UNIT 12
数据依赖公理与无损分解

本讲主要目标

学完本讲后，你应该能够了解：

1、函数依赖的逻辑蕴含的定义；

2、Armstrong推理规则系统；

3、函数依赖集的闭包的概念，及属性集闭包的计算方法；

4、函数依赖集等价的定义，并能判断两个函数依赖集是否等价；

5、每个函数依赖集都等价于一个最小函数依赖集，并能计算最小函数依赖集；

6、关系模式的规范化过程就是关系模式的分解过程；是将一个关系模式分解成一组等价的关系模式的过程；

7、模式分解的等价包括了属性的等价、无损连接性和函数依赖的保持性

8、将关系模式分解为3NF，且既具有无损连接性又具有函数依赖保持性的算法；

9、将关系模式分解为BCNF，且具有无损连接性的算法。

本讲主要内容

一、公理及其推论

二、闭包的概念及其计算

三、函数依赖集的等价

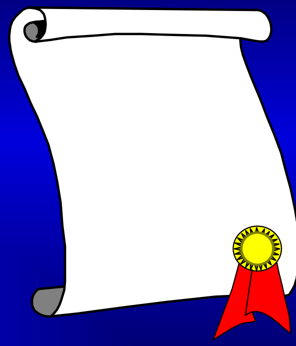
四、最小函数依赖集

五、等价模式分解的定义

六、无损连接性

七、函数依赖保持性

八、模式分解的算法



一、公理及其推论

1. 阿氏公理

在进行函数依赖的检查时，需要判断另外一些函数依赖是否成立，通过函数依赖的逻辑蕴涵即可做到。

定义4.11 设 F 是关系模式R的函数依赖集，X、Y是R的属性子集，如果从 F 的函数依赖中能够推出X→Y，则称F**逻辑蕴涵**X→Y。

一、公理及其推论

Armstrong公理（阿氏公理）：

对R<U, F> 有：

- A1自反律：若Y⊆X，则X→Y。
- A2增广律：若X→Y，则XZ→YZ。
- A3传递律：若X→Y、Y→Z，则X→Z。

- 由自反律：所有的平凡函数依赖都是成立的
- 由增广律：函数依赖两边增加相同属性也成立
- 由传递律：由已知函数依赖可以推导出新依赖

一、公理及其推论

2、公理的推论：

（1）合并规则：若X→Y 、 X→Z，则X→YZ。

（2）分解规则：若X→YZ，则X→Y, X→Z。

（3）伪传递规则：若X→Y 、WY→Z，则WX→Z。

（4）复合规则：若X→Y 、W→V，则XW→YV。

一、公理及其推论

3、公理系统的特性

建立公理体系的目的在于有效而准确的从已知的函数依赖推出未知的函数依赖，公理系统满足两个特性：

(1) **正确性**：按阿氏公理推出的依赖都是正确的

(2) **完备性**：能推出所有的依赖

- **完备性等价于**：所有不能用公理系统推出的依赖都不成立，即不能由F逻辑蕴涵
- **或者**存在一个具体关系r，F中所有函数依赖都满足r，但不能用公理推出的$X \rightarrow Y$却不满足r。

本讲主要内容

一、公理及其推论

二、闭包的概念及其计算

三、函数依赖集的等价

四、最小函数依赖集

五、等价模式分解的定义

六、无损连接性

七、函数依赖保持性

八、模式分解的算法

二、闭包的概念及其计算

1、函数依赖集F的闭包

定义4.12 设有关系模式R<U, F>, X、Y为U的属性子集，函数依赖集F的闭包定义为：

$F^+=\{X\rightarrow Y \mid X\rightarrow Y$ 基于阿氏公理推出$\}$

即F^+为F所逻辑蕴含的函数依赖全体。

$F^+=\{$ ① F中的函数依赖，由属性语义决定；

② 由F推出的非平凡的函数依赖；

③ 由F推出的平凡的函数依赖：$A\rightarrow\varphi$、$A\rightarrow A$、$AB\rightarrow A$、….这一类函数依赖与F无关，对R中任何属性都成立

