第四章 关系规范化理论

4.3 数据依赖的公理系统

- 逻辑蕴含
- Armstrong公理系统
- 属性集闭包
- 函数依赖集的等价和覆盖

一、逻辑蕴含

定义：对于R(U, F)，如果X→Y不在F中，但是对于其任何一个关系r，X→Y都成立，则称F逻辑蕴含X→Y。

[或者说：X→Y可以由F导出]

例：关系模式R(U，F)

其中U(A，B，C，D，E，F，G)

F(A→B，C→D，AB→E，F→G)

问：F是否逻辑蕴含A→E

二、Armstrong公理系统

对于关系模式R(U，F)，有

- **公理1**：**自反律**(Reflexivity)

  若$Y \subseteq X \subseteq U$，则$X \to Y$为F所蕴含。

- **公理2**：**增广律**(Augmentation)

  若$X \to Y$为F所蕴含，且$Z \subseteq U$，则$XZ \to YZ$为F所蕴含。

- **公理3**：**传递律**(Transitivity)

  若$X \to Y$，$Y \to Z$为F所蕴含，则$X \to Z$为F所蕴含。

自反律、增广律、传递律是最基本的**Armstrong**公理。

由自反律、增广律、传递律可以导出下面三条推理规则：

公理4：合并规则

由X→Y，X→Z，有X→YZ。

公理5：伪传递规则

由X→Y，WY→Z，有XW→Z

公理6：分解规则

由X→Y及 $Z \subseteq Y$，有X→Z。

定理： Armstrong公理系统是有效的(正确性）、完备的。

正确性： 指公理1、2、3是正确的。

有效性： 指由F出发根据Armstrong公理推导出来的每一个函数依赖一定在F^+。

完备性： 指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。

引理

根据合并规则和分解规则可以得到：

定理　若$A_i(i=1,2,\ldots,n)$是关系模式R的属性，所以$X\rightarrow\{A_1,A_2,\ldots,A_n\}$成立的充分必要条件是$X\rightarrow A_i$均成立。

即：$X\rightarrow A_1A_2\ldots A_n$成立的充分必要条件是$X\rightarrow A_i$成立（ $i=1,2,\ldots,n$ ）。

例：关系模式R(U，F)

其中U(A，B，C，D，E，F，G)

F(A→B，C→D，AB→E，F→G)

问：F是否逻辑蕴含A→E

解：

∵ A→B （已知）

∴ A→AB（增广率）

∵ AB→E （已知）

∴ A→E （传递率）

例：证明：对R（A，B，C，G，H，I），F＝{A→B，A→C，CG→H，CG→I，B→H}，存在：A→H，CG→HI，AG→I

求证：

1. 由于A→B，B→H，依传递律，可得A→H
2. 由于CG→H，CG→I，依合并规则，可得CG→HI
3. 由于A→C，CG→I，依伪传递律，可得AG→I。

也可另证为： 由A→C，依增广律，得AG→CG，又CG→I，依传递律，得：AG→I

三、属性集闭包

1、F的闭包

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

F^+={X→Y|F以及能由F根据Armstrong公理导出}

所有被一个已知函数依赖集F及F逻辑蕴涵的那些函数依赖的集合为F的闭包（Closure），记为F^+。

2、X关于函数依赖集F的闭包

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

【例】设关系模式R（A、B、C）的函数依赖集为 $F=\{A \rightarrow B, B \rightarrow C\}$，分别求$A$、$B$、$C$的闭包。

