- 6.1 问题的提出
- 6.2 规范化
- 6.3 数据依赖的公理系统

6.3 数据依赖的公理系统

定义6.11 对于满足一组函数依赖F的关系模式 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t、s，若t[X]=s[X]，则 t[Y]=s[Y]），则称F逻辑蕴涵X →Y。

- Armstrong公理系统
  - 一套推理规则，是模式分解的理论基础
  - 用于求给定关系的码
  - 用于从一组函数依赖求得蕴涵的函数依赖

Armstrong公理系统 设U为属性集总体，F是U上的一组函数依赖，于是有关系模式R <U,F >。对R <U,F> 来说有以下的推理规则：

A1 自反律：若$Y \subseteq X \subseteq U$，则$X \rightarrow Y$ 为F所蕴涵。

A2 增广律：若$X \rightarrow Y$为F所蕴涵，且$Z \subseteq U$，则$XZ \rightarrow YZ$ 为F所蕴涵。

A3 传递律：若$X \rightarrow Y$及$Y \rightarrow Z$为F所蕴涵，则$X \rightarrow Z$ 为F所蕴涵。

- 定理6.1 Armstrong推理规则是正确的。

证明

■A1 自反律

设$Y \subseteq X \subseteq U$ 。

对R <U,F> 的任一关系r中的任意两个元组t、s：

若t[X]=s[X]，由于$Y \subseteq X$，有t[Y]=s[Y]，

所以X→Y成立，

自反律得证。

■A2 增广律

设X→Y为F所蕴涵，且Z $\subseteq$ U。

对R<U,F> 的任一关系r中任意的两个元组t、s：

若t[XZ]=s[XZ]，则有t[X]=s[X]和t[Z]=s[Z]；

由X→Y，于是有t[Y]=s[Y]，

所以t[YZ]=s[YZ]，XZ→YZ为F所蕴涵，

增广律得证。

■A3 传递律

设$X \rightarrow Y$及$Y \rightarrow Z$为F所蕴涵。

对R<U,F> 的任一关系r中的任意两个元组t、s：

若t[X]=s[X]，由于$X \rightarrow Y$，有t[Y]=s[Y]；

再由$Y \rightarrow Z$，有t[Z]=s[Z]，

所以$X \rightarrow Z$为F所蕴涵，

传递律得证。

根据A1，A2，A3三条推理规则可以得到三条推理规则：

- 合并规则：

由X→Y，X→Z，有X→YZ。

- 伪传递规则：

由X→Y，WY→Z，有XW→Z。

- 分解规则：

由X→Y及Z⊆Y，有X→Z。

根据合并规则和分解规则，可得引理6.1

引理6.1 $X \rightarrow A_1 A_2 \ldots A_k$成立的充分必要条件是$X \rightarrow A_i$成立（i=1，2，…，k）。

- 定义6.12 在关系模式$R<U,F>$中为F所逻辑蕴涵的函数依赖的全体叫作F的闭包，记为F^+。

- 定义6.13 设F为属性集U上的一组函数依赖，X、Y $\subseteq U$， $X_F^+=\{ A|X\rightarrow A$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。

引理6.2 设F为属性集U上的一组函数依赖，X、$Y \subseteq U$，$X \rightarrow Y$能由F根据Armstrong公理导出的充分必要条件是$Y \subseteq X_F^+$。

- 引理6.2的用途

  判定$X \rightarrow Y$是否能由F根据Armstrong公理导出的问题，就转化为求出X_F^+，判定Y是否为X_F^+的子集的问题。

求闭包的算法

算法6.1 求属性集X（$X \subseteq U$）关于U上的函数依赖集F的闭包X_F^+

- 输入：X，F
- 输出：X_F^+
- 步骤：

①令$X^{(0)}=X$，$i=0$

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

③ $X^{(i+1)}=B \cup X^{(i)}$

④判断$X^{(i+1)}=X^{(i)}$

⑤若$X^{(i+1)}$与$X^{(i)}$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6.11] 已知关系模式$R<U, F>$，其中

$U=\{A, B, C, D, E\}$；

$F=\{AB\to C, B\to D, C\to E, EC\to B, AC\to B\}$。

求$(AB)_F^+$。

解：由算法6.1，设$X^{(0)}=AB$。

计算$X^{(1)}$：逐一的扫描F集合中各个函数依赖，找左部为A、B或AB的函数依赖。得到两个：$AB\rightarrow C$，$B\rightarrow D$。于是$X^{(1)}=AB\cup CD=ABCD$。

因为$X^{(0)}\neq X^{(1)}$，所以再找出左部为$ABCD$子集的那些函数依赖，又得到$C\rightarrow E$，$AC\rightarrow B$，于是$X^{(2)}=X^{(1)}\cup BE=ABCDE$。

因为$X^{(2)}$已等于全部属性集合，所以$(AB)_F^+=ABCDE$。

有效性与完备性

- 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

定理6.2 Armstrong公理系统是有效的、完备的。

证明：

1. 有效性
   - 有效性实际上是“正确性”
   - 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$X \to Y$不能由F从Armstrong公理导出，那么它必然不为F所蕴涵
- 分三步证明：

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

证：因为 $V \subseteq X_F^+$，所以有$X \to V$成立；

因为$X \to V$，$V \to W$，于是$X \to W$ 成立；

所以$W \subseteq X_F^+$。

（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。

$$\begin{array}{cc} \overbrace{11\ldots\ldots1}^{X_F^+} & \overbrace{00\ldots\ldots1}^{U\text{-}X_F^+} \\ 11\ldots\ldots1 & 11\ldots\ldots1 \end{array}$$