$\}$

二、闭包的概念及其计算

例：有关系模式R(A, B, C)，它的函依赖集F={A→B, B→C}，计算F的闭包。

解：

(1) F中的函数依赖： A→B，B→C

(2) 由F推出的非平凡函数依赖： A→C，AC→B…

(3) 由F推出的平凡函数依赖： A→A，AC→A，…

F^+的计算很麻烦，可能会非常大

$R\langle U,F\rangle$， $U=(X,Y,Z)$， $F=\{X\to Y, Y\to Z\}$,

$F^+ = \{$

$X\to X$, $Y\to Y$, $Z\to Z$, $XY\to X$, $XZ\to X$, $YZ\to Y$, $XYZ\to X$,

$X\to Y$, $Y\to Z$, $XY\to Y$, $XZ\to Y$, $YZ\to Z$, $XYZ\to Y$,

$X\to Z$, $Y\to YZ$, $XY\to Z$, $XZ\to Z$, $YZ\to YZ$, $XYZ\to Z$,

$X\to XY$, $XY\to XY$, $XZ\to XY$, $XYZ\to XY$,

$X\to XZ$, $XY\to YZ$, $XZ\to XZ$, $XYZ\to YZ$

$X\to YZ$, $XY\to XZ$, $XZ\to XY$, $XYZ\to XZ$,

$XY\to XYZ$, $XZ\to XYZ$, $XYZ\to XYZ$

$\}$

二、闭包的概念及其计算

2、属性集X的闭包

定义4.13 设F是属性集合U上的一个函数依赖集，$X \subseteq U$：

$X_F^+ = \{A \mid X \rightarrow A$ 能由F用阿氏公理导出$\}$

X_F^+称为属性集X关于F的闭包，也可简写为X^+。

X_F^+是由X从F中推出的所有函数依赖右部的集合。

例如：R(A, B, C)中，$F=\{A \rightarrow B, B \rightarrow C\}$，则

$A_F^+=ABC$，$B_F^+=BC$，$C_F^+=C$

属性闭包的计算比F的闭包要容易地多。

二、闭包的概念及其计算

定理4.6：X→Y能从F中用阿氏公理导出的充要条件是：$Y \subseteq X_F^+$

3、属性集X闭包的计算

F的闭包F^+计算起来相当麻烦，且其中有很多冗余信息，因此计算F^+是不现实的。

判断X →Y在不在F^+中，只要判断Y是否属于X^+。因此计算F^+的问题可以变换成计算X^+的问题，而X^+的计算相对简单。

二、闭包的概念及其计算

算法4.1 ： 求属性集X关于F的闭包X_F^+(X^+)。

- 算法：

设 R<U, F>，A为U中属性(集)。

(1) $X^{(0)}=X$

(2) $X^{(i+1)}=X^{(i)}\bigcup A$

其中A：

对F中任一个没有使用过的$Y_j \rightarrow Z_j$，且$Y_j \subseteq X^{(i)}$；

所有的这些Z_j构成A

求得$X^{(i+1)}$ 后，对$Y_j \rightarrow Z_j$做已使用的标记。

如果找不到这样的A，则转(4)

(3) 若$X^{(i+1)}=X^{(i)}$ 或 $X^{(i+1)}=U$则转(4)，否则转(2)。

(4) 结束，输出$X^{(i)}$，即为X^+

算法会终止吗？

最多|U-X|步

二、闭包的概念及其计算

- 例如：设有关系模式R<U, F>，其中U={A, B, C, D, E, I}，

  F={A→D, AB →E, BI →E, CD →I, E →C}

  计算$(AE)^+$

解：

(1) $X^0=\{AE\}$

(2) $X^1=X^0 \cup \{DC\}=\{ACDE\}$，用了A→D和E→C

(3) $X^2=X^1 \cup \{I\}=\{ACDEI\}$，用了CD →I

(4) F中没有左部属于X^2的函数依赖，X^2为$(AE)^+$

二、闭包的概念及其计算

- 例如：设有关系模式R<U, F>，其中U={A, B, C, D, E, G}，F={AB→C, C→A, BC→D, ACD→B, D→EG, BE →C, CG →BD, CE →AG}，求$(BD)^+$

