq W \subseteq U$) 上一定成立。反之则不然，即 $X \twoheadrightarrow Y$ 在 W ($W \subset U$) 上成立，在 U 上并不一定成立
 - 原因：多值依赖不仅涉及属性组 X 和 Y ，而且涉及 U 中其余属性 Z
- 一般地，在 $R(U)$ 上若有 $X \twoheadrightarrow Y$ 在 W ($W \subset U$) 上成立，则称 $X \twoheadrightarrow Y$ 为 $R(U)$ 的嵌入型多值依赖
- 函数依赖 $X \rightarrow Y$ 的有效性仅决定于 X 、 Y 这两个属性集的值
- 只要在 $R(U)$ 的任何一个关系 r 中，元组在 X 和 Y 上的值满足定义6.1，则函数依赖 $X \rightarrow Y$ 在任何属性集 W ($XY \subseteq W \subseteq U$) 上成立

定义6.1 函数依赖



多值依赖与函数依赖的区别

(2) 若函数依赖 $X \rightarrow Y$ 在 $R(U)$ 上成立, 则任何 $Y' \subset Y$ 均有 $X \rightarrow Y'$ 成立

多值依赖 $X \twoheadrightarrow Y$ 若在 $R(U)$ 上成立, 不能断言对于任何 $Y' \subset Y$ 有 $X \twoheadrightarrow Y'$ 成立

- 例如, 关系 $R(A,B,C,D)$, $A \twoheadrightarrow BC$ 成立, (也有 $A \twoheadrightarrow D$ 成立)。 $A \twoheadrightarrow B$ 不一定成立, 因为 B 的取值不仅取决于属性 A , 还取决于属性 C,D 。
- 表6.6: 有 R 的一个关系实例, 在此实例上 $A \twoheadrightarrow B$ 是不成立的

表6.6 R的一个实例

	A	B	C	D
t	a_1	b_1	c_1	d_1
w	a_1	b_1	c_1	d_2
v	a_1	b_2	c_2	d_1
s	a_1	b_2	c_2	d_2



多值依赖与函数依赖的区别

(2) $A \twoheadrightarrow B$ 成立的情况

	A	B	C	D
	a1	b1	c1	d1
	a1	b1	c1	d2
<i>t</i>	a1	b2	c2	d1
<i>w</i>	a1	b2	c2	d2
<i>v</i>	a1	b1	c2	d1
<i>s</i>	a1	b1	c2	d2
	?	?	?	?
	?	?	?	?



练习

关系U中有属性ABCD，包含如下记录，请问最后两个元组中的A和D的值分别为多少时，多值依赖 $BC \twoheadrightarrow A$ 才有可能成立？

	Y	X	Z	
	A	B	C	D
t	2	4	3	3
	2	3	5	3
s	1	4	3	2
	3	1	2	2
w	1	4	3	3
v	2	4	3	2



6.2 规范化

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

6.2.8 4NF

6.2.9 规范化小结



6.2.8 4NF

111

- 定义6.10 关系模式 $R\langle U, F \rangle \in 1NF$ ，如果对于R的每个非平凡多值依赖 $X \twoheadrightarrow Y$ ($Y \subsetneq X$)，**X都含有码**，则 $R \in 4NF$ 。
 - 4NF：限制关系模式的**属性之间不允许有非平凡且非函数依赖的多值依赖**。
 - 4NF所允许的非平凡多值依赖实际上是函数依赖
 - $X \twoheadrightarrow Y$ ， Z 不为空， X 含有码
 - 4NF消除了非主属性对候选码以外属性的多值依赖



6.2.8 4NF

112

- 定义6.10 关系模式 $R\langle U, F \rangle \in 1NF$ ，如果对于R的每个非平凡多值依赖 $X \twoheadrightarrow Y$ ($Y \not\subseteq X$)，X都含有码，则 $R \in 4NF$ 。
- 如果 $R \in 4NF$ ，则 $R \in BCNF$
 - 不允许有非平凡且非函数依赖（非1:1）的多值依赖（1:n）
 - 不允许像Teaching（教员，课程，参考书）关系中的多值依赖存在
 - 即，允许的非平凡多值依赖是函数依赖（1:1）
 - 不会有冗余

NOTE: 若 $X \twoheadrightarrow Y$ ，而 $Z = \varphi$ （空），则称 $X \twoheadrightarrow Y$ 为平凡的多值依赖

