数据管理基础

第6章 关系数据理论

(复习思考题参考答案)

智能软件与工程学院

复习思考题 1 -- 参考答案

- 1. 请举例说明, 不好的关系模式设计会带来哪些方面的问题?
- □ [例6.1] 再关系Student中存在的问题
 - 数据冗余度高:学生就读的院系、院系系主任的姓名等信息会存在重复存储,浪费大量的存储空间。
 - ▶更新异常:因为前述的数据冗余存储,在修改学生就读院系、系主任姓名时,对应的update操作需要在相关的元组中被重复执行,既增加了执行上的时间开销,也可能因为修改不彻底而面临数据不一致的危险。

Sno	Sdept	Mname	Cno	Grade
S 1	计算机	张明	C1	95
S2	计算机	张明	C1	90
S3	计算机	张明	C1	88
S4	计算机	张明	C1	70
S5	计算机	张明	C1	78
S1	计算机	张明	C2	86
S 3	计算机	张明	C2	94
S6	数学	李刚	C2	92
S7	数学	李刚	C3	85

- ▶ 插入异常:在插入还没有开始招生的新院系、或者插入还没有开始选课的新学生元组时,会因为违反'实体完整性'约束而导致新元组插入失败。
- ▶ 删除异常:在删除学生S7对应的选课元组时,会连带删除掉本不该被删除的课程信息(课程C3)。

复习思考题 2 -- 参考答案

- 1. 请给出以下函数依赖的定义:
 - ① 什么是函数依赖?
 - ② 什么是平凡函数依赖?什么是非平凡函数依赖?
 - ③ 什么是完全函数依赖?什么是部分函数依赖?
 - ④ 什么是传递函数依赖?
- □ [定义6.1] 设R(U)是一个属性集U上的关系模式,X和Y是U的子集。若对于R(U)的任意一个可能的关系r, r中不可能存在两个元组在X上的属性值相等,而在Y上的属性值不等,则称"X函数确定Y"或"Y函数依赖于X",记作 $X \to Y$ 。
- □ 在关系模式R(U)中,对于U的子集X和Y,
 - \triangleright 如果 $X \to Y$ 但 $Y \nsubseteq X$, 则称 $X \to Y$ 是 '非平凡函数依赖'
 - > 如果 $X \to Y$ 但 $Y \subset X$, 则称 $X \to Y$ 是 '平凡函数依赖'
- □ [定义6.2] $\alpha R(U)$ 中,如果 $X \to Y$,并且对于X的任何一个真子集X,都有 $X \to Y$,则称Y对X完全函数依赖,记作 $X \xrightarrow{F} Y$
 - \rightarrow 若 $X \rightarrow Y$,但Y不完全函数依赖于X,则称Y对X部分函数依赖,记作 $X \xrightarrow{P} Y$

复习思考题 3 -- 参考答案

- 3. 码
 - ① 请用函数依赖给出关系的码(key)和超码(superkey)的定义;
 - ② 在关系模式中,什么是主属性?什么是非主属性?请设计一个例子关系,并写出该关系的关系模式、函数依赖集、候选码、主属性集,非主属性集;
 - ③ 请简要说明下列概念的相互关系:码,主码,主属性,非主属性。
- ① [定义6.4] 设K为R(U,F)中的属性或属性组合。若 $K \xrightarrow{F} U$,则K称为R的一个候选码,简称'码'。
 - \triangleright 如果 $K \to U$ (可能是部分函数依赖,也可能是完全函数依赖),则K称为R的一个超码。
 - > 候选码的超集是超码; 候选码是最小的超码。
- ② 在一个关系中,所有候选码中的属性都是该关系的主属性;除了主属性外,其他属性是这个关系的非主属性。
 - 》 例如: 在关系R(U,F)中, $U = \{A,B,C,D,E\}$, $F = \{A \rightarrow BCD,BD \rightarrow A,C \rightarrow E\}$,该关系有两个候选码: $A \cap BD$,主属性集为 $\{A,B,D\}$,非主属性集为 $\{C,E\}$ 。
- ③ 在一个关系中,可能存在若干个候选码。在关系模式R的所有候选码中,可以选择其中的一个做为该关系的主码。码、主码、主属性、非主属性之间的关系可以描述如下:
 - > 一个关系可能存在若干个候选码,但主码最多只能定义一个;
 - \triangleright 如果A是关系R的一个主属性,则至少存在关系R的一个码K,满足 $A \in K$
 - \triangleright 如果A是关系R的一个非主属性,则对于关系R的所有候选码K,都有 $A \notin K$

