

数据库系统之三

--数据建模与数据库设计



第16讲 模式分解存在什么问题

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

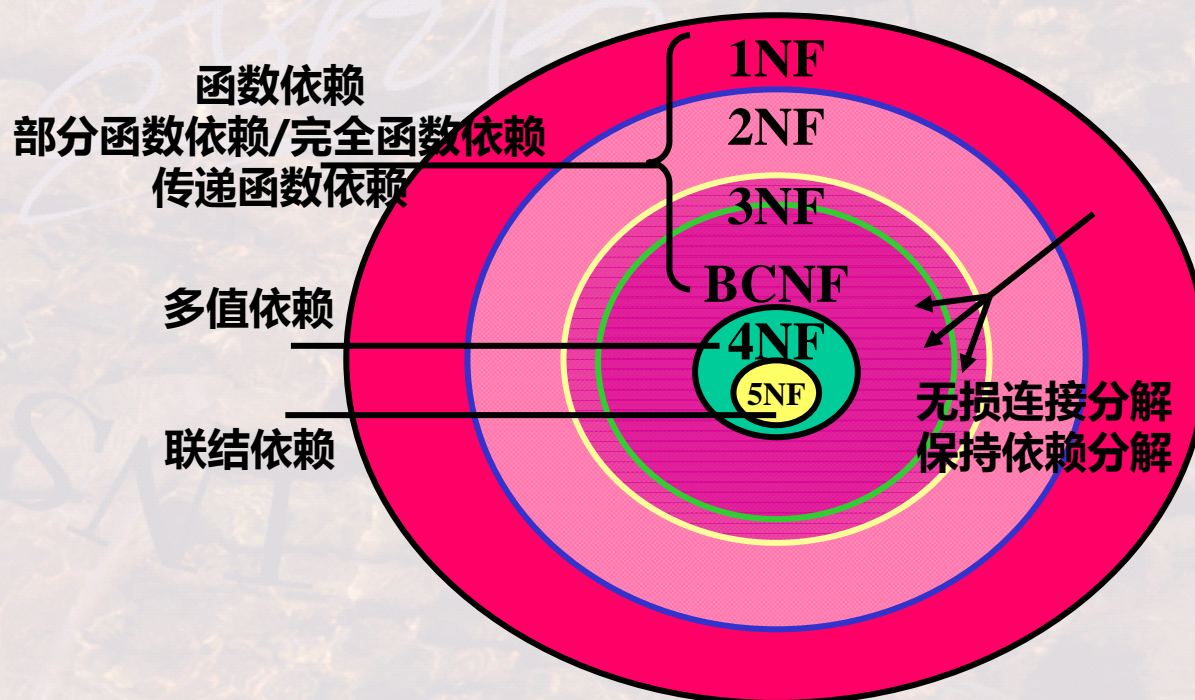
本讲学习什么？

如何避免数据库的一致性问题——数据库的规范性设计

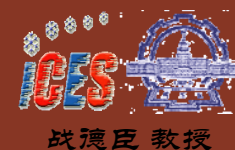
数据库的规范性设计需要分析数据库Table中的属性在取值方面有什么依存关系？数据库设计过程中应遵循什么样的原则

➤数据库设计理论

- 数据依赖理论
- 关系范式理论
- 模式分解理论



本讲学习什么？



基本内容

1. 模式分解存在什么问题
2. 无损连接分解及其检验算法
3. 保持依赖分解及其检验算法
4. 关系模式无损连接或保持依赖的分解算法
5. 数据库设计需要知道的？

重点与难点

- 理解模式分解存在的问题是什么，怎样解决
- 两个概念：无损连接分解和保持依赖分解
- 关系模式如何进行无损连接或保持依赖地分解
- 五个算法：2个检验算法，3个分解算法(无损连接分解，保持依赖分解，既无损连接又保持依赖分解)

模式分解存在的问题

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

模式分解存在的问题

(1)模式分解的概念



[Definition] 模式分解

关系模式 $R(U)$ 的分解是指用 R 的一组子集 $\rho = \{R_1(U_1), \dots, R_k(U_k)\}$ 来代替它。

其中 $U = U_1 \cup U_2 \cup \dots \cup U_k$; $U_i \not\subseteq U_j$ ($i \neq j$)。

注：为便于后面叙述，我们用 R_i 代替 $R_i(U_i)$ ， R 代替 $R(U)$ 。

➤ 对于关系模式 R 的任一关系 r ，它向 ρ 的投影连接记为 $m_\rho(r)$ ：

$$m_\rho(r) = \pi_{R_1}(r) \bowtie \dots \bowtie \pi_{R_k}(r) = \bowtie_{(i=1, \dots, k)} \pi_{R_i}(r)$$

