数据库系统原理

教程：数据库系统概论（第5版）

结合：CMU 15-445/645 INTRO TO DATABASE SYSTEMS

华中科技大学 计算机学院

左琼

第六章　关系数据理论

Principles of Database Systems

第六章 关系数据理论

6.3、Armstrong公理系统

1. 逻辑蕴涵，闭包
2. Armstrong公理系统（3+3）
3. 属性闭包
4. 等价、覆盖和最小函数依赖集

问题：

- 对于给定的一组函数依赖，如何判断另外的一些函数依赖是否成立？
- 如何找出R上所有的函数依赖？
- 码如何求？

——一套有效而完备的公理推理系统

——Armstrong公理系统。

6.3 数据依赖的公理系统

定义6.11（逻辑蕴含）对于关系模式R(U, F)，其任何一个关系r，若函数依赖X → Y都成立（即r中任意两元组s，t，若t[X]=s[X]，则t[Y]=s[Y]），则称函数依赖集F逻辑蕴含X→Y，或X→Y从F推导出来的，或X→Y 逻辑蕴含于F 。

定义6.12 在关系模式R(U, F)中为F所逻辑蕴含的函数依赖的全体叫做F的闭包，记作F^+。

问题：

1）如何从已知的F出发，推出F^+中的所有函数依赖?

2）已知F和X、Y，如何判断X→Y是否在F^+中?

根据已知的F出发推导出新的函数依赖，需要使用一些推理规则。1974年，W.W.Armstrong总结了各种规则，形成了著名的Armstrong公理系统。

6.3 数据依赖的公理系统

- Armstrong公理的内容：

设有关系模式R(U, F)，U为属性全集，F是U上的函数依赖集，X，Y，Z $\subseteq$ U。则有：

- A1 自反律：若$Y \subseteq X \subseteq U$，则$X \rightarrow Y$为F所蕴含(给出平凡的函数依赖)。
- A2 增广律：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：如$X \rightarrow Y$及$Y \rightarrow Z$为F 所蕴含，则$X \rightarrow Z$为F所蕴含。

$$\left.\begin{array}{c} X \rightarrow Y \\ t[X] = s[X] \end{array}\right\} \Rightarrow \left.\begin{array}{c} t[Y] = s[Y] \\ Y \rightarrow Z \end{array}\right\} \Rightarrow t[Z] = s[Z] \Rightarrow X \rightarrow Z$$

传递律证明

Armstrong公理的正确性

- Armstrong公理的推论
  - 合并规则：若$X \rightarrow Y$，$X \rightarrow Z$，则$X \rightarrow YZ$
  - 分解规则：若$X \rightarrow YZ$，则$X \rightarrow Y$，$X \rightarrow Z$
  - 伪传递规则：若$X \rightarrow Y$，$YW \rightarrow Z$，则$XW \rightarrow Z$

由合并规则和分解规则可得：

引理6.1：如果A_i（$i=1,2,3,\ldots,n$）是关系模式R的属性，则

$X \rightarrow A_1A_2\ldots A_n$的充要条件是$X \rightarrow A_i$（$i=1,2,\ldots,n$）均成立。

课堂练习

- 课堂练习：已知关系模式R(U, F)，U = (A, B, C, G, H, I)，F = {A→B, A→C, CG→H, CG→I, B→H}，判断下列函数依赖是否为F的逻辑蕴涵?
  - A→ H　　是
  - CG → HI　　是
  - AG → I　　是

问题：能不能用一种一般性的算法来判定X→Y是否是F的逻辑蕴涵?

