

第7章 关系数据库设计理论

7.1 学习要点

关系数据库设计理论是关系数据库的又一个重点。关系数据库的逻辑设计主要是设计关系模式，而深入了解函数依赖和码的概念则是设计和分解关系模式的基础。学习本章的目的有两个。一个是理论方面的，本章用更加形式化的关系数据理论来描述和研究关系模型。另一个是实践方面的，关系数据库设计理论是我们进行数据库设计的有力工具。

- 1、知道什么是函数依赖、完全函数依赖、部分函数依赖和传递函数依赖，能确定两个或多个属性间的函数依赖，计算属性的闭集；
- 2、理解关系的码和超码、主属性和非主属性；
- 3、理解 1NF、2NF、3NF 和 BCNF 的定义，并能辨别某关系属于哪种范式类型；
- 4、掌握规范化一个关系模式的原则方法，能够将某 1NF 关系规范化为 3NF 或 BCNF；
- 5、理解多值依赖和连接依赖，初步掌握分解成第四范式的方法。

7.2 习题讲解

1. 理解并给出下列名词的涵义：

函数依赖、部分函数依赖、传递函数依赖、超码、多值依赖。

答：

函数依赖是数据库中两个属性集之间的约束。设 $R(U)$ 是属性集 U 上的关系模式， X 、 Y 是 U 的子集， r 是 R 的任一具体关系。设 t_1 、 t_2 是关系 r 中的任意两个元组，如果 $t_1[X]=t_2[X]$ ，有 $t_1[Y]=t_2[Y]$ ，则称 X 函数决定 Y ，或 Y 函数依赖于 X ，记作 $X \rightarrow Y$ 。

在关系模式 $R(U)$ 中， X 、 Y 是 U 的子集，若 $X \rightarrow Y$ ，且存在 $X' \subset X$ ，使 $X' \rightarrow Y$ ，则称 $X \rightarrow Y$ 是部分函数依赖（partial functional dependency），记作 $X \xrightarrow{P} Y$ 。

在关系模式 $R(U)$ 中， X 、 Y 是 U 的子集，若 $X \rightarrow Y$ ， $Y \rightarrow Z$ ，并且 Y 不函数依赖于 X ，则称 Z 传递函数依赖于 X 。

包含候选码的属性或属性组称为超码(Super key)。

设有关系模式 $R(U)$ ， X 、 Y 为 U 的子集， $Z=U-XY$ ， r 是 R 的任一关系，如果 r 中存在两个元组 t_1 、 t_2 满足 $t_1[X]=t_2[X]$ ，则 r 中必然存在两个元组 t_3 、 t_4 ，使得

(1) $t_3[X]=t_4[X]=t_1[X]=t_2[X]$

(2) $t_3[Y]=t_1[Y]$ 且 $t_4[Y]=t_2[Y]$

(3) $t_3[Z]=t_2[Z]$ 且 $t_4[Z]=t_1[Z]$

则称 $X \twoheadrightarrow Y$ 是多值依赖（multivalued dependency, MVD）， X 多值决定 Y 。

2. 设有关系模式 $R(ABCDE)$ ，有函数依赖集 $F=\{A \rightarrow B, AB \rightarrow D, E \rightarrow AD, E \rightarrow C\}$ 和 $G=\{A \rightarrow BD, E \rightarrow AC\}$ ，判断 F 和 G 是否等价。

答：

$A_G^+=ABD$ ， $E_G^+=ABCDE$ ，可知 F 中的函数依赖 $A \rightarrow B$ 、 $AB \rightarrow D$ 、 $E \rightarrow AD$ 、 $E \rightarrow C$ 都属于 G^+ ，所以 $F \subseteq G^+$ ； $A_F^+=ABD$ ， $E_F^+=ABCDE$ ，可知 G 中的函数依赖 $A \rightarrow BD$ 、 $E \rightarrow AC$ 都属于 F^+ ，所以

$G \subseteq F^+$ 。

根据引理 5.3, F 与 G 等价。

3. 设有一关系模式 $R(ABCD)$, 其函数依赖集 $F = \{A \rightarrow BC, B \rightarrow C, AB \rightarrow C, AC \rightarrow D\}$, 求 F 的最小依赖集 F_{\min} 。

答:

(1) 首先用分解规则将 F 中所有的函数依赖的右部属性单一化。得 $F = \{A \rightarrow B, A \rightarrow C, B \rightarrow C, AB \rightarrow C, AC \rightarrow D\}$ 。

(2) 去掉 F 中多余的依赖。具体做法是: 从第一个函数依赖(假设为 $X \rightarrow Y$)开始, 把它从 F 中去掉, 求 X^+ , 若 X^+ 包含 Y , 则 $X \rightarrow Y$ 是多余的, 要去掉; 若 X^+ 不包含 Y , 则不能去掉 $X \rightarrow Y$ 。检查全部依赖后可得 G 。显然 G 符合最小函数依赖集定义 5.6 的条件(2)。

这里, 对于 $A \rightarrow B$, 由于 $(A)^+ = ACD$ 不包含 B , 所以不能去掉; 而由于从 F 中去掉 $A \rightarrow C$ 后, $A^+ = ABCD$, 包含了 C , 所以 $A \rightarrow C$ 是多余的, 从 F 中去掉; 接下去 $B \rightarrow C$ 不能去掉, 而且 $AB \rightarrow C$ 明显多余, 从 F 中去掉; $(AC)^+ = ABC$ 不包含 D , 所以 $AC \rightarrow D$ 不能去掉, 最后得 $G = \{A \rightarrow B, B \rightarrow C, AC \rightarrow D\}$ 。