解：

(1) X^0={BD}

(2) X^1={BDEG}　　D→EG

(3) X^2={BCDEG}　　BE →C

(4) X^3={ABCDEG}　　C→A

(5) X^3=U，所以$(BD)^+=X^3$

二、闭包的概念及其计算

- 判断闭包计算结束的方法：

（1）$X^{(i+1)}=X^i$

（2）当发现X^i中包含了所有的属性

（3）当F中的函数依赖的右边再也找不到X^i中未出现过的属性

（4）在 F 中未用过的函数依赖的左边已没有X^i的子集

二、闭包的概念及其计算

4、总结：属性集闭包的作用

- **测试超键**

　　如果X^+包含所有R的所有属性，那么X是R的超键

- **检测函数依赖**

　　判断X->Y 是否成立，只需判断$Y \subseteq X^+$ 。计算X^+，然后判断这个属性集闭包是否包含Y

- **计算F的函数依赖集闭包F^+**

　　计算所有可能的属性子集的属性集闭包，综合得到函数依赖集闭包*

*https://www.icourse163.org/learn/CQU-1450125165?tid=1465870482#/learn/content?type=detail&id=1245867219&sm=1

课堂练习

1. 关系模式R<U, F>， U={A, B, C, D}， F={A→B, AC→D , AB →C}， 给出R的候选键？

2. U={A, B, C, D, E}， F={A→B, B →C, CD →E}判断A →E是否成立？

本讲主要内容

一、 公理及其推论

二、 闭包的概念及其计算

三、 函数依赖集的等价

四、 最小函数依赖集

五、 等价模式分解的定义

六、 无损连接性

七、 函数依赖保持性

八、 模式分解的算法

三、函数依赖集的等价

1、等价定义

定义4.14：如果$F^+=G^+$，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。

2、有关性质

(1) 若$G\subseteq F$，则$G^+\subseteq F^+$

(2) $(F^+)^+=F^+$

三、函数依赖集的等价

3、判断方法

定理4.9：F与G等价的充分必要条件是$F \subseteq G^+$，和$G \subseteq F^+$

。

- 该定理给出了判断函数依赖集F和G是否等价的可行算。

根据定理4.9：只需$F \subseteq G^+$和$G \subseteq F^+$，即证

对每个$T \in F$，有$T \in G^+$；对每个$S \in G$，有$S \in F^+$，T和S是形如$X \rightarrow Y$的属性依赖。

而验证$X \rightarrow Y \in G^+$，根据定理4.6：只需$Y \subseteq X_G^+$

转为计算X_G^+

三、函数依赖集的等价

例如：$F=\{A\rightarrow B, B\rightarrow C\}$，$G=\{A\rightarrow BC, B\rightarrow C\}$，判断F和G是否等价。

解：（1）先检查F中的每一个函数依赖是否属于G^+。

$\because A_G^+=ABC$，$\therefore B\subseteq A_G^+$，$\therefore A\rightarrow B\in G^+$（定理4.6）

又$\because B_G^+=BC$，$\therefore C\subseteq B_G^+$，$\therefore B\rightarrow C\in G^+$

$\therefore F\subseteq G^+$

（2）然后检查G中的每一个函数依赖是否属于F^+。

$\because A_F^+=ABC$，$\therefore BC\subseteq A_F^+$，$\therefore A\rightarrow BC\in F^+$

又$\because B_F^+=BC$，$\therefore C\subseteq B_F^+$，$\therefore B\rightarrow C\in F^+$