F的闭包

- 例如：从F = {X→A1，X→A2，…，X→An}出发可推导出2^n个不同的函数依赖。

F={X→Y, Y→Z}

F^+={

X→φ, Y→φ, Z→φ, XY→φ, XZ→φ, YZ→φ, XYZ→φ,

X→X, Y→Y, Z→Z, XY→X, XZ→X, YZ→Y, XYZ→X,

X→Y, Y →Z, XY→Y, XZ→Y, YZ→Z, XYZ→Y,

X→Z, Y→YZ, XY→Z, XZ→Z, YZ→YZ,XYZ→Z,

X→XY, XY→XY,XZ→XY, XYZ→XY,

X→XZ, XY→YZ,XZ→XZ, XYZ→YZ,

X→YZ, XY→XZ,XZ→XY, XYZ→XZ,

X→XYZ, XY→XYZ,XZ→XYZ, XYZ→XYZ }

F={X→A1, …… , X→An}的闭包F^+计算是一个NP完全问题

属性闭包

- □ 定义6.13：设有关系模式R(U, F)，U = {A1，A2，…，An}，$X \subseteq U$，F是U上的一个函数依赖集，则称所有用Armstrong公理从F推导出的函数依赖$X \rightarrow Ai$中所有Ai的属性集合为属性集X关于F的闭包，记为X_F^+。即：

  X_F^+ = { A | $X \rightarrow A$能由F根据Armstrong公理推导出 }

【例】在关系模式R(U, F)中，U={A, B, C}，F={$A \rightarrow B$, $B \rightarrow C$}，则A、B、C关于F的闭包为：

$A_F^+ = ABC$

$B_F^+ = BC$

$C_F^+ = C$

属性闭包

引理6.2：函数依赖X→Y能由F根据Armstrong公理推导出来的充要条件是$Y \subseteq X_F^+$。（证明略）

由该定理可知，判定X→Y是否能由F根据Armstrong公理导出，可转化为：求X_F^+，判定$Y \subseteq X_F^+$是否成立。

- 算法6.1 求属性集X ($X \subseteq U$) 关于U上的函数依赖集F的属性闭包。
  - 输入：X，F　输出：X_F^+
  - 方法：

（1）$X(0) = \phi$，$X(1) = X$；

（2）如果$X(0) \neq X(1)$，置$X(0) = X(1)$，否则转（4）；

（3）对F中的每个函数依赖Y→Z，若$Y \subseteq X(0)$，置$X(1) = X(1) \cup Z$，即将Y的右部并入X(1)中。转（2）；

（4）输出X(1)，即为X_F^+。

属性闭包计算举例

【例】设有关系模式R(U, F)，U= (A, B, C, D, E), F = {AB→C, B→D, C→E, EC→B, AC→B}，计算$(AB)_F^+$。

	所用依赖	$X^{(0)}$	$X^{(1)}$
初始值		ϕ	AB
第一遍	AB→C, B→D	AB	ABCD
第二遍	C→E, AC→B	ABCD	ABCDE
第三遍	EC→B	ABCDE	ABCDE

计算结果：$(AB)_F^+$ = ABCDE

通过计算属性闭包可以判断一个属性组是否为关系的码。

如本例中：∵ $(AB)_F^+$ = ABCDE =U, 且$A_F^+ \neq U$，$B_F^+ \neq U$

∴AB为R的一个码。

课堂练习

【例】设关系模式R(B，O，I，S，Q，D)，函数依赖集F={S→D，I→S，IS→Q，B→Q，S→O，D→I}。找出R的所有候选码，并指出R最高属于第几范式。

输入：关系模式R的属性集U，及其函数依赖集F

输出：R的所有候选码集合K

步骤：

（1）令$K=\phi$；

（2）求从未在F中函数依赖的右部出现过的属性集X；

（3）求X_F^+，若$X_F^+=U$，则令$K=\{X\}$，转（7）；

（4）求在F中函数依赖左右部都出现过的属性集Y；

（5）依次取Y中每个属性 (设为A)，求$(XA)_F^+$，若$(XA)_F^+=U$， 则令$K=K \cup \{XA\}$；

（6）依次取Y中每两个、三个…（设为Z），若XZ不包含K中的任一候选码，则求$(XZ)_F^+$，若$(XZ)_F^+=U$，则令$K=K\cup \{XZ\}$；

（7）输出K中所有候选码。

$B_F^+=BQ$

$BS_F^+=U$	$BI_F^+=U$	$BD_F^+=U$

Armstrong公理

- 定理：Armstrong公理是有效的、完备的。
  - 有效性：由F出发，根据Armstrong公理推导出来的每一个函数依赖一定在F^+中。
  - 完备性：F^+中的每一个函数依赖，必定可以由F出发，根据Armstrong公理推导出来。
- 证明思路：

