第十一章 关系数据理论

1. 函数依赖
2. 范式
3. 模式分解

模式的分解

- 关系模式的规范化过程是通过对关系模式的投影分解来实现的，但是投影分解方法不是唯一的，不同的投影分解会得到不同的结果。
- 在这些分解方法中，只有能够保证分解后的关系模式与原关系模式等价的方法才是有意义的。

模式的分解

- 对于关系模式SD(SNO,SN,AGE,DEPT,MN)，规范到3NF，可以有下面三种效果不同的分解方法

- 第一种： 分解为 S(SNO,SN,AGE,DEPT)， D(DEPT，MN)
  如此分解的结果，根据前节分析，可以消除数据异常，是好的分解方法

模式的分解

- 第二种：分解为S1(SNO,SN,AGE,DEPT)， D1(SNO，MN) 分解后的关系实例如下图
- 分解以后，两个关系的主键都为SNO，也不存在非主属性对主键的传递函数依赖，所以两个关系均属于3NF

S1 D1

SNO	SN	AGE	DEPT		SNO	MN
S1	赵亦	17	计算机		S1	刘伟
S2	钱尔	18	信息		S2	王平
S3	孙珊	20	信息		S3	王平
S4	李思	21	自动化		S4	刘伟

模式的分解

- 验证可知SD=S1⋈D1，关系模式SD等于S1和D1在SNO上的自然连接，分解具有无损连接性。
- 但是，SD中的函数依赖DEPT→MN在S1和D1中没有保留
- 这种分解结果，仍然存在着一些问题：
  - ✓ 1．数据冗余。每个系名和系主任的名字存储的次数等于该系的学生人数。
  - ✓ 2．插入异常。当一个新系没有招生时，系主任的名字则无法插入。
  - ✓ 3．删除异常。某系学生全部毕业而没有招生时，要删除全部学生的记录，两个关系都要涉及，有关该系的信息将被删除。
  - ✓ 4．更新异常。更换系主任时，需改动较多的学生记录。另外，某个学生要转系，还必须修改两个关系。

模式的分解

- 之所以存在上述问题，是因为分解得到的两个关系模式不是相互独立的。
- SD中的函数依赖DEPT→MN既没有投影到关系模式S1上，也没有投影到关系模式D1上，而是跨在这两个关系模式上，也就是说这种分解方法没有保持原关系中的直接函数依赖，却用了原关系隐含的传递函数依赖SNO $\xrightarrow{t}$ MN。
- 分解只具有无损连接性，但没有“保持函数依赖”。

模式的分解

- 第三种：分解为S2(SNO,SN,AGE,MN), D2(DEPT，MN)
- 分解以后，两个关系均为3NF 。
- 但验证发现S2⋈D2比原来的关系SD多了两个元组（S1，赵亦，17，自动化，刘伟）和（S4，李思，21，计算机，刘伟），因此此分解方法是不可恢复的。
- 另外，这种分解方法只保持了原来的SD中的DEPT→MN这个完全函数依赖，但却丢失了SNO→DEPT完全依赖，所以分解也没有保持原有关系中的函数依赖。

模式的分解

- 第三种：分解为S2(SNO,SN,AGE,MN), D2(DEPT，MN)
- 此种分解也存在着很多的数据异常（哪些？）

S2　　　　D2

SNO	SN	AGE	MN		DEPT	MN
S1	赵亦	17	刘伟		计算机	刘伟
S2	钱尔	18	王平		信息	王平
S3	孙珊	20	王平		自动化	刘伟
S4	李思	21	刘伟			

模式的分解

- 有哪些分解方法?
- 不同的模式分解方法，为何会产生不同的分解效果?
- 该怎样正确的进行模式分解?

- 特别是，怎么回答下面的问题?

设U={ABCDEF}， F={ABC→DE, DE→ABC, AB→D, E→C, DE→F}， 该怎样分解到3NF，或BCNF?