(3) 去掉 F_2 中的函数依赖左边的多余的属性。具体做法是: 检查 F 中所有左边是非单属性的函数依赖, 如 $XY \rightarrow A$, 要判断 Y 是否为多余属性, 只要在 F 中求 X^+ , 若 X^+ 包含 A , 则 Y 是多余属性, 否则 Y 不是多余属性。该题 G 中 $AC \rightarrow D$ 的 C 属性多余, 去掉后得到的函数依赖集 $F_{\min} = \{A \rightarrow B, B \rightarrow C, A \rightarrow D\}$ 。

4. 设有关系模式 $R(ABCDE)$, 其函数依赖 $F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A\}$, 试求

(1) 计算 B^+ 。

(2) 求 R 的所有码。

答:

(1) 根据算法 5.1, 令 $X^{(0)} = B$

在 F 中找出左边是 B 的子集的函数依赖, 有 $B \rightarrow D$;

$X^{(1)} = BD$

因为 $X^{(1)} \neq X^{(0)}$,

继续在 F 中找出左边是 BD 的子集的函数依赖, 由于不存在这样的函数依赖, 所以不必再计算下去了。

结果为: $B^+ = BD$ 。

(2) 因为 $A^+ = ABCDE$, $E^+ = ABCDE$, $(BC)^+ = ABCDE$, $(CD)^+ = ABCDE$, $B^+ = BD$, $C^+ = C$, $D^+ = D$, 所以候选码为 A 、 E 、 BC 、 CD 。

5. 试验证上题 4 中的 R 的以下两个分解是否具有无损连接性:

$\rho_1 = \{R_1(ABC), R_2(ADE)\}$

$\rho_2 = \{R_1(ABC), R_2(CDE)\}$ 。

答:

对于 $\rho_1 = \{R_1(ABC), R_2(ADE)\}$, $R_1 \cap R_2 = A$, $R_1 - R_2 = BC$, $R_2 - R_1 = DE$, 根据 $A^+ = ABCDE$ 知 $(R_1 \cap R_2) \rightarrow (R_1 - R_2)$ 和 $(R_1 \cap R_2) \rightarrow (R_2 - R_1)$ 都满足, 根据定理 5.4, 分解具有无损连接性。

对于 $\rho_2 = \{R_1(ABC), R_2(CDE)\}$, $R_1 \cap R_2 = C$, $R_1 - R_2 = AB$, $R_2 - R_1 = DE$, 根据 $C^+ = C$ 知 $(R_1 \cap R_2) \rightarrow (R_1 - R_2)$ 和 $(R_1 \cap R_2) \rightarrow (R_2 - R_1)$ 都不满足, 根据定理 5.4, 分解不具有无损连接性。

6. 设有关系模式 $R(ABCDEFGHI)$, 其函数依赖 $F = \{AB \rightarrow C, A \rightarrow DE, B \rightarrow GH, D \rightarrow I\}$, 试求

(1) 求 R 的所有码。

(2) 将 R 分解成 2NF。

(3) 将 R 无损分解成 3NF, 并且具有保持依赖性。

答:

(1) 因为 $A^+ = ADEI$, $B^+ = BGH$, $(AB)^+ = ABCDEGHI$, A、B 没有在函数依赖的右边出现, 所以候选码为 AB。

(2) 因为 2NF 要求每一个非主属性完全函数依赖于码, 而 AB 为码, 非主属性 C 完全函数依赖于码, 非主属性 DEGH 都部分依赖于码, 所以将 R 分解为三个模式 R1、R2 和 R3:

R1 (ABC); R2 (ADEI); R3 (BGH)。

这样, R1 中的非主属性 C 完全函数依赖于码 AB; R2 中的非主属性 DEI 完全函数依赖于码 A; R3 中的非主属性 GH 完全函数依赖于码 B。因此, R1、R2 和 R3 都满足 2NF。

(3) 首先求出 F 的最小依赖集 F_{min} , 即 $F_{min} = \{AB \rightarrow C, A \rightarrow DE, B \rightarrow GH, D \rightarrow I\}$ 。

根据算法 5.3, 得 $\rho = \{ABC, ADE, BGH, DI\}$, 由于 R 的码是 AB, 所以根据算法 5.4 可得, $\sigma = \rho \cup \{AB\} = \{ABC, ADE, BGH, DI, AB\}$, 由于 $AB \subseteq ABC$, 因此 $\sigma = \{ABC, ADE, BGH, DI\}$, σ 是 R 的同时具有无损连接性和依赖保持性的 3NF 分解。

7. 设有关系模式 R(ABCD), 其函数依赖 $F = \{A \rightarrow C, C \rightarrow A, B \rightarrow AC, D \rightarrow AC\}$, 试求

(1) 求 R 的所有码。

(2) 将 R 无损分解成 BCNF。

(3) 将 R 无损分解 3NF, 并且具有保持依赖性。

答:

(1) 因为 $(BD)^+ = ABCD$, B、D 没有在函数依赖的右边出现, 所以候选码为 BD。

(2) 根据算法 5.5:

初始化 $\rho = \{R\}$;

