

## Formas Normales Avanzadas: Dependencias Multivaluadas

Teoría y Diseño de Base de Datos  
Departamento de Informática y Estadística

### Problemas presentados

- Existen esquemas que aún en FNBC no están lo suficiente normalizados porque se encuentran dependencias entre los datos
- Consideremos la base de datos:

**Clases (materia, profesor, bibliografía)**

- *Cualquier tupla de la relación significa que los profesores dictan materias con su bibliografía.*
- Hay tantos profesores como docentes puedan dictar la materia, y los libros que pueden usarse como bibliografía en la misma

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

4

### Desarrollo de la clase

- Definición de Dependencias Multivaluadas
- Problemas que se encuentran con DMV
- Cuarta Forma Normal
- *Ullman, Capítulo 7*
- *Covadonga, Capítulo 4*
- *Cualquier otra bibliografía con los temas*

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

2

### Problemas presentados

**Clases (materia, profesor, bibliografía)**

Materia	Profesor	Bibliografía
Base de Datos	Sánchez	Conceptos de Base de Datos
Base de Datos	Sánchez	Fundamentos de Base de Datos
Base de Datos	Hidalgo	Conceptos de Base de Datos
Base de Datos	Hidalgo	Fundamentos de Base de Datos
Base de Datos Avanzadas	Sánchez	Diseño Avanzado de Bases de Datos
Base de Datos Avanzadas	Martín	Diseño Avanzado de Bases de Datos

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

5

### Problemas presentados

- Hasta ahora se ha estudiado un tipo especial de restricciones semánticas sobre los datos, las dependencias funcionales, que permiten diseñar esquemas en 2FN, 3FN y FNBC.
- Pero **aún en FNBC**, a veces siguen presentándose **redundancias** que provocan anomalías de actualización. La razón de ello es que existen otros tipos de dependencias entre los datos, distintas de las funcionales, que también hay que tener en cuenta en el diseño relacional.

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

3

### Atributos independientes

- No hay razón para relacionar a los profesores con los libros, son atributos independientes
- Pero como son valores multivaluados, deberán repetirse con cada profesor y cada materia en todas las combinaciones
- Genera gran redundancia
- Sin embargo, