

## Chapter8 关系数据库的规范化理论

### 8.2 规范化理论

#### 8.2.1 函数依赖

设有关系模式 $R(U)$ ， $U$ 是关系模式 $R$ 的属性集合， $X$ 、 $Y$ 是 $U$ 的子集。若对于任一个符合关系模式 $R(U)$ 的关系 $r$ 中的任一元组 $t$ 在 $X$ 中的属性值确定后，则元组 $t$ 在 $Y$ 中的属性值也必确定，则称 $Y$ 函数依赖于 $X$ ，或者称 $X$ 函数决定 $Y$ ，并记为 $X \rightarrow Y$ 。

#### 平凡/非平凡函数依赖

- 一个函数依赖关系 $X \rightarrow Y$ 如满足 $Y \not\subseteq X$ ，则称此函数依赖是非平凡的函数依赖。否则，我们称其为平凡函数依赖。

#### 完全函数依赖

- 在关系模式 $R(U)$ 中，如有 $X \subseteq U$ ， $Y \subseteq U$ ，满足 $X \rightarrow Y$ ，且对任何 $X$ 的真子集 $X'$ 都有 $X' \not\rightarrow Y$ ，则称 $Y$ 完全函数依赖于 $X$ ，并记作： $X \xrightarrow{f} Y$

#### 部分函数依赖

- 在关系模式 $R(U)$ 中，如有 $X \subseteq U$ ， $Y \subseteq U$ ，满足 $X \rightarrow Y$ ，但 $Y$ 不完全函数依赖于 $X$ ，则称 $Y$ 部分依赖于 $X$ ，并记作： $X \xrightarrow{p} Y$

#### 传递函数依赖

- 在关系模式 $R(U)$ 中，如有 $X \subseteq U$ ， $Y \subseteq U$ ， $Z \subseteq U$ 且满足： $X \rightarrow Y$ ， $Y \not\subseteq X$ ， $Y \not\rightarrow X$ ， $Y \rightarrow Z$ ，则称 $Z$ 传递函数依赖于 $X$ ；否则，称为非传递函数依赖（直接函数依赖）。

#### Armstrong公理系统

- 规则 $R_1$ ：自反规则

- 如果 $Y$ 是 $X$ 的子集，则： $X \rightarrow Y$

- 规则 $R_2$ ：增广规则

- 如果 $X \rightarrow Y$ ，则： $XZ \rightarrow YZ$

- 规则 $R_3$ ：传递规则

- 如果 $X \rightarrow Y$ ， $Y \rightarrow Z$ ，则： $X \rightarrow Z$

- 规则 $R_4$ ：分解规则

- 如果 $X \rightarrow YZ$ ，则： $X \rightarrow Y$

➤规则R<sub>5</sub>: 合并规则

▪ 如果  $X \rightarrow Y$  且  $X \rightarrow Z$ , 则:  $X \rightarrow YZ$

➤规则R<sub>6</sub>: 伪传递规则

▪ 如果  $X \rightarrow Y$  且  $WY \rightarrow Z$ , 则:  $WX \rightarrow Z$

关键字 (也称为 码 或 key)

➤在关系模式  $R(U, F)$  中, 如有  $K \subseteq U$  且满足:

$$K \xrightarrow{f} U$$

则称  $K$  为  $R$  的关键字。

关键字的所有子集都不能确定U!

主属性和非主属性

**定义8.7:** 主属性(集)

- 由关系模式  $R$  的所有关键字中的属性所构成的集合被称为关系模式  $R$  的 **主属性集**
- 主属性集中的属性被称为关系模式  $R$  的 **主属性**

**定义8.8:** 非主属性(集)

- 由主属性集之外的其它属性所构成的集合被称为关系模式  $R$  的 **非主属性集**
- 非主属性集中的属性被称为关系模式  $R$  的 **非主属性**

寻找关键字

➤**方法一:** 利用Armstrong公理系统中的推导规则, 从已知的函数依赖集F中推导得到如下的函数依赖关系:

$$K \xrightarrow{f} U$$

**算法8-2:** 寻找关系模式  $R(U, F)$  的关键字  $K$

1. set  $K := R$ ;
2. for each attribute  $A$  in  $K$ 
  - {
  - compute  $(K - A)_F^+$ ;
  - if  $(K - A)_F^+$  contains all the attributes in  $R$ , then
    - {
    - set  $K := K - \{A\}$ ;
    - }
  - }

## 8.2.2 与函数依赖有关的范式

### 第一范式 (1NF)

➤ 如果关系模式  $R(U)$  中的每个属性值都是一个不可分割的数据量，则称该关系模式满足第一范式，并记为： $R \in 1NF$

### 第二范式 (2NF)

▪ 设有关系模式  $R(U) \in 1NF$ ，且其每个非主属性都完全函数依赖于关键字，则称关系模式  $R(U)$  满足第二范式，并记作： $R \in 2NF$