  1. 有效性：可由引理6.1得证；

  2. 完备性：只需证明逆否命题：若函数依赖$X \to Y$不能由F从Armstrong公理导出，那么它必然不为F所蕴含。

引理6.1：如果A_i（i=1,2,3,…,n）是关系模式R的属性，则$X \to A_1A_2 \ldots A_n$的充要条件是$X \to A_i$（i=1,2,…,n）均成立。

函数依赖的等价和覆盖

定义6.14：

- 等价：若F和G是R的两个函数依赖集，如果$F^+ = G^+$, 则称F等价于G。
- 覆盖：若F和G是R的两个等价的函数依赖集$F^+ = G^+$，则称F覆盖G，同时G也覆盖F。

引理6.3：设F和G是R的两个函数依赖集，则F和G等价的充分必要条件是$F \subseteq G^+$ 且 $G \subseteq F^+$。

证明:

1) 必要性

若$F^+ = G^+$，显然有$F \subseteq F^+ \subseteq G^+$，$G \subseteq G^+ \subseteq F^+$。

2) 充分性

若$F \subseteq G^+$，$G \subseteq F^+$

则 $F^+ \subseteq (G^+)^+$，即$F^+ \subseteq G^+$；$G^+ \subseteq (F^+)^+$，即$G^+ \subseteq F^+$

$\therefore G^+ = F^+$。 (证毕)

最小函数依赖集

定义6.15 最小函数依赖集

若函数依赖F满足下列条件，则称F为一个最小函数依赖集，记为F_m。

1）F中每个函数依赖的右部都是单属性；

2）对于F中的任一函数依赖X→A，F-{X→A}与F是不等价的（即F中不存在多余的依赖）

3）对于F中的任一函数依赖X→A，不存在X的子集Z，使得F与(F-{X→A})∪{Z→A} 等价（即左部无多余属性）

□ F的最小依赖集求解算法:

1）用分解规则将F中所有函数依赖的右部分解为单属性的函数依赖，去掉重复依赖；

2）去掉多余依赖：对每个依赖X→Y，令$G = F - \{X \to Y\}$，求X_G^+，若$Y \subseteq X_G^+$，则X→Y为多余依赖，将其从F中去掉；

3）去掉依赖左部的多余属性：对每个左部为多属性的依赖，如X→A，设$X = B_1B_2 \ldots B_m$，逐一考察B_i，若$A \in (X - B_i)_F^+$，则B_i是多余属性，用$X - B_i$代替X。

重复23，直到Fm不再改变

最小函数依赖集计算举例

【例1】已知F = { AB→C, C→A, BC→D, ACD→B, D→EG, BE→C, CG→BD, CE→AG }，求F的最小依赖集F_m。

解:

1）将F中所有函数依赖右部不为单属性的化为单属性

F_1 = { AB→C, C→A, BC→D, ACD→B, D→E, D→G, BE→C, CG→B, CG→D, CE→A, CE→G }

2）去掉F_1中多余的函数依赖

- 对AB→C，令G = F_1-{AB→C}，计算$(AB)_G^+$=AB,
  ∵$C \not\subset (AB)_G^+$，∴AB→C不能去掉；
- 对C→A，令G = F_1-{C→A}，计算C_G^+=C，∵$A \not\subset C^+$，∴C→A不能去掉；
- 对于ACD→B, …, $(ACD)_G^+$=ABCDEG，…，∴ACD→G去掉 ……