这里： $\pi_{R_i}(r) = \{t[R_i] \mid t \in r, i=1, \dots, k\}$

➤ 模式分解需要关注：

□ R 与 ρ 在数据内容方面是否等价：分解的无损连接性；

□ R 与 ρ 在数据依赖方面是否等价：分解的保持依赖性。

模式分解存在的问题

(2)关于模式分解的一些特性：数据内容的等价性

➤ [引理1] 设 R 为一关系模式， $\rho = \{R_1, \dots, R_k\}$ 是 R 的一个分解， r 是 R 的任一个关系， $r_i = \pi_{R_i}(r)$ ，则有规则成立：

(rule 1) $r \subseteq m_\rho(r)$

(rule 2) 若 $s = m_\rho(r)$ ，则 $\pi_{R_i}(s) = r_i$ (即： $\pi_{R_i}(m_\rho(r)) = \pi_{R_i}(r)$)

(rule 3) $m_\rho(m_\rho(r)) = m_\rho(r)$

r	A	B	C	D
	1	2	A	D
	2	2	A	E
	2	3	B	C
	1	3	B	D

r1	A	B	C
	1	2	A
	2	2	A
	2	3	B
	1	3	B

r2	B	C	D
	2	A	D
	2	A	E
	3	B	C
	3	B	D

$m_\rho(r)$	A	B	C	D
1	2	A	D	
1	2	A	E	
2	2	A	D	
2	2	A	E	
1	3	B	C	
1	3	B	D	
2	3	B	C	
2	3	B	D	

模式分解存在的问题

(3)关于模式分解的一些特性：数据约束的等价性

[示例] $R(C, S, Z)$, C 是城市, S 是街区, Z 是邮政编码

$F = \{ CS \rightarrow Z, Z \rightarrow C \}$

$\rho = \{ R_1(SZ), R_2(CZ) \}$

分解后是否有
约束丢失了?

当模式不符合关系范式时，进行模式分解

示例： $R(A, B, C, D, E, F, G)$

函数依赖集合 $\{ A \rightarrow B, A \rightarrow C, C \rightarrow D, C \rightarrow E, E \rightarrow FG \}$

候选键：A；有传递依赖，R不满足3NF。

分解规则：

将每一个函数依赖单独组成一个关系

$\rho = \{ R1(A, B), R2(A, C), R3(C, D), R4(C, E), R5(E, F, G) \}$

可以看出：每一个模式都属于3NF

也可以合并一些关系：

$\rho = \{ R12(A, B, C), R34(C, D, E), R5(E, F, G) \}$

分解后的关系的
连接与分解前关
系的等价性？

分解前的约束，
在分解后是否还
存在？

无损连接分解及其检验算法

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

[Definition] 无损连接分解

对于关系模式 $R(U, F)$, U 是属性全集, F 是函数依赖集合, $\rho = \{R_1, \dots, R_k\}$ 是 R 的一个分解, 如果对于 R 的任何满足函数依赖集 F 的关系 r , 有

$$r = m_\rho(r),$$

则称 ρ 是 R 相对于 F 的一个无损连接分解, 其中:

$$m_\rho(r) = \pi_{R_1}(r) \bowtie \dots \bowtie \pi_{R_k}(r) = \bigbowtie_{(i=1, \dots, k)} \pi_{R_i}(r)$$

无损连接分解及其检验算法

(2)无损连接分解的检验算法

➤ [Algorithm] 无损连接性检验算法

Input : 关系模式 $R = A_1 A_2 \dots A_n$, 函数依赖集 F , 分解 $\rho = \{R_1, \dots, R_k\}$

Output : ρ 是否是无损连接的判断

Method : (1) 构造一 k 行 n 列的表, 可称为 R_ρ 表。其中第 j 列对应于 A_j , 第 i 行对应于 R_i , 若 $A_j \in R_i$, 则 R_ρ 表中第 i 行第 j 列位置填写符号 a_j , 否则填写 b_{ij} 。

	A_1	...	A_i	...	A_n
R_1					
...					
R_i			a_i 或 b_{ij}		
...					
R_k					

	A	B	C	D	E
R_1	a_1	b_{12}	b_{13}	a_4	b_{15}
R_2	a_1	a_2	b_{23}	b_{24}	b_{25}
R_3	b_{31}	a_2	b_{33}	b_{34}	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	b_{53}	b_{54}	a_5

无损连接分解及其检验算法

(2)无损连接分解的检验算法

(2)根据 $\forall (X \rightarrow Y) \in F$, 对 R_ρ 表进行修改 :

给定 $X \rightarrow Y$, 在表中寻找对应于 X 中所有属性分量之列上符号全相同的行。若能找到, 则在这些行的对应于 Y 中属性的那些列上置相同符号: 若其中有一个行之相应列上为 a_j , 则使其它行同一列上置 a_j ; 若相应列上均为 b_{ij} (或 b_{ij}), 则使其它行同一列上置某一个 b_{ij} (或 b_{ij} , 任一个都可, 只要相同);

(3)在上述修改的表中, 如果发现有一行变成 a_1, a_2, \dots, a_n (全 a), 则 ρ 是无损连接分解, 否则 ρ 是有损连接分解。

	A_1	...	A_i	...	A_n
R_1					
...					
R_j			a_j 或 b_{ij}		
...					
R_k					

为什么可以
用 $X \rightarrow Y$
修改呢?

	A	B	C	D	E
R_1	a_1	b_{12}	b_{13}	a_4	b_{15}
R_2	a_1	a_2	b_{23}	b_{24}	b_{25}
R_3	b_{31}	a_2	b_{33}	b_{34}	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	b_{53}	b_{54}	a_5

示例：已知 $R = \{ ABCDE \}$

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

$\rho = \{ R_1(AD), R_2(AB), R_3(BE), R_4(CDE), R_5(AE) \}$

问： ρ 是否具有无损连接性

解：(1)构造 R_ρ 表。

	A	B	C	D	E
R_1	a_1	b_{12}	b_{13}	a_4	b_{15}
R_2	a_1	a_2	b_{23}	b_{24}	b_{25}
R_3	b_{31}	a_2	b_{33}	b_{34}	a_5
R_4	b_{41}	b_{42}	a_3	a_4	a_5
R_5	a_1	b_{52}	b_{53}	b_{54}	a_5

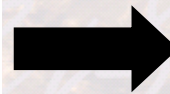
无损连接分解及其检验算法

(3)无损连接分解检验算法的应用示例

(2)用每一个函数依赖修改 R_ρ 表

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

	A	B	C	D	E
R ₁	a ₁	b ₁₂	b ₁₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	b ₂₃	b ₂₄	b ₂₅
R ₃	b ₃₁	a ₂	b ₃₃	b ₃₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	b ₅₃	b ₅₄	a ₅



用 $A \rightarrow C$
修改

	A	B	C	D	E
R ₁	a ₁	b ₁₂	b ₁₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	b ₁₃	b ₂₄	b ₂₅
R ₃	b ₃₁	a ₂	b ₃₃	b ₃₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	b ₁₃	b ₅₄	a ₅



用 $B \rightarrow C$
修改

	A	B	C	D	E
R ₁	a ₁	b ₁₂	b ₁₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	b ₁₃	b ₂₄	b ₂₅
R ₃	b ₃₁	a ₂	b ₁₃	b ₃₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	b ₁₃	b ₅₄	a ₅

无损连接分解及其检验算法

(3)无损连接分解检验算法的应用示例

(2)用每一个函数依赖修改 R_ρ 表

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

	A	B	C	D	E
R ₁	a ₁	b ₁₂	b ₁₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	b ₁₃	b ₂₄	b ₂₅
R ₃	b ₃₁	a ₂	b ₁₃	b ₃₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	b ₁₃	b ₅₄	a ₅



用 $CE \rightarrow A$
修改

	A	B	C	D	E
R ₁	a ₁	b ₁₂	b ₁₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	b ₁₃	b ₂₄	b ₂₅
R ₃	a ₁	a ₂	b ₁₃	b ₃₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	b ₁₃	b ₅₄	a ₅



用
 $C \rightarrow D, DE \rightarrow C$
修改

	A	B	C	D	E
R ₁	a ₁	b ₁₂	a ₃	a ₄	b ₁₅
R ₂	a ₁	a ₂	a ₃	a ₄	b ₂₅
R ₃	a ₁	a ₂	a ₃	a ₄	a ₅
R ₄	b ₄₁	b ₄₂	a ₃	a ₄	a ₅
R ₅	a ₁	b ₅₂	a ₃	a ₄	a ₅

(3)检查是否有一行变成 a_1, a_2, \dots, a_n (全a)

有：无损连接；无：有损连接

[定理]算法正确判断出了一个分解是否是无损连接的分解。

证明：(略)。

[定理]设 F 是关系模式 R 上的一个函数依赖集合。 $\rho=\{R_1, R_2\}$ 是 R 的一个分解，则：当且仅当 $R_1 \cap R_2 \rightarrow R_1 - R_2$ 或者 $R_1 \cap R_2 \rightarrow R_2 - R_1$ 属于 F^+ 时， ρ 是关于 F 无损连接的。

证明：(略，此定理可由前述算法证明)。

	$R_1 \cap R_2$	$R_1 - R_2$	$R_2 - R_1$
R_1	a_1	a_2	b_{13}
R_2	a_1	b_{22}	a_3

前述算法是对多个关系，这个定理是针对两个关系的

[引理]设关系模式R具有函数依赖集F， $\rho = \{R_1, \dots, R_k\}$ 是R的一个分解，且是关于F无损连接的分解，则有：

□(a)对特定的i，设 $F_i = \pi_{R_i}(F)$ ， $\sigma = (S_1, S_2, \dots, S_m)$ 是 R_i 关于 F_i 的无损连接分解，则R到 $\{R_1, \dots, R_{i-1}, S_1, S_2, \dots, S_m, R_{i+1}, \dots, R_k\}$ 的分解是关于F无损连接的。

□(b)设 $\tau = \{R_1, \dots, R_k, R_{k+1}, \dots, R_n\}$ 是R的另一分解，且 τ 包括了 ρ 中的关系模式，则 τ 也是关于F无损连接的。

证明：(略)

保持依赖分解及其检验算法

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

➤[Definition]保持依赖分解

对于关系模式 $R(U, F)$, U 是属性全集, F 是函数依赖集合, $\rho=\{R_1, \dots, R_k\}$ 是 R 的一个分解, 如在 $\pi_{R_i}(F)$ 中的所有依赖之并集($i=1, \dots, k$)逻辑蕴涵 F 的每个依赖, 则称分解 ρ 保持依赖集 F 。其中 $\pi_{R_i}(F)$ 是 F 在 R_i 上的投影, 即 F 中的任一投影 $X \rightarrow Y$, 如果 X, Y 均包含于 R_i , 则 $X \rightarrow Y \in \pi_{R_i}(F)$ 。 R_i 指 R_i 的属性集。

注: (1)保持依赖的分解可能不是无损连接的。

(2)无损连接的分解可能不是保持依赖的。

示例: $R(CSZ)$, $F=\{CS \rightarrow Z, Z \rightarrow C\}$, C 是城市, S 是街区, Z 是邮政编码, $\rho=\{R_1(SZ), R_2(CZ)\}$ 为一无损连接分解, 但却不保持依赖;

示例: $R(ABCD)$, $F=\{A \rightarrow B, C \rightarrow D\}$, $\rho=\{R_1(AB), R_2(CD)\}$ 为一保持依赖分解, 但不是无损连接分解。

保持依赖分解及其检验算法

(2)保持依赖分解的检验算法



[Algorithm]保持依赖性检验算法

Input : 关系模式 $R=A_1A_2...A_n$, R 上的函数依赖集 F , 分解 $\rho=\{R_1,...,R_k\}$

Output : ρ 是否是保持依赖的判断

Method : 令 $G = \cup_{(i=1 \text{ to } k)} \pi_{R_i}(F)$, 只需检查 G 是否覆盖 F 即可。具体算法如下：

□首先对每个 $X \rightarrow Y \in F$ 计算 G 中的 X^+_G : (如果 X 不包含于 R_i 则不需计算了)

$Z = X$

WHILE Z 的变化发生 DO

FOR $i = 1$ to k DO

$Z = Z \cup ((Z \cap R_i)^+ \cap R_i)$

□判断 G 是否逻辑蕴涵 $X \rightarrow Y$: 前面计算的结果 Z 便是 X^+ , 如果 Z 包含 Y , 则 G 逻辑蕴涵 $X \rightarrow Y$, 否则便不逻辑蕴涵。

□判断 ρ 是否保持依赖 : 如果 G 逻辑蕴涵 F 中的每一个函数依赖, 则说 ρ 是保持依赖的分解, 否则便不是保持依赖的分解。

[定理]算法正确确定了 $X \rightarrow Y$ 是否在 G^+ 中。

证明 : (略)。

应用示例

Input : $R(A, B, C, D, E)$

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

$\rho = \{ R_1(AC), R_2(BC), R_3(CDE) \}$

Output : ρ 是否是保持依赖的判断

Method : 依据题意

$\pi_{R_1}(F) = \{ A \rightarrow C \}, \pi_{R_2}(F) = \{ B \rightarrow C \}, \pi_{R_3}(F) = \{ C \rightarrow D, DE \rightarrow C \}$

$G = \{ A \rightarrow C, B \rightarrow C, C \rightarrow D, DE \rightarrow C \}$, 显然不保持依赖。

□对函数依赖 $A \rightarrow C \in F$ 计算 G 中的 X^+_G :

$Z = \{A\} \cup \{C\} \cup \{\} \cup \{\} = \{A, C\}$, C 包含于 Z 中 , 所以 $A \rightarrow C$ 被 G 逻辑蕴涵

□对函数依赖 $DE \rightarrow C \in F$ 计算 G 中的 X^+_G :

$Z = \{D, E\} \cup \{\} \cup \{\} \cup \{C, D\} = \{C, D, E\}$, C 包含于 Z 中 , 所以 $A \rightarrow C$ 被 G 逻辑蕴涵

□对函数依赖 $CE \rightarrow A \in F$ 计算 G 中的 X^+_G :

$Z = \{C, E\} \cup \{\} \cup \{\} \cup \{D\} = \{C, E, D\}$, A 不包含于 Z 中 , 所以不被 G 逻辑蕴涵

$$Z = Z \cup ((Z \cap R_i)^+_G \cap R_i)$$

左端属性必须是某一个关系模式的属性

求关于 G 的属性闭包, 结果属性也必须是此关系模式的属性

将关系模式分解成3NF或BCNF

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

关系模式分解成BCNF

示例： $R(A, B, C, D, E, F, G)$

函数依赖集合 $\{ A \rightarrow B, A \rightarrow C, C \rightarrow D, C \rightarrow E, E \rightarrow FG \}$

候选键： A; 有不依赖于候选键的其他函数依赖，R不满足BCNF。

分解规则：

将左侧不含候选键的函数依赖单独组成一个关系, 将包含候选键的组成一关系

$\rho = \{ R_1(C, D), R_2(C, E), R_3(E, F, G), R_4(A, B, C) \}$

可以看出： $R_1 \in \text{BCNF}; R_2 \in \text{BCNF}; R_3 \in \text{BCNF}; R_4 \in \text{BCNF};$

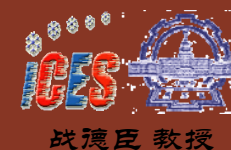
也可以将R1和R2合并：

$\rho = \{ R_{12}(C, D, E), R_3(E, F, G), R_4(A, B, C) \}$

**ρ 满足无损连接性?
 ρ 满足保持依赖性吗?**

将关系模式分解成3NF或BCNF

(2)关系模式无损连接地分解成BCNF算法



[Algorithm]无损连接分解成BCNF的算法。

Input : 关系模式 $R(U, F)$

Output : R 的一个无损连接分解 ρ , ρ 中的每个关系模式都是 F 在该模式上投影的BCNF。

Method : (1) 令 $\rho = \{ R \}$ 。

(2)对每个模式 $s \in \rho$, 若 $s \notin \text{BCNF}$, 则 s 上必有 $X \rightarrow A$ 成立且 X 不是 s 的超键且 $A \notin X$, 此时用模式 s_1, s_2 替代 ρ 中的模式 s , 其中 s_1 由 A 和 X 构成, s_2 由 s 中除 A 以外的所有属性构成(可以发现, $s_1 \in \text{BCNF}$)。

(3)重复步骤(2), 直至 ρ 中全部关系模式达到BCNF。

注: 本算法不能保证一关系模式分解成BCNF而又保持依赖。

[定理]算法正确地给出了关系模式到BCNF的无损连接分解。

证明: (略)。

$S_1 \cap S_2$	$S_1 - S_2$	$S_2 - S_1$
X	A	...

关系模式分解成3NF

示例： $R(A, B, C, D, E, F, G)$

函数依赖集合 $\{ A \rightarrow B, A \rightarrow C, C \rightarrow D, C \rightarrow E, E \rightarrow FG \}$

候选键： A ；有传递依赖， R 不满足3NF。

分解规则：

将每一个函数依赖单独组成一个关系

$\rho = \{ R_1(A, B), R_2(A, C), R_3(C, D), R_4(C, E), R_5(E, F, G) \}$

可以看出：每一个模式都属于3NF，且 ρ 是保持依赖的

也可以合并一些关系：

$\rho = \{ R_{12}(A, B, C), R_{34}(C, D, E), R_5(E, F, G) \}$

ρ 满足保持依赖性吗？ ρ 满足无损连接性？

[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 ρ , ρ 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 ρ 为上述模式之集合, 则 ρ 即为所求之分解。

[定理] 算法正确地给出了关系模式到第三范式的保持依赖分解。

证明: (略)。

既保持依赖，又无损连接的分解

[定理] 设 σ 是按前述算法构造的R的一个第三范式分解，X是R的候选键，则： $\tau = \sigma \cup \{ X \}$ 将是R的一个分解，且该分解中的所有关系模式是第三范式的， τ 有保持依赖和无损连接性。

注： τ 并不一定具有定理性质的最小可能关系模式的集合。我们可以依次去掉一个关系模式，只要所要求的性质仍具备，直至求得上述最小集合。

示例： $R(A, B, C, D, E, F, G)$

函数依赖： $A \rightarrow B, A \rightarrow C, C \rightarrow D, C \rightarrow E, E \rightarrow FG$