复习思考题 4 -- 参考答案

- 4. 设有一个图书借阅关系R(借书证号,身份证号,书号,书名,借阅日期,归还日期),其中的数据约束如下:
 - ① 借书证号、身份证号分别是借书证、读者的标识属性;
 - ② 一个读者只能办理一张借书证,一张借书证只能对应一个读者;
 - ③ 每一本图书都有一个唯一的书号,不同的图书可能有相同的书名;
 - ④ 一个读者可以同时借阅多本图书,也可以在不同时候借阅同一本图书;
 - ⑤ 系统需要记录一本图书每一次被借阅的借阅日期和归还日期,并保存所有的借阅历史。

请写出在借阅关系R上的非平凡的完全函数依赖。

□ 关系R上的函数依赖集如下(最小覆盖,结果不唯一)

```
{借书证号→身份证号,
身份证号→借书证号,
书号→书名,
(书号,借阅日期)→(借书证号,归还日期),
(书号,归还日期)→借阅日期
```

复习思考题 5 -- 参考答案

- 5. 设有一个期末考试监考安排关系R, 其中的属性有:课程的课程号(cno)和课程名(cname),授课教师的工作证编号(tno)和姓名(tname),监考老师的工作证编号(in_no),每一场考试的开始时间(s_date)、结束时间(e_date)和考试教室(room)。其中:课程号和工作证编号分别是课程及教师的标识属性,开始时间和结束时间是date类型(含日期和时间)的字段,并且规定:
 - 每一门课程至少有一位授课教师,也可能安排多位授课教师;
 - ② 一位老师也可以担任多门课程的授课任务;
 - ③ 每一门课的期末考试只安排一场,可分在多个教室中同时进行,除了授课教师外,在每一间考试教室中都必须安排一位或多位监考老师;
 - ④ 同一时间段、同一间教室中只能安排一门课程的考试;
 - ⑤ 一位老师可以担任多门课程的监考任务,但在同一时间段内,一位老师只能在指定的一间 教室中监考一门课;
 - ⑥ 授课教师必须参加自己承担授课任务的课程监考(不限定教室)。 R(cno, cname, tno, tname, in_no, s_date, e_date, room)

请找出该关系中的所有函数依赖(非平凡的完全函数依赖)。

□ 发现的函数依赖如下:

F = { tno \rightarrow tname, cno \rightarrow (cname, s_date, e_date), (cno, in_no) \rightarrow room, (tno, s_date) \rightarrow cno, (tno, e_date) \rightarrow cno, (in_no, s_date) \rightarrow cno, (in_no, s_date) \rightarrow cno, (room, e_date) \rightarrow cno }

复习思考题 6 -- 参考答案

- 6. 范式定义
 - ① 请写出下列范式的定义: 1NF, 2NF, 3NF, BCNF
 - ② 简要说明,上述各级范式之间的相互关系。
- ① 各级范式的定义如下:
- \square 如果一个关系模式R的所有属性都是不可分的基本数据项,则 $R \in 1NF$.
- □ [定义6.6] 若关系 $R \in 1NF$,并且每一个非主属性都完全函数依赖于任何一个候选码,则 $R \in 2NF$.
- □ [定义6.7] 设关系模式 $R(U,F) \in 1NF$,若R中不存在这样的码X、属性组Y及非主属性Z ($Z \nsubseteq Y$),使得 $X \to Y$, $Y \to Z$ 成立, $Y \nrightarrow X$,则称 $R(U,F) \in 3NF$.
- □ [定义6.8] 设关系模式 $R(U,F) \in 1NF$,若 $X \to Y$ 且 $Y \nsubseteq X$ 时X必含有码,则 $R(U,F) \in BCNF$.
- ② 如果由所有满足第一范式的关系组成集合,由所有满足第二范式的关系组成集合,,那么这些范式相互之间的关系可以表示如下:

 $BCNF \subseteq 3NF \subseteq 2NF \subseteq 1NF$

复习思考题 7 -- 参考答案 (1)

- 7. 在一个关系模式中,会因为存在下列函数依赖而产生不合理的数据冗余存储和操作异常。请利 用函数依赖、候选码等概念的定义来解释其原因。
 - ① 存在非主属性A对候选码K的部分函数依赖: $K \stackrel{p}{\rightarrow} A$
 - ② 存在非主属性A对候选码K的传递函数依赖: $K \to X, X \not\subset K, X \to K, X \to A$
 - ③ 存在主属性B对候选码K的部分函数依赖: $K \xrightarrow{p} B$;
 - ④ 存在主属性B对候选码K的传递函数依赖: $K \to X, X \not\subset K, X \to K, X \to B$
- \triangleright 设K是关系R的候选码, t_1 、 t_2 是关系R中的元组。根据候选码的定义可知:
 - a) If $t_1[k] = t_2[k]$ then $t_1 = t_2$
 - b) 对于候选码K的任意一个真子集 K_0 ($K_0 \subset K$),都有 $K_0 \rightarrow head(R)$,即:在关系R中允许存在一对不同的元组 $t_1 \neq t_2$ 满足 $t_1[k_0] = t_2[k_0]$
- ① 如果存在非主属性A对候选码K的部分函数依赖 $K \stackrel{p}{\rightarrow} A$, 那么:
 - 因为 $K \xrightarrow{p} A$, 所以至少存在K的一个真子集 $K_0 \subset K$ 并且 $K_0 \to A$
 - 根据候选码的定义可知,在关系R中允许存在一对不同的元组 $t_1 \neq t_2$ 满足 $t_1[k_0] = t_2[k_0]$
 - 由 $K_0 \to A$ 和 $t_1[k_0] = t_2[k_0]$ 可知: $t_1[A] = t_2[A]$
 - 所以,允许 $t_1[k_0]$ 与 $t_1[A]$ 之间的取值对应关系重复出现在不同的元组中,因此产生了数据冗余存储。

复习思考题 7 -- 参考答案 (2)

- ② 存在非主属性A对候选码K的传递函数依赖: $K \to X, X \not\subset K, X \to K, X \to A$
- ▶ 根据传递函数依赖的定义可知,必存在一个属性集合X,满足: $K \to X$, $X \nsubseteq K$, $X \to K$, $X \to A$
- ▶ 那么:
 - ① 因为 $X \rightarrow K$, 所以在关系R中,允许存在一对元组 t_1 、 t_2 满足 $t_1[X] = t_2[X]$ 但 $t_1[K] \neq t_2[K]$
 - ② 由 $X \to A$ 和 $t_1[X] = t_2[X]$ 可知: $t_1[A] = t_2[A]$
 - ③ 由 $t_1[K] \neq t_2[K]$ 可知, t_1 、 t_2 是关系R中的两个不同的元组
 - ④ 所以,允许 $t_1[X]$ 与 $t_1[A]$ 之间的取值对应关系重复出现在不同的元组中,因此产生了数据冗余存储。
- \triangleright 在这里就可以理解,在传递函数依赖的定义里,为什么要有 $X \rightarrow K$
- \triangleright 如果没有这个限制,也就是说,如果也存在函数依赖 $X \to K$,那么:
 - 如果 $t_1[X] = t_2[X]$, 则 $t_1[K] = t_2[K]$, 即 t_1 、 t_2 是关系R中的同一个元组
 - 这样, $t_1[X]$ 与 $t_1[A]$ 之间的取值对应关系只会出现确定的一个元组中,不会出现重复存储
- □ 同理, 可以解释③和④。

复习思考题 8/9/10 -- 参考答案 (略)

- 8. Armstrong公理系统
 - ① 请写出三条基本规则: 自反律, 增广律, 传递律
 - ② 请写出以下两条扩充规则并证明:分解规则,合并规则
 - ③ 请举例说明:利用Armstrong公理系统中的传递律推导得到的函数依赖不一定是传递函数依赖。
- 9. '函数依赖集闭包'与'属性集闭包'
 - ① 什么是函数依赖的逻辑蕴涵?
 - 2 什么是函数依赖集闭包?
 - ③ 什么是属性集闭包? 请写出属性集闭包的计算算法。
- 10.'函数依赖集等价'与'极小函数依赖集'
 - ① 请写出下列概念的定义:函数依赖集等价,极小函数依赖集
 - ② 极小函数依赖集的判定方法是什么?请写出极小函数依赖集的计算算法。

