第四章 关系规范化理论

4.1 问题的提出

4.2 函数依赖和范式

4.3 数据依赖的公理系统

4.4 关系模式的分解方法

问 题

1. 函数依赖集的等价定义
2. <u>最小函数依赖集的定义和求解算法</u>
3. 模式分解的原则
4. <u>无损连接性如何判断</u>

4.3.5 函数依赖集的等价和最小函数依赖集

定义4.17

如果$G^+=F^+$，则称F与G等价，记为$F \equiv G$。

定理 4.7

$F^+=G^+$的充分必要条件是$F \subseteq G^+$且$G \subseteq F^+$

$F \equiv G$?

① $F \subseteq G^+$

② $G \subseteq F^+$

两种方法:

① Amstrong公理

② 公理4.2

定理4.2: 设F为属性集U上的一组函数依赖关系，$X,Y \subseteq U$， $X \to Y$能由F根据Armstrong公理导出的充分必要条件是$Y \subseteq X_F^+$。

例：R(U)　U=ABC

F={A→B，B→C，A→C，AB→C,A → BC}

G={A→B，B→C}

F和G是否等价？

F与G等价

证明：

1：A→B，B→C 传递规则 A→C

2： A→B，扩展AB→BB 即 AB → B 再由 B→C 所以 AB→C

3： A→B，B→C 扩展 B→BC 所以 A → BC

定义 4.18： 最小依赖集定义：

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，也称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→A，使得F与F-{X →A}等价。[不存在冗余]

3) F中不存在这样的函数依赖X→A，X有真子集Z使得F-{X →A}∪{Z→A}与F等价。[决定因素不存在冗余]

算法 4.2

求F_m（F的最小依赖集）的算法

(1)将$X \to A_1A_2\ldots A_k(k>2)$转换为$X \to A_i(i=1,2,\ldots,k)$

[将右部属性分解为单个属性]

(2)逐个检查函数依赖$X \to A$，令$G=F-\{X \to A\}$，若$A \in (X)_G^+$，则从F中去掉$X \to A$。[逐个检查F中的每一项，看是否$F-\{X \to A\}$与F等价]

(3)逐个检查函数依赖$X \to A$，若$X=B_1B_2\ldots B_m$，逐个考查$B_i(i=1,2,\ldots,m)$，若$A \in (X-B_i)_F^+$，则以$X-B_i$取代X。

[判每个函数依赖左部是否有冗余属性]

例：将下列函数依赖集F划为最小函数依赖集。

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

解：1：分解为单个属性F1=F

2：消去F中冗余的函数依赖

考察A→B：令X=A 求X^+=？ $X^{(0)}=A$ $X^{(1)}=AC=X^+$

因为B不属于X^+ 所以A→B不冗余。

考察B→A：令X=B 求X^+=？ $X^{(0)}=B$ $X^{(1)}=BC$ $X^{(2)}=ABC=X^+$ 因为A属于X^+ 所以B→A冗余。

考察B→C：令X=B 求X^+=？ $X^{(0)}=B$ $X^{(1)}=B=X^+$ 因为C不属于X^+ 所以B→C不冗余。

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

考察 $A\to C$：令 $X=A$ 求 $X^+=$？ $X^{(0)}=A$ $X^{(1)}=AB$
$X^{(2)}=ABC=X^+$ 因为 C 属于 X^+ 所以 $A\to C$ 冗余。

考察 $C\to A$：令 $X=C$ 求 $X^+=$？ 因为 A 不属于 X^+
所以 $C\to A$ 不冗余。

$F2=\{A\to B，B\to C，C\to A\}$

3：判每个函数依赖左部是否有冗余属性

$Fm=\{A\to B，B\to C，C\to A\}$

练习1：

设有关系模式R（A，B，C，D），其上的函数依赖集为：

F＝{A →C , C →A，B → AC，D → AC}，求F的最小覆盖。

例：求下列函数依赖集F的最小函数依赖集。

F={A→C ，C→A，B→AC，D→AC}

解：　1：右侧分解为单个属性

F1={A→C ,C→A，B→A， B→C，D→A ，D→C}

2：消去F1中冗余的函数依赖

F2={A→C ,C→A，B→C，D→C}

3：判每个函数依赖左部是否有冗余属性

Fm={A→C，C→A，B→C，D→C}

但是结果不唯一

假定我们要构造一个数据库，属性集为｛A,B,C,D,E,F,G｝，
给定的函数依赖集F如下：
F=｛BCD→A，BC→E，A→F，F→G，C→D，A→G｝.
找出这个函数依赖集的最小覆盖G

