



第六章 关系数据理论

Principles of Database Systems

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

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（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 $X \rightarrow Y$ 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 $X \rightarrow Y$ ，或 $X \rightarrow Y$ 从 F 推导出来的，或 $X \rightarrow Y$ 逻辑蕴含于 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 $X \rightarrow Y$ 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:

- **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)。
- **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。
- **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。



Armstrong公理的正确性

- 设 r 是 $R(U, F)$ 上的任一关系, $t, s \in r$

- 自反律:** 若 $Y \subseteq X$, 则 $X \rightarrow Y$

$$\left. \begin{array}{l} t[X] = s[X] \\ Y \subseteq X \end{array} \right\} \rightarrow t[Y] = s[Y] \rightarrow X \rightarrow Y$$

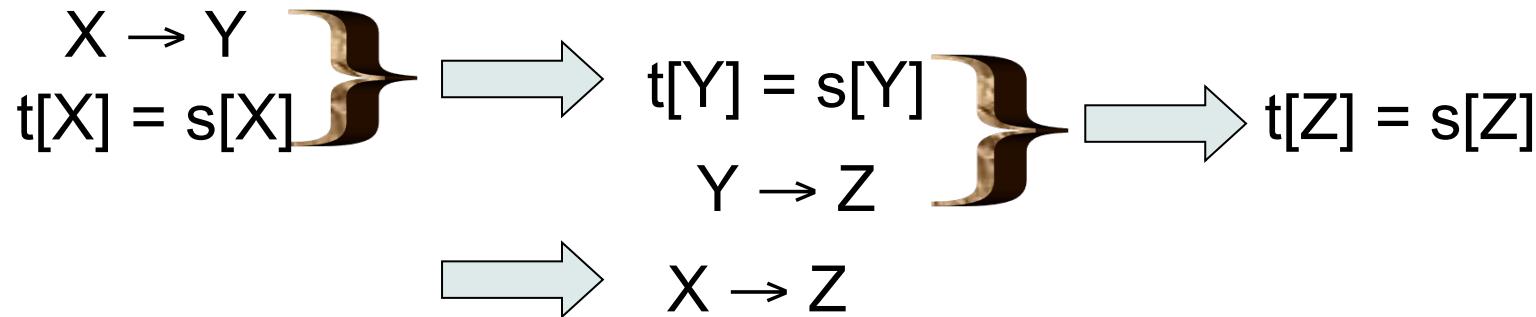
- 增广律:** 若 $X \rightarrow Y$, 则 $XZ \rightarrow YZ$

$$\begin{aligned} t[XZ] = s[XZ] &\rightarrow t[X] = s[X] \\ &\quad X \rightarrow Y \\ &\quad \left. \begin{array}{l} t[X] = s[X] \\ t[Y] = s[Y] \end{array} \right\} \rightarrow t[Y] = s[Y] \\ t[XZ] = s[XZ] &\rightarrow t[Z] = s[Z] \\ &\quad \left. \begin{array}{l} t[Z] = s[Z] \\ t[YZ] = s[YZ] \end{array} \right\} \rightarrow t[YZ] = s[YZ] \\ &\rightarrow XZ \rightarrow YZ \end{aligned}$$

Armstrong公理的正确性



- **传递律:** 若 $X \rightarrow Y$, $Y \rightarrow Z$, 则 $X \rightarrow Z$



- Armstrong公理的推论
 - ❖ **合成规则:** 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 $XW \rightarrow Z$
- 思考: 如何使用Armstrong公理证明推论的正确性?



Armstrong公理推论的证明

1) 合成规则 (若 $X \rightarrow Y$, $X \rightarrow Z$, 则 $X \rightarrow YZ$)

$\because X \rightarrow Z$, \therefore 有 $XX \rightarrow XZ$, 即 $X \rightarrow XZ$ (增广律)

又 $\because X \rightarrow Y$, $\therefore XZ \rightarrow YZ$ (增广律)

$\therefore X \rightarrow YZ$ (传递律)

2) 分解规则 (若 $X \rightarrow YZ$, 则 $X \rightarrow Y$, $X \rightarrow Z$)

$\because Y \subseteq YZ \subseteq U$, $\therefore YZ \rightarrow Y$ (自反律)

同理 $YZ \rightarrow Z$ (自反律)