F_2 = {AB→C, C→A, BC→D, D→E, D→G, BE→C, CG→B, CE→G }

最小函数依赖集计算举例

$F_2 = \{AB\to C, C\to A, BC\to D, D\to E, D\to G, BE\to C, CG\to B, CE\to G\}$

3）去掉F_2中函数依赖左部多余属性

- 对$AB\to C$，在F_2中分别计算：

  对A，求$B_{F2}^+ = B$，因为$C \not\subset B_{F2}^+$，所以A不是多余属性。

  对B，求$A_{F2}^+ = A$，因为$C \not\subset A_{F2}^+$，所以B不是多余属性。

- 对$BC\to D$，在 F_2中分别计算：

  对B，求$C_{F2}^+ = CA$，因为$D \not\subset C_{F2}^+$，所以B不是多余属性。

  对C，求$B_{F2}^+ = B$，因为$D \not\subset B_{F2}^+$，所以C不是多余属性。

……

$\because$ F_2中函数依赖左部无多余属性，$\therefore F_3 = F_2$ $\therefore F_m = F_2$

注意：F的最小函数依赖集不是唯一的，<u>与计算顺序有关</u>。

最小函数依赖集

【例2】设U = {ABCG}，F = {C→A，A→G，CG→B，B→A}, 求F_m。

解:

1) 右部都是单属性;

2) F是无冗余的;

3) 去除左边多余属性:

CG→B:

对于C：G_F^+= G，所以C不是多余属性;

对于G：C_F^+ = CAGB，G是多余属性， C→B 成立。得到:

F_m = {C→A，A→G，C→B，B→A}

4) 重复2、3步骤:

去冗余：令F'= F_m - {C→A} , A⊆ $C_{F'}^+$= {CBAG}, 去掉C→A， …， ∴ F_m=F'.

最小函数依赖集

【例3】设$F=\{AB\rightarrow C，A\rightarrow B\}$, 求$F_m$。

解:

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

对不对？为什么？

不对！上面的F与F_m根本不等价。

- 去除左边多余属性，对于 $AB\rightarrow C$
  - 对于A，$B_F^+ = B$，不能去掉A；
  - 对于B，$A_F^+ = ABC$，去掉B；（即：有：$A\rightarrow AB\rightarrow C$）

$\therefore F_m=\{A\rightarrow C，A\rightarrow B\}$

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 保持函数依赖的模式分解

*6.5 无损连接的模式分解

本章小结

怎样的分解才是正确的、恰当的?

——本节给出模式分解的理论基础。

6.4 保持函数依赖的模式分解

- 定义6.16：设有关系模式R(U,F)，称$\rho=\{R_1(U_1, F_1), R_2(U_2, F_2), \ldots, R_n(U_n, F_n)\}$为R的一个分解，其中：
  1) $U = U_1 \cup U_2 \cup \ldots U_n$ （分解后各个关系模式所含属性的并集等于U）
  2) U_i与U_j可以相交，但不允许 $U_i \subseteq U_j$ （$1 \le i, j \le n$）
  3) F_i是F在U_i上的投影(也可记作$\prod_{R_i}(F)$)

- 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$F_i$称为F在属性集$U_i$上的投影。若$F^+ = F_1^+ \cup F_2^+ \cup \ldots F_n^+$，则$\rho=\{R_1(U_1, F_1), R_2(U_2, F_2), \ldots, R_n(U_n, F_n)\}$是关系模式R(U,F)的一个保持函数依赖的分解。

- 分解的目标：
  - 达到更高级范式
  - 分解后数据可以还原
  - 分解后属性间的依赖关系保持不变

模式分解带来的问题(1)

【例1】

$R(A, B, C)$

A	B	C
1	1	2
2	2	1

$\Pi_{A,B}(R)$

A	B
1	1
2	2

$\Pi_{B,C}(R)$

B	C
1	2
2	1

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

A	B	C
1	1	2
2	2	1

数据可以还原

$R(A, B, C)$

A	B	C
1	1	1
2	1	2

$\Pi_{A,B}(R)$

A	B
1	1
2	1

$\Pi_{B,C}(R)$

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

数据无法还原

模式分解带来的问题(2)

【例2】 $\{A \to B, B \to C\}$

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

⇒

A	B
a1	b1
a2	b1
a3	b2
a4	b3

A	C
a1	c1
a2	c1
a3	c2
a4	c1

⇒

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

$\bowtie$

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

插入

违反 $B \to C$

分解的正确性标准

1) 无损连接性 —— 分解所得到的各个关系模式经过自然连接可以还原成被分解的关系模式，既不增加原来没有的元组也不丢失原有的元组。

2) 依赖保持性 —— 分解所得到的各个关系模式上的函数依赖的集合与被分解关系模式原有的函数依赖集等价，没有被丢失的现象。

- 对于一个分解，必然有下面四种可能的结果：
  - 具有无损连接性，不具有依赖保持性
  - 不具有无损连接性，具有依赖保持性
  - 既有无损连接性，又有依赖保持性（理想情况）
  - 既没有无损连接性，又没有依赖保持性