模式的分解

- 模式分解的定义

  关系模式R<U,F>的一个分解：

  $\rho=\{R_1<U_1,F_1>，R_2<U_2,F_2>，\ldots，R_n<U_n,F_n>\}$

  $U=U_1\cup U_2\cup\ldots\cup U_n$，且不存在 $U_i\subseteq U_j$，F_i 为 F在 U_i 上的投影

- 函数依赖集合$\{X\rightarrow Y \mid X\rightarrow Y \in F^+ \wedge XY\subseteq U_i\}$ 的一个覆盖 F_i 叫作 F 在属性 U_i 上的投影

那我们需要先了解函数依赖的公理系统

数据依赖的公理系统

- Armstrong公理系统
  - 一套推理规则，是模式分解算法的理论基础
  - 用途
    - ✓ 从一组函数依赖求得蕴含的函数依赖
    - ✓ 求给定关系模式的候选键

已知A→B, BC→D， 候选键是？

- 逻辑蕴含

对于满足一组函数依赖 F 的关系模式R <U， F>， 其任何一个关系r，若函数依赖X→Y都成立, 则称

F 逻辑蕴含 X →Y

数据依赖的公理系统

- 对关系模式R <U，F >，有以下的推理规则：
  - Al. 自反律（Reflexivity）：
    若Y $\subseteq$ X $\subseteq$ U，则X →Y为 F 所蕴含。
  - A2.增广律（Augmentation）：若X→Y为F所蕴含，且Z $\subseteq$ U，则XZ→YZ为 F 所蕴含。
  - A3.传递律（Transitivity）：若X→Y及Y→Z为 F 所蕴含，则X→Z为 F 所蕴含。

注意：由自反律所得到的函数依赖均是平凡的函数依赖

数据依赖的公理系统

■ 根据A1，A2，A3这三条推理规则可以得到下面三条推理规则：

- 合并规则：由X→Y，X→Z，有X→YZ。
- 伪传递规则：由X→Y，WY→Z，有XW→Z。
- 分解规则：由X→Y及 $Z \subseteq Y$，有X→Z。

■ 根据合并规则和分解规则，可得引理：

$X \rightarrow A_1 A_2 \ldots A_k$ 成立的充分必要条件是 $X \rightarrow A_i$ 成立

数据依赖的公理系统

■ 函数依赖闭包

- 在关系模式R<U，F>中为 F 所逻辑蕴含的函数依赖的全体叫作 F的闭包，记为F^+。
  - ✓ 蕴含==导出
- F^+的计算是NP完全问题，怎么办?

转换思路

■ 定义

设F为属性集U上的一组函数依赖，$X \subseteq U$，

则 $X_F^+ = \{ A \mid X \rightarrow A$能由 F 根据Armstrong公理导出$\}$

X_F^+称为属性集 X 关于函数依赖集 F 的闭包

数据依赖的公理系统

- 引理

设F为属性集U上的一组函数依赖，X，Y $\subseteq$ U，X→Y能由F 根据Armstrong公理导出的充分必要条件是Y $\subseteq$ X_F^+

- 用途

将判定 X→Y 是否能由 F 根据Armstrong公理导出的问题，转化为求出X_F^+，判定Y是否为X_F^+的子集的问题

数据依赖的公理系统

■ 求属性集X（$X \subseteq U$）关于U上的函数依赖集F 的闭包X_F^+的算法

- 输入：X，F
- 输出： X_F^+

■ 步骤：

（1）令X（0）=X，i=0

（2）求B，这里$B = \{ A \mid (\exists V)(\exists W)(V \to W \in F \wedge V \subseteq X(i) \wedge A \in W)\}$；

（3）X（i+1）=B∪X（i）

数据依赖的公理系统

■ 求属性集X（X $\subseteq$ U）关于U上的函数依赖集F 的闭包X_F^+的算法

- 输入：X，F
- 输出： X_F^+

■ 步骤：

（4）判断X（i+1）= X （i）吗?

