

# 第六章 关系数据理论

**6.1 问题的提出**

**6.2 规范化**

**6.3 数据依赖的公理系统**

**6.4 模式的分解**

**6.5 小结**



# 函数依赖的推理规则

- 如何判断关系模式的范式级别？
- 如何求关系模式的候选码？



# 函数依赖的公理系统

- 逻辑蕴涵
- Amstrong公理的内容及正确性
- Amstrong公理的推论
- 闭包计算
- 公理的完备性
- 函数依赖集的等价和最小化



# 逻辑蕴涵

有时需要根据给定的一组函数依赖来判断另外一些函数依赖是否成立，这就是函数依赖逻辑蕴涵所要研究的内容。

比如有关系模式 $R(U, F)$ ， $U = \{A, B, C\}$ ， $F = \{A \rightarrow B, B \rightarrow C\}$ ，问 $A \rightarrow C$ 是否也成立？



# 逻辑蕴涵

定义：设有关系模式 $R(U, F)$ ， $X \subseteq U$ 、 $Y \subseteq U$ ，  
如果从 $F$ 中的函数依赖能够推导出 $X \rightarrow Y$ ，则称 $F$ 逻辑  
蕴涵 $X \rightarrow Y$ ，或称 $X \rightarrow Y$ 是 $F$ 的逻辑蕴涵。



# Armstrong公理系统

设 $U$ 为属性集总体， $F$ 是 $U$ 上的一组函数依赖，于是有关系模式 $R(U, F)$ 。对于 $R(U, F)$ 来说有下面的推理规则：

- A1.自反律（Reflexivity）：若 $Y \subseteq X \subseteq U$ ，则 $X \rightarrow Y$ 为 $F$ 所蕴含。
- A2.增广律（Augmentation）：若 $X \rightarrow Y$ 为 $F$ 所蕴含，且 $Z \subseteq U$ ，则 $XZ \rightarrow YZ$ 为 $F$ 所蕴含。
- A3.传递律（Transitivity）：若 $X \rightarrow Y$ 及 $Y \rightarrow Z$ 为 $F$ 所蕴含，则 $X \rightarrow Z$ 为 $F$ 所蕴含。



# 导出规则

1. 根据A1, A2, A3这三条推理规则可以得到下面三条推理规则:

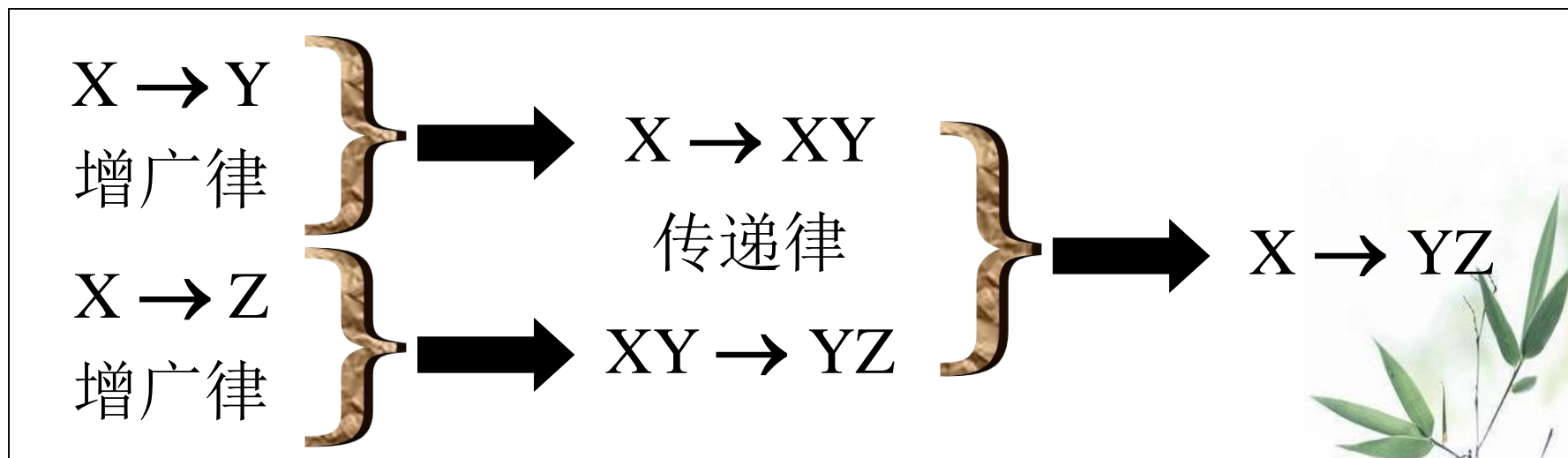
- 合并规则: 由 $X \rightarrow Y$ ,  $X \rightarrow Z$ , 有 $X \rightarrow YZ$ 。
- 伪传递规则: 由 $X \rightarrow Y$ ,  $WY \rightarrow Z$ , 有 $XW \rightarrow Z$ 。
- 分解规则: 由 $X \rightarrow Y$ 及  $Z \subseteq Y$ , 有 $X \rightarrow Z$ 。