否则称 $X \twoheadrightarrow Y$ 为非平凡的多值依赖

不用管平凡多值依赖，因为它是合理的（平凡多值依赖是4NF）



4NF (续)

例: Teaching(C,T,B) \notin 4NF

存在非平凡的多值依赖 $C \twoheadrightarrow T$, 且C不是码 (C只是码的一部分), 这是存在不好性质的根源

- ❑ 数据冗余
- ❑ 插入复杂
- ❑ 删除操作复杂
- ❑ 修改操作复杂

课程C	教师T	参考书B
物理	李勇	普通物理学
物理	李勇	光学原理
物理	李勇	物理习题集
物理	王军	普通物理学
物理	王军	光学原理
物理	王军	物理习题集
数学	李勇	数学分析
数学	李勇	微分方程
数学	李勇	高等代数
数学	张平	数学分析
数学	张平	微分方程
数学	张平	高等代数

... ..



4NF

- 用投影分解法把Teaching分解为如下两个关系模式:

$$CT(C, T) \in 4NF, \quad CB(C, B) \in 4NF$$

- $C \twoheadrightarrow T, C \twoheadrightarrow B$ 是平凡多值依赖

解决上述问题

- 数据冗余: 参考书只需要CB中存储一次
- 插入操作: 课程增加一名教师时, 只需CT插入
- 删除操作: 某课程去掉一本参考书, 只需CB中删除一元组
- 修改操作: 某课程修改一本参考书, 只需修改CB



4NF (续)

115

- ❖ [例6.10]的WSC中,
 - ❖ 关系模式WSC的码是(W,S,C), 即All-key
 - ❖ 存在 $W \twoheadrightarrow S$, $W \twoheadrightarrow C$,他们都是非平凡多值依赖。而W不是码,
 - ❖ 因此 $WSC \notin 4NF$ 。

- ❖ 可以把WSC分解成 $WS(W,S)$, $WC(W,C)$
 - ❖ $WS \in 4NF$
 - ❖ $WC \in 4NF$



4NF (续)

116

确定一个关系模式是否为4NF (一般可以分为三个步骤) :

1. 确定该关系模式的码
2. 判断关系模式里的多值依赖是否是平凡多值依赖, 如果是平凡多值依赖, 则该范式是4NF
3. 如果多值依赖是非平凡多值依赖, 判断该多值依赖的决定因素 (左半部分) 是否包含码。如果包含码, 则该范式是4NF, 否则, 该范式不是4NF



6.2 规范化

6.2.1 函数依赖

6.2.2 码

6.2.3 范式

6.2.4 2NF

6.2.5 3NF

6.2.6 BCNF

6.2.7 多值依赖

6.2.8 4NF

6.2.9 规范化小结



6.2.9 规范化小结

118

- 在关系数据库中
 - 对关系模式的基本要求是满足第一范式。
- 规范化程度过低的关系不一定能够很好地描述现实世界
 - 可能存在插入异常、删除异常、修改复杂、数据冗余等问题
 - 解决方法就是对其进行规范化，转换成高级范式。

- 一个低一级范式的关系模式，通过模式分解可以转换为若干个高一级范式的关系模式集合，这种过程就叫关系模式的规范化。



6.2.9 规范化小结

119

- 关系数据库的规范化理论是数据库逻辑设计的工具
- 目的：尽量消除插入、删除异常，修改复杂，数据冗余
- 规范化的基本思想
 - 逐步消除数据依赖中不合适的部分，使模式中的各关系模式达到某种程度的“分离”
 - 即采用“一事一地”的模式设计原则
 - 让一个关系描述一个概念、一个实体或者实体间的一种联系。若多于一个概念就把它“分离”出去。
 - 实质：概念的单一化



规范化小结（续）

□ 关系模式规范化的基本步骤

消除一些非平凡函数依赖，这些非平凡函数依赖的**决定属性集**不是码

1NF

↓ 消除非主属性对码的部分函数依赖

消除决定属性

2NF

集非码的非平

↓ 消除非主属性对码的传递函数依赖

凡函数依赖

3NF

↓ 消除主属性对码的部分和传递函数依赖

BCNF

↓ 消除非平凡且非函数依赖的多值依赖

4NF



规范化小结（续）

121

- 不能说规范化程度越高的关系模式就一定越好
- 在设计数据库模式结构时，必须对现实世界的实际情况和用户应用需求作进一步分析，确定一个合适的、能够反映现实世界的模式
- 上面的规范化步骤可以在其中任何一步终止