（5）若相等或X（i）=U，则X（i）就是X_F^+，算法终止。

（6）若否，则 i=i+l，返回第（2）步。

数据依赖的公理系统

■ 已知关系模式R<U，F>，其中 U={A，B，C，D，E}；F={AB→C，B→D，C→E，EC→B，AC→B}。求（AB）$_F^+$ 。

- 解 设X（0）=AB；

(1) 计算X（1）：逐一的扫描F集合中各个函数依赖，

找左部为A，B或AB的函数依赖。得到两个：

AB→C，B→D。

于是X（1）=AB∪CD=ABCD。

数据依赖的公理系统

■ 已知关系模式R<U，F>，其中 U={A，B，C，D，E}；F={AB→C，B→D，C→E，EC→B，AC→B}。求（AB）$_F^+$ 。

- 解 设X（0）=AB；

(2)因为X（0）≠ X（1） ，所以再找出左部为ABCD子集的那些函数依赖，又得到AB→C，B→D， C→E，AC→B，

于是X（2）=X（1）∪BCDE=ABCDE。

(3)因为X（2）=U，算法终止

所以（AB）$_F^+$ =ABCDE。

数据依赖的公理系统

■ 练一练：

- $U=\{A, B, C, D\}$; $F=\{A \rightarrow B, BC \rightarrow D\}$;
  - ✓ $A+ = AB$.
  - ✓ $C+ = C$.
  - ✓ $(AC)+ = ABCD$.

数据依赖的公理系统

- 建立公理系统体系目的：从已知的 f 推导出未知的f
- 有效性：由F出发根据Armstrong公理推导出来的每一个函数依赖一定在F+中

  /* Armstrong公理系统是正确的
- 完备性：F+中的每一个函数依赖，必定可以由F出发根据Armstrong公理推导出来

  /* Armstrong公理是完备的
  - 完备性：所有不能用Armstrong公理推导出来f, 都不为真；若 f 不能用Armstrong公理推导出来， $f \notin F+$

数据依赖的公理系统

- 我们在模式分解的时候常常要确定给定关系模式的最小依赖集，为此需了解下面的概念和方法：

- 函数依赖集等价
  - 如果G+=F+，就说函数依赖集 F 覆盖G（F是G的覆盖，或G是F的覆盖），或F与G等价。
  - F+ = G+ 的充分必要条件是

    $F \subseteq G+$，和 $G \subseteq F+$

  要判定$F \subseteq G+$，只须逐一对 F 中的函数依赖X→Y，考察 Y 是否属于X_{G+}^{+} 就行了。

数据依赖的公理系统

- 最小依赖集
  - 如果函数依赖集 F 满足下列条件，则称 F 为一个极小函数依赖集。亦称为最小依赖集或最小覆盖。

(1) F中任一函数依赖的右部仅含有一个属性。

(2) F中不存在这样的函数依赖$X \to A$，使得F 与 $F-\{X \to A\}$等价。

(3) F中不存在这样的函数依赖$X \to A$， X有真子集Z 使得$F-\{X \to A\} \cup \{Z \to A\}$与F等价。

数据依赖的公理系统

- 例：对于关系模式S<U，F>，其中

U={ SNO，Sdept，MN，Cname，G }，

F={ SNO→Sdept，Sdept→MN，（SNO，Cname）→G }

- 设F′={SNO→SDEPT，SNO→MN，Sdept→MN，

(SNO，Cname)→G，(SNO，Sdept)→Sdept }

- 则F′不是最小覆盖集，而F 是最小覆盖集。

  因为：F′- {SNO→MN}与F′等价

  F′- {(SNO，SDEPT)→SDEPT}也与F′等价

  F′- {(SNO，SDEPT)→SDEPT}

  ∪{SNO→SDEPT}也与F′等价

数据依赖的公理系统

- 每一个函数依赖集F 均等价于一个极小函数依赖集F_m。此F_m称为F 的最小依赖集
- 最小依赖集构造方法：
  - 依据定义分三步对F 进行“极小化处理”，找出F的一个最小依赖集。