R 的码是 BD, 因此 A、C、B 和 D 不是超码。首先由左边不包含码的非平凡函数依赖 $A \rightarrow C$, 从 R 中分出 $AC(A, C)$, 得 $\rho = \{R1(ABD), AC(A, C)\}$ 。

然后, R1 的码是 BD, 因此 B 和 D 不是超码。由左边不包含码的非平凡函数依赖 $B \rightarrow AC$, 从 R1 中分出 $BA(B, A)$, 得 $\rho = \{BA(B, A), BD(B, D), AC(A, C)\}$ 。

BA、BD 和 AC 都是 BCNF, 并且具有无损连接性。 ρ 不具有依赖保持性。

另外, $\sigma = \{DA(D, A), BD(B, D), AC(A, C)\}$ 也是具有无损连接性的分解, 但同样不具有依赖保持性。

(3) 首先求出 F 的最小依赖集 F_{min} , 即 $F_{min} = \{A \rightarrow C, C \rightarrow A, B \rightarrow A, D \rightarrow A\}$ 。

根据算法 5.3, 得 $\rho = \{AC, CA, BA, DA\}$, 由于 R 的码是 BD, 所以根据算法 5.4 可得, $\sigma = \rho \cup \{BD\} = \{AC, CA, BA, DA, BD\}$, 由于 $AC = CA$, 因此 $\sigma = \{AC, BA, DA, BD\}$, σ 是 R 的同时具有无损连接性和依赖保持性的 3NF 分解。另外 $\rho = \{AC, BC, DC, BD\}$ 也是 R 的同时具有无损连接性和依赖保持性的 3NF 分解。

8. 在 1NF~BCNF 范围内, 指出下列关系模式是第几范式, 并说明理由。

(1) R(ABC), $F = \{A \rightarrow B, B \rightarrow A, A \rightarrow C\}$ 。

(2) R(ABC), $F = \{A \rightarrow B, B \rightarrow A, C \rightarrow A\}$ 。

(3) R(ABCD), $F = \{B \rightarrow D, AB \rightarrow C\}$ 。

(4) R(ABC), $F = \{B \rightarrow C, AC \rightarrow B\}$ 。

(5) R(ABCD), $F = \{B \rightarrow D, D \rightarrow B, AB \rightarrow C\}$ 。

(6) R(ABCDE), $F = \{AB \rightarrow CE, E \rightarrow AB, C \rightarrow D\}$ 。

答:

(1) 候选码为 A、B, 可知 F 中每个非平凡函数依赖的左边都包含码, 所以 $R \in BCNF$ 。

(2) 候选码为 C, 可知非主属性都完全函数依赖于码, 但存在非主属性 B 传递函数依

依赖于码 C，所以 $R \in 2NF$ 。

(3) 候选码为 AB，可知非主属性 D 部分函数依赖于码，所以 $R \in 1NF$ 。

(4) 候选码为 AB、AC，可知 A、B、C 都是主属性，不存在非主属性部分函数依赖或传递对函数依赖于码，另外非平凡函数依赖 $B \rightarrow C$ 的左边不包含码，所以 $R \in 3NF$ 。

(5) 候选码为 AB、AD，可知不存在非主属性 C 部分函数依赖或传递对函数依赖于码，另外非平凡函数依赖 $B \rightarrow D$ 的左边不包含码，所以 $R \in 3NF$ 。

(6) 候选码为 E、AB，可知非主属性都完全函数依赖于码，但存在非主属性 D 传递函数依赖于码，所以 $R \in 2NF$ 。

9. 设关系模式 $R(ABC)$ 上有一多值依赖 $A \twoheadrightarrow B$ ，已知 R 的当前关系存在三个元组，如图 7.1 所示，试问该关系中至少还应存在哪些元组才能满足多值依赖的要求。

答：

A	B	C
a1	b1	c1
a1	b2	c2
a1	b3	c3

图 7.1 R 的当前关系

根据多值依赖的定义，可知当前关系中还应该存在 6 个元组： $(a1, b1, c2)$ ， $(a1, b1, c3)$ ， $(a1, b2, c1)$ ， $(a1, b2, c3)$ ， $(a1, b3, c1)$ ， $(a1, b3, c2)$ 。

7.3 自测题

一、选择题

1. 设学生关系模式为：学生（学号，姓名，年龄，性别，成绩，专业），则该关系模式的主码是（ ）。

- A. 姓名 B. 学号，姓名 C. 学号 D. 学号，姓名，年龄

2. 设一关系模式为 $R(A, B, C, D, E)$ 及函数依赖 $F=\{A \rightarrow B, B \rightarrow E, E \rightarrow A, D \rightarrow E\}$ ，则关系模式的 R 候选码是（ ）。

- A. AD B. CD C. EB D. EC

3. 下列有关范式的叙述中正确的是（ ）。

- A. 如果关系模式 $R \in 1NF$ ，且 R 中主属性完全函数依赖于主码，则 R 是 $2NF$
B. 如果关系模式 $R \in 3NF$ ， $X, Y \in U$ ，若 $X \rightarrow Y$ ，则 R 是 $BCNF$
C. 如果关系模式 $R \in BCNF$ ，若 $X \twoheadrightarrow Y$ ($Y \notin X$) 是平凡的多值依赖，则 R 是 $4NF$
D. 一个关系模式如果属于 $4NF$ ，则一定属于 $BCNF$ ；反之不成立

4. 给定关系模式 $SCP(Snum, Cnum, P)$ ，其中 $Snum$ 表示学号， $Cnum$ 表示课程号， P 表示名次。若每一名学生每门课程都有一定的名次，而每门课程每一名次只有一名学生，则以下叙述中错误的是（ ）。

- A. $(Snum, Cnum)$ 和 $(Cnum, P)$ 都可以作为候选码
B. $(Snum, Cnum)$ 是唯一的候选码
C. 关系模式 SCP 既属于 $3NF$ 也属于 $BCNF$
D. 关系模式 SCP 没有非主属性

5. 消除多值依赖所引起的冗余是属于（ ）。

A. 2NF B. 3NF C. 4NF D. BCNF

6. 下列叙述中正确的是 ()。

A. 3NF 不能保持多值依赖 B. 4NF 肯定能保持多值依赖
C. BCNF 可能保持函数依赖 D. 4NF 不能保持函数依赖

7. 若关系模式 R 的函数依赖集 F 中的所有候选码都是决定因素, 则 R 的最高范式是 ()。

A. 2NF B. 3NF C. BCNF D. 4NF

8. 关系模式 R 包含属性 {A1, A2, A3, A4, A5}, 其中 {A1, A2} 为码, 则下面说法正确的是 ()。

A. {A1} 或 {A2} 有可能单独成为 R 的码
B. {A1, A2, A3} 必然也是 R 的码
C. R 中绝不可能出现两个在 A1, A2 上取值完全相同的元组
D. R 的所有元组中, A1 或者 A2 的值都是不能重复的

9. 设一关系模式为 R (A, B, C) 及函数依赖 $F = \{AB \rightarrow C, BC \rightarrow A, B \rightarrow C\}$, 则下列说法 () 是正确的。

A. R 一定消除了插入和删除异常 B. R 属于 BCNF
C. R 属于 3NF D. R 一定消除了冗余

10. 在关系模式 R 中, 函数依赖 $X \rightarrow Y$ 的语义是 ()

A. 在 R 的任意两个关系中, 若 X 值相等, 则 Y 值也相等
B. 在 R 的当前关系中, 若两个元组的 X 值相等, 则 Y 值也相等
C. 在 R 的任意关系中, Y 值应与 X 值相等
D. 在 R 的当前关系中, Y 值应与 X 值相等

11. 函数依赖 $X \rightarrow Y$ 能从 Armstrong 推理规则推出的充分必要条件是 ()

A. $Y \rightarrow X$ B. $Y \subseteq X^+$ C. $X \subseteq Y^+$ D. $X^+ = Y^+$

12. 在关系模式 R(U) 中, 包含属性 {A, B, C, D}, R 的码为 {A, B}, 则下面的选项中有 R 的超码的有 ()

A. {A} B. {C, D} C. {A, B, C, D} D. {B, C, D}

13. 对于 FD $X \rightarrow Y$, 如果有 $Y \subseteq X$, 那么称 $X \rightarrow Y$ 是一个 ()

A. 包含函数依赖 B. 增广的函数依赖
C. 传递的函数依赖 D. 平凡的函数依赖

14. 设有关系模式 R (A, B, C, D), F 是 R 上成立的 FD 集, $F = \{A \rightarrow B, B \rightarrow C\}$, B 的闭包 B^+ 为由 B 函数决定的属性集, 则 B^+ 为 ()

A. ABC B. BCD C. BC D. C

15. 设有关系模式 R (A, B, C, D), F 是 R 上成立的 FD 集, $F = \{B \rightarrow C, D \rightarrow C\}$, 属性集 AB 的闭包 $(AB)^+$ 为 ()

A. ABCD B. ABC C. CD D. BCD

16. 现在只知道关系模式包含的属性和码 (用下划线表示), 则一定是第二范式的关系是 ()

A. R1 {A1, A2, A3} B. R2 {B1, B2, B3}
C. R3 {C1, C2, C3} D. 以上都不是

二、填空题

1. 在关系模式 R 中, 若每个数据项都是不可再分割的, 那么 R 一定属于第____范式。

2. 通过模式分解把属于低级范式的关系模式转换为几个属于高级范式的关系模式的集合,

这一过程称为_____。

3、关系数据库设计理论，主要包括三个方面内容：_____、_____和_____。

4、对于函数依赖 $X \rightarrow Y$ ，如果 $Y \subseteq X$ ，则称 $X \rightarrow Y$ 是一个_____。

5、如果 $Y \subseteq X \subseteq U$ ，则 $X \rightarrow Y$ 成立。这条推理规则称为_____。

6、如果 $X \rightarrow Y$ 和 $Y \rightarrow Z$ 成立，则 $X \rightarrow Z$ 成立。这条推理规则称为_____。

7、如果 $X \rightarrow Y$ 和 $Z \subseteq Y$ 成立，则 $X \rightarrow Z$ 成立。这条推理规则称为_____。

三、问答题

1、关系规范化一般应遵循的原则是什么？

2、多值依赖与函数依赖有哪些主要的区别？

3、试叙述 "无损联接" 的定义。

4、试叙述 "保持函数依赖" 的定义。

四、综合题

1、建立一个关于系、学生、班级、学会等诸信息的关系数据库。其中描述学生的属性有：学号、姓名、出生年月、系名、班号、宿舍号；描述班级的属性有：班号、专业名、系名、人数、入校年份；描述系的属性有：系名、系号、系办公地点、人数；描述学会的属性有：学会名、成立年份、地点、人数。