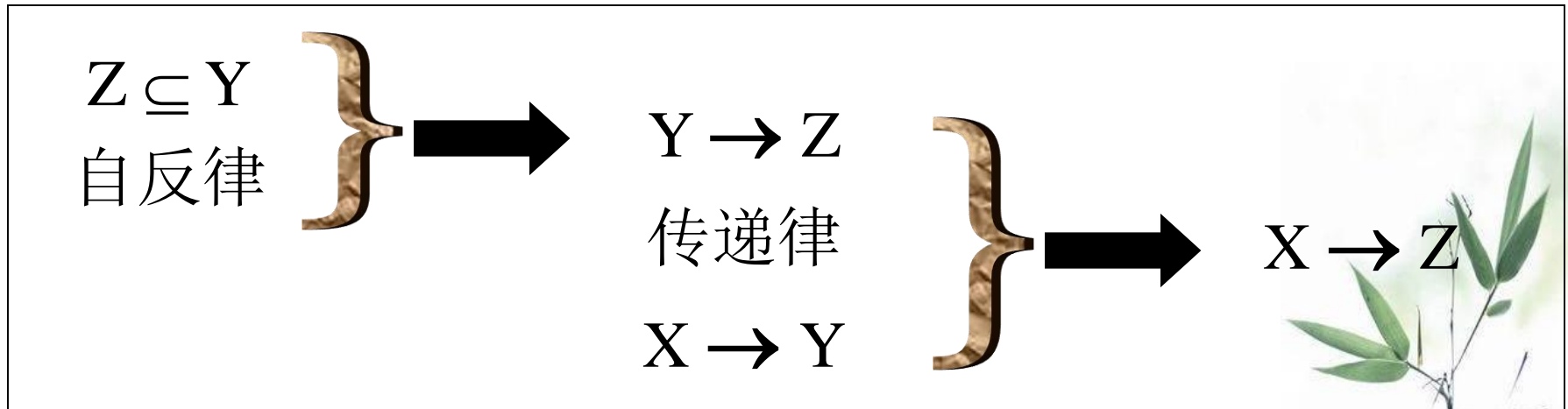
# 导出规则

- 由Armstrong公理导出的推理规则
  - 合并律(union rule)
    - 若 $X \rightarrow Y$ ,  $X \rightarrow Z$ , 则 $X \rightarrow YZ$



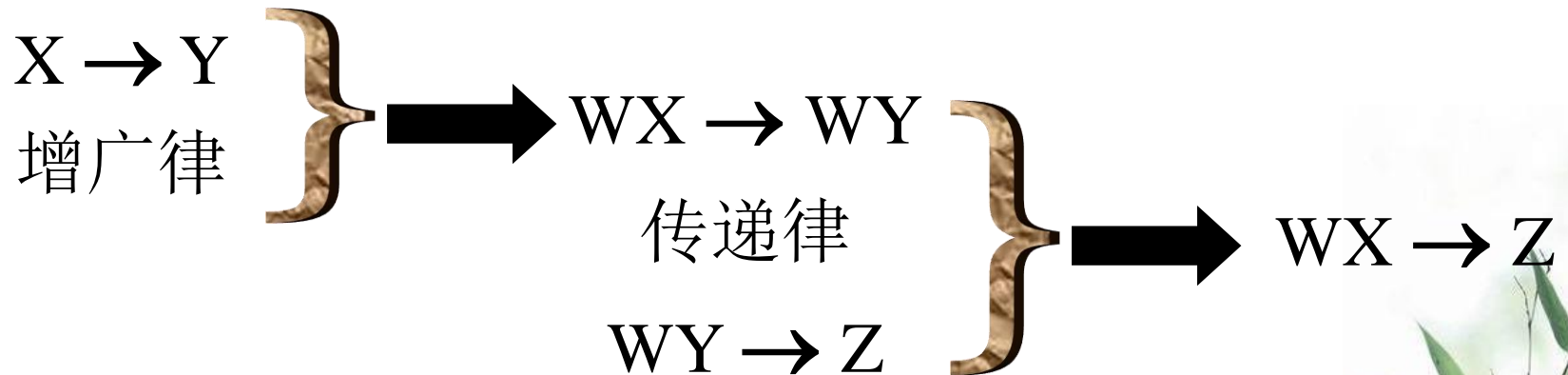
# 导出规则

- 由Armstrong公理导出的推理规则
  - 分解律(decomposition rule)
    - 若 $X \rightarrow Y$ , 及  $Z \subseteq Y$  则  $X \rightarrow Z$



# 导出规则

- 由Armstrong公理导出的推理规则
  - 伪传递律(pseudotransitivity rule)
    - 若 $X \rightarrow Y$ ,  $WY \rightarrow Z$ , 则 $WX \rightarrow Z$



**例：**已知关系模式 $R(U, F)$ ， $U = (A, B, C, G, H, I)$ ， $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$ ，判断下列函数依赖是否为 $F$ 的逻辑蕴涵？

- $A \rightarrow H$        $A \rightarrow B, B \rightarrow H \Rightarrow A \rightarrow H$  (传递律)
- $AG \rightarrow HI$        $CG \rightarrow H, CG \rightarrow I \Rightarrow CG \rightarrow HI$  (合成规则)  
                          $A \rightarrow C, CG \rightarrow HI \Rightarrow AG \rightarrow HI$  (伪传递规则)
- $AB \rightarrow CHG$        $AB \rightarrow A, A \rightarrow C \Rightarrow AB \rightarrow C$  (自反律, 传递律)  
                          $AB \rightarrow B, B \rightarrow H \Rightarrow AB \rightarrow H$  (自反律, 传递律)  
                          **$AB \rightarrow G ?$**



# 导出规则

2. 根据合并规则和分解规则，可得引理6.1

引理6.1  $X \rightarrow A_1 A_2 \dots A_k$  成立的充分必要条件是  
 $X \rightarrow A_i$  成立 ( $i=1, 2, \dots, k$ )



# 函数依赖闭包

定义：在关系模式  $R\langle U, F \rangle$  中为  $F$  所逻辑蕴含的函数依赖的全体叫作  $F$  的闭包，记为  $F^+$ 。