(1)逐一检查F中各函数依赖FD_i：$X \rightarrow Y$，

若$Y = A_1A_2 \ldots A_k$，$k > 2$，

则用 $\{ X \rightarrow A_j \mid j=1, 2, \ldots, k\}$ 来取代$X \rightarrow Y$。

数据依赖的公理系统

- 最小依赖集构造方法：
  - 依据定义分三步对F 进行“极小化处理”，找出F的一个最小依赖集。

(2)逐一检查F 中各函数依赖FD_i：$X\rightarrow A$，令$G=F-\{X\rightarrow A\}$，

若$A\in X_G^+$，则从F中去掉此函数依赖。

(3)逐一取出F 中各函数依赖FD_i：$X\rightarrow A$，设$X=B_1B_2\ldots B_m$，

逐一考查B_i（i=l，2，…，m），

若$A\in(X-B_i)_F+$，则以$X-B_i$取代X。

经此处理，最后剩下的F就一定是极小依赖集。

数据依赖的公理系统

- 练一练：$F = \{A\to B，B\to A，B\to C，A\to C，C\to A\}$

  - 以下F_m1、F_m2都是F的最小依赖集：

    $F_m1 = \{A\to B，B\to C，C\to A\}$

    $F_m2 = \{A\to B，B\to A，A\to C，C\to A\}$

- F的最小依赖集F_m不一定是唯一的，它与对各函数依赖FD_i及$X\to A$中X各属性的处置顺序有关
- 若改造后的F 与原来的F 相同，说明F 本身就是一个最小依赖集

假设有一个关系模式 R(A,B,C,D)，以及以下的函数依赖集 *F*：A→B B→C D→B D→C 我们需要求出 *F* 的最小依赖集 F′。 A→B, B→C, D→B.

给定函数依赖集F如下： F = {A→B, BC→D, BD→E, AB→C, ACD→E}， 求F的最小依赖集

$F'=\{A \to B, BC \to D, BD \to E, AB \to C\}$

模式的分解

- 再来理解模式分解的定义

  关系模式R<U,F>的一个分解：

  $\rho=\{R_1<U_1,F_1>, R_2<U_2,F_2>, \ldots, R_n<U_n,F_n>\}$

  $U=U_1\cup U_2\cup\ldots\cup U_n$，且不存在 $U_i\subseteq U_j$，F_i 为 F在 U_i 上的投影

- 函数依赖集合$\{X\rightarrow Y \mid X\rightarrow Y\in F^+\wedge XY\subseteq U_i\}$ 的一个覆盖 F_i 叫作 F 在属性 U_i 上的投影
  - 覆盖即等价的意思

模式的分解

- 求F在U_i上的投影的计算方法：
  - 找到U_i上所有可能的函数依赖
  - 验证函数依赖是否可由F推导而来
    - 对于每个X->Y, 检查Y是否属于X_{ui}^+
  - 最小化依赖集

假设有一个关系模式 R(A,B,C,D,E)，及其函数依赖集 F 如下

(A→B，BC→D, BD→E, AB→C, ACD→E)

计算F在属性子集 U={A,B,C} 上的投影 F_U

F_U={A→B, A→C}

假设有一个关系模式 R(A,B,C,D,E)，及其函数依赖集 *F* 如下：A→B BC→D BD→E AB→C ACD→E

计算F在属性子集 U={A,B,C} 上的投影 FU

我们要找到 U={A,B,C} 上所有可能的函数依赖。这些候选依赖包括但不限于 A→B、A→C 、B→A 、B→C 、C→A 、C→B 、AB→C 、BC→A 等等。

验证候选依赖

对于每个候选依赖 X→Y，检查它是否能从 *F* 推导出来

得到函数依赖集 *F* 在属性子集 U上的投影 FU 是：
FU={A→B,A→C,AB→C}

去除冗余后，是A→B, A→C

模式的分解

- 模式分解是不是正确的，有下面两个判定依据：
  - 1 分解具有无损连接性
    - ✓ R与ρ在数据内容方面是否等价
  - 2 分解保持函数依赖
    - ✓ R与ρ在函数依赖方面是否等价

模式的分解

有信息损失的分解：

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

$\bowtie$

ρ

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

模式的分解

■ 分解的无损连接性

- 关系模式R<U,F>的一个分解 $\rho=\{ R_1<U_1,F_1>，R_2<U_2,F_2>，\ldots, R_n<U_n,F_n>\}$
- 若R与R1、R2、…、Rn自然连接的结果相等，则称关系模式R的这个分解ρ 具有无损连接性（Lossless join）
- 具有无损连接性的分解保证不丢失信息
- 无损连接性不一定能解决插入异常、删除异常、修改复杂、数据冗余等问题

模式的分解

■ 无损连接性分解的判定方法---1. 列表法

（1）构造一个k行n列的二维表T，第i 行对应于一个关系模式R_i，第j列对应于属性A_j，令：

$t_{ij}=a_j$，若A_j属于R_i；否则$t_{ij}=b_{ij}$。

（2）对于F中一个FD：X→Y，如果表格中有两行在X分量上相等，在Y分量上不相等，那么把这两行在Y分量上改成相等。如果Y的分量中有一个是a_j，那么另一个也改成a_j； 如果没有a_j，那么用其中的一个b_{ij}替换另一个(尽量把ij改成较小的数，亦即取i 值较小的那个)