若有关语义为：一个系有若干专业，每个专业每年只招一个班，每个班有若干学生；一个系的学生住在同一宿舍区；每个学生可参加若干学会，每个学会会有若干学生；学生参加某学会会有一个入会年份。

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

指出各关系模式的候选码、外部码，并说明有没有全码存在？

2、已知关系模式 $R(U, F)$ ， $U=\{A, B, C, D, E\}$ ， $F=\{A \rightarrow B, D \rightarrow C, BC \rightarrow E, AC \rightarrow B\}$ 。求 $(AE)_F^+$ 、 $(AD)_F^+$ 。

3、已知关系模式 $R(U, F)$ ， $U=\{A, B, C, D, E, G\}$ ， $F=\{AB \rightarrow C, D \rightarrow EG, C \rightarrow A, BE \rightarrow C, BC \rightarrow D, CG \rightarrow BD, ACD \rightarrow B, CE \rightarrow AG\}$ 。请将 F 化为最小函数依赖集。

4、已知关系模式 $R(U, F)$ ， $U=\{A, B, C\}$ ； $F=\{A \rightarrow B, B \rightarrow A, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$ 。求 F 的最小函数依赖集。

5、对给定的关系模式 $R(U, F)$ ， $U=\{A, B, C\}$ ， $F=\{A \rightarrow B\}$ ，有如下的两个分解：

(1) $P_1=\{AB, BC\}$

(2) $P_2=\{AB, AC\}$

判断这两个分解是否无损。

6、关系 $R(U, F)$ ， $U=\{A, B, C, D, E, F, G, H, I, J\}$ ， $F=\{ABD \rightarrow E, AB \rightarrow G, B \rightarrow F, C \rightarrow J, CJ \rightarrow I, G \rightarrow H\}$ 。

(1) 该函数依赖集 F 是最小函数依赖集吗？

(2) 求出该关系的候选码。

7、关系 $R(A, B, C, D, E)$ 满足下列函数依赖：

$F=\{A \rightarrow C, C \rightarrow D, B \rightarrow C, DE \rightarrow C, CE \rightarrow A\}$

(1) 给出关系 R 的候选关键字。

(2) 判断 $P=\{AD, AB, BC, CDE, AE\}$ 是否无损连接分解。

(3) 将 R 分解为 BCNF，并具有无损连接性。

8、设关系模式 $R(S\#, C\#, GRADE, TNAME, TADDR)$ ，其属性分别表示学生学号、选修课程和编号，成绩，任课教师姓名，教师地址等意义。

如果规定，每个学生每学一门课程只有一个成绩；每门课只有一个任课教师；每个教师有一个地址（此处不允许教师同名同姓）。

(1) 试写出关系模式 R 基本的 FD 和候选码。

(2) 试把 R 分解成 2NF 模式集，并说明理由。

(3) 试把 R 分解成 3NF 模式集，并说明理由。

9、设关系模式 R (A, B, C, D) 满足下列函数依赖: $F = \{AB \rightarrow C, C \rightarrow D, D \rightarrow A\}$, 试写出关系模式 R 的候选码。

10、设关系模式 R (U, F), $U = \{A, B, C, D, E\}$, $F = \{AB \rightarrow DE, C \rightarrow E, D \rightarrow C, E \rightarrow A\}$, 现把 R 投影到 S (A, B, C), 求出 S 中的函数依赖集。

11、设关系模式 R (U, F), $U = \{A, B, C, D, E, F, G\}$, $F = \{A \rightarrow B, BC \rightarrow DE, AEF \rightarrow G\}$, 计算 $\{A, C\}$ 关于这个函数依赖集的闭包。该函数依赖集是否蕴涵函数依赖 $ACF \rightarrow DG$?

7.4 自测题参考答案

一、选择题

1. C 2. C 3. D 4. B 5. C 6. C 7. D 8. C 9. C 10. B
11. B 12. C 13. D 14. C 15. B 16. B 17. D 18. D

二、填空题

1. 1
2. 规范化
3. 数据依赖 范式 模式设计方法（可交换位置）
4. 平凡的函数依赖
5. 自反律
6. 传递律
7. 分解律
8. 保持函数依赖

三、问答题

1、关系规范化一般应遵循的原则如下

(1) 将关系模式进行无损连接分解，在关系模式分解的过程中，数据不能丢失或增加，要保证数据的完整性。

(2) 合理地选择规范化的程度。在规范化时，既要考虑到低级范式造成的冗余度高，数据的不一致性，又要考虑到高级范式查询效率低的矛盾。

(3) 正确性和可实现性原则。

2、多值依赖与函数依赖的主要区别如下：

(1) 在关系模式 R (U) 中，函数依赖 $X \rightarrow Y$ 的有效性仅决定 X、Y 这两个属性集的值。只要在关系 R (U) 的任一关系 r 中，元组在 X 和 Y 上的值满足函数依赖的定义，则函数依赖 $X \rightarrow Y$ 在任何属性集 W ($XY \subseteq W \subseteq U$) 上都成立。

对于多值依赖，若 $X \twoheadrightarrow Y$ 在 W ($W \subset U$) 上成立，但在 U 上不一定成立，所以多值依赖的有效性与属性集的范围有关。 $X \twoheadrightarrow Y$ 在 U 上成立，则在 W ($XY \subseteq W \subseteq U$) 上成立，反之则不然。

