

## ● 第六章 关系理论

### 2. 建立一个关于系，学生，班级，学会等诸信息的关系数据库

描述学生的属性有：学号，姓名，出身年月，系名，班号，宿舍区

描述班级的属性有：班号，专业号，系名，人数，入校年份

描述系的属性有：系名，系号，系办公室地点，人数

描述学会的属性有：学会名，成立年份，地点，人数

有关语义如下：一个系有若干专业，每个专业每年只招收一个班，每个班有若干学生，一个系的学生住在同一个宿舍区，每个学生可以参加若干学会，每个学会会有若干学生，学生参加某学会有一个入会年份。

1) 请给出关系模式，写出每个关系模式的极小函数依赖集，指出是否存在传递函数依赖，对于函数依赖左部是多属性的情况，讨论函数依赖是完全函数依赖，还是部分函数依赖。

2) 列出每个关系模式所属的范式并说明原因。

3) 指出各关系的候选码，外部码，有没有全码存在？

说明：

- 1) 所有的关系名和属性名使用题目给的中文名称即可
- 2) 极小函数依赖集是指不含部分依赖和传递依赖等多余的依赖。

答：

1) 关系模式与极小函数依赖集：

◆ 学生（学号，姓名，出生年月，班号，宿舍区）

$F_m = \{ \text{学号} \rightarrow \text{姓名}, \text{学号} \rightarrow \text{出生年月}, \text{学号} \rightarrow \text{班号}, \text{班号} \rightarrow \text{系名}, \text{系名} \rightarrow \text{宿舍区} \}$

码：学号

传递函数依赖：学号 $\rightarrow$ 班号，班号 $\rightarrow$ 系名 $\Rightarrow$ 学号 $\rightarrow$ 系名（非主属性对码的传递依赖）

班号 $\rightarrow$ 系名，系名 $\rightarrow$ 宿舍区 $\Rightarrow$ 班号 $\rightarrow$ 宿舍区（非主属性之间的传递依赖，而不是非主属性对码的传递依赖）

◆ 班级（班号，专业名，系名，班级人数，入校年份）

$F_m = \{ \text{班号} \rightarrow \text{专业名}, \text{班号} \rightarrow \text{班级人数}, \text{班号} \rightarrow \text{入校年份},$

$\text{专业名} \rightarrow \text{系名}, (\text{专业名}, \text{入校年份}) \rightarrow \text{班号} \}$

主属性：班号，专业名，入校年份 非主属性：系名，班级人数，

码：班号 或者 （专业名，入校年份）

非主属性对码的传递依赖：班号→专业名，专业名→系名 ⇒ 班号→系名

（专业名，入校年份） → 班号：完全函数依赖

（专业名，入校年份） → 系名：部分函数依赖

#### ◆ 系（系号，系名，系办公地点，系人数）

$F_m = \{\text{系号} \rightarrow \text{系名}, \text{系名} \rightarrow \text{系号}, \text{系号} \rightarrow \text{系办公地点}, \text{系号} \rightarrow \text{系人数}\}$

不存在部分依赖也没有传递依赖。

#### ◆ 学会（学会名，成立年月，地点，学会人数）

$F_m = \{\text{学会名} \rightarrow \text{成立年月}, \text{学会名} \rightarrow \text{地点}, \text{学会名} \rightarrow \text{学会人数}\}$