复习思考题 11 -- 参考答案 (1)

11. 请利用Armstrong公理系统证明下面的推导过程是否成立?如果不成立,请给出具体的例子关系。

1	$\{ W \rightarrow Y, X \rightarrow Z \} \Rightarrow \{ WX \rightarrow Y \}$	✓
2	$\{ X \rightarrow Y \} \text{ and } Z \subseteq Y \Rightarrow \{X \rightarrow Z \}$	✓
3	$\{ X \rightarrow Y, X \rightarrow W, WY \rightarrow Z \} \Rightarrow \{ X \rightarrow Z \}$	✓
4	$\{ XY \rightarrow Z, Y \rightarrow W \} \Rightarrow \{ XW \rightarrow Z \}$	×
5	$\{ X \rightarrow Z, Y \rightarrow Z \} \Rightarrow \{ X \rightarrow Y \}$	×
6	$\{ X \rightarrow Y, XY \rightarrow Z \} \Rightarrow \{ X \rightarrow Z \}$	✓
7	$\{ X \rightarrow Y, Z \rightarrow W \} \Rightarrow \{ XZ \rightarrow YW \}$	√
8	$\{ XY \rightarrow Z, Z \rightarrow X \} \Rightarrow \{ Z \rightarrow Y \}$	×
9	$\{ X \rightarrow Y, Y \rightarrow Z \} \Rightarrow \{ X \rightarrow YZ \}$	✓
10	$\{ XY \rightarrow Z, Z \rightarrow W \} \Rightarrow \{ X \rightarrow W \}$	×

复习思考题 11 -- 参考答案 (2)

1. 答: W→Y, X→Z ⇒ WX→Y 成立

增广律: W→Y ⇒ WX→YX

自反律: YX→Y

传递律: WX→YX, YX→Y ⇒ WX→Y

6. 答: X→Y, XY→Z ⇒ X→Z 成立

增广律: X→Y ⇒ XX→XY 即 X→XY

传递律: X→XY, XY→Z ⇒ X→Z

2. 答: X→Y and Z ⊆ Y ⇒ X→Z 成立

自反律: Z ⊂ Y ⇒ Y→Z

传递律: $X \rightarrow Y, Y \rightarrow Z \Rightarrow X \rightarrow Z$

7. 答: X→Y, Z→W ⇒ XZ→YW 成立

增广律: X→Y ⇒ XZ→YZ

增广律: Z→W ⇒ YZ→YW

传递律: XZ→YZ, YZ→YW ⇒ XZ→YW

3. 答: X→Y, X→W, WY→Z ⇒ X→Z 成立

合并规则: X→W, X→Y ⇒ X→WY

传递律: X→WY, WY→Z ⇒ X→Z

9. 答: X→Y, Y→Z ⇒ X→YZ 成立

传递律: X→Y, Y→Z ⇒ X→Z

合并规则: $X \rightarrow Y$, $X \rightarrow Z \Rightarrow X \rightarrow YZ$

复习思考题 11 -- 参考答案 (3)

4. 答: XY→Z, Y→W ⇒ XW→Z 不成立

8. 答:	$XY \rightarrow Z, Z \rightarrow X$	$X \Rightarrow Z \rightarrow Y$	不成立
	/		

X	Υ	Z	W
1	1	1	1
1	2	2	1

Х	Υ	Z
1	1	1
1	4	1

5. 答: X→Z, Y→Z ⇒ X→Y 不成立

10. 2	F: X	(→Z.	Z→W	\Rightarrow	$X\rightarrow W$	不成立
-------	------	------	-----	---------------	------------------	-----