$\therefore G\subseteq F^+$

由（1）和（2）可得F和G等价。（定理4.9）

本讲主要内容

一、 公理及其推论

二、 闭包的概念及其计算

三、 函数依赖集的等价

四、 最小函数依赖集

五、 等价模式分解的定义

六、 无损连接性

七、 函数依赖保持性

八、 模式分解的算法

四、最小函数依赖集

定义4.15：若F满足下列条件，则称其为一个最小函数依赖集Fm。

（1）F中每个函数依赖的右部都是单属性；

（2）对于F的任一函数依赖X→A，F-{X→A}与F都不等价，即无多余函数依赖；

（3）对于F中的任一函数依赖X→A和X的真子集X'，(F-(X→A))U{X'→A}与F都不等价，即左部无多余属性。

最小：（1）F中每个函数依赖的右部没有多余的属性；
（2）F中不存在多余的函数依赖；
（3）F中每个函数依赖的左部没有多余的属性。

四、最小函数依赖集

例如：哪个是最小依赖集？

(1) $F1=\{A \to D, BD \to C, C \to AD\}$

(2) $F2=\{AB \to C, B \to A, B \to C\}$

(3) $F3=\{BC \to D, D \to A, A \to D\}$

F1中第三个依赖的右部不是单属性

F2中第一个依赖左部有多余属性A

F3满足最小依赖集的三个条件

四、最小函数依赖集

定理4.10：每个F与Fm等价。

算法4.2 计算最小函数依赖集Fm

(1) 分解：使F中任一函数依赖的右部仅含有单属性。

(2) 最小化左边的多余属性：

方法：对F中任一$XY \to A$，在F中求X^+，

若$A \subseteq X^+$，则Y为多余的。

(3) 删除冗余的函数依赖：

方法：对F中任一$X \to A$，在$F-\{X \to A\}$中求X^+，

若$A \subseteq X^+$，则$X \to A$为多余的。

四、最小函数依赖集

例：设有=$\{B \to C,\ C \to AB,\ BC \to A\}$，求与F等价的最小函数依赖集。

- 分解$C \to AB$，$F=\{B \to C,\ C \to A,\ C \to B,\ BC \to A\}$
- 判断$BC \to A$。$B^+ = ABC$，$A \subseteq B+$，则C在$BC \to A$中是多余的。

  $F=\{B \to C,\ C \to A,\ C \to B,\ B \to A\}$
- 判断$B \to C$是否冗余，$F' =\{C \to A,\ C \to B,\ B \to A\}$

  $B^+ = BA$，$C \not\subseteq B^+$，$B \to C$非冗余。$F=\{B \to C,\ C \to A,\ C \to B,\ B \to A\}$

判断$C \to A$是否冗余，$F' =\{B \to C,\ C \to B,\ B \to A\}$

$C^+ = ABC$，$A \subseteq C^+$，$C \to A$冗余。$F=\{B \to C,\ C \to B,\ B \to A\}$

判断$C \to B$是否冗余，$F' =\{B \to C,\ B \to A\}$

$C^+ = C$，$B \not\subseteq C^+$，$C \to B$非冗余。$F=\{B \to C,\ C \to B,\ B \to A\}$

判断$B \to A$是否冗余，$F' =\{B \to C,\ C \to B\}$

$B^+ = BC$，$A \not\subseteq B^+$，$B \to A$非冗余。$Fm=\{B \to C,\ C \to B,\ B \to A\}$

四、最小函数依赖集

例如：设有函数依赖集

$F=\{A\rightarrow B, ABCD\rightarrow E, EF\rightarrow G, EF\rightarrow H, ACDF\rightarrow EG\}$

求与F等价的最小函数依赖集。

$Fm = \{A\rightarrow B, ACD\rightarrow E, EF\rightarrow G, EF\rightarrow H\}$

注意：一个函数依赖集的最小集不是惟一的。

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

$Fm1=\{A\rightarrow B, B\rightarrow C, C\rightarrow A\}$，

$Fm2=\{A\rightarrow B, B\rightarrow A, A\rightarrow C, C\rightarrow A\}$。

方法1：无多余属性；依次判断B->A，A->C是否冗余；

方法2：无多余属性；依次判断B->C是否冗余。

本讲主要内容

一．公理及其推论

二．闭包的概念及其计算

三．函数依赖集的等价

四．最小函数依赖集

五．等价模式分解的定义

六．无损连接性

七．函数依赖保持性

八．模式分解的算法

五、等价模式分解的定义

1、模式分解

- **关系模式的分解过程**实际上就是将一个关系模式分解成一组**等价**的关系子模式的过程。
- 关系模式是五元组R（U，D，dom，F）。
  - U 属性名集合
  - D 为属性所来自的域　比如 SexType
  - DOM 为属性向域的映像集合 比如 DOM（sex）=SexType
  - F 属性间数据依赖关系集合
- 模式的分解包括三个方面：
  - 属性的分解
  - 关系的分解
  - 函数依赖的分解