➤问题：如何求出  $F^+$  中的一切函数依赖？

假设有关系模式  $R(U, F)$ ,  $U = \{X, Y, Z\}$ ,  $F = \{X \rightarrow Y, Y \rightarrow Z\}$ , 则  $F^+$  为：

$$\left\{ \begin{array}{l} X \rightarrow \Phi, XY \rightarrow \Phi, XZ \rightarrow \Phi, XYZ \rightarrow \Phi, Y \rightarrow \Phi, YZ \rightarrow \Phi, Z \rightarrow \Phi \\ X \rightarrow X, XY \rightarrow X, XZ \rightarrow X, XYZ \rightarrow X, Y \rightarrow Y, YZ \rightarrow Y, Z \rightarrow Z \\ X \rightarrow Y, XY \rightarrow Y, XZ \rightarrow Y, XYZ \rightarrow Y, Y \rightarrow Z, YZ \rightarrow Z, \\ X \rightarrow Z, XY \rightarrow Z, XZ \rightarrow Z, XYZ \rightarrow Z, Y \rightarrow YZ, YZ \rightarrow YZ, \\ X \rightarrow XY, XY \rightarrow XY, XZ \rightarrow XY, XYZ \rightarrow XY, \\ X \rightarrow XZ, XY \rightarrow XZ, XZ \rightarrow XZ, XYZ \rightarrow XZ, \\ X \rightarrow YZ, XY \rightarrow YZ, XZ \rightarrow YZ, XYZ \rightarrow YZ, \\ X \rightarrow XYZ, XY \rightarrow XYZ, XZ \rightarrow XYZ, XYZ \rightarrow XYZ, \end{array} \right.$$



➤问题：如何求出  $F^+$  中的一切函数依赖？

例如，  $F=\{X \rightarrow A_1, \dots, X \rightarrow A_n\}$ ，至少可以推导出多少个函数依赖？

至少可以推导出  $2^n$  个函数依赖

这个问题属于计算复杂性问题：NP完全问题。

能否找到判断  $X \rightarrow Z \in F^+$  是否成立的有效办法？





## 属性集闭包

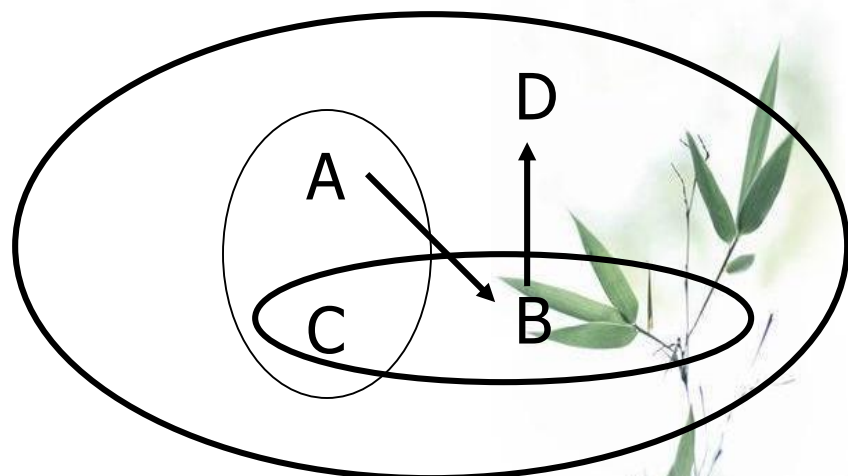
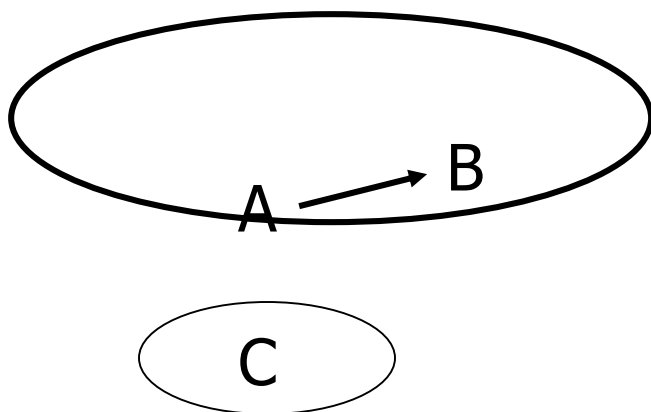
**定义6.13** 设 $F$ 为属性集 $U$ 上的一组函数依赖,  $X$ 、 $Y \subseteq U$ ,  $X_F^+ = \{ A_i | X \rightarrow A_i \text{ 能由 } F \text{ 根据Armstrong公理导出} \}$ ,  $X_F^+$ 称为属性 $X$ 关于函数依赖集 $F$ 的闭包。



例：关系 $R\langle U, F \rangle$ 中：

$U=ABCD$ ,  $F=\{A \rightarrow B, BC \rightarrow D\}$ ;

- $A_F^+ = AB$
- $C_F^+ = C$
- $(AC)_F^+ = ABCD$



## 关于闭包的引理

- 引理6.2

设 $F$ 为属性集 $U$ 上的一组函数依赖,  $X, Y \subseteq U$ ,  $X \rightarrow Y$ 能由 $F$ 根据Armstrong公理导出的充分必要条件是 $Y \subseteq X_F^+$

- 用途

将判定 $X \rightarrow Y$ 是否能由 $F$ 根据Armstrong公理导出的问题, 转化为求出 $X_F^+$ 、判定 $Y$ 是否为 $X_F^+$ 的子集的问题



# 求闭包的算法

**算法6.1** 求属性集 $X$  ( $X \subseteq U$ ) 关于 $U$ 上的函数依赖集 $F$ 的闭包 $X_F^+$

输入:  $X, F$

输出:  $X_F^+$

步骤:

(1) 令 $X^{(0)} = X, i=0$

(2) 求 $B$ , 这里 $B = \{ A \mid (\exists V)(\exists W)(V \rightarrow W \in F \wedge V \subseteq X^{(i)} \wedge A \in W) \}$ ;

(3)  $X^{(i+1)} = B \cup X^{(i)}$

(4) 判断 $X^{(i+1)} = X^{(i)}$  吗?