(2) 若函数依赖 $X \rightarrow Y$ 在 $R(U)$ 上成立, 则对于任何 $Y' \subset U$, 均有 $X \rightarrow Y$ 成立。

对于多值依赖 $X \twoheadrightarrow Y$ 在 $R(U)$ 上成立, 但不能断言对于任何 $Y' \subset U$, 有 $X \twoheadrightarrow Y$ 成立

3、答: 设 R 是关系模式, 分解成数据库模式 $\rho = \{R_1, \dots, R_K\}$, F 是 R 上的 FD 集. 如果对 R 中满足 F 的每个关系 r 都有:

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

则称分解 ρ 相对于 F 是“无损联接分解”。

4、答: 设 R 是关系模式, 分解成数据库模式 $\rho = \{R_1, \dots, R_K\}$, F 是 R 上的 FD 集, 如果 $F^+ = (\cup \pi_{R_i}(F))^+$, 则称分解 ρ 保持函数依赖集 F 。

四、综合题

1、解:

(1) 关系模式如下:

学生: $S(Snum, Sname, Sbirth, Dept, Class, Rno)$

班级: $C(Class, Pname, Dept, Cnum, Cyear)$

系: $D(Dept, Dno, Office, Dnum)$

学会: $M(Mname, Myear, Maddr, Mnum)$

学生学会: $SM(Snum, Mname, SMyear)$

(2) 每个关系模式的最小函数依赖集如下:

学生 S 的最小函数依赖集如下:

$Snum \rightarrow Sname, Snum \rightarrow Sbirth, Class \rightarrow Dept, Snum \rightarrow Class, Dept \rightarrow Rno$

因为 $Snum \rightarrow Dept, Dept \rightarrow Rno$, 所以 $Snum$ 与 Rno 之间存在着传递依赖

因为 $Class \rightarrow Dept, Dept \rightarrow Rno$, 所以 $Class$ 与 Rno 之间存在着传递依赖

因为 $Snum \rightarrow Class, Class \rightarrow Dept$, 所以 $Snum$ 与 $Dept$ 之间存在着传递依赖

班级 C 的最小函数依赖集如下:

$Class \rightarrow Pname, Class \rightarrow Cnum, Class \rightarrow Cyear, Pname \rightarrow Dept$

因为 $Class \rightarrow Pname, Pname \rightarrow Dept$, 所以 $Class$ 与 $Dept$ 之间存在着传递依赖

系 D 的最小函数依赖集如下:

$Dept \rightarrow Dno, Dno \rightarrow Office, Dno \rightarrow Dnum$

因为 $Dept \rightarrow Dno, Dno \rightarrow Office$, 所以 $Dept$ 与 $Office$ 之间存在着传递依赖

因为 $Dept \rightarrow Dno, Dno \rightarrow Dnum$, 所以 $Dept$ 与 $Dnum$ 之间存在着传递依赖

学会 M 的最小函数依赖集如下:

$Mname \rightarrow Myear, Mname \rightarrow Maddr, Mname \rightarrow Mnum$

该模式不存在传递依赖

学生学会 SM 的最小函数依赖集如下:

$(Snum, Mname) \rightarrow SMyear$ 是完全函数依赖

(3) 各关系模式的候选码、外部码、全码如下:

学生 S 候选码: $Snum$; 外部码: $Dept, Class$; 无全码

班级 C 候选码: $Class$; 外部码: $Dept$; 无全码

系 D 候选码: $Dept$ 或 Dno ; 无外部码; 无全码

学会 M 候选码: Mname; 无外部码; 无全码

学生学会 SM 候选码: (Snum, Mname); 外部码: Snum, Mname; 无全码

2、解:

(1) 求 $(AE)_F^+$, 根据上述算法, 设 $X^{(0)} = AE$ 。

计算 $X^{(1)}$ 。逐一扫描 F 中的各个函数依赖, 寻找左部为 A、E 或 AE 的函数依赖, 找到一个 $A \rightarrow B$, 故有 $X^{(1)} = AE \cup B$ 。

计算 $X^{(2)}$ 。逐一扫描 F 中的各个函数依赖, 寻找左部为 ABE 或 ABE 子集的函数依赖, 因为找不到这样的函数依赖, 所以, $X^{(1)} = X^{(2)}$, 算法终止。 $(AE)_F^+ = ABE$ 。

(2) 求 $(AD)_F^+$, 根据上述算法, 设 $X^{(0)} = AD$ 。

计算 $X^{(1)}$ 。逐一扫描 F 中的各个函数依赖, 寻找左部为 A、D 或 AD 的函数依赖, 找到两个: $A \rightarrow B$, $D \rightarrow C$ 函数依赖。故有 $X^{(1)} = AD \cup BC$ 。

计算 $X^{(2)}$ 。逐一扫描 F 中的各个函数依赖, 寻找左部为 ADBC 或 ADBC 子集的函数依赖, 找到两个: $BC \rightarrow E$, $AC \rightarrow B$ 函数依赖。故有 $X^{(2)} = ABCD \cup E$ 。所以, $X^{(2)} = ABCDE = U$, 算法终止。 $(AD)_F^+ = ABCDE$ 。

3、解:

(1) 假设 $CG \rightarrow B$ 为冗余的函数依赖, 那么, 从 F 中去掉它后能根据 Armstrong 公理系统的推理规则导出。

因为 $CG \rightarrow D$, $C \rightarrow A$ (已知)