若X=A，

$\because A\rightarrow B$，$B\rightarrow C$ （已知）

$\therefore A\rightarrow C$ （传递律）

$\because A\rightarrow A$ （自反律）

$\therefore X_F^+=\{A，B，C\}$

若X=B，

$\because B\rightarrow B$

$B\rightarrow C$

$\therefore X_F^+=\{B，C\}$

若X=C，

$\because C\rightarrow C$

$\therefore X_F^+=\{C\}$

求属性集闭包的算法

算法：输入：X，F　　　输出：X_F^+

(1)令$X^{(0)}=X$，$i=0$

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

(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。

例1： 已知关系模式R(U，F)，其中

U={A，B，C，D，E}；

F={AB→C，B→D，C→E，EC→B，AC→B}

求$(AB)_F^+$。

解：1：$X^{(0)}$=AB

找出左部为A，B或AB的函数依赖

2：计算$X^{(1)} = X^{(0)} \cup C \cup D = ABCD$

3：求$X^{(2)} = X^{(1)} \cup E \cup B = ABCDE$

4：由于$X^{(2)}$ 已经等于全部属性集合所以

$(AB)_F^+ = ABCDE$

例2: 已知关系模式R(U，F)，其中

U={A，B，C，D，E，F，G，H}；

F={A→D，AB→E，BH→E，CD→H，E→C}

令X=AE，求X_F^+。

解：1：$X^{(0)}$=AE

2：$X^{(1)} = X^{(0)} \cup D \cup C = ACDE$

3：$X^{(2)} = X^{(1)} \cup D \cup H \cup C = ACDEH$

4：$X^{(3)} = ACDEH$不变，即$X^{(3)} = X^{(2)}$

所以 $X_F^+ = (AE)_F^+ = ACDEH$

对于属性闭包算法的终止条件，

下列四种方法是等价的：

1、$X^{(i+1)} = X^{(i)}$

2、当发现$X^{(i)}$包含了全部属性时；

3、在F中的函数依赖的右部属性中，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部属性中已没有$X^{(i)}$的子集。

F逻辑蕴含的充要条件

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$Y \subseteq X_F^+$。

练习：

设关系模式R(U, F)，其中U＝{A，B，C，D，E，I}，F＝{A→D，AB→C，BI→C，ED→I，C→E}，试证明AC→I被F所逻辑蕴含。

四、码值理论

对于给定的关系R（A1，A2，……，An）和函数依赖集F，可将其属性分为4类：

- L类 仅出现在F的函数依赖左部的属性
- R类 仅出现在F的函数依赖右部的属性
- N类 在F的函数依赖左右两边均未出现的属性
- LR类 在F的函数依赖左右两边均出现的属性

定理1：对于给定的关系模式R及其函数依赖集F，若X（X∈R）是L类属性，则X一定是R的候选码的成员。

定理2：对于给定的关系模式R及其函数依赖集F，若X（X∈R）是N类属性，则X一定是R的候选码的成员。

定理3：对于给定的关系模式R及其函数依赖集F，若X（X∈R）是R类属性，则X必不在任何候选码中。

定理4：对于给定的关系模式R及其函数依赖集F，若X（X∈R）是LR类属性，则X可能是R的候选码的成员。

定理4.6

对于给定的关系模式R(U, F)，若X(X∈R)是R的L类和N类属性组成的属性集，且X_F^+包含R的全部属性，则X是R的唯一候选码。

例：设有关系模式R(A, B, C, D, E, P), R 的函数依赖集为：

F={A→D, E→D, D→B, BC →D, DC →A}, 求R的候选码。

因为 C，E是L类属性，P是N类属性，所以CEP包含在所有候选码中；

因为 $(CEP)^{+}$ = ABCDEP;

所以 CEP是R的唯一候选码。

候选关键字求解算法

1) 将R的所有属性分为L、R、N和LR两类，令X代表L和N类，Y代表LR类

2) 求X^+。若X^+包含了R的全部属性，则X为R的惟一候选关键字，转5；否则转3

3) 在Y中取一属性A，求$(XA)^+$。若它包含R的全部属性，则转4；否则，调换一属性反复进行这一过程，直到试完所有Y中的属性。

4) 如果已找出所有候选关键字，则转5；否则在Y中依次取两个、三个、……，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

例：设关系模式R=(O, B, I, S, Q, D)，其上的函数依赖集F=(S→D, D→S, I→B, B→I, B→O, O→B)，求R的所有候选关键字。

分析：$X^+=Q$

$(QB)^+=(QO)^+=(QI)^+$

$(QD)^+=(QS)^+$

候选码：QBS, QBD, QOD, QOS, QID, QIS

五、函数依赖集的等价和覆盖

等价定义：

如果$G^+=F^+$，则称F与G等价，记为$F \equiv G$。

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

例：R(U)　　U=ABC

F={A→B，B→C，A→C，AB→C,A → BC}

可以写成：

G={A→B，B→C}

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

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$X\rightarrow A$，使得F与$F-\{X\rightarrow A\}$等价。[不存在冗余FD]

3)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。[决定因素不存在冗余]

例： U={SNO，SDEPT，MN，CNAME，G}

F={SNO → SDEPT，SDEPT → MN，{SNO，CNAME} → G}

设F’ ={SNO → SDEPT，SNO → MN，SDEPT → MN，(SNO，CNAME) → G，(SNO，SDEPT) →SDEPT}

结论： F与 F’ 等价

F是最小覆盖，F’不是。

求Fm（F的最小依赖集）的算法

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

[将右部属性分解为单个属性]

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

(3)逐个检查函数依赖$X \rightarrow A$，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。

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

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→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→C冗余。

考察C→A：令X=C 求X^+=？ 因为A不属于X^+ 所以C→A不冗余。

因此 $F2=\{A\to B，B\to C，C\to A\}$

3：判每个函数依赖左部是否有冗余属性

练习：

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。

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

设关系模式R(U, F)，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

作答