(5) 若相等或 $X^{(i)} = U$ , 则 $X^{(i)}$  就是 $X_F^+$ , 算法终止。

(6) 若否, 则  $i=i+1$ , 返回第 (2) 步。



# 函数依赖闭包

[例1] 已知关系模式  $R\langle U, F\rangle$ , 其中

$U=\{A, B, C, D, E\};$

$F=\{AB\rightarrow C, B\rightarrow D, C\rightarrow E, EC\rightarrow B, AC\rightarrow B\}。$

求  $(AB)_{F^+}$ 。

(所有可以推出来的属性)

解 设  $X^{(0)}=AB;$

(1)  $X^{(1)}=AB\cup CD=ABCD \neq U。$

(2)  $X^{(0)} \neq X^{(1)}$

$X^{(2)}=X^{(1)} \cup BE=ABCDE。$

(3)  $X^{(2)}=U$ , 算法终止

$\rightarrow (AB)_{F^+}=ABCDE。$



# 闭包的计算

- 示例1

$R \langle U, F \rangle$ ,  $U = (A, B, C, G, H, I)$ ,  $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$ , 计算  $(AG)_F^+$

所用依赖	$(AG)_F^+$
$A \rightarrow B$	AGB
$A \rightarrow C$	AGBC
$CG \rightarrow H$	AGBCH
$CG \rightarrow I$	AGBCHI
$(AG)_F^+$	= AGBCHI



对于算法6.1， 令  $a_i = |X^{(i)}|$ ，  $\{a_i\}$  形成一个步长大于1的严格递增的序列， 序列的上界是  $|U|$ ， 因此该算法最多  $|U| - |X|$  次循环就会终止。



判断下列关系模式最高属于几范式,说明理由。

1)  $R(A,B,C,D)$   $F=\{B \rightarrow D, AB \rightarrow C\}$

2)  $R(A,B,C)$   $F=\{A \rightarrow B, C \rightarrow A\}$





# 候选关键字的求解理论

## ➤ 属性闭包判别思想

根据属性在函数依赖中出现的位置，将属性分为左**L**,右**R**和非**N**等几类，分别对这几类属性是否满足候选关键字的定义做判断。



# 候选关键字的求解理论

□ 给定 $R(A_1, A_2, A_3 \dots A_M)$  和函数依赖集 $F$ ，则属性分为四类：

- ✓ **L类**：仅出现在 $F$ 中左部的属性
- ✓ **R类**：仅出现在 $F$ 中右部的属性
- ✓ **N类**：在 $F$ 的左部、右部均不出现的属性
- ✓ **LR类**：在 $F$ 的左部、右部均出现的属性



# 候选关键字的求解理论

- 定理1: 对给定的 $R(U, F)$ , 若 $X \subseteq U$ 是L类属性, 则 $X$ 必为 $R$ 的任一候选关键字的成员。  
推论1: 若 $X^+$ 包含了 $R$ 的全部属性, 则 $X$ 为 $R$ 的唯一候选关键字。
- 例:  $R(A, B, C, D)$ ,  
 $F = \{D \rightarrow B, B \rightarrow D, AD \rightarrow B, AC \rightarrow D\}$ ,  
 $A, C$ 是L类属性, 是候选关键字的成员,  
 $(AC)^+ = ABCD$ ,  $AC$ 为 $R$ 的唯一候选关键字。



# 候选关键字的求解理论

- 定理2：对给定的 $R(U, F)$ ，若 $X \subseteq U$ 是 $N$ 类属性，则 $X$ 必为 $R$ 的任一候选关键字的成员。
- 定理3：对给定的 $R(U, F)$ ，若 $X \subseteq U$ 是 $R$ 类属性，则 $X$ 不在 $R$ 的任一候选关键字中。
- ❖ 推论2：对给定的 $R(U, F)$ ，若 $X \subseteq U$ 是 $N$ 和 $L$ 类属性的集合，且 $X^+$ 包含了 $R$ 的全部属性，则 $X$ 为 $R$ 的唯一候选关键字。



# 候选关键字的求解理论


- 定理1-3确定了L类和N类属性必是任何候选关键字的成员，而R类属性必不是任何候选关键字的成员，简化了求解候选关键字的方法，从而降低了计算候选关键字的复杂度。



## 候选码求解算法

输入：关系模式R及其函数依赖集F。

输出：R的所有候选码。

- (1) 将R的所有属性分为L、R、N、LR四类，令X代表L、N两类属性，Y代表LR类属性。
  - (2) 求 $X_F^+$ 。若 $X_F^+$ 包含R的全部属性，则X即为R的**唯一**候选码，转(5)；否则，转(3)。
  - (3) 在Y中任取一属性A，求 $(AX)_F^+$ 。若它包含了R的全部属性，则转(4)；否则，调换一属性，反复进行这一过程，直到试完Y中的所有属性。
  - (4) 如果已找出所有候选码，则转(5)；否则，在Y中依次取2个、3个、...，求它们的属性闭包，直到其闭包包含R的全部属性。
  - (5) 停止，输出结果。
- 

# 候选码的求解：例1

- 设关系模式 $R(A, B, C, D, E, P)$ ，其函数依赖集：

$$F = \{A \rightarrow D, E \rightarrow D, D \rightarrow B, BC \rightarrow D, DC \rightarrow A\}$$

求 $R$ 的所有候选码。

- 解： L类： C, E

R类：

N类： P

LR类： A, B, D

CEP

- 因为 $(CEP)_F^+ = CEPDBA$ ，所以 $CEP$ 是 $R$ 的唯一候选码。



## 候选码的求解：例2

- 设关系模式 $R(S, D, I, B, O, Q)$ , 其函数依赖集:

$$F = \{ S \rightarrow D, I \rightarrow B, B \rightarrow O, O \rightarrow Q, Q \rightarrow I \}$$

求 $R$ 的所有候选码。

- 解: L类( $S$ ); R类( $D$ ) ; N类(无); LR类( $I, B, O, Q$ )

因为 $S^+ = SD$ , 所以 $S$ 不是 $R$ 的候选码;

因为 $(SI)^+ = SIDBOQ$ , 所以 $SI$ 是一个候选码;

因为 $(SB)^+ = SBDOQI$ , 所以 $SB$ 也是一个候选码;

因为 $(SO)^+ = SODQIB$ , 所以 $SO$ 也是一个候选码;