模式的分解

■ 无损连接性分解的判定方法---1. 列表法

（3）若在修改的过程中，发现表格中有一行全是a，即$a_1,a_2,\ldots,a_n$，那么可立即断定ρ 相对于F 是无损连接分解，此时不必再继续修改。若经过多次修改直到表格不能修改之后，发现表格中不存在有一行全是a的情况，那么分解就是有损的。特别要注意，这里有个循环反复修改的过程，因为一次修改可能导致表格能继续修改。

模式的分解

- 现有关系模式R<U,F>,U={A,B,C}，F={B→C}

  R的一个分解为R_1(A,B), R_2(B,C)，该分解是否无损连接性?

1) 构造初始表格

A	B	C
a1	a2	b13
b11	a2	a3

2) 根据B→C修改表格

A	B	C
a1	a2	a3
b11	a2	a3

第一行符合a1,a2,...,an的情况，判断该分解无损连接性

模式的分解

- 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

  ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

  分解是否具有无损连接性?

模式的分解

■ 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

1）初始表格

	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

模式的分解

- 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

  ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

用A → C 修改

2）用每个函数依赖修改表格

	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

模式的分解

- 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

  ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

2）用每个函数依赖修改表格

用B → 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

模式的分解

■ 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

用CE→A, C→D, DE→C依次修改，最后得到：

2）用每个函数依赖修改表格

	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

模式的分解

- 现有关系模式 R(A,B,C,D,E)， 其F={A → C, B → C, C → D, DE → C, CE → A}。将其分解为：

  ρ ={R_1(A,D), R_2(A,B), R_3(B,E), R_4(C,D,E), R_5(A,E)}

3）检查是否有一行变成全a
有：无损连接
无：有损连接

	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

模式的分解

■ 无损连接性分解的判定方法---2. 定理法

(适合关系模式R分解为两个关系模式R1、R2时)

定理：若关系模式R<U,F>中，被分解为ρ ={R_1<U_1,F_1>, R_2<U_2,F_2>}是R的一个分解，若$R_1 \cap R_2 \rightarrow R_1 - R_2$或者 $R_1 \cap R_2 \rightarrow R_2 - R_1$，则称$\rho$ 具有无损连接性

示例：已知R<U,F>,U={A,B,C}，F={$B \rightarrow C$}

R的一个分解为R1(A,B), R2(B,C)

解：$R_1 \cap R_2$ 为B，R_2-R_1为C, 因为$B \rightarrow C$，因此分解是无损连接

模式的分解

■ 无损连接性分解的判定方法---2. 定理法

引理： 设关系模式R 具有函数依赖集F， ρ={ R_1， R_2， …， R_n} 是R的一个分解，且是关于F无损连接性的分解，则有：

对特定i，设$F_i=\pi_{Ri}$（F）（F_i是F在R_i上的投影），σ= （S1, S2, ..., Sm）是R_i关于F_i的无损连接分解，则R到{ R_1， ..., R_{i-1}， S1, S2, …， Sm, R_{i+1}, …, R_n} 的分解是关于F 无损连接的。

用投影分解方法逐步分解，

可以保证无损连接性

模式的分解

- 分解的函数依赖保持性
  - 设关系模式R(U，F)被分解为若干个关系模式$R_1(U_1，F_1)$，$R_2(U_2，F_2)$，..., $R_n(U_n，F_n)$，其中$U=U_1 \cup U_2 \cup \ldots \cup U_N$，且不存在$U_i \subseteq U_j$，$F_i$为F在$U_i$上的投影。如果F 所蕴含的任意一个函数依赖一定也由 $(F_1 \cup F_2 \ldots \cup F_n)$所蕴含，则称关系模式R的分解具有函数依赖保持性。