不存在部分依赖也没有传递依赖。

#### ◆ 学生-学会（学号，学会名，入会年份）

$F_m = \{(\text{学号}, \text{学会名}) \rightarrow \text{入会年月}\}$

（学号，学会名）→入会年月：完全函数依赖

不存在部分依赖也没有传递依赖。

**举例求解极小函数依赖集的过程：**

##### ● 学生关系

学号→姓名，学号→出生年月，学号→班号，班号→系名，系名→宿舍

学号→系名，学号→宿舍

$F_m$  求解：（依照  $F_m$  求解算法）

①②省略（因为左右都是单一属性）

③消除冗余函数依赖

(1) 学号<sub>F</sub><sup>+</sup> - {学号→姓名} = {学号，出生年月，班号，系名，宿舍} = F0

姓名 不属于 F0， 所以不能删。

(2) 学号<sub>F</sub><sup>+</sup> - {学号→出生年月} = {姓名，班号，系名，宿舍} = F0

出生年月不属于 F0， 所以不能删。

$$(3) \{学号\}_F^+ - \{学号 \rightarrow 班号\} = \{姓名, 出生年月, 系名, 宿舍\} = F0$$

班号不属于 F0， 所以不能删。

$$(4) \{班号\}_F^+ - \{班号 \rightarrow 系名\} = \{班号\} = F0$$

系名不属于 F0， 所以 班号 $\rightarrow$ 系名 不能删。

$$(5) \{系名\}_F^+ - \{系名 \rightarrow 宿舍\} = \{系名\} = F0$$

宿舍不属于 F0， 所以不能删。

$$(6) \{学号\}_F^+ - \{学号 \rightarrow 系名\} = \{姓名, 出生年月, 班号, 系名, 宿舍\} = F0$$

系名属于 F0， 所以 学号 $\rightarrow$ 系名 可以删。

$$(7) \{学号\}_F^+ - \{学号 \rightarrow 宿舍\} = \{姓名, 出生年月, 班号, 系名, 宿舍\} = F0$$

宿舍属于 F0， 所以可以删。

因此 Fm 如下：

学号 $\rightarrow$ 姓名, 学号 $\rightarrow$ 出生年月, 学号 $\rightarrow$ 班号, 班号 $\rightarrow$ 系名, 系名 $\rightarrow$ 宿舍

#### ● 班级关系

班号 $\rightarrow$ 专业名, 班号 $\rightarrow$ 系名, 班号 $\rightarrow$ 班人数, 班号 $\rightarrow$ 入学年份,

专业名 $\rightarrow$ 系名, (专业名, 入学年份) $\rightarrow$  班号

Fm 求解: (依照 Fm 求解算法)

①省略 (因为右边是单一属性)

②消除冗余属性: (专业名, 入学年份) $\rightarrow$  班号

(1) 去掉入学年份:

$$专业名_F^+ = \{专业名, 系名\} = F0, 班号 \text{ 不属于 } F0, \text{ 所以入学年份不多余。}$$

(2) 去掉专业名:

$$入学年份_F^+ = \{入学年份\} = F0, 班号 \text{ 不属于 } F0, \text{ 所以专业名不多余。}$$

③消除冗余函数依赖

(1) 班号<sub>F</sub><sup>+</sup> - {班号->专业名} = {系名, 人数, 入学年份} = F0

专业名 不属于 F0, 所以不能删。

(2) 班号<sub>F</sub><sup>+</sup> - {班号->系名} = {班号, 专业名, 系名, 人数, 入学年份} = F0

系名 属于 F0, 所以可以删。

。。。。

其余均不能删除

因此 Fm 如下：

班号->专业名, 班号->班人数, 班号->入学年份, 专业名->系名,  
(专业名, 入学年份) -> 班号

2) 码:

关系	候选码	外部码	全码
学生	学号	班号, 系名	无
班级	班号, (专业名, 入校年份)	系名	无
系	系号, 系名	无	无
学会	学会名	无	无
学生-学会	(学号, 学会名)	学号, 学会名	无

### 3) 关系模式与极小函数依赖集

关系	所属范式	原因
学生	2NF	不存在部分依赖, 但是有非主属性对码传递依赖, 所以属于 2NF。
班级	1NF	存在非主属性对码部分依赖和递依赖, 所以属于 1NF。
系	4NF	不存在部分依赖也没有传递依赖, 并且每一个决定因素都包含码, 所以属于 BCNF。由于其上没有多值依赖, 所以是 4NF。
学会	4NF	不存在部分依赖也没有传递依赖, 并且每一个决定因素都包含码, 所以属于 BCNF。由于其上没有多值依赖, 所以是 4NF。
学生-学会	4NF	不存在部分依赖也没有传递依赖, 并且每一个决定因素都包含码, 所以属于 BCNF。由于其上没有多值依赖, 所以是 4NF。

### 6. 考虑关系模式 $R(A,B,C,D,E)$ , 回答下面问题。

1) 若 A 是 R 的候选码, 具有函数依赖  $BC \rightarrow DE$ , 那么在什么条件下 R 是 BCNF

答: 属性 BC 包含码。

2) 如果存在依赖:  $A \rightarrow B$ ,  $BC \rightarrow D$ ,  $DE \rightarrow A$ , 列出 R 的所有码

答: ACE, DEC, BCE

3) 如果存在依赖:  $A \rightarrow B$ ,  $BC \rightarrow D$ ,  $DE \rightarrow A$ , R 属于 3NF 还是 BCNF?