因为 $(SQ)^+ = SQDIBO$ , 所以 $SQ$ 也是一个候选码。





# 函数依赖集等价

**定义6.14** 如果  $G^+ = F^+$ ，就说函数依赖集  $F$  **覆盖**  $G$ （ $F$  是  $G$  的覆盖，或  $G$  是  $F$  的覆盖），或  $F$  与  $G$  **等价**。

**引理6.3**  $F^+ = G^+$  的充分必要条件是  $F \subseteq G^+$ ，和  $G \subseteq F^+$



# 最小依赖集

**定义6.15** 如果函数依赖集 $F$ 满足下列条件，则称 $F$ 为一个极小函数依赖集。亦称为最小依赖集或最小覆盖。

(1)  $F$ 中任一函数依赖的右部仅含有一个属性。

(2)  $F$ 中不存在这样的函数依赖 $X \rightarrow A$ ， $X$ 有真子集 $Z$ 使得 $F - \{X \rightarrow A\} \cup \{Z \rightarrow A\}$ 与 $F$ 等价。

(3)  $F$ 中不存在这样的函数依赖 $X \rightarrow A$ ，使得 $F$ 与 $F - \{X \rightarrow A\}$ 等价。



## 最小依赖集

[例2] 关系模式  $S\langle U, F\rangle$ , 其中:

$U=\{ \text{Sno}, \text{Sdept}, \text{Mname}, \text{Cno}, \text{Grade} \},$

$F=\{ \text{Sno} \rightarrow \text{Sdept}, \text{Sdept} \rightarrow \text{Mname}, (\text{Sno}, \text{Cno}) \rightarrow \text{Grade} \}$

设  $F'=\{ \text{Sno} \rightarrow \text{Sdept}, \text{Sno} \rightarrow \text{Mname}, \text{Sdept} \rightarrow \text{Mname},$

$(\text{Sno}, \text{Cno}) \rightarrow \text{Grade}, (\text{Sno}, \text{Sdept}) \rightarrow \text{Sdept} \}$

$F$  是最小覆盖, 而  $F'$  不是。

因为:  $F' - \{ \text{Sno} \rightarrow \text{Mname} \}$  与  $F'$  等价

$F' - \{ (\text{Sno}, \text{Sdept}) \rightarrow \text{Sdept} \}$  也与  $F'$  等价



# 极小化过程

**定理6.3** 每一个函数依赖集 $F$ 均等价于一个极小函数依赖集 $F_m$ 。此 $F_m$ 称为 $F$ 的最小依赖集。



## 极小化过程（续）

(1) 逐一检查  $F$  中各函数依赖  $FD_i: X \rightarrow Y$ , 若  $Y = A_1 A_2 \dots A_k$ ,  $k > 2$ , 则用  $\{X \rightarrow A_j | j=1, 2, \dots, k\}$  来取代  $X \rightarrow Y$ 。

(2) 逐一取出  $F$  中各函数依赖  $FD_i: X \rightarrow A$ , 设  $X = B_1 B_2 \dots B_m$ , 逐一考查  $B_i$  ( $i=1, 2, \dots, m$ ), 若  $A \in (X - B_i)_F^+$ , 则以  $X - B_i$  取代  $X$ 。

(3) 逐一检查  $F$  中各函数依赖  $FD_i: X \rightarrow A$ , 令  $G = F - \{X \rightarrow A\}$ , 若  $A \in X_G^+$ , 则从  $F$  中去掉此函数依赖。

即：去掉  $X \rightarrow A$  后， $X$  是否仍可推出  $A$



## 极小化过程（续）

[例3]  $F = \{A \rightarrow B, B \rightarrow A, \cancel{B \rightarrow C}, A \rightarrow C, C \rightarrow A\}$

$F_{m1}$ 、 $F_{m2}$ 都是 $F$ 的最小依赖集：

$$F_{m1} = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\} \checkmark$$

$$F_{m2} = \{A \rightarrow B, B \rightarrow A, A \rightarrow C, C \rightarrow A\} \checkmark$$

- $F$ 的最小依赖集 $F_m$ 不唯一
- 极小化过程也是检验 $F$ 是否为极小依赖集的一个算法



## 极小化过程（续）

$$F = \{A \rightarrow B, B \rightarrow A, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$$

- 检查  $A \rightarrow B$ ,  $G = F - \{A \rightarrow B\} = \{B \rightarrow A, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$

$$(A)_G^+ = \{A, C\}, B \notin \{A, C\}$$

- 检查  $B \rightarrow A$ ,  $G = F - \{B \rightarrow A\} = \{A \rightarrow B, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$

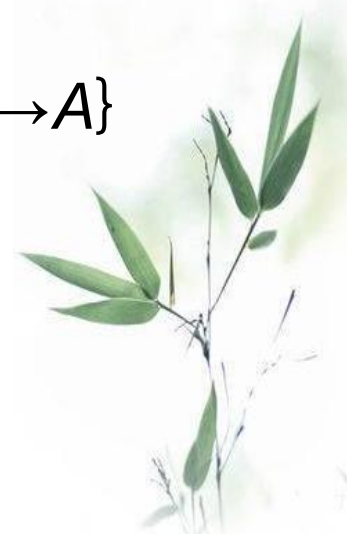
$$(B)_G^+ = \{A, B, C\}, A \in \{A, B, C\}$$

所以从F中删除  $B \rightarrow A$ .

- 检查  $A \rightarrow C$ ,  $G = F - \{A \rightarrow C\} = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$

$$(A)_G^+ = \{A, B, C\}, C \in \{A, B, C\}$$

所以从F中删除  $A \rightarrow C$ .