模式的分解

■ 如果一个分解具有无损连接性，则它能够保证不丢失信息。

■ 如果一个分解保持了函数依赖，则它可以减轻或解决各种异常情况。

■ 分解具有无损连接性和分解保持函数依赖是两个互相独立的标准。具有无损连接性的分解不一定能够保持函数依赖。同样，保持函数依赖的分解也不一定具有无损连接性。

模式的分解

- 示例：R（CSZ），F={ CS → Z, Z → C}，则：

ρ={R_1（SZ），R_2（CZ）} 为一无损分解，但不保持依赖

- 又：R（ABCD），F={A → B, C → D },
- ρ={ R_1（AB），R_2（CD）} 保持了函数依赖，但不是无损分解。

模式的分解

■ 函数依赖保持性的判定方法

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是否保持函数依赖的判断

Method: 令$G=\cup_{(i=1,...,k)}\pi_{Ri}(F)$， 只需检查G是否覆盖F

■ 具体算法

1. 对每个$X\rightarrow Y\in F$， 计算G中的X_G^+（如果X不包含于R_i就不需计算了）：

   Z=X

   While Z发生变化 DO

   For i=1 to k Do

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

求出的Z就是X_G^+

模式的分解

- 函数依赖保持性的判定方法

- 具体算法

2. 判断G是否逻辑蕴含 X → Y： 如果Z 包含Y，则G 逻辑蕴含X→Y, 否则不是。
3. 判断 ρ 是否保持依赖： 如果G 逻辑蕴含F 中的每一个函数依赖，则说ρ是保持函数依赖的分解，否则不是。

模式的分解

■ 函数依赖保持性判断示例：

- R(A, B, C, D, E), F={A→C, B→C, C→D, DE→C, CE→A}

ρ={ R_1(AC), R_2(BC), R_3(CDE) }

- 分析：π_{R1}(F)={A→C}， π_{R2}(F)={B→C}， π_{R3}(F)={C→D, DE→C} , 则 G={A→C, B→C, C→D, DE→C}， 初步判断分解没有保持函数依赖。但怎样通过算法证明呢？：

□ 1. 对函数依赖A→C∈F 计算G中的X_G^+:

z={A} ∪ {C} ∪{} ∪{}={A, C}， C 包含于Z中，所以A→C被G逻辑蕴含。

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

左端属性是某个关系模式(R_i)的属性

求关于G的属性闭包，结果属性也须是R_i的属性

模式的分解

- 函数依赖保持性判断示例:
  - R(A, B, C, D, E), F={A→C, B→C, C→D, DE→C, CE→A}

  ρ={ R_1(AC), R_2(BC), R_3(CDE) }
  - 分析：π_{R1}(F)={A→C}， π_{R2}(F)={B→C}， π_{R3}(F)={C→D, DE→C} ,

    则G={A→C, B→C, C→D, DE→C}， 初步判断分解没有保持函数依赖

    。但怎样通过算法证明呢？：

□ 1. 对函数依赖A→C∈F 计算G中的X_G^+:

z={A} ∪ {C} ∪{} ∪{}={A, C}， C 包含于Z中，所以A→C被G逻辑蕴含。

□ 2. 对函数依赖DE→C ∈F 计算G中的X_G^+:

z={D,E} ∪ {} ∪{} ∪{C,D}={C,D,E}， C 包含于Z中，所以DE→C被G逻辑蕴含

模式的分解

- 函数依赖保持性判断示例:
  - R(A, B, C, D, E), F={A→C, B→C, C→D, DE→C, CE→A}

ρ={ R_1(AC), R_2(BC), R_3(CDE) }

  - 分析:π_{R1}(F)={A→C}, π_{R2}(F)={B→C}, π_{R3}(F)={C→D, DE→C} ,则G={A→C, B→C, C→D, DE→C}, 初步判断分解没有保持函数依赖。但怎样通过算法证明呢? :

- 1. 对函数依赖A→C∈F 计算G中的X_G^+:

A→C被G逻辑蕴含。

- 2. 对函数依赖DE→C ∈F 计算G中的X_G^+:

DE→C被G逻辑蕴含

- 3. 对函数依赖CE→A ∈F 计算G中的X_G^+:

z={C,E} ∪ {} ∪{} ∪{D}={C,D,E}, A不包含于Z中,所以不被G逻辑蕴含

所以G 没有逻辑蕴含F中的每一个函数依赖,没有保持函数依赖

给定一个关系模式 R(A,B,C,D) 及其函数依赖集 F （A→B， B→C， C→D）

将关系模式 R 分解成以下两个子模式：R1(A,B)， R2(B,C,D)

判断这个分解是否具有函数依赖保持性

解：

先确定G： G(A→B, B→C, C→D)

针对F中的A→B，计算A_G^+ ，得(A,B)，包含B，所以被G蕴含

针对B→C, 得B_G^+=(B，C), 包含C，所以被G蕴含

针对C→D，得C_G^+=(C，D)，包含D，所以被G蕴含

因此这个分解是具有函数依赖保持性的

模式的分解

- 怎样正确的分解到3NF和BCNF?