解：　　1：右侧分解为单个属性

F1=F={BCD→A，BC→E，A→F，F→G，C→D，A→G}

2：消去F1中冗余的函数依赖

F2={BCD→A，BC→E，A→F，F→G，C→D}

3：判每个函数依赖左部是否有冗余属性

1) 考察BCD→A：

判断B是否冗余，
令X=CD 求X^+=？ $X^{(0)}$=CD $X^{(1)}$=CD
因为A不属于X^+ 所以B不冗余。

判断C是否冗余，
令X=BD 求X^+=？ $X^{(0)}$=BD $X^{(1)}$=BD
因为A不属于X^+ 所以C不冗余。

F2={BCD→A，BC→E，A→F，F→G，C→D}

3：判每个函数依赖左部是否有冗余属性

1) 考察BCD→A：

判断D是否冗余，

令X=BC 求X^+=？ $X^{(0)}$=BC X^+=BCEDAFG

因为A属于X^+ 所以D冗余，去掉D。

2) 考察BC→E：

判断B是否冗余，

令X=C 求X^+=？ $X^{(0)}$=C $X^{(1)}$=CD

因为E不属于X^+ 所以B不冗余。

判断C是否冗余，

令X=B 求X^+=？ $X^{(0)}$=B $X^{(1)}$=B

因为E不属于X^+ 所以C不冗余。

Gm={BC→A，BC→E，A→F，F→G，C→D}

学习通04-3

设有关系模式R（A，B，C，D，E），R的函数依赖集F＝｛AB→D，B→CD，DE→B，C→D，D→A｝

⑴ 计算$(AB)^+$

⑵ 求R的所有候选码

⑶ 求F的最小覆盖

（按照字母顺序排序）

4.1 问题的提出

4.2 函数依赖和范式

4.3 数据依赖的公理系统

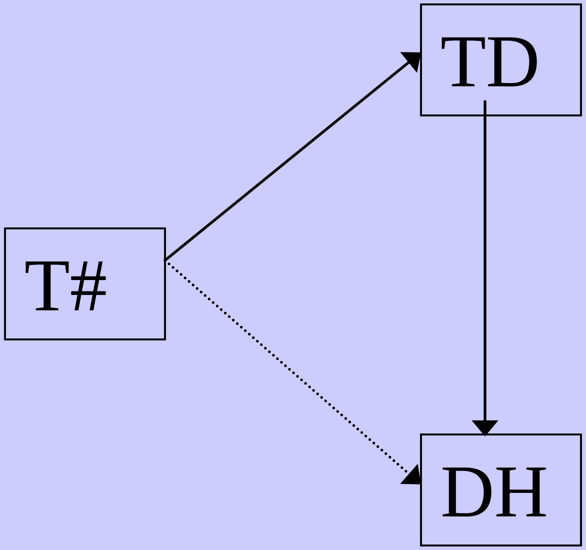
4.4 关系模式的分解方法

1. 模式分解的原则

2. 无损连接性如何判断

设一关系模式R(T#, TD, DH), 其中T#表示教师编号, TD表示教师所属系部, DH表示系主任名。假定每位教师只能在一个系任教, 每个系只有一位系主任。