$\therefore X \rightarrow YZ$, $\therefore X \rightarrow Y$ (传递律)

同理 $X \rightarrow Z$ (传递律)

3) 伪传递规则 (若 $X \rightarrow Y$, $YW \rightarrow Z$, 则 $XW \rightarrow Z$)

$\because X \rightarrow Y$, $\therefore XW \rightarrow YW$ (增广律, 两边扩充W)

$\therefore YW \rightarrow Z$, $\therefore XW \rightarrow Z$ (传递律)

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

Armstrong公理

- 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^+ 。
- 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：
 1. 有效性：可由定理6.1得证
 2. 完备性：只需证明**逆否命题**：若函数依赖 $X \rightarrow Y$ 不能由F从Armstrong公理导出，那么它必然不为F所蕴含。



Armstrong公理完备性证明

即：存在一个关系 r ， F 中所有的函数依赖都满足 r ，而不能用公理推出的 $X \rightarrow Y$ 却不能满足 r ，即 F 不能逻辑蕴涵 $X \rightarrow Y$ 。证明就是要找到这个 r 。

证明：

(1) 引理：若 $V \rightarrow W$ 成立，且 $V \subseteq X_F^+$ ，则 $W \subseteq X_F^+$

(2) 构造一张二维表 r ，它由下列两个元组构成，可以证明 r 必是 $R(U,F)$ 的一个关系，即 F^+ 中的全部函数依赖在 r 上成立。

X_F^+	$U-X_F^+$
11.....1	00.....0
11.....1	11.....1

如果一个属性集不完全属于 X_F^+ ，则该属性集在2个元组上的属性值必不相等。

(3) 若 $X \rightarrow Y$ 不能由 F 从Armstrong公理导出，则 Y 不是 X_F^+ 的子集。

$X \subseteq X_F^+$ ，而 $Y \not\subseteq X_F^+$ ， X 在 r 上2个元组上一致，而 Y 在 r 的2个元组上不一致。

课堂练习



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

- ❖ $A \rightarrow H$ 是
- ❖ $CG \rightarrow HI$ 是
- ❖ $AG \rightarrow I$ 是

问题：能不能用一种一般性的算法来判定 $X \rightarrow Y$ 是否是 F 的逻辑蕴涵？



F的闭包

- **例如：**从 $F = \{X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_n\}$ 出发可推导出 2^n 个不同的函数依赖。

$$F = \{X \rightarrow Y, Y \rightarrow Z\}$$

$$F^+ = \{$$

$$\begin{array}{llllll} X \rightarrow \varphi, & Y \rightarrow \varphi, & Z \rightarrow \varphi, & XY \rightarrow \varphi, & XZ \rightarrow \varphi, & YZ \rightarrow \varphi, \\ X \rightarrow X, & Y \rightarrow Y, & Z \rightarrow Z, & XY \rightarrow X, & XZ \rightarrow X, & YZ \rightarrow Y, \\ X \rightarrow Y, & Y \rightarrow Z, & & XY \rightarrow Y, & XZ \rightarrow Y, & YZ \rightarrow Z, \\ X \rightarrow Z, & Y \rightarrow YZ, & & XY \rightarrow Z, & XZ \rightarrow Z, & YZ \rightarrow YZ, \\ X \rightarrow XY, & & & XY \rightarrow XY, & XZ \rightarrow XY, & XYZ \rightarrow Y, \\ X \rightarrow XZ, & & & XY \rightarrow YZ, & XZ \rightarrow XZ, & XYZ \rightarrow YZ, \\ X \rightarrow YZ, & & & XY \rightarrow XZ, & XZ \rightarrow XY, & XYZ \rightarrow XZ, \\ X \rightarrow XYZ, & & & XY \rightarrow XYZ, & XZ \rightarrow XYZ, & XYZ \rightarrow XYZ \} \end{array}$$

$F = \{X \rightarrow A_1, \dots, X \rightarrow A_n\}$ 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 $R(U, F)$, $U = \{A_1, A_2, \dots, A_n\}$, $X \subseteq U$, 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 $X \rightarrow A_i$ 中所有 A_i 的属性集合为**属性集 X 关于 F 的闭包**, 记为 X_F^+ 。即:

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

- **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:

$$A_F^+ = ABC, \quad B_F^+ = BC, \quad C_F^+ = C$$



属性闭包

引理6.2：函数依赖 $X \rightarrow Y$ 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 $X \rightarrow Y$ 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。