- 1. 分解到BCNF的无损连接分解法
  - 关系模式R（U，F）
  - 令ρ={ R1, R2, ..Rn }
  - 对每个模式S∈ ρ， 若s ∉ BCNF，则s 上必有X → A成立且X 不是s的超键（A ∉ X）。此时用模式S1, S2代替S, 其中S1由A和X构成（可知S1 ∈ BCNF），S2由S中除了A以外所有属性构成
  - 重复上面一步，直至ρ中所有关系模式达到BCNF
- 本算法无法保证保持所有函数依赖
- 换言之，可以无损连接的分解成BCNF，但有可能会丢失部分函数依赖

模式的分解

- 2. 分解到3NF的保持依赖分解法
  - 关系模式R（U，F），其中F是函数依赖集最小覆盖（保证分解时候不会有冗余的依赖）
  - 分解：

1. 把R 中不出现在F 中的属性去掉并单独组成一个模式U0
2. 对于所有X→A ∈F，则以XA组成一模式Ui； 若有X→A1, X→A2,…X→Am都属于F，则以XA1A2…Am组成一模式Uj（即将n个模式合并为一个模式）
3. 取ρ为上述模式集合，则ρ为所求分解。

此方法不能保证无损连接

模式的分解

2. 分解到3NF的保持依赖分解法

- 例： SL(Sno, Sdept, Sloc)

  F= {Sno→Sdept, Sdept→Sloc, Sno→Sloc }

  极小化处理：F’ ={Sno→Sdept, Sdept→Sloc}

  得到保持依赖的3NF分解：R1{Sno, Sdept}和R2{Sdept, Sloc}

- 例： S{CTHRSG}， 若最小依赖集F’ ={C→T, HR→C, CS→G, HS→R, HT→R}， 则：
- 保持函数依赖的3NF分解为：{CT}, {CHR}, {CSG}, {HSR}, {HTR}

把F’ 中所有函数依赖按左边属性分组，形成各关系模式，再加上F’ 中未出现的属性组成的关系模式，就是分解的结果

模式的分解

- 3. 分解到3NF的既保持依赖，又无损连接的分解法
  - 设ρ为按前述方法构造的R的一个分解，X是R的候选键。则：

$\sigma = \rho \cup \{X\}$ 是R的一个分解， 若有某个Ui, X⊆Ui, 则将{X}从σ中去掉；若Ui ⊆X， 则将Ui从σ中去掉。这样处理得到的σ中的所有关系模式是3NF的，并且保持了函数依赖和无损连接性

- 示例：

R(A, B, C, D, E, F, G)， 函数依赖A→B, A→C, C→D, C→E, E→FG

保持函数依赖分解成3NF的关系模式集合：

ρ={R1(A, B, C), R2(C, D, E), R3(E, F, G)}

因为A是候选键，且ρ中包含R1（A， B, C)，所以ρ是无损连接的

模式的分解

- 3. 分解到3NF的既保持依赖，又无损连接的分解法
  - 例： S{CTHRSG}， 若最小依赖集F’ ={C→T, HR→C, CS→G, HS→R, HT→R}， 则候选键为HS
  - 依前述分析，保持函数依赖的3NF分解为：{CT}, {CHR}, {CSG}, {HSR}, {HTR}
  - 则具有无损连接和函数依赖保持性的3NF分解为：