五、等价模式分解的定义

2、等价模式分解

对于一个关系模式的分解是多种多样的，但是分解后产生的模式应与原来的模式等价：

（1）分解后子模式的属性集与原模式属性集相同

（2）无损连接性

查询时的连接操作是否会丢失某些信息或多出某些信息。

（3）保持函数依赖

分解后的模式满足的函数依赖集是否能蕴含分解前的依赖集。

五、等价模式分解的定义

- (1) 模式分解的**属性**等价性 —— 如果关系模式$R(A)$被分解为关系模式$R_1(A_1), R_2(A_2), \ldots, R_n(A_n)$，且

  $$A = A_1 \cup A_2 \cup \ldots \cup A_n$$

  则该分解是属性等价的分解

- (2) 模式分解的无损连接性 —— 如果关系模式$R(A)$被分解为关系模式$R_1(A_1), R_2(A_2), \ldots, R_n(A_n)$，且

  $$R = R_1 \bowtie R_2 \bowtie \ldots \bowtie R_n$$

  则该分解是**无损连接**的分解

五、等价模式分解的定义

例如：关系模式R(A,B,C)，分解为

ρ1={R1(A),R2(B),R3(C)}

满足属性等价，但无法恢复r，不是无损的。

在对关系模式进行分解的过程中，满足属性等价且有冗余属性的分解一定具有无损连接性吗？

五、等价模式分解的定义

关系模式R(A,B,C)，分解为ρ2={R1(AC)， R2(BC)}

A	B	C
a_1	b_1	c_1
a_2	b_2	c_1

分解 ⇨

A	C
a_1	c_1
a_2	c_1

B	C
b_1	c_1
b_2	c_1

连接 ⇨

A	B	C
a_1	b_1	c_1
a_1	b_2	c_1
a_2	b_1	c_1
a_2	b_2	c_1

结论：满足属性等价且有冗余属性的分解不一定具有无损连接性

五、等价模式分解的定义

- (3)模式分解的函数依赖的保持性——如果关系模式R(A, F)被分解为关系模式$R_1(A_1, F_1)$，$R_2(A_2, F_2)$, …, $R_n(A_n, F_n)$, 且

$$F^+ = \left(\bigcup_{i=1}^{i=n} F_i\right)^+$$

则该分解是具有**函数依赖保持性**的分解

五、等价模式分解的定义

3、实例

有一个关系模式R（A，B，C），存在函数依赖 $F=\{A\rightarrow B,\ B\rightarrow C\}$，下面的几个分解中，哪一个最好？

$\rho_1=\{R_1(A), R_2(B), R_3(C)\}$

$\rho_2=\{R_4(A, B), R_5(A, C)\}$

$\rho_3=\{R_4(A, B), R_6(B, C)\}$

$\rho_4=\{R_5(A, C), R_6(B, C)\}$

分解	属性等价	保持无损连接性	保持函数依赖性
ρ_1	√	×	×
ρ_2	√	√	×
ρ_3	√	√	√
ρ_4	√	×	×

本讲主要内容

一、 公理及其推论

二、 闭包的概念及其计算

三、 函数依赖集的等价

四、 最小函数依赖集

五、 等价模式分解的定义

六、 无损连接性

七、 函数依赖保持性

八、 模式分解的算法

六、无损连接性

定义4.17

R<U，F>，若R的分解$\rho=\{R_1, R_2 \cdots . R_k\}$对R中任何一个关系r，有：

$$r=\Pi_{R1}(r) \bowtie \Pi_{R2}(r) \bowtie \cdots \bowtie \Pi_{Rk}(r)$$

称分解ρ具有**无损连接性**

$\Pi_{R1}(r)$ 表示关系r在模式R_1的属性上的投影

六、无损连接性

如何判断分解是否具有无损连接性？

- 算法4.3　无损连接性检验

  输入：关系模式$R(A_1, A_2, \cdots, A_n)$，它的函数依赖集F，以及分解$\rho=\{R_1, R_2, \cdots, R_k\}$。

  输出：确定ρ是否具有无损连接性。

方法：

(1)构造一个k行n列的表，第i行对应于关系模式R_i，第j列对应于属性A_j。如果$A_j \in R_i$，则在第i行第j列上放符号a_j，否则放符号b_{ij}。（属于用a代表，且位置信息用j表示；不属于用b代表,且位置信息用ij表示。）