T#	TD	DH
T1	D1	AA
T2	D1	AA
T3	D2	BB
T4	D3	CC



分解1：

ρ_1={R_1(T#)， R_2(TD)， R_3(DH)}

分解2：

ρ_2={R_1(T#， TD)， R_2(T#， DH)}

分解3：

ρ_3={R_1(T#， TD)， R_2(TD， DH)}

分析：这三种分解那一个最好？

分解 1：

R_1

T#
T1
T2
T3
T4

R_2

TD
D1
D2
D3

R_3

DH
AA
BB
CC

问题：T1是哪一个系的教师？无法回答。R1，R2，R3也无法恢复到原来的R。

分解 2：

R_1

T#	TD
T1	D1
T2	D1
T3	D2
T4	D3

R_2

T#	DH
T1	AA
T2	AA
T3	BB
T4	CC

此时，R1，R2的分解是可恢复的，但仍然存在操作异常。原因：TD→DH 在R1，R2中没有体现。

分解 3：

R_1

T#	TD
T1	D1
T2	D1
T3	D2
T4	D3

R_2

TD	DH
D1	AA
D2	BB
D3	CC

此时，R1，R2的分解是可恢复的，并且消除了操作异常。

1. 分解等价性的判定准则

分解前的关系模式R和分解后的关系子模式集合ρ，是否表示同样的数据

分解的无损连接性

分解前的关系模式R和分解后的关系子模式集合ρ，是否保持相同的函数依赖

分解的函数依赖保持性

4.4.2 分解的无损连接性判定

1. 分解的无损连接性：

如果一个关系模式分解后，可以通过自然连接恢复原模式的信息，这一特性称为分解的无损连接性。

算法 4.4 判定分解的无损连接性

$\rho=\{R_1(U_1,F_1),\ldots,R_k(U_k,F_k)\}$是$R(U,F)$的一个分解，$U=\{A_1,\ldots,A_n\}$，$F=\{FD_1,FD_2,\ldots,FD_p\}$。

1) 构造一个n列k行的二维表T。

$$T_{ij}=\begin{cases} a_j, & \text{如果} A_j \in R_i \\ b_{ij}, & \text{如果} A_j \notin R_i \end{cases}$$

2) 根据F中函数依赖修改表T的内容。

修改规则： 逐个考察F中的每个函数依赖$X \rightarrow Y$，在属性X所在的那些列上找出具有相同符号的行，在这些行上使对应于Y的各属性列位置上的符号改为相同，如果其中有一个符号为a_j，则把其它符号也改为a_j，否则改为b_{mj}，其中m是这些行的最小行号。直至在表中发现一行已变成$a_1a_2 \ldots a_k$，或表不能再进行修改为止。

3) 反复进行2），如果发现表中有一行已变成$a_1a_2\ldots a_k$，则表示该分解具有无损连接性，否则分解不是无损连接的。

例: 已知 R(U，F) U={A，B，C，D，E，F}，F={AB→C, C→D，A→F，D→E，D→F}

R的一个分解为：

$\rho=\{R_1(A，B，C)，R_2(C，D)，R_3(D，E，F)\}$

判断ρ是否具有无损连接性。

第一步：建T

A	B	C	D	E	F
a_1	a_2	a_3	b_{14}	b_{15}	b_{16}
b_{21}	b_{22}	a_3	a_4	b_{25}	b_{26}
b_{31}	b_{32}	b_{33}	a_4	a_5	a_6

$\rho=\{R_1(A，B，C)$
$R_2(C，D)$
$R_3(D，E，F)\}$

$$T_{ij}=\begin{cases} a_j, & \text{如果} A_j \in R_i \\ b_{ij}, & \text{如果} A_j \notin R_i \end{cases}$$

第二步：逐个考察函数依赖，并修改表。

(1)AB→C, (2)C→D, (3)A→F, (4)D→E, (5)D→F

A	B	C	D	E	F
a_1	a_2	a_3	a_4	a_5	a_6
b_{21}	b_{22}	a_3	a_4	a_5	a_6
b_{31}	b_{32}	b_{33}	a_4	a_5	a_6