$A \rightarrow B, A \rightarrow C, B \rightarrow A$   
例：已知关系 $R(ABC)$ ， $F = \{A \rightarrow BC, B \rightarrow C, BC \rightarrow A\}$ ，求 $F$ 的最小依赖集。

① 将 $F$ 中所有函数依赖右部不为单属性的化为单属性

$$F = \{A \rightarrow B, A \rightarrow C, B \rightarrow C, BC \rightarrow A\}$$

$$\{A \rightarrow B, B \rightarrow C, B \rightarrow A\}$$

② 去掉 $F$ 中多余的函数依赖

- 对于 $A \rightarrow B$ ，令 $G = \{A \rightarrow C, B \rightarrow C, BC \rightarrow A\}$ ，求 $A_G^+ = AC$ ， $\because B \notin AC$ ， $\therefore A \rightarrow B$ 不能去掉；
  - 对于 $A \rightarrow C$ ，令 $G = \{A \rightarrow B, B \rightarrow C, BC \rightarrow A\}$ ，求 $A_G^+ = ABC$ ， $\because C \subseteq ABC$ ， $\therefore$ 从 $F$ 中**去掉** $A \rightarrow C$ ，  
 $F = \{A \rightarrow B, B \rightarrow C, BC \rightarrow A\}$ ；
  - 对于 $B \rightarrow C$ ，令 $G = \{A \rightarrow B, BC \rightarrow A\}$ ，求 $B_G^+ = B$ ， $\because C \notin B$ ， $\therefore B \rightarrow C$ 不能去掉；
  - 对于 $BC \rightarrow A$ ，令 $G = \{B \rightarrow C, A \rightarrow B\}$ ，求 $BC_G^+ = BC$ ， $\because A \notin BC$ ， $\therefore BC \rightarrow A$ 不能去掉；
- $$F = \{A \rightarrow B, B \rightarrow C, BC \rightarrow A\}$$

③ 去掉 $F$ 中函数依赖左部多余属性

- 对于 $BC \rightarrow A$ 
  - 对于 $B$ ，求 $C_F^+ = C$ ， $\because A \notin C_F^+$ ， $\therefore B$ 不是多余属性；
  - 对于 $C$ ，求 $B_F^+ = ABC$ ， $\because A \subseteq B_F^+$ ， $\therefore C$ 是多余属性，将 $BC \rightarrow A$ 替换为 $B \rightarrow A$ 。

$$F = \{A \rightarrow B, B \rightarrow C, B \rightarrow A\}$$

④ 重复步骤②③， $F$ 不再变化， $\therefore F = \{A \rightarrow B, B \rightarrow C, B \rightarrow A\}$



练习：将下列函数依赖集F化为最小依赖集

$$F = \left( \begin{array}{ll} AB \rightarrow C & D \rightarrow EG \\ C \rightarrow A & BE \rightarrow C \\ BC \rightarrow D & CG \rightarrow BD \\ ACD \rightarrow B & CE \rightarrow AG \end{array} \right)$$



# 第六章 关系数据理论

**6.1 问题的提出**

**6.2 规范化**

**6.3 数据依赖的公理系统**

**\*6.4 模式的分解**

**6.5 小结**



## 6.4 模式的分解

- 把低一级的关系模式分解为若干个高一级的关系模式的方法不是唯一的
- 只有能够保证分解后的关系模式与原关系模式等价，分解方法才有意义



## 模式的分解（续）

**定义6.16** 关系模式 $R\langle U, F \rangle$ 的一个分解：

$$\rho = \{ R_1\langle U_1, F_1 \rangle, R_2\langle U_2, F_2 \rangle, \dots, R_n\langle U_n, F_n \rangle \}$$

$U = \bigcup_{i=1}^n U_i$ ，且不存在  $U_i \subseteq U_j$ ， $F_i$  为  $F$  在  $U_i$  上的投影

**定义6.17** 函数依赖集合  $\{X \rightarrow Y \mid X \rightarrow Y \in F^+ \wedge XY \subseteq U_i\}$  的一个覆盖  $F_i$  叫作

$F$  在属性  $U_i$  上的投影



# 模式分解的定义

- 分解的基本代数运算
  - 投影
  - 自然连接
- 分解的目标
  - 无损连接分解
  - 保持函数依赖
  - 达到更高级范式



例: SL (Sno, Sdept, Sloc)

$F = \{ Sno \rightarrow Sdept, Sdept \rightarrow Sloc, Sno \rightarrow Sloc \}$

$SL \in 2NF$

存在插入异常、删除异常、冗余度大和修改复杂等问题

SL

Sno	Sdept	Sloc
95001	CS	A
95002	IS	B
95003	MA	C
95004	IS	B
95005	PH	B



# 1. SL分解为下面三个关系模式

SN(Sno)

SD(Sdept)

SO(Sloc)

SN	SD	SO
Sno	Sdept	Sloc
95001	CS	A
95002	IS	B
95003	MA	C
95004	PH	
95005		



2. SL分解为下面二个关系模式:

NL(Sno, Sloc)

DL(Sdept, Sloc)

分解后的关系为:

NL

Sno

Sloc

95001

A

95002

B

95003

C

95004

B

95005

B

DL

Sdept

Sloc

CS

A

IS

B

MA

C

PH

B





NL  $\bowtie$  DL

Sno	Sloc	Sdept
95001	A	CS
95002	B	IS
95002	B	PH
95003	C	MA
95004	B	IS
95004	B	PH
95005	B	IS
95005	B	PH

NL DL比原来的SL关系  
多了3个元组

无法知道95002、95004、  
95005

究竟是哪个系的学生

元组增加了，信息丢失了



3. 将SL分解为下面二个关系模式:

ND(Sno, Sdept)    NL(Sno, Sloc)

分解后的关系为:

ND		NL	
Sno	Sdept	Sno	Sloc
95001	CS	95001	A
95002	IS	95002	B
95003	MA	95003	C
95004	IS	95004	B
95005	PH	95005	B