六、无损连接性

(2) 重复考察F中的每一个函数依赖，并修改表中的元素。

其方法如下：取F中一个函数依赖X→Y，在X的分量中寻找相同的行，然后将这些行中Y的分量改为相同的符号，如果其中有a_j，则将b_{ij}改为a_j；若其中无a_j，则全部改为b_{ij}（i是这些行的行号最小值）。

(3) 如果发现表中某一行变成了a_1，a_2，…，a_n，则分解ρ具有无损连接性；

如果F中所有函数依赖都不能再修改表中的内容，且没有发现这样的行，则分解ρ不具有无损连接性。

六、无损连接性

- 示例一：U={A, B, C, D, E}， F={AB→C， C→D, D→E}
  ρ ={ABC, CD, DE}

	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

AB→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

C→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→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

六、无损连接性

- **示例二：** U={A, B, C, D, E}，
  F={A→C， B→C， C→D, DE→C ，CE→A}
  ρ ={AD, AB, BE, CDE, AE}

	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_{32}	b_{33}	b_{54}	a_5

A→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_{32}	b_{13}	b_{54}	a_5

六、无损连接性

$B \rightarrow C$

	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_{32}	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_{32}	b_{13}	a_4	a_5

六、无损连接性

$DE \rightarrow C$

	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_{32}	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_{32}	a_3	a_4	a_5

六、无损连接性

特殊情况（只有两个子模式）下的无损分解的判定：

- **定理4.11** 设ρ=(R1，R2)是R的一个分解，F是R上的函数 依赖集，分解ρ具有无损连接性的充分必要条件是：

$$R1 \cap R2 \rightarrow (R1-R2) \in F^+$$

或 $$R1 \cap R2 \rightarrow (R2-R1) \in F^+$$

只能用于判断分解为2个子模式的情况。

六、无损连接性

例:下列分解是否具有无损连接性和函数依赖保持性。

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

(1)　$\rho_1=\{AB, BC\}$

(2)　$\rho_2=\{AC, BC\}$

可以用两种方法检验

- 方法1：无损连接性检验
- 方法2：特殊情况方法检验

六、无损连接性

◆解：方法1 无损连接性检验

(1) $\rho_1=\{AB, BC\}$
$F=\{A\to B, C\to B\}$

对ρ_1和F构造表：

Ri	A	B	C
AB	a_1	a_2	b_{13}
BC	b_{21}	a_2	a_3

检查$F=\{A\to B, C\to B\}$
对$A\to B$，A列中无相同的行；
对$C\to B$，C列中无相同的行。
ρ_1不具有无损连接性。

六、无损连接性

(2) $\rho_2=\{AC,\ BC\}$
$F=\{A\rightarrow B,\ C\rightarrow B\}$

对ρ_2和F构造表：

Ri	A	B	C
AC	a_1	b_{12}	a_3
BC	b_{21}	a_2	a_3

检查$F=\{A\rightarrow B,\ C\rightarrow B\}$
对$C\rightarrow B$，C列有相同的行，改写B列的相异符号为a，下标为列号2。第一行变为$a_1a_2a_3$，ρ_2具有无损连接性。