所以 $CGA \rightarrow ACD$ (增广律)

又因为 $ACD \rightarrow B$ (已知)

所以 $CGA \rightarrow B$ (传递律)

又因为 $C \rightarrow A$ (已知)

所以 $CG \rightarrow B$ (蕴含)

(2) 同理可证: $CE \rightarrow A$ 是冗余的。

(3) 又因为 $C \rightarrow A$, $ACD \rightarrow B$ 可知, 所以去掉左边多余的属性得 $CD \rightarrow B$ 。

4、解:

答案 1: 设 $B \rightarrow A$ 是冗余的, 将其从 F 中去掉, 看能否根据 Armstrong 公理系统的推理规则导出。

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

所以 $B \rightarrow A$ (传递律)

故 $B \rightarrow A$ 是冗余的, 将其从 F 中去掉, 得 F_1 , $F_1 = \{A \rightarrow B, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$ 。

又设 $A \rightarrow C$ 为冗余, 将其从 F_1 中去掉,

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

所以 $A \rightarrow C$ (传递律)

故 $A \rightarrow C$ 是冗余的, 将其从 F_1 中去掉, 得 F_{M1} , $F_{M1} = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$ 。

通过分析可以发现, 在 F_{M1} 中的其它函数依赖是非冗余的, 所以, F_{M1} 为最小函数依赖集。

答案 2: 设 $B \rightarrow C$ 是冗余的, 将其从 F 中去掉, 看能否根据 Armstrong 公理系统的推理规则导出。

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

所以 $B \rightarrow C$ (传递律)

故 $B \rightarrow C$ 是冗余的, 将其从 F 中去掉, 得 F_{M2} , $F_{M2} = \{A \rightarrow B, B \rightarrow C, A \rightarrow C, C \rightarrow A\}$ 。

因为在 F_{M2} 中的其它函数依赖是非冗余的, 所以, F_{M2} 为最小函数依赖集。

5、解:

(1) 可根据无损连接定理判断本题。

因为 $AB \cap BC = B$

$AB - BC = A$

$BC - AB = C$

所以 $B \rightarrow A \notin F^+$

$B \rightarrow C \notin F^+$

故 P1 为有损连接。

(2) 根据无损连接定理判断本题。

因为 $AB \cap AC = A$

$AB - AC = B$

$AC - AB = C$

所以 $B \rightarrow A \notin F^+$

故 P2 为无损连接。

6、解：

(1) 该函数依赖集不是最小函数依赖集。因为 $CJ \rightarrow I$ 中的 J 为冗余属性。证明如下：

因为 $C \rightarrow J$, $CJ \rightarrow I$ (已知)

所以 $C \rightarrow I$ 逻辑蕴含 $CJ \rightarrow I$ 。

(2) 该关系的候选码为 ABCD。

因为 {ABCDGJ} 是一个超码。所谓超码是指所有出现在函数依赖左边的属性以及没有在 F 中出现的属性的集合。

因为 $C \rightarrow J$ (已知)

所以可将 J 从超码中去掉

又因为 $AB \rightarrow G$ (已知)

所以可将 G 从超码中去掉

此时，超码中只剩下 ABCD，由于它们都没有在函数依赖集的任何一个函数依赖的右边出现，即它们也是 L 属性，所以它们都不能从超码中去掉。故候选码为 ABCD。

所谓超码是指能唯一标识关系中的每一个元组，但超码可能有冗余属性，不含冗余属性的是候选码。

7、解：

(1) 从函数依赖集 F 中看出，候选关键字至少包含 BE，因为 BE 为 L 类属性。在关系模式 R 中，无 NLR 属性，BE 不依赖于谁，而 $(BE)^+ = ABCDE$ ，所以，BE 是 R 的候选关键字。

(2) 构造一个二维表，如下表所示。

属性 模式	A	B	C	D	E
AD	a1	b12	b13	a4	b15
AB	a1	a2	b23	b24	b25
BC	b31	a2	a3	b34	b35
CDE	b41	b42	a3	a4	a5
AE	a1	b52	b53	b54	a5

①根据 $A \rightarrow C$ ，对上表进行处理，由于属性列 A 上第一、二、五行相同，但属性列 C 上对应的一、二、五行上无 a_i 元素，所以，只能将 b13、b24、b53 改为行号最小值 b13。又

根据 $C \rightarrow D$ ，将属性列 D 上 b34 改为 a4，修改后的表如下表所示。

模式 \ 属性	A	B	C	D	E
AD	a1	b12	b13	a4	b15
AB	a1	a2	b13	b24	b25
BC	b31	a2	a3	a4	b35
CDE	b41	b42	a3	a4	a5
AE	a1	b52	b13	b54	a5

②根据 $B \rightarrow C$ ，对上表进行处理，由于属性列 B 上第二、三行相同，但属性列 C 上对应的三行为 a3 元素，所以，只能将第二行 b13 改为 a3。又根据 $DE \rightarrow C$ ， $CE \rightarrow A$ 不能修改此表，所以修改后的表如下表所示。

模式 \ 属性	A	B	C	D	E
AD	a1	b12	b13	a4	b15
AB	a1	a2	a3	b24	b25
BC	b31	a2	a3	a4	b35
CDE	b41	b42	a3	a4	a5
AE	a1	b52	b13	b54	a5