保持依赖的分解成3NF的集合：

$\rho = \{ R_{12}(A, B, C), R_{34}(C, D, E), R_5(E, F, G) \}$

ρ 满足无损
连接性？

[Algorithm]无损连接分解成4NF

Input : 关系模式 $R(U, D)$, D 为 R 上的一个依赖集(多值、函数依赖)。

Output : R 的一个无损连接分解 ρ , ρ 中的每个关系模式都是 D 在该模式上投影的4NF。

Method : (1) 令 $\rho = \{ R \}$ 。

(2)对每个模式 $s \in \rho$, 若 $s \notin 4NF$, 则 s 上必有一依赖 $X \twoheadrightarrow Y$ 成立且 X 不是 s 的超键且 $Y - X \neq \emptyset$, $XY \neq s$, 令 $Z = Y - X$, 显然有 $X \twoheadrightarrow Z$, 此时用模式 s_1, s_2 替代 ρ 中的模式 s , 其中 s_1 由 Y 和 X 构成, s_2 由 $Y - X$ 构成。

(3)重复步骤(2), 直至 ρ 中全部关系模式达到4NF。

[定理]算法正确地给出了关系模式到4NF的无损连接分解。

证明 : (略)。

连接依赖与第5NF

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

[Definition]连接依赖

设 R 为一关系模式， $\rho = \{ R_1, \dots, R_n \}$ 为 R 的一个分解，若对 R 的任一关系 r 均有： $r_{n目} = \pi_{R_1}(r) \bowtie \pi_{R_2}(r) \bowtie \dots \bowtie \pi_{R_n}(r)$ ，则称 R 满足 n 目连接依赖，记为 $JD[R_1, \dots, R_n]$ ，或记为 n -JD。

注意：

- 多值依赖性是多值依赖的特例，2-JD: 多值依赖，后者被认为是数据依赖的最一般形式；
- 关系模式 $R(U)$ ， $X \twoheadrightarrow Y$ ， $Z = U - X - Y$ ，对 R 的任一关系 r ，按多值依赖分解成两个关系模式 $\rho = \{ R_1(XZ), R_2(YZ) \}$ ，则 ρ 满足无损联结，即 $r = m_\rho(r)$ 。即2-JD。
- JD是对关系的隐含约束，将影响到元组的插入和删除。

[Definition] 5NF

当且仅当关系模式R的每个连接依赖均按其候选键进行连接运算时(均由R的候选键所隐含)，则称R是第五范式的，记为 $R \in 5NF$ 。

注意：

- 第五范式消除了不按候选键连接的连接依赖(R的无损连接分解中各模式必含有一个候选键)，但其语义背景抽象。
- $5NF \subseteq 4NF$ 。第五范式也称投影连接范式，即PJNF。
- 虽然总能把一个关系无损分解成多个5NF的关系，但由于目前尚不清楚如何找到关系的所有JD，故不清楚如何确定5NF关系(只能用穷举法)。