答: 因为 ABCDE 都是主属性, 所以 R 是 3NF。

主属性之间存在部分依赖 ( $A \rightarrow B$ ), 所以不是 BCNF。

候选码计算:

$LR = \{A, B, D\}$   $L = CE$

$(CEA)^+ = \{CEABD\} = U$        $(CEB)^+ = \{CEBDA\} = U$        $(CED)^+ = \{CEDAB\} = U$

所以 CEA, CEB, CED 是候选码。

## 8. 证明

### 1) 如果 R 是 BCNF，则 R 是 3NF，反之则不然。

证明：

#### a) 如果 R 是 BCNF，则 R 是 3NF

反正法：假设 R 是 BCNF，但 R 不是 3NF。

由于不是 3NF，所以存在非主属性对码的传递依赖，即 R 中存在候选码 X、属性组 Y 和非主属性 Z，满足： $X \rightarrow Y$ ， $Y \rightarrow Z$ ，Y 不决定 X。由于 Y 不决定 X，因此 Y 不包含候选码，即  $Y \rightarrow Z$  的决定因素不包含码，这与 R 是 BCNF 矛盾，所以结论得正。

#### b) 如果 R 是 3NF，则 R 不一定是 BCNF

证明：STJ 数据库 S 表示学生，T 表示教师，J 表示课程，根据语义有  $(S, T) \rightarrow J$ ， $(S, J) \rightarrow T$ ， $T \rightarrow J$ ，此时 R 是 3NF，但是不是 BCNF ( $T \rightarrow J$  中，T 是决定因素但不包含码)。

### 2) 如果 R 是 3NF，则 R 一定是 2NF。

反正法：R 是 3NF，但不是 2NF。可知，必然存在一个非主属性 Z 不完全函数依赖于码，因此存在候选码 X 的真子集 Y，使得  $Y \rightarrow Z$  成立，且 Y 不能决定 X；

同时 Y 是主属性，Z 不是主属性，因此 Y 不包含 Z。

综上，存在候选码 X，属性组 Y，非主属性组 Z，有  $X \rightarrow Y$  (Y 不能决定 X)， $Y \rightarrow Z$ ，出现了传递依赖，这与 3NF 的定义矛盾，所以原题得正。

## 附加作业

针对课本算法 6.2（保持函数依赖的 3NF 转换分解算法），用例子说明该算法的运行过程。

示例如下：

### ➤ 算法6.2（合成法）转换为3NF的保持函数依赖的分解

[Algorithm]保持依赖分解成3NF的算法。

Input：关系模式R(U, F)，F是函数依赖集最小覆盖。

Output：R的一个保持依赖分解 $\rho$ ， $\rho$ 中的每个关系模式都是F在该模式上投影的3NF。

Method：(1)把R中不出现在F中的属性去掉并单独组成一模式。

(2)对 $\forall X \rightarrow A \in F$ ，则以XA组成一模式；若有 $X \rightarrow A_1, X \rightarrow A_2, \dots, X \rightarrow A_m$ 都属于F，则以 $XA_1A_2 \dots A_m$ 组成一模式(即将n个模式合并为一个模式)。

(3)取 $\rho$ 为上述模式之集合，则 $\rho$ 即为所求之分解。

例关系R:  $U = A, B, C, D, E, F, G$

$F = \{A \rightarrow B, A \rightarrow C, C \rightarrow D, C \rightarrow E, E \rightarrow FG\}$

保持函数依赖的分解：

1)  $R_1(A, B), R_2(A, C), R_3(C, D), R_4(C, E), R_5(E, F, G)$  【相同左部为一组】

2)  $R_{12}(A, B, C), R_{34}(C, D, E), R_5(E, F, G)$  【合并相同的左部】