由于没有相同的分量，所以表不改变

因此，该分解具有无损连接性。

$\rho_1=\{R_1(T\#),\ R_2(TD),\ R_3(DH)\}$

$\rho_2=\{R_1(T\#,\ TD),\ R_2(T\#,\ DH)\}$

$\rho_3=\{R_1(T\#,\ TD),\ R_2(TD,\ DH)\}$

ρ_1

T#	TD	DH
a_1	b_{12}	b_{13}
b_{21}	a_2	b_{23}
b_{31}	b_{32}	a_3

ρ_2

T#	TD	DH
a_1	a_2	b_{13}
a_1	b_{22}	a_3

具有无损连接性

ρ_3

T#	TD	DH
a_1	a_2	b_{13}
b_{21}	a_2	a_3

具有无损连接性

定理4.8:

设$\rho=\{R_1，R_2\}$是关系模式R的一个分解，F是R的一个函数依赖集，则对于F，ρ具有无损连接性的充分必要条件是：

$$R_1 \cap R_2 \rightarrow R_1 - R_2 \in F^+$$

或 $R_1 \cap R_2 \rightarrow R_2 - R_1 \in F^+$。

如果两个关系模式间的公共属性集至少包含其中一个关系模式的码，则此分解必定具有无损连接性。

$\rho 1 = \{R1(T\#), R2(TD), R3(DH)\}$

$\rho 2 = \{R1(T\#, TD), R2(T\#, DH)\}$

$F = (T\# \rightarrow TD, TD \rightarrow DH)$

$\rho 3 = \{R1(T\#, TD), R2(TD, DH)\}$

ρ_2

$R1 \cap R2 = T\#$

$R1 - R2 = TD$

$R2 - R1 = DH$

$T\# \rightarrow TD$

$T\# \rightarrow DH$

ρ_3

$R1 \cap R2 = TD$

$R1 - R2 = T\#$

$R2 - R1 = DH$

$TD \rightarrow T\#$

$TD \rightarrow DH$

由于$T\# \rightarrow TD \in F^+$，所以ρ2是无损连接

由于$TD \rightarrow DH \in F^+$，所以ρ3是无损连接

例：设有关系模式R(U, F)　　U={A, B, C}, 函数依赖集F={A→B, C→B}，分解ρ ={R1(A,B), R2(B,C)}，检验是否具有无损连接性。

解：

（R1∩R2）→（R1－R2）

＝（AB∩BC）→（AB－BC）

＝B→A $\notin$ F$^+$

（R1∩R2）→（R2－R1）

＝（AB∩BC）→（BC－AB）

＝B→C $\notin$ F$^+$

A	B	C
a_1	a_2	b_{13}
b_{21}	a_2	a_3

不具有无损连接性

学习通04-4

设有关系模式R=ABCDE，F={A→C，B→D，C→D，DE→C，CE→A}
现有如下分解：ρ={AD，AB，BC，CDE，AE}
（1）求出R的所有候选码。（按照字母顺序）
（2）验证上述分解是否无损连接。（填是或者否）

设有关系模式R=ABCDE，F={A→C，B→D，C→D，DE→C，CE→A}

现有如下分解：ρ={AD，AB，BC，CDE，AE}

（1）求出R的所有候选码。

（2）验证上述分解是否无损连接。

（1）求出R的所有候选关键字

L: B, E

LR: A, C, D

$(BE)^+ = BE \cup D \cup C \cup A = ABCDE$

所以BE是候选关键字

（2）验证上述分解是否无损连接。

F={A→C，B→D，C→D，DE→C，CE→A}，

ρ ={AD, AB, BC, CDE, AE}

第一步：建T **第二步：修改T**

A	B	C	D	E
a_1	b_{12}	a3	a_4	b_{15}
a_1	a_2	a3	a4	b_{25}
b_{31}	a_2	a_3	a4	b_{35}
a1	b_{42}	a_3	a_4	a_5
a_1	b_{52}	a3	a4	a_5

第三步：判断 不具有无损连接性

总　结

1. 函数依赖集的等价定义
2. 最小函数依赖集的定义和求解算法
3. 模式分解的原则
4. 无损连接性如何判断