数据库设计需要知道的

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

数据库设计需要知道的

(1)问题：怎样保证数据库设计的正确性？

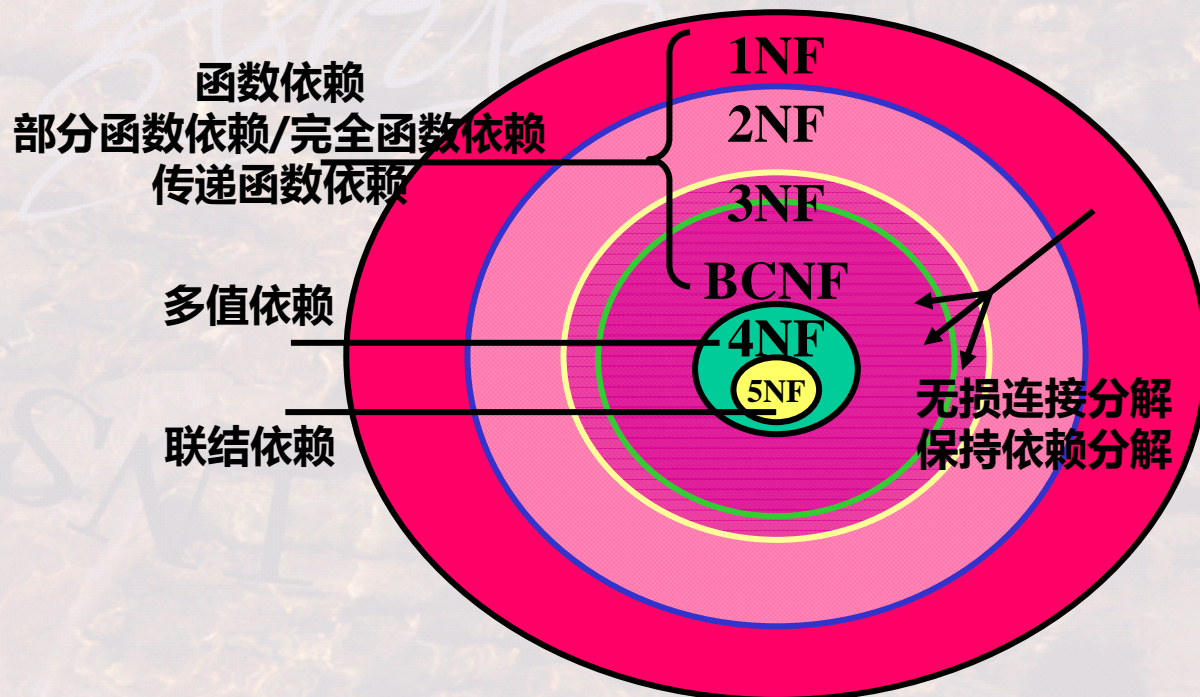


如何避免数据库的一致性问题——数据库的规范性设计

数据库的规范性设计需要分析数据库Table中的属性在取值方面有什么依存关系？数据库设计过程中应遵循什么样的原则

➤数据库设计理论

- 数据依赖理论
- 关系范式理论
- 模式分解理论



数据库设计需要知道的

(3)数据库设计理论要解决的根本问题



根本的问题

- 哪些属性被组织成一个关系？
- 是一个**大关系模式**呢，还是若干**小关系模式**？
- 大关系模式存在什么问题？

学号	姓名	班级	班主任	班主任职称
2003510101	张三	035101	张林	讲师
2003510102	李四	035101	张林	讲师
2003510103	王五	035101	张林	讲师
2003510104	李六	035101	张林	讲师
2003510105	张四	035101	张林	讲师
2003510106	张五	035101	张林	讲师
2003510107	张小三	035101	张林	讲师
2003510108	张小四	035101	张林	讲师
2003510109	李小三	035101	张林	讲师
2003510110	李小四	035101	张林	讲师
2003520201	周三	035202	郑东	副教授
2003520202	赵四	035202	郑东	副教授
2003520203	赵五	035202	郑东	副教授
2003520204	赵六	035202	郑东	副教授
2003520205	钱四	035202	郑东	副教授
2003520206	强五	035202	郑东	副教授
2003520207	梁小三	035202	郑东	副教授
2003520208	梁小四	035202	郑东	副教授
2003520209	王小三	035202	郑东	副教授
2003520210	王小四	035202	郑东	副教授

学号	姓名	班级
2003510101	张三	035101
2003510102	李四	035101
2003510103	王五	035101
2003510104	李六	035101
2003510105	张四	035101
2003510106	张五	035101
2003510107	张小三	035101
2003510108	张小四	035101
2003510109	李小三	035101
2003510110	李小四	035101
2003520201	周三	035202
2003520202	赵四	035202
2003520203	赵五	035202
2003520204	赵六	035202
2003520205	钱四	035202
2003520206	强五	035202
2003520207	梁小三	035202
2003520208	梁小四	035202
2003520209	王小三	035202
2003520210	王小四	035202

班级	班主任	班主任职称
035101	张林	讲师
035202	郑东	副教授

数据库设计需要知道的

(4)关系模式设计的折中

➤关系模式设计需要折中

- 遵循关系范式原则，则需要将一个关系模式，拆解成两个或多个小的模式；而查询时，需要将这两个或多个小的模式联结成一个模式；
- 遵循关系范式原则避免了冗余、插入异常、删除异常等问题，但由于联结运算的低效率，使得查询速度很慢。因此需要折中。

学号	姓名	班级	班主任	班主任职称
2003510101	张三	035101	张林	讲师
2003510102	李四	035101	张林	讲师
2003510103	王五	035101	张林	讲师
2003510104	李六	035101	张林	讲师
2003510105	张四	035101	张林	讲师
2003510106	张五	035101	张林	讲师
2003510107	张小三	035101	张林	讲师
2003510108	张小四	035101	张林	讲师
2003510109	李小三	035101	张林	讲师
2003510110	李小四	035101	张林	讲师
2003520201	周三	035202	郑东	副教授
2003520202	赵四	035202	郑东	副教授
2003520203	赵五	035202	郑东	副教授
2003520204	赵六	035202	郑东	副教授
2003520205	钱四	035202	郑东	副教授
2003520206	强五	035202	郑东	副教授
2003520207	梁小三	035202	郑东	副教授
2003520208	梁小四	035202	郑东	副教授
2003520209	王小三	035202	郑东	副教授
2003520210	王小四	035202	郑东	副教授