### 模式分解的方法

- 1) 找出所有不满足范式M要求的函数依赖关系
- 2) 选择一个不符合要求的函数依赖关系作如下的分解：
  - 假设  $X \xrightarrow{f} Y \in F^+$  且不满足范式M的要求，则我们将关系模式R分解为如下的两个子关系：
    - »  $R_1 (X \cup Y, \{X \rightarrow Y\})$
    - »  $R_2 (Head(R) - Y, F_2)$ ，其中：  
 $F_2 = \{A \rightarrow B \mid A \rightarrow B \in F^+ \text{ 且 } (A \cup B) \subseteq Head(R_2)\}$
- 3) 对于分解得到的子关系模式 $R_2$ 重复上述的步骤1)和步骤2)，直到所有的子关系模式都能满足范式M的要求
- 4) 合并那些具有相同关键字的子关系模式

### 第三范式 (3NF)

➤ 设有关系模式  $R(U) \in 2NF$ ，且其每个非主属性都不传递函数依赖于关键字，则称关系模式  $R(U)$  满足第三范式，并记作： $R \in 3NF$

➤ 不符合3NF要求的函数依赖关系具有如下特征：

▪  $X \xrightarrow{f} Y$ ，Y是关系SCG的非主属性，而X并不是关键字，或者X只是关键字的一个真子集

### BCNF

▪ 设关系模式  $R(U)$  满足 1NF，且若  $X \rightarrow Y$  时 X 必包含有该关系模式的关键字，则称关系模式  $R(U)$  满足BCNF范式，并记作： $R \in BCNF$



## 8.3 规范化所引起的一些问题

### 8.3.1 函数依赖理论

#### 函数依赖集的等价

如果F中的每一个FD都是G所逻辑蕴涵的，且G中的每一个FD也都是F所逻辑蕴涵的，则F和G逻辑等价。

#### 最小函数依赖集

- 与函数依赖集 **F** 相等价的所有函数依赖集中的最小者被称为函数依赖集 **F** 的**最小函数依赖集**。
- ‘**最小函数依赖集**’也被称为‘**最小覆盖**’ (Minimal Cover)

- 对于 **F** 中的每一个FD关系  $X \rightarrow A$  均作如下判断：
1. 依赖因素 **A** 为单个属性；
  2. 令：  $F_1 = F - \{X \rightarrow A\}$ ，则  $F_1^+ \neq F^+$ ；  
– 不存在冗余的函数依赖关系
  3. 对于决定因素**X**的每一个真子集**Y**( $Y \subset X$ ) 均作如下判断：令  $F_2 = (F - \{X \rightarrow A\}) \cup \{Y \rightarrow A\}$ ，则  $F_2^+ \neq F^+$   
– 不存在部分函数依赖关系

### 8.3.2 模式分解的研究

#### 无损连接性

- 设**R**是一个关系模式，**F**是关系模式**R**上的一个函数依赖集， $\rho = \{R_1, R_2, \dots, R_k\}$  是对**R**的一个分解。
- 如果对**R**中满足**F**的每一个关系实例 **r** 都有：

$$r = \pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \dots \bowtie \pi_{R_k}(r)$$

其中：  $\pi_{R_i}(r)$  为关系实例**r**在关系模式**R<sub>i</sub>**上的投影，即根据**R<sub>i</sub>**中的属性对**r**执行投影运算后的结果。

- 则称该分解  $\rho$  相对于**F**是“无损联接分解”，或称分解  $\rho$  具有无损联接性。

**定理8.3.2.1：** 如果**R**的分解为  $\rho = \{R_1, R_2\}$ ，**F**为**R**所满足的函数依赖集合，分解**ρ**具有无损联接性的充分必要条件是：

$$R_1 \cap R_2 \rightarrow (R_1 - R_2) \text{ 或 } R_1 \cap R_2 \rightarrow (R_2 - R_1)$$

## 依赖保持性

- 设  $F$  是属性集  $U$  上的函数依赖集,  $Z$  是  $U$  上的一个子集,  $F$  在  $Z$  上的投影用  $\pi_Z(F)$  来表示
- 设存在关系模式  $R$  的一个分解  $\rho = \{ R_1, R_2, \dots, R_k \}$ ,  $F$  是  $R$  上的函数依赖集。如果:  
$$F^+ = (\pi_{R_1}(F) \cup \pi_{R_2}(F) \cup \dots \cup \pi_{R_k}(F))^+$$
- 则称分解  $\rho$  具有依赖保持性

## 分解算法

假设存在一个关系模式  $R$ , 其函数依赖集为  $F$ , 利用该算法可以将其分解到满足第三范式, 同时该分解满足无损联接性和依赖保持性

### ➤ 算法8-4:

- 1) 计算  $F$  的最小函数依赖集, 并用来代替  $F$  进行下面的模式分解;
- 2)  $S = \Phi$ ;
- 3) 对  $F$  中的每一个函数依赖  $X \rightarrow Y$  做如下处理:
  - ❖ 如果: 对于集合  $S$  中的每一个子关系模式  $Z$  都有:  
 $X \cup Y \not\subseteq Z$
  - ❖ 则由  $X$  和  $Y$  构成一个新的子关系加入到集合  $S$  中,  
即:  $S = S \cup \text{Heading}(X \cup Y)$
- 4) 如果关系  $R$  的每一个候选关键字  $K$  都没有出现在分解后的子关系模式中, 即: 对集合  $S$  中的每一个子关系模式  $Z$  都有:  $K \not\subseteq Z$   
则: 从关系  $R$  中任选一个候选关键字  $K$ , 并由  $K$  中的属性单独构成一个子关系模式并加入到集合  $S$  中去,  
即:  $S = S \cup \text{Heading}(K)$