X	Υ	Z
2	2	2
2	3	2
3	4	1

X	Υ	Z	W
1	1	1	1
1	2	2	2

复习思考题 12/13 -- 参考答案

- 12. 设 $F = \{A \rightarrow C, AC \rightarrow D, E \rightarrow AD, E \rightarrow H\}$, 请给出 F 的极小函数依赖集。
- □ 极小函数依赖集的计算结果是:

$$\{A \rightarrow CD, E \rightarrow AH\}$$

- 13. $M = \{ABD \rightarrow AC, C \rightarrow BE, AD \rightarrow BF, B \rightarrow E\}$, 请计算M的极小函数依赖集。
 - □ 极小函数依赖集的计算结果是:

$$\{AD \rightarrow CF, C \rightarrow B, B \rightarrow E\}$$

复习思考题 14 -- 参考答案

- 14. 给定关系模式R(A,B,C,D,E,F) 及其上的函数依赖集: $S = \{A \rightarrow E, B \rightarrow ADE, DF \rightarrow AC, ADF \rightarrow B\}$
 - ① 请直接写出与S等价的极小函数依赖集;
 - ② 请直接写出关系模式R的所有候选码、主属性集、非主属性集;
 - ③ 请使用3NF模式分解算法对关系R进行分解,并满足无损联接性和依赖保持性;
 - ④ 上题分解结果是否满足BCNF?如果不满足,请将其继续分解到满足BCNF并说明理由。
- ① 答: $\{A \rightarrow E, B \rightarrow AD, DF \rightarrow BC\}$
- ② 答: 有两个候选码: {B,F} 和 {D,F}
- ③ 答:分解结果是 $R_1(A, E)$ $R_2(A, B, D)$ $R_3(B, C, D, F)$
- ④ 答:在上述分解结果中, R_1 和 R_2 满足BCNF,但 R_3 不满足BCNF。到BCNF的分解方法如下: R_3 的候选码是 $\{B,F\}$ 和 $\{D,F\}$,函数依赖 $B \to D$ 不满足BCNF,故将原关系 R_3 分解为 $R_{31}(B,D)$ 和 $R_{32}(B,C,F)$ 。又因为关系 R_2 包含有关系 R_{31} 中的所有属性,不需要保留新分解得到的关系 R_{31} ,用 关系 R_{32} 替代上一步分解中的关系 R_3 ,最终得到满足BCNF的分解结果如下:
 - $ightharpoonup R_1(A, E)$ **FDs:** $\{A \rightarrow E\}$
 - $ightharpoonup R_2(A, B, D)$ **FDs:** $\{B \rightarrow AD\}$
 - $ightharpoonup R_3(B, C, F)$ **FDs:** { BF \rightarrow C }

该分解满足无损连接性, 但不保持函数依赖。

复习思考题 15 -- 参考答案

- 15. 给定关系R(A,B,C,D,E,F,G)及其上的函数依赖集: (不需要写计算过程) M = { ABC→DEF, AC→BG, D→F, E→GC }
 - ① 请直接写出与M等价的极小函数依赖集;
 - ② 请直接写出关系R的所有候选码、主属性集、非主属性集;
 - ③ 请使用3NF模式分解算法对关系R进行分解,并满足无损联接性和依赖保持性;
 - ④ 上述的分解结果是否满足BCNF?如满足BCNF,请简单说明理由;否则,请将其继续分解 到满足BCNF。
- ① 答: $\{AC \rightarrow BDE, D \rightarrow F, E \rightarrow GC\}$
- ② 答: 有两个候选码: {A, C} 和 {A, E}
- ③ 答: $R_1(\{A, B, C, D, E\}, \{AC \to BDE, E \to C\})$ $R_2(\{D, F\}, \{D \to F\})$ $R_3(\{C, E, G\}, \{E \to GC\})$
- ④ 答:在上述分解结果中, R_2 和 R_3 满足BCNF,但 R_1 不满足BCNF。到BCNF的分解方法如下: R_1 的候选码是 $\{A,C\}$ 和 $\{A,E\}$, R_1 中的函数依赖 $E \to C$ 不满足BCNF,故将原关系 R_1 分解为 $R_{11}(\{C,E\},\{E \to C\})$ 和 $R_{12}(\{A,B,D,E\},\{AE \to BD\})$ 。