分解的无损连接性和保持函数依赖性

- 定义6.18：任给关系模式R(U, F)，$\rho=\{R_1, R_2, \ldots, R_n\}$是R的一个分解。若对R的任一关系 r 都有：

$$\rho = \Pi_{R1}(r) \bowtie \Pi_{R2}(r) \bowtie \ldots\ldots \bowtie \Pi_{Rn}(r)$$

则称分解ρ是R的一个无损分解。

即：无损分解可通过自然连接运算还原。

- 定义：任给R（U，F），ρ={R1，R2，......Rn}是R的一个分解，若 $F \Leftrightarrow \Pi_{R1}(F1) \cup \Pi_{R2}(F2) \cup \ldots\ldots \cup \Pi_{Rn}(Fn)$，

则称ρ具有函数依赖保持性。

模式分解带来的问题

【例】对于关系模式SC(Sno,Sdept,Sloc)，F={Sno→Sdept, Sdept→Sloc}.

Sno	Sdept	Sloc
99001	计算机	D1
99002	电信	D2
99003	自控	D3
99004	电信	D2
99005	管理	D2

哪些分解具有无损连接性?

R1 ⋈ R2=R

哪些保持了FD?

ρ1：√

R1

Sno	Sdept
99001	计算机
99002	电信
99003	自控
99004	电信
99005	管理

R2

Sno	Sloc
99001	D1
99002	D2
99003	D3
99004	D2
99005	D2

未保持fd（丢失了fd: sdept→sloc ）

ρ2：✗

R3

Sno	Sloc
99001	D1
99002	D2
99003	D3
99004	D2
99005	D2

R4

Sdept	Sloc
计算机	D1
电信	D2
自控	D3
管理	D2

未保持fd（丢失了fd: sno→sdept）

模式分解带来的问题

【例】对于关系模式SC(Sno,Sdept,Sloc)，F={Sno→Sdept, Sdept→Sloc}

Sno	Sdept	Sloc
99001	计算机	D1
99002	电信	D2
99003	自控	D3
99004	电信	D2
99005	管理	D2

哪些分解具有无损连接性?

R1 ⋈ R2=R

哪些保持了FD?

ρ3： √

R5

Sno	Sdept
99001	计算机
99002	电信
99003	自控
99004	电信
99005	管理

R6

Sdept	Sloc
计算机	D1
电信	D2
自控	D3
管理	D2

保持了fd

ρ4： ✗

R7

Sno
99001
99002
99003
99004
99005

R8

Sdept
计算机
电信
自控
管理

R9

Sloc
D1
D2
D3

未保持fd

6.4 保持函数依赖的模式分解

算法6.2（合成法） 达到3NF且保持函数依赖的分解算法。

输入: 给定关系模式$R\langle U, F\rangle$

输出: $\rho = \{R_1,\ldots, R_k\}$, $R_i \in 3NF$, $i = 1,2,\ldots,k$

步骤:

1. 求R 的最小函数依赖集F';
2. 找出不在F'中出现的属性U_0，将它们构成一个关系模式$R_0\langle U_0, F_0\rangle$，并从$U$中去掉它们(剩余属性仍记为$U$);
3. 若有$X\to A$,且$XA=U$, 则$\rho = \{R\}$, 算法终止;
4. 否则，对于F'中的每个$X\to A$,构成一个关系模式XA. 如果F'中有（左部相同）$X\to A_1$, $X\to A_2$, …,$X\to A_n$,则可以用$XA_1 A_2\ldots A_n$代替n个模式XA_1, XA_2,…, Xa_n, $U_i=\{XA_1\ldots A_n\}$；如发现某个$U_i \subseteq U_j$, 则应将U_i去掉。
5. $\rho = \{R_1\langle U_1, F_1\rangle,\ \ldots, R_k\langle U_k, F_k\rangle\}\cup R_0\langle U_0, F_0\rangle$