③根据 $A \rightarrow C$ ，对上表进行处理，由于属性列 A 上第一、二、五行相同，但属性列 C 上对应的一、二、五行上有 a3 元素，所以只能将第一、五行 b13 改为 a3。又根据 $C \rightarrow D$ 将属性列 D 上 b24、b24 改为 a4，修改后的表如下表所示。

模式 \ 属性	A	B	C	D	E
AD	a1	b12	a3	a4	b15
AB	a1	a2	a3	a4	b25
BC	b31	a2	a3	a4	b35
CDE	b41	b42	a3	a4	a5
AE	a1	b52	a3	a4	a5

④根据 $CE \rightarrow A$ ，对上表进行处理，由于属性列 CE 上第四、五行相同，但属性列 A 上对应的第五行为 a1 元素，所以，将第四行 b41 改为 a1。所以修改后的表如下表所示。

模式 \ 属性	A	B	C	D	E
AD	a1	b12	a3	a4	b15
AB	a1	a2	a3	a4	b25
BC	b31	a2	a3	a4	b35
CDE	a1	b42	a3	a4	a5
AE	a1	b52	a3	a4	a5

⑤继续扫描 F 不能修改此表，由于找不到一行全为 a，所以该分解是有损的。
 (3) 考虑 $A \rightarrow C$ ，因为 AC 不包含候选关键字，所以 AC 不是 BCNF，故将 ABCDE 分解为两个子模式：R1 (AC) 和 R2 (ABDE)。此时， $R1 \in BCNF$ 。

继续分解 R2，考虑 $B \rightarrow D$ ，将 ABDE 分解为两个子模式：R21 (BD) 和 R22 (ABE)，此时 R21 和 R22 均属于 BCNF。所以 R 分解为 BCNF，并具有无损连接性的分解如下：
 $P = \{AC, BD, ABE\}$

8、解：(1)根据“每个学生每学一门课只有一个成绩”的语义，可写出 $FD(S\#, C\#) \rightarrow GRADE$ ，根据“每门课只有一个任课教师”的语义，可写出 $FD C\# \rightarrow TNAME$ ，根据“每个教师有一个地址”的语义，可写出 $FD TNAME \rightarrow TADDR$ （其他据推理规则推出的FD就不必写出），候选码是 $(S\#, C\#)$ 。因为从 $(S\#, C\#)$ 可函数决定全部属性。

(2)由于 R 中存在下列两个 FD：

$$(S\#, C\#) \rightarrow TNAME \text{ 和 } C\# \rightarrow TNAME$$

因此 $(S\#, C\#) \rightarrow TNAME$ 是一个局部 FD，即 TNAME局部依赖候选码 $(S\#, C\#)$ 。也就是 R 不是 2NF。此时，R 中就会存在数据冗余和数据异常，如果一门课程有 50 个学生选修，那么在关系中就会出现 50 个元组，即这门课程的任课老师姓名和地址就要重复出现 50 次。这就是数据冗余，随之就会出现各种操作异常。

如果把 R 分解成 $R_1(S\#, C\#, GRADE)$ 和 $R_2(C\#, TNAME, TADDR)$ ，就能消除上面提到的局部 FD 和数据冗余问题，并且 R_1 和 R_2 都是 2NF 模式。

(3)前面 R_1 已是 3NF 了，但 R_2 还不是 3NF。因为在 R_2 中候选码是 $C\#$ ，并且存在下列两个 FD：

$$C\# \rightarrow TNAME \text{ 和 } TNAME \rightarrow TADDR$$

即 TADDR 传递依赖于候选码 $C\#$ ，所以也不是 3NF。

此时， R_2 中也会出现数据冗余和操作异常。如果一个教师开设 5 门课，那么在关系中就会出现 5 个元组，即这个教师的地址就要重复出现 5 次。这也是数据冗余，随之也会产生各种操作异常。

如果把 R_2 分解成 $R_{21}(C\#, TNAME)$ 和 $R_{22}(TNAME, TADDR)$ ，就会消除上面提到的传递依赖和数据冗余问题，并且 R_{21} 和 R_{22} 都是 3NF 模式。因此 R 分解成的 3NF 模式集是 $\{R_1, R_{21}, R_{22}\}$ 。

9、解： $\{A, B\}^+ = \{A, B, C, D\}$ ， $\{B, C\}^+ = \{A, B, C, D\}$ ， $\{B, D\}^+ = \{A, B, C, D\}$ 所以候选码为 AB,BC,BD。

10、解：先计算 $\{A, B, C\}$ 的所有非空的真子集的闭包：

$$A^+ = A$$

$$B^+ = B$$

$$C^+ = ACE$$

$$(AB)^+ = ABCDE$$

$$(AC)^+ = ACE$$

$$(BC)^+ = ABCDE$$

忽略 D, E 得到 S 的函数依赖集 $\{C \rightarrow A, AB \rightarrow C, BC \rightarrow A\}$ ，由 $C \rightarrow A$ 可推出 $BC \rightarrow A$ ，所以函数依赖集也可以是 $\{C \rightarrow A, AB \rightarrow C\}$

11、 $\{A, C\}^+ = \{A, B, C, D, E\}$ ；

$\{A, C, F\}^+ = \{A, B, C, D, E, F, G\}$ ，所以函数依赖集 F 蕴涵函数依赖 $ACF \rightarrow DG$ 。