{CT}, {CHR}, {CSG}, {HSR}, {HTR}, {HS}

因{HS}与{HSR}重复，且{HS} ⊆ {HSR}, 因此去掉{HS}，最终结果为：

{CT}, {CHR}, {CSG}, {HSR}, {HTR}

模式的分解

- 练习题：
- 设U={ABCDEF}， F={ABC→DE, DE→ABC, AB→D, E→C, DE→F}， 求最小依赖集，并分解到3NF
- 解：

1. F’ ={ABC→D, ABC→E, DE→A, DE→B, DE→C, AB→D, E→C, DE→F}
2. 发现AB→D, ABC→D，去掉ABC→D；同理去掉DE→C
3. 得到Fm={ABC→E, DE→A, DE→B, AB→D, E→C, DE→F}
4. 其候选键是ABC, ABE, DE
5. 先获得保持函数依赖的分解：{(ABCE), (DEABF), (ABD), (EC)}
6. （EC）包含在（ABCE）中，去掉（EC）； 同理去掉（ABD）
7. 由于三个候选键都已经被包含在已有分解中，所以不用再增加分解
8. 分解结果: R1(ABCE), R2(DEABF)

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D

E →B

AD →C

- 求R的候选键。

（AEC）(AED).

- 将R分解为BCNF，给出所有可能的分解方案

$X^+ \neq R$

Solution: There are two solutions:

Solution 1:

Table	$X^+ = ?$	New table 1	New table 2
$R(A,B,C,D,E)$	$ABC+ = ABCD$	$R_1(A,B,C,D)$	$R_2(A,B,C,E)$
$R_1(A,B,C,D)$	$AD+ = ACD$	$R_3(A,C,D)$	$R_4(A,B,D)$
$R_2(A,B,C,E)$	$E+ = BE$	$R_5(B,E)$	$R_6(A,C,E)$

Answer: $R_3(A,C,D), R_4(A,B,D), R_5(B,E), R_6(A,C,E)$.

Solution 2:

Table	$X^+ = ?$	New table 1	New table 2
$R(A,B,C,D,E)$	$E+ = BE$	$R_1(B,E)$	$R_2(A,C,D,E)$
$R_2(A,C,D,E)$	$AD+ = ACD$	$R_3(A,C,D)$	$R_4(A,D,E)$

Answer: $R_1(B,E), R_3(A,C,D), R_4(A,D,E)$.

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D

E →B

AD →C

- 分解结果有没有保持函数依赖？请用算法证明你的判断

选择soluton1 R3(ACD), R4(ABD), R5(BE), R6(ACE)进行证明。

G=（AD→C, E→B）

对于ABC→D, 有z={ABC}U{}U{}={ABC}，D不包含于z,所以ABC→D不被逻辑蕴含。所以没有保持函数依赖。

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D

E →B

AD →C

- 分解结果是不是无损连接的？请证明

对于soluton2 R1(BE), R3(ACD), R4(ADE)，用列表法证明

初始表格

	A	B	C	D	E
R1	B11	A2	B13	B14	A5
R3	A1	B22	A3	A4	B25
R4	A1	B32	B33	A4	A5

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D

E →B

AD →C

- 分解结果是不是无损连接的？请证明

对于soluton2 R1(BE), R3(ACD), R4(ADE)，用列表法证明

针对E→B修改表格

	A	B	C	D	E
R1	B11	A2	B13	B14	A5
R3	A1	B22	A3	A4	B25
R4	A1	A2	B33	A4	A5

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D

E →B

AD →C

- 分解结果是不是无损连接的？请证明

对于soluton2 R1(BE), R3(ACD), R4(ADE)，用列表法证明

针对AD→C修改表格

	A	B	C	D	E
R1	B11	A2	B13	B14	A5
R3	A1	B22	A3	A4	B25
R4	A1	A2	A3	A4	A5

最后一行是A，得证是无损连接性分解

练习题

有关系模式R(A, B, C, D, E)及其上的函数依赖F：

ABC →D
E →B
AD →C

- 分解结果是不是无损连接的？请证明

也可以用定理法证明，

对于solution2， 第一次分解后R1（B，E），R2(A，C，D,E)， 此时$R_1 \cap R_2$ =（E），R1-R2=(B)， 因为F中E→B成立，所以此次分解无损连接。

第二次分解R3（A，C，D），R4(A,D，E)， $R_1 \cap R_2$ =（A，C），R1-R2=(C)，（A，C）→C是平凡的函数依赖，成立，所以此次分解无损连接。

没有后续分解了，此前的分解都无损连接，因此整个分解是无损连接的，得证。

作业

- 课本203页第2题， 第7题（第5，6，7，8小题)
- 下周上课前交