又因为关系 R_3 包含有关系 R_{11} 中的所有属性,不需要保留新分解得到的关系 R_{11} ,用关系 R_{12} 替代上一步分解中的关系 R_4 ,最终得到满足BCNF的分解结果如下:

 $R_1(\{A,B,D,E\}, \{AE \to BD\})$ $R_2(\{D,F\}, \{D \to F\})$ $R_3(\{C,E,G\}, \{E \to GC\})$ 该分解满足无损连接性,但不保持函数依赖。

复习思考题 16 -- 参考答案

- 16. 设有一个机场跑道使用调度管理系统,其关系模式如下: 机场编号 跑道编号 飞机编号 使用开始时间 使用结束时间 R(ano, Ino, pno, s_time, e_time) 其中:
 - ▶ 机场编号ano和飞机编号pno分别是机场和飞机的标识属性;
 - 在一个机场中,可能有多条用于飞机起降的跑道,每条跑道都有一个唯一的编号Ino;分属于不同机场的跑道,可能有相同的跑道编号;
 - 每一条跑道每次只能供一架飞机使用(供飞机起降用);
 - ▶ 使用开始时间和使用结束时间的数据类型是时间戳(timestamp)。
 - ① 根据上述描述,请写出关系R上的极小函数依赖集。(不需要写计算过程)
 - ② 关系R最高能够满足哪个范式的定义?请简单说明理由。
- ① 答: $\{(sno, lno, s_time) \rightarrow (pno, e_time), (sno, lno, e_time) \rightarrow s_time, (pno, s_time) \rightarrow (sno, lno, e_time) (pno, e_time) \rightarrow s_time \}$
- ① 答:满足BCNF。理由:该关系有四个候选码(sno,lno,s_time), (sno,lno,e_time), (pno,s time)和 (pno,e time),所有的函数依赖都满足BCNF的定义。

复习思考题 17 -- 参考答案 (1)

17. 设有一个大学生创新项目管理关系P(pno, mgrno, sno, pyear, ps)

其属性包括:项目的编号pno、执行年份pyear、验收等级ps,项目负责学生的学号mgrno,项目参与学生的学号sno。其中:

- ①项目编号和学号分别是项目和学生的标识属性;②每个项目都有唯一的一名负责的学生,以及可能的若干名参与的学生;③每个项目的执行周期只有一年;④一个学生可以负责或参与过若干个项目,但每一年最多只能负责或参与一个项目。
- ① 请写出该关系上的极小函数依赖集(直接写出结果)
- ② 该关系最高能够满足到第几范式?请简单说明理由;
- ③ 关系P是否满足3NF? 如不满足, 请用到3NF的模式分解算法直接对其进行模式分解;
- ④ 上述分解结果是否满足BCNF? 如不满足,请将其进一步分解到满足BCNF。

复习思考题 17 -- 参考答案 (2)

- ① 答: { pno \rightarrow (mgrno, pyear, ps), (mgrno, pyear) \rightarrow pno, (sno, pyear) \rightarrow pno }
- ② 答: 首先寻找该关系上的候选码, 共找到2个候选码: (sno, pyear) 和 (sno, pno) 存在非主属性 ps 对候选码(sno,pno) 的部分函数依赖, 所以不满足2NF, 最高只能满足1NF。
- ③ 答:不满足3NF。利用到3NF的模式分解算法直接分解的结果如下:

关系模式	函数依赖集	候选码
R ₁ (pno, mgrno, pyear, ps)	pno→(mgrno, pyear, ps) (mgrno, pyear)→pno	pno (mgrno, pyear)
R ₂ (pno, pyear, sno)	pno→pyear (sno, pyear)→pno	(pno, sno) (sno, pyear)

④ 答: $R_1 \in BCNF$, $R_2 \notin BCNF$ 。可将 R_2 进一步分解为 R_{21} (pno, pyear)和 R_{22} (pno, sno)这两个子关系,由于head(R_{21})⊆head(R_1),新分解出来的关系 R_{21} 是多余的,可以略去。最终到BCNF的分解结

果如下:

关系模式	函数依赖集	候选码
na ibilo, iligilio, bveat, bsi	pno→(mgrno, pyear, ps) (mgrno, pyear)→pno	pno (mgrno, pyear)
R ₂₂ (pno, sno)		(pno, sno)