Ri	A	B	C
AC	a_1	a_2	a_3
BC	b_{21}	a_2	a_3

六、无损连接性

◆解：方法2 特殊情况方法检验

(1) ρ_1={AB，BC}

F={A→B，C→B}

$R1 \cap R2 = B$

$(R1-R2) = A$ ，$(R2-R1) = C$

$R1 \cap R2 \rightarrow (R1-R2) \notin F^+$

$R1 \cap R2 \rightarrow (R2-R1) \notin F^+$

ρ_1不是无损连接分解。

六、无损连接性

(2) $\rho_2=\{AC，BC\}$

$F=\{A\rightarrow B，C\rightarrow B\}$

$R1\cap R2 = C$

$(R1-R2) = A$；$(R2-R1) = B$；

$R1\cap R2\rightarrow(R2-R1) \in F^+$

ρ_2是无损连接分解。

本讲主要内容

一. 公理及其推论

二. 闭包的概念及其计算

三. 函数依赖集的等价

四. 最小函数依赖集

五. 等价模式分解的定义

六. 无损连接性

七. 函数依赖保持性

八. 模式分解的算法

七、函数依赖保持性

函数依赖保持性：

- 关系模式分解等价性的另一个要求是分解后的模式满足的函数依赖集应能蕴含分解前的依赖集。

定义4.20设F是关系模式 R的函数依赖集，Z是R的一个属性集合，则称Z所涉及到的F^+中所有的函数依赖为F在Z上的投影，记为$\Pi_Z(F)$，有：

$$\Pi_Z(F)=\{X\rightarrow Y \mid X\rightarrow Y\in F^+ 且 XY\subseteq Z\}$$

例如：$F=\{A\rightarrow B, C\rightarrow B, B\rightarrow D, D\rightarrow C\}$，设Z=CD，则

$\Pi_{CD}(F)=\{C\rightarrow D, D\rightarrow C\}$

七、函数依赖保持性

定义4.21 设关系模式R的一个分解 $\rho=\{R_1, R_2, \cdots, R_k\}$，F是R的函数依赖集，如果F等价于 $\Pi_{R1}(F)\cup\Pi_{R2}(F)\cup\cdots\cup\Pi_{Rk}(F)$，则称分解ρ具有函数依赖保持性。

保持函数依赖的分解就是指：当一个关系模式R分解后，无语义丢失，且原来的函数依赖关系都分散在分解后的子模式中。

七、函数依赖保持性

例如：　分析以下分解是否具有无损连接性和函数依赖保持性

(1) 设S1(A, B, C)，F1={A →B}在R上成立，ρ1={AB, AC}

令R1=AB, R2=AC，　R1 $\cap$ R2=A，　R1−R2=B

因为(R1 $\cap$ R2) →R1−R2即A→B $\in$ F1$^+$，所以ρ1相对于F1是无损连接

又因为Π_{R1}(F1) $\cup$ Π_{R2}(F1)={A→B}与F1等价，所以ρ1相对于F1是函数依赖保持的分解

七、函数依赖保持性

(2) 设 $S2(A, B, C)$，$F2=\{A \to C,\ B \to C\}$ 在R上成立，$\rho 2=\{AB, AC\}$

令 $R1=AB$, $R2=AC$，$R1 \cap R2=A$，$R2-R1=C$

因为 $(R1 \cap R2) \to R2-R1$ 即 $A \to C \in F2^+$，所以 $\rho 2$ 相对于F2是无损连接

又因为 $\Pi_{R1}(F2) \cup \Pi_{R2}(F2)=\{A \to C\}$ 与F2不等价，所以 $\rho 2$ 相对于F2不保持函数依赖

七、函数依赖保持性

(3) 设S3(A, B, C)，F3={A →B, B→C} 在R上成立，ρ3={AC, BC}

令R1=AC, R2=BC

因为 $(R1 \cap R2)$=C， R2-R1=B， R1-R2=A， C→A或C→B都不成立，所以ρ3相对于F3不是无损连接

又因为 $\Pi_{R1}(F3) \cup \Pi_{R2}(F3)$={A→C, B →C} 与F3不等价，所以$\rho$3相对于F3不保持函数依赖

本讲主要内容

一. 公理及其推论

二. 闭包的概念及其计算

三. 函数依赖集的等价

四. 最小函数依赖集

五. 等价模式分解的定义

六. 无损连接性

七. 函数依赖保持性

八. 模式分解的算法

八、模式分解的算法

- 对关系模式进行分解，使它的模式成为3NF或BCNF，但这样的分解不一定都能保证具有无损连接性和函数依赖保持性。
- 对于任一关系模式，可找到一个分解达到3NF，且具有无损连接性和函数依赖保持性。（算法。4.4）
- 而对模式的BCNF分解，可以保证无损连接，但不一定能保证保持函数依赖集。（算法4.5）