ND⋈NL

Sno	Sdept	Sloc
95001	CS	A
95002	IS	B
95003	MA	C
95004	CS	A
95005	PH	B

与SL关系一样，因此没有丢失信息



## 模式分解中存在的问题

**R(A, B, C)**

A	B	C
1	1	2
2	2	1

$\Pi_{AB}(R)$

A	B
1	1
2	2

$\Pi_{BC}(R)$

B	C
1	2
2	1

$\Pi_{AB}(R) \bowtie \Pi_{BC}(R)$

A	B	C
1	1	2
2	2	1

无损分解

**R(A, B, C)**

A	B	C
1	1	1
2	1	2

$\Pi_{AB}(R)$

A	B
1	1
2	1

$\Pi_{BC}(R)$

B	C
1	1
1	2

$\Pi_{AB}(R) \bowtie \Pi_{BC}(R)$

A	B	C
1	1	1
1	1	2
2	1	1
2	1	2

有损分解

# 模式分解中存在的问题

$\{A \rightarrow B, B \rightarrow C\}$

A	B	C
a1	b1	c1
a2	b1	c1
a3	b2	c2
a4	b3	c1



插入

A	B
a1	b1
a2	b1
a3	b2
a4	b3
a5	b3



A	C
a1	c1
a2	c1
a3	c2
a4	c1
a5	c3

=

A	B	C
a1	b1	c1
a2	b1	c1
a3	b2	c2
a4	b3	c1
a5	b3	c3

违反  
 $B \rightarrow C$

# 无损连接分解

- 算法：（判别一个分解的无损连接性）

$$U = \{A_1, A_2, \dots, A_n\}$$

$$\rho = \{R_1 \langle U_1, F_1 \rangle, R_2 \langle U_2, F_2 \rangle, \dots, R_k \langle U_k, F_k \rangle\}$$

1. 建立一个n列k行的矩阵

$$TB = \{C_{ij} \mid \text{若 } A_j \in U_i, C_{ij} = a_j, \text{ 否则 } C_{ij} = b_{ij}\}$$

	$A_1$	$A_2$	$\dots$	$A_n$
$U_1$				
$\dots$				
$U_k$				

# 无损连接分解

2.对F中每一个函数依赖 $X \rightarrow Y$ ，若TB中存在元组 $t_1, t_2$ ，使得 $t_1[X]=t_2[X]$ ， $t_1[Y] \neq t_2[Y]$ ，则对每一个 $A_i \in Y$ ：

①若 $t_1[A_i], t_2[A_i]$ 中有一个等于 $a_j$ ，则另一个也改为 $a_j$ ；

②若①不成立，则取 $t_1[A_i] = t_2[A_i]$ （ $t_2$ 的行号小于 $t_1$ ）。

3.反复执行 2.，直至：

①TB中出现一行为 $a_1, a_2, \dots, a_n$ 的一行。

② TB不再发生变化，且没有一行为 $a_1, \dots, a_n$ 。

在①情况下， $\rho$ 为无损分解，否则为有损分解。



# 无损连接分解

- 示例一：  $U=\{A,B,C,D,E\}$ ,  $F=\{AB\rightarrow C, C\rightarrow D, D\rightarrow E\}$   
 $\rho =\{(A, B, C), (C, D), (D, E)\}$

$AB\rightarrow C$

	A	B	C	D	E
ABC	$a_1$	$a_2$	$a_3$	$b_{14}$	$b_{15}$
CD	$b_{21}$	$b_{22}$	$a_3$	$a_4$	$b_{25}$
DE	$b_{31}$	$b_{32}$	$b_{33}$	$a_4$	$a_5$

	A	B	C	D	E
ABC	$a_1$	$a_2$	$a_3$	$b_{14}$	$b_{15}$
CD	$b_{21}$	$b_{22}$	$a_3$	$a_4$	$b_{25}$
DE	$b_{31}$	$b_{32}$	$b_{33}$	$a_4$	$a_5$

$C\rightarrow D$

	A	B	C	D	E
ABC	$a_1$	$a_2$	$a_3$	$a_4$	$b_{15}$
CD	$b_{21}$	$b_{22}$	$a_3$	$a_4$	$b_{25}$
DE	$b_{31}$	$b_{32}$	$b_{33}$	$a_4$	$a_5$

$D\rightarrow E$

	A	B	C	D	E
ABC	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
CD	$b_{21}$	$b_{22}$	$a_3$	$a_4$	$a_5$
DE	$b_{31}$	$b_{32}$	$b_{33}$	$a_4$	$a_5$



# 无损连接分解

- 示例二：  $U=\{A,B,C,D,E\}$ ,  
 $F=\{A\rightarrow C, B\rightarrow C, C\rightarrow D, DE\rightarrow C, CE\rightarrow A\}$   
 $\rho =\{(A, D), (A, B), (B, E), (C, D, E), (A, E)\}$



# 无损连接分解

- 示例二：  $U=\{A,B,C,D,E\}$ ,  
 $F=\{A\rightarrow C, B\rightarrow C, C\rightarrow D, DE\rightarrow C, CE\rightarrow A\}$   
 $\rho =\{(A, D), (A, B), (B, E), (C, D, E), (A, E)\}$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{23}$	$b_{24}$	$b_{25}$
BE	$b_{31}$	$a_2$	$b_{33}$	$b_{34}$	$a_5$
CDE	$b_{41}$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$b_{53}$	$b_{54}$	$a_5$

$A\rightarrow C$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{13}$	$b_{24}$	$b_{25}$
BE	$b_{31}$	$a_2$	$b_{33}$	$b_{34}$	$a_5$
CDE	$b_{41}$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$b_{13}$	$b_{54}$	$a_5$

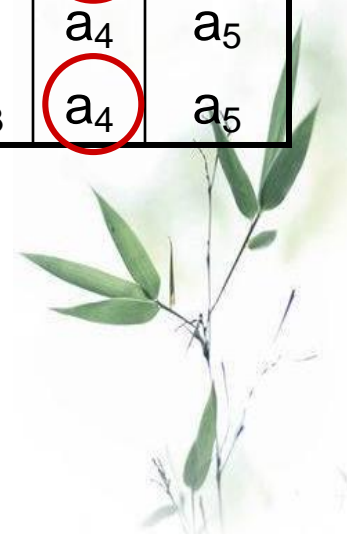
# 无损连接分解

$B \rightarrow C$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{13}$	$b_{24}$	$b_{25}$
BE	$b_{31}$	$a_2$	$b_{13}$	$b_{34}$	$a_5$
CDE	$b_{41}$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$b_{13}$	$b_{54}$	$a_5$