若r 不是$R<U,F>$ 的关系，则必由于F中有某一个函数依赖$V \rightarrow W$在r上不成立所致。由r 的构成可知，V 必定是X_F^+ 的子集，而W不是 X_F^+ 的子集，可是由第（1）步，$W \subseteq X_F^+$，矛盾。

所以r 必是$R<U,F>$的一个关系。

（3）若$X \rightarrow Y$不能由F从Armstrong公理导出，则Y不是X_F^+的子集。（引理6.2）

因此必有Y的子集Y'满足$Y' \subseteq U - X_F^+$，

则$X \rightarrow Y$在r中不成立，

即$X \rightarrow Y$必不为$R<U,F>$蕴涵。

*

Armstrong公理的完备性及有效性说明:

- “导出”与“蕴涵”是两个完全等价的概念
- F^+ ：为F所逻辑蕴涵的函数依赖的全体（定义6.12 ）

↓

- F^+ ：可以说成由F出发借助Armstrong公理导出的函数依赖的集合

*

- 定义6.14 如果$G^+=F^+$，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。

两个函数依赖集等价是指它们的闭包等价

*

函数依赖集等价的充要条件

引理6.3 $F^+ = G^+$ 的充分必要条件是$F \subseteq G^+$和$G \subseteq F^+$。

证：必要性显然，只证充分性。

（1）若$F \subseteq G^+$，则$X_F^+ \subseteq X_{G^+}^+$。

（2）任取$X \rightarrow Y \in F^+$ 则有 $Y \subseteq X_F^+ \subseteq X_{G^+}^+$。

所以$X \rightarrow Y \in (G^+)^+ = G^+$。即$F^+ \subseteq G^+$。

（3）同理可证$G^+ \subseteq F^+$，所以$F^+ = G^+$。

引理6.3给出了判断两个函数依赖集等价的可行算法

如何判定$F \subseteq G^+$？

只需逐一对F中的函数依赖$X \rightarrow Y$考察 Y是否属于$X_{G^+}^+$

定义6.15 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，亦称为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$X \to A$，使得F与F-$\{X \to A\}$等价。

（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。

即F中的函数依赖均不能由F中其他函数依赖导出

F中各函数依赖左部均为最小属性集（不存在冗余属性）

*

[例6.12] 考察6.1节中的关系模式$S<U,F>$， 其中：

U={Sno, Sdept, Mname, Cno, Grade}，

F={Sno→Sdept, Sdept→Mname, (Sno,Cno)→Grade}

F是最小覆盖

F'={Sno→Sdept, Sno→Mname, Sdept→Mname, (Sno,Cno)→Grade, (Sno,Sdept)→Sdept}

F'不是最小覆盖

因为：F'- {Sno→Mname} 与 F'等价

F'- {(Sno,Sdept)→Sdept} 也与F'等价

定理6.3 每一个函数依赖集F均等价于一个极小函数依赖集F_m。此F_m称为F的最小依赖集。

证：构造性证明，分三步对F进行“极小化处理”，找出F的一个最小依赖集。

（1）逐一检查F中各函数依赖FD_i：$X\rightarrow Y$，

若$Y=A_1A_2 \ldots A_k$，$k\geq 2$，

则用$\{X\rightarrow A_j \mid j=1,2,\ldots,k\}$来取代$X\rightarrow Y$。

引理6.1保证了F变换前后的等价性。

（2）逐一检查F中各函数依赖FD_i：$X \rightarrow A$,

令$G=F-\{X \rightarrow A\}$,

若$A \in X_G^+$，则从F中去掉此函数依赖。

由于F与G 等价的充要条件是$A \in X_G^+$

因此F变换前后是等价的。

*

（3）逐一取出F中各函数依赖FD_i：$X \to A$,

设$X=B_1B_2 \ldots B_m$，$m \geq 2$，逐一考查B_i（$i=1$，2，…，m），

若$A \in (X-B_i)_F^+$，则以$X-B_i$取代X。

由于F与$F-\{X \to A\} \cup \{Z \to A\}$等价的充要条件是$A \in Z_F^+$，其中$Z=X-B_i$，因此$F$变换前后是等价的。

最后剩下的F就一定是极小依赖集。

因为对F的每一次"改造"都保证了改造前后的两个函数依赖集等价，因此剩下的F与原来的F等价。

证毕

*

定理6.3的证明过程是求F极小依赖集的过程,也是检验F是否为极小依赖集的一个算法

若改造后的F与原来的F相同，说明F就是一个最小依赖集

[例6.13] $F=\{A\rightarrow B, B\rightarrow A, B\rightarrow C, A\rightarrow C, C\rightarrow A\}$

F的最小依赖集：

$$F_{m1}= \{A\rightarrow B, B\rightarrow C, C\rightarrow A\}$$

*

F的最小依赖集F_m不一定是唯一的，它与对各函数依赖FD_i及$X \rightarrow A$中X各属性的处置顺序有关。

[例6.13]（续）

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

F_{m1}、F_{m2}都是F的最小依赖集：

$$F_{m1}=\{A\rightarrow B, B\rightarrow C, C\rightarrow A\}$$

$$F_{m2}=\{A\rightarrow B, B\rightarrow A, A\rightarrow C, C\rightarrow A\}$$

*

在$R<U,F>$中可以用与F等价的依赖集G来取代F

原因：两个关系模式$R_1<U,F>$，$R_2<U,G>$，如果F与G等价，那么R_1的关系一定是R_2的关系。反过来，R_2的关系也一定是R_1的关系。

*

小结

- 规范化理论为数据库设计提供理论的指南和工具
  - 仅仅是指南和工具
- 并不是规范化程度越高，模式就越好
  - 必须结合应用环境和现实世界的具体情况合理地选择数据库模式