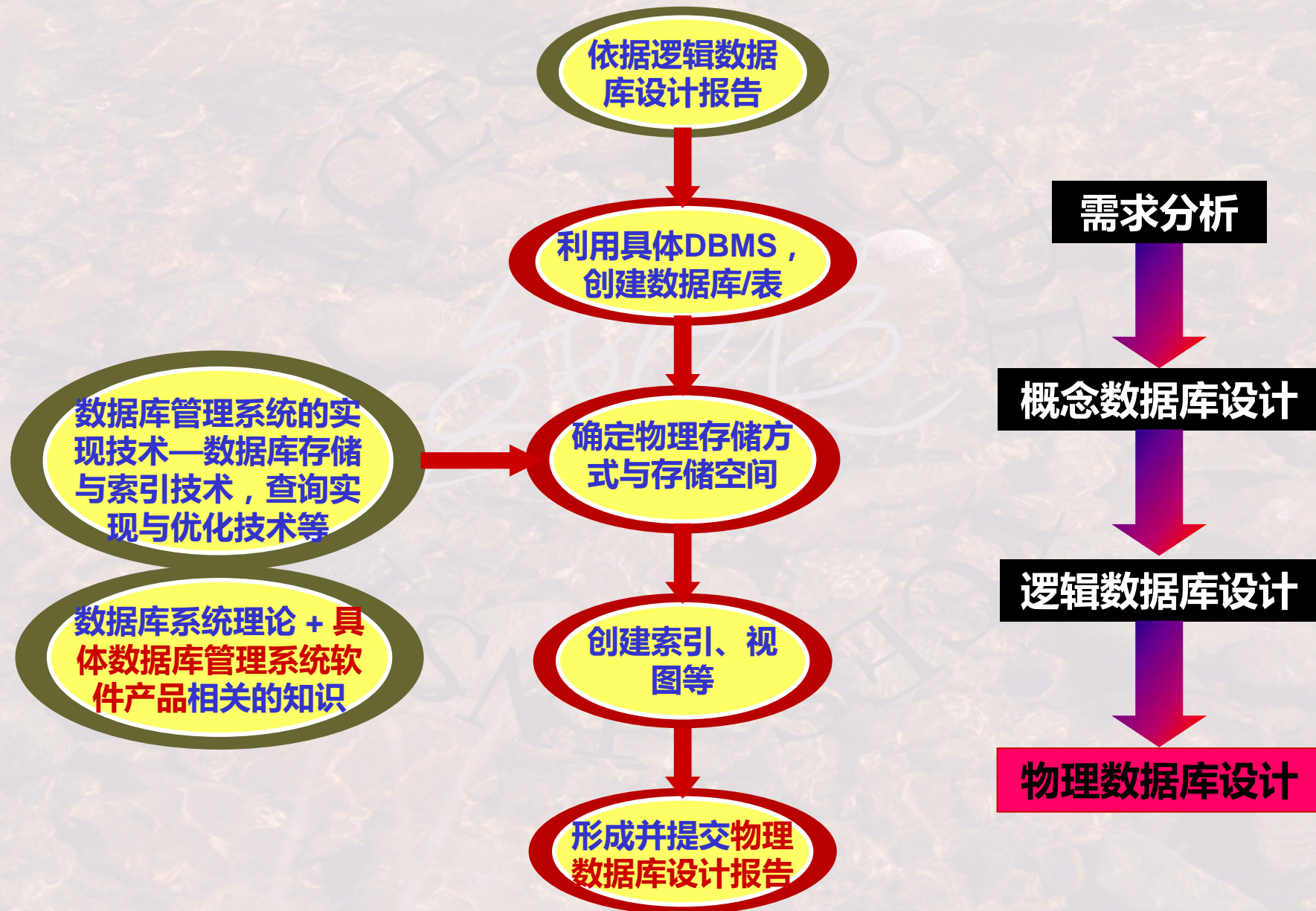
学号	姓名	班级
2003510101	张三	035101
2003510102	李四	035101
2003510103	王五	035101
2003510104	李六	035101
2003510105	张四	035101
2003510106	张五	035101
2003510107	张小三	035101
2003510108	张小四	035101
2003510109	李小三	035101
2003510110	李小四	035101
2003520201	周三	035202
2003520202	赵四	035202
2003520203	赵五	035202
2003520204	赵六	035202
2003520205	钱四	035202
2003520206	强五	035202
2003520207	梁小三	035202
2003520208	梁小四	035202
2003520209	王小三	035202
2003520210	王小四	035202

班级	班主任	班主任职称
035101	张林	讲师
035202	郑东	副教授

通常建议：
关系模式符合BCNF

数据库设计需要知道的

(5)数据库设计还包括物理数据库设计—需要继续学习



回顾本讲学了什么？

战德臣

哈尔滨工业大学 教授·博士生导师

黑龙江省教学名师

教育部大学计算机课程教学指导委员会委员

Research Center on **I**ntelligent
Computing for **E**nterprises & **S**ervices,
Harbin **I**nstitute of **T**echnology

回顾本讲学习了什么？