$C \rightarrow D$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{13}$	$a_4$	$b_{25}$
BE	$b_{31}$	$a_2$	$b_{13}$	$a_4$	$a_5$
CDE	$b_{41}$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$b_{13}$	$a_4$	$a_5$



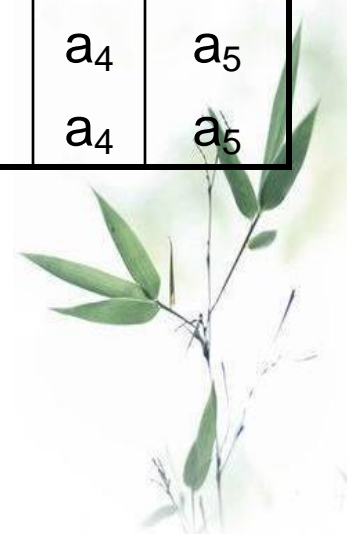
# 无损连接分解

$DE \rightarrow C$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{13}$	$a_4$	$b_{25}$
BE	$b_{31}$	$a_2$	$a_3$	$a_4$	$a_5$
CDE	$b_{41}$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$a_3$	$a_4$	$a_5$

$CE \rightarrow A$

	A	B	C	D	E
AD	$a_1$	$b_{12}$	$b_{13}$	$a_4$	$b_{15}$
AB	$a_1$	$a_2$	$b_{13}$	$a_4$	$b_{25}$
BE	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$
CDE	$a_1$	$b_{42}$	$a_3$	$a_4$	$a_5$
AE	$a_1$	$b_{52}$	$a_3$	$a_4$	$a_5$



# 无损连接分解

## - 定理

关系模式 $R(U)$ 的分解 $\rho = \{R_1, R_2\}$ ，则 $\rho$ 是一个无损连接分解的充要条件是  $R_1 \cap R_2 \rightarrow R_1 - R_2$ （或 $R_1 \cap R_2 \rightarrow R_2 - R_1$ ）成立

	$R_1 \cap R_2$	$R_1 - R_2$	$R_2 - R_1$
$R_1$	a...a	a...a	b...b
$R_2$	a...a	b...b	a...a

$$R_1 \cap R_2 \rightarrow R_1 - R_2$$

	$R_1 \cap R_2$	$R_1 - R_2$	$R_2 - R_1$
$R_1$	a...a	a...a	b...b
$R_2$	a...a	a...a	a...a



# 无损连接分解

$R=ABC$ ,  $F=\{A \rightarrow B\}$ ,

$\rho_1=\{R_1(AB), R_2(AC)\}$

$R_1 \cap R_2 = A$ ,  $R_1 - R_2 = B$

由  $A \rightarrow B$  , 得到  $\rho_1$  是无损连接分解

$\rho_2=\{R_1(AB), R_2(BC)\}$

$R_1 \cap R_2 = B$ ,  $R_1 - R_2 = A$ ,  $R_2 - R_1 = C$

$B \rightarrow A$ ,  $B \rightarrow C$  均不成立, 所以  $\rho_2$  不是无损连接分解



## 举例

- 设关系模式  $R(A, B, C, D, E, P)$ ,  $F = \{A \rightarrow B, C \rightarrow P, E \rightarrow A, CE \rightarrow D\}$ 。设有分解:

$$\rho = \{R_1(A, B, E), R_2(C, D, E, P)\}$$

判断分解  $\rho$  是否无损连接分解。

- 解: 因为  $R_1 \cap R_2 = E$ ,  $R_1 - R_2 = AB$ , 而  $E \rightarrow A$  和  $E \rightarrow B$  均成立 (即  $R_1 \cap R_2 \rightarrow R_1 - R_2$  成立), 所以  $\rho$  是无损连接分解。



# 保持函数依赖的模式分解

设关系模式 $R\langle U, F \rangle$ 被分解为若干个关系模式

$$R_1\langle U_1, F_1 \rangle, R_2\langle U_2, F_2 \rangle, \dots, R_n\langle U_n, F_n \rangle$$

（其中 $U = U_1 \cup U_2 \cup \dots \cup U_n$ ，且不存在 $U_i \subseteq U_j$ ， $F_i$ 为 $F$ 在 $U_i$ 上的投影），若 $F$ 所逻辑蕴含的函数依赖一定也由分解得到的某个关系模式中的函数依赖 $F_i$ 所逻辑蕴含，则称关系模式 $R$ 的这个分解是保持函数依赖的（Preserve dependency）





## 举例

- **【例】**关系模式 $R=\{CITY, ST, ZIP\}$ ，其中 $CITY$ 为城市， $ST$ 为街道， $ZIP$ 为邮政编码， $F=\{(CITY, ST) \rightarrow ZIP, ZIP \rightarrow CITY\}$ 。
- 如果将 $R$ 分解成 $R_1$ 和 $R_2$ ， $R_1=\{ST, ZIP\}$ ，  
 $R_2=\{CITY, ZIP\}$ ，  
检查分解是否具有无损连接和保持函数依赖。
- 解：1) 检查无损连接性。  
求得： $R_1 \cap R_2 = \{ZIP\}$ ； $R_2 - R_1 = \{CITY\}$ 。  
 $\because (ZIP \rightarrow CITY) \in F^+ \therefore$  分解具有无损连接性  
2) 检查分解是否保持函数依赖  
求得： $\pi R_1(F) = \Phi$ ； $\pi R_2(F) = \{ZIP \rightarrow CITY\}$ 。  
 $\therefore \pi R_1(F) \cup \pi R_2(F) = \{ZIP \rightarrow CITY\} \neq F^+$ 。
- $\therefore$  该分解不保持函数依赖。