八、模式分解的算法

- **算法4.4：转换为3NF的保持无损连接及函数依赖的分解：**

设： R<U，F>

①**最小化**：求F的最小函数依赖集F_m

②**排除**：对F_m中任一X→A，若XA=U则不分解（R已为3NF），结束。

③**独立**：若R中Z属性在F_m中未出现，则所有Z为一个子模式，令U=U-Z。

④**分组**：对F_m中 X→A1，…. X→An，用合成规则合成一个，再对F_m中每个X→A，令R_i=XA。

R的分解为{R_1，R_2，….R_K}

⑤**添键**：如果分解中没有一个子模式含R的侯选码X，则将分解变成{R_1，R_2，….R_K，X}，如果存在R_i属于R_j，则删去R_i

八、模式分解的算法

例：设关系模式R<U, F>，U={E, G, H, I, J}，
$F=\{E\rightarrow I, J\rightarrow I, I\rightarrow G, GH\rightarrow I, IH\rightarrow E\}$，将其分解为3NF且同时具有无损连接性和函数依赖保持性

解：求出最小依赖集为

$Fm=\{E\rightarrow I, J\rightarrow I, I\rightarrow G, GH\rightarrow I, IH\rightarrow E\}$

得到分解为：{EI, JI, IG, GHI, IHE}

由侯选码的定义和属性闭包的求解算法可以得到R的侯选码中至少包含J和H，且$(JH)^+=IJHGE=U$，所以JH是R的唯一侯选码

上面的分解中没有子模式含有JH，加上侯选码并去重后得到最终的分解：{JI, GHI, IHE, JH}

八、模式分解的算法

- **算法4.5：** 转换为BCNF的保持无损连接的分解：

  设： R<U, F>

  ① 令ρ={R}。

  ② 如果ρ中的所有关系模式都是BCNF，结束。

  ③ Ri<Ui,Fi>为ρ中不是BCNF的一个关系模式，则Ri中必有X→A ∈Fi+ (A ∈ X) 且X不是Ri的键。将Ri分解为S1=XA, S2=Ui-A，更新ρ， 转2）。

八、模式分解的算法

例：设关系模式R<U, F>，U={C, T, H, R, S, G}，F={CS→G, C→T, TH→R, HR→C, HS→R}，将其无损连接地分解为BCNF。

解：经分析知道R上的候选键为HS

令ρ ={CTHRSG}，ρ不是BCNF

选择不符合BCNF条件CS→G，将CTHRSG 分解为CSG和CTHRS；模式CSG的候选键为CS，其上只有一个依赖CS→G，是BCNF；模式CTHRS的候选键为HS，不是BCNF，选择C → T，把CTHRS 分解为CT和CHRS； CT是BCNF；CHRS侯选键是HS，不是BCNF，选择HR →C，把CHRS分解为HRC和HRS ， HRC和HRS 都是BCNF。

最终的分解为{CSG，CT，HRC，HRS}

Questions?

本讲主要目标

学完本讲后，你应该能够了解：

1、函数依赖的逻辑蕴含的定义；

2、Armstrong推理规则系统；

3、函数依赖集的闭包的概念，及属性集闭包的计算方法；

4、函数依赖集等价的定义，并能判断两个函数依赖集是否等价；

5、每个函数依赖集都等价于一个最小函数依赖集，并能计算最小函数依赖集；

6、关系模式的规范化过程就是关系模式的分解过程；是将一个关系模式分解成一组等价的关系模式的过程；

7、模式分解的等价包括了属性的等价、无损连接性和函数依赖的保持性

8、将关系模式分解为3NF，且既具有无损连接性又具有函数依赖保持性的算法；

9、将关系模式分解为BCNF，且具有无损连接性的算法。

问题讨论

1、在关系模式R（A，B，C，D）中，存在函数依赖关系{ A B→C，B→D}，则其候选键是什么？最高达到第几范式?

2、在关系模式R（A，B，C，D）中，存在函数依赖关系{A → B，A → C，A → D，BC→A}，则其候选键是什么？最高达到第几范式?

3、在关系模式R（D，E，G）中，存在函数依赖关系{E → D，DG→E }，则其候选键是什么？最高达到第几范式?