6.4 保持函数依赖的模式分解

[例] R<U,F>,U={S#,D,M,C#,G},F={S#→ D,S# → M,D→ M,(S#,C#)→ G}. 试将R分解3NF,并保持函数依赖。

解: (1)最小依赖集:

$$F' = \{S\# \rightarrow D, D \rightarrow M, (S\#, C\#) \rightarrow G\}$$

(2)分解

R_1(S#， D)， {S#→D}

R_2(D， M)， { D→M}

R_3(S#， C#， G)， {(S#,C#)→G}

6.4 保持函数依赖的模式分解

- [例] R(C, T, H, R, S, G)，其中C—课程，T—教师，H—时间，R—教室，S—学生，G—成绩，满足如下语义：

  1．C→T (每门课仅一名教师上)

  2．HR→C (任一时间，一个教室只能上一门课)

  F = 3．HT→R (一个时间，一个教师只能在一个教室上课)

  4．CS→G (一个学生，一门课只有一个成绩)

  5．HS→R (一个时间，一个学生只能在一个教室上课)

  求R的3NF分解，并要求保持函数依赖。

解：(1)最小依赖集

$$F' = \{C \to T,\ HR \to C,\ HT \to R,\ CS \to G,\ HS \to R\}$$

(2)分解

$CT(C,T), \{C \to T\}$ 授课老师安排表

$HRC(H,R,C), \{HR \to C\}$ 教室安排表

$HTR(H,T,R), \{HT \to R\}$ 教师课表

$CSG(C,S,G), \{CS \to G\}$ 成绩表

$HSR(H,S,R), \{HS \to R\}$ 学生课表

*6.5 无损连接的模式分解

- 问题：如何验证一个分解ρ是否为无损分解？
- 算法6.3 判定一个分解的无损连接性：
  - 输入: R(A1,A2,…An)，F，ρ={R1,R2,…Rk}， F={FD1, FD2,…, FDρ}

    设：FDi为Xi→Ali；
  - 输出：分解ρ是否具有无损连接性
  - 步骤：

    (1) 建立k*n的矩阵S，列j 对应属性Aj，行i对应分解中的一个关系模式Ri;

    若属性Aj属于Ui，则在j列i行处填aj，否则填bij。

示例1：第1步

U = { A, B, C, D, E }

F = { AB→C, C→D, D→E }

ρ = { (A, B, C), (C, D), (D, E) }

	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

无损连接的模式分解

(2) 逐个检查F中的每个函数依赖，利用fd数据间的等值关系，修改表中元素。

(3) 如果S中存在一行全为“a”类符号，则ρ具有无损连接性，否则不具有。

示例1：
第2，3步

$F = \{ AB \to C, C \to D, D \to E \}$

	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

$AB \to C$

	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

$C \to D$

	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

$D \to E$

	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

无损连接的模式分解

- 示例2：

  U={A,B,C,D,E},

  F={A→C, B→C, C→D,DE→C ,CE→A}

  ρ ={(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→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

示例2：F={A→C, B→C, C→D, DE→C , CE→A}

B→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→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

DE→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→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

无损连接的模式分解

判断分解无损连接性的简单算法

- 定理6.5：设R（U, F），$\rho=\{R_1, R_2\}$是R的一个分解，F是R上的函数依赖集, ρ具有无损连接性的充要条件是：

$$(U_1 \cap U_2) \to (U_1 - U_2) \in F^+ \quad \text{或} \quad (U_1 \cap U_2) \to (U_2 - U_1) \in F^+$$

- 示例：

R=ABC, F={A→ B}， 判断以下分解是否具有无损连接性。

1） $\rho_1=\{R_1(AB), R_2(AC)\}$

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

由A → B， 得到ρ_1是无损连接分解

2） $\rho_2=\{R_1(AB), R_2(BC)\}$

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

B→A, B→C均不成立，所以ρ_2不是无损连接分解

例：对于关系模式 SC(Sno,Sdept,Sloc)，F={Sno→Sdept, Sdept→Sloc}.