计算属性闭包的算法

输入： X, F

输出： X_F^+

方法：

- (1) $X^{(0)} = \emptyset, X^{(1)} = X;$
- (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 $Y \rightarrow Z$, 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 $R(U, F)$, $U = \{A, B, C, D, E\}$, $F = \{AB \rightarrow C, B \rightarrow D, C \rightarrow E, EC \rightarrow B, AC \rightarrow B\}$, 计算 $(AB)_F^+$ 。

所用依赖	$X^{(0)}$	$X^{(1)}$
初始值	\emptyset	AB
第一遍 $AB \rightarrow C, B \rightarrow D$	AB	ABCD
第二遍 $C \rightarrow E, AC \rightarrow B$	ABCD	ABCDE
第三遍 $EC \rightarrow B$	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： $\because (AB)_F^+ = ABCDE = U$, 且 $A_F^+ \neq U$, $B_F^+ \neq U$
 $\therefore AB$ 为R的一个码。

课堂练习



- 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=Φ;
- (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 ∪ {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 ∪ {XZ};
- (7) 输出K中所有候选码。



函数依赖的等价和覆盖

- ❖ **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，则称F覆盖G，同时G也覆盖F。

定理：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

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 $X \rightarrow A$, $F - \{X \rightarrow A\}$ 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 $X \rightarrow A$, 不存在 X 的子集 Z , 使得 F 与 $(F - \{X \rightarrow A\}) \cup \{Z \rightarrow A\}$ 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 $X \rightarrow Y$, 令 $G = F - \{X \rightarrow Y\}$, 求 X_G^+ , 若 $Y \subseteq X_G^+$, 则 $X \rightarrow Y$ 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 $X \rightarrow A$, 设 $X = B_1 B_2 \dots B_m$, 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 。

最小函数依赖集计算举例

例：已知 $F = \{ AB \rightarrow C, C \rightarrow A, BC \rightarrow D, ACD \rightarrow B, D \rightarrow EG, BE \rightarrow C, CG \rightarrow BD, CE \rightarrow AG \}$, 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

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

2) 去掉 F_1 中多余的函数依赖

对 $AB \rightarrow C$, 令 $G = F_1 - \{AB \rightarrow C\}$, 计算 $(AB)_G^+ = AB$

$\because C \not\subset (AB)_G^+$, $\therefore AB \rightarrow C$ 不能去掉;

对 $C \rightarrow A$, 令 $G = F_1 - \{C \rightarrow A\}$, 计算 $C_G^+ = C$

$\because A \not\subset C^+$, $\therefore C \rightarrow A$ 不能去掉;

.....

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

最小函数依赖集计算举例

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

3) 去掉 F_2 中函数依赖左部多余属性

对 $AB \rightarrow C$, 在 F_2 中分别计算:

对A, 求 $B_{F_2}^+ = B$, 因为 $C \not\subset B_{F_2}^+$, 所以A不是多余属性。

对B, 求 $A_{F_2}^+ = A$, 因为 $C \not\subset A_{F_2}^+$, 所以B不是多余属性。

对 $BC \rightarrow D$, 在 F_2 中分别计算:

对B, 求 $C_{F_2}^+ = CA$, 因为 $D \not\subset C_{F_2}^+$, 所以B不是多余属性。

对C, 求 $B_{F_2}^+ = B$, 因为 $D \not\subset B_{F_2}^+$, 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, 与计算顺序有关。



最小函数依赖集

例: $U = \{ Sno, SDept, Manager, Cname, G \}$

$F = \{ Sno \rightarrow SDept, SDept \rightarrow Manager, (Sno, Cname) \rightarrow G \}$

设 $F' = \{ Sno \rightarrow SDept, Sno \rightarrow Manager,$
 $SDept \rightarrow Manager, (Sno, Cname) \rightarrow G,$
 $(Sno, SDept) \rightarrow SDept \}$

F 和 F' 是最小依赖集吗?

解:

F 是最小依赖集,

F' 不是。

因为: $F' - \{ Sno \rightarrow Manager \}$ 与 F' 等价,
 $F' - \{ (Sno, SDept) \rightarrow SDept \}$ 与 F' 等价

最小函数依赖集



例：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。

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