ρ_1={ R_1(Sno, Sdept), R_2(Sno, Sloc) }

课堂练习

设关系模式R(A,B,C,D)，F是R上成立的FD集，F={B→D,AD→C},那么ρ={ABC,BCD}相对于F(　　)

A.是无损联接分解，也是保持FD的分解

B.是无损联接分解，但不保持FD的分解

C.不是无损联接分解，但保持FD的分解

D.既不是无损联接分解，也不保持FD的分解

6.5.2 分解的无损连接性和保持函数依赖性

- 要记住的三个事实：
  - 要求分解保持函数依赖，模式分离总可以达到3NF，不一定能达到BCNF。
  - 要求分解既保持函数依赖又具有无损连接性，可以达到3NF，不一定能达到BCNF。
  - 要求分解具有无损连接性，一定可以达到4NF。

- 算法6.2（合成法）转换为3NF的保持函数依赖的分解。
- 算法6.3 判别一个分解的无损连接性。
- 算法6.4 转换为3NF既有无损连接性又保持函数依赖的分解。

6.5.3 既无损连接又保持函数依赖的模式分解算法

- 算法6.4 达到3NF既保持函数依赖又无损连接的分解。

（1）设$\rho = \{R_1\langle U_1, F_1\rangle, \ldots, R_k\langle U_k, F_k\rangle\}$是$R\langle U, F\rangle$的一个保持函数依赖的3NF分解(可由前一算法求得)

（2）设X为$R\langle U, F\rangle$的码，

若有某个U_i，$X \subseteq U_i$，则ρ即为所求，

否则令$\tau = \rho \cup \{R^*\langle X, F_X\rangle\}$，$\tau$即为所求

(如发现某个$U_i \subseteq X$, 则应将U_i去掉)

6.5.3 既无损连接又保持函数依赖的模式分解算法

- 示例：R(C, T, H, R, S, G)，其中C—课程，T—教师，H—时间，R—教室，S—学生，G—成绩，满足如下语义：

  1．C→T (每门课仅一名教师上)
  2．HR→C (任一时间，一个教室只能上一门课)
F = 3．HT→R (一个时间，一个教师只能在一个教室上课)
  4．CS→G (一个学生，一门课只有一个成绩)
  5．HS→R (一个时间，一个学生只能在一个教室上课)

求R的3NF分解，并要求保持函数依赖。

解：(1)最小依赖集

$F' = \{C\to T,\ HR\to C,\ HT\to R,\ CS\to G,\ HS\to R\}$

(2)分解

$CT(C,T)$, $\{C\to T\}$ 授课老师安排表

$HRC(H,R,C)$, $\{HR\to C\}$ 教室安排表

$HTR(H,T,R)$, $\{HT\to R\}$ 教师课表

$CSG(C,S,G)$, $\{CS\to G\}$ 成绩表

$HSR(H,S,R)$, $\{HS\to R\}$ 学生课表

这个分解是无损连接的分解吗？

模式分解算法的内涵

设计最优的原则：

- 表达性：
  - 数据等价（无损联接性）
  - 依赖等价（依赖保持性）
- 分离性：指属性间的独立联系—基本信息单位。

  目的：

  a.消除操作异常

  b.消除数据冗余

  方法：规范化——用范式表达结果
- 最小冗余性：

  要求优化后的DB能表达原DB的所有信息。

本章小结

一、函数依赖

函数依赖的定义、类型、码，主属性和非主属性

二、关系模式的规范化

1NF，2NF，3NF，BCNF，4NF

三、Armstrong公理系统

1. 逻辑蕴涵，闭包
2. Armstrong公理系统（三条基本规则和三个推论）
3. 属性闭包（求关键字）
4. 等价、覆盖和最小函数依赖集

四、关系模式分解

1. 无损连接性（检验方法）
2. 依赖保持性
3. 分解为3NF的算法（具有依赖保持性及无损连接性）

小结

- 若要求分解具有无损连接性，那么模式分解一定能够达到4NF。
- 若要求分解保持函数依赖，那么模式分解一定能够达到3NF，但不一定能够达到BCNF。
- 若要求分解既具有无损连接性，又保持函数依赖，则模式分解一定能够达到3NF，但不一定能够达到BCNF。

- 本章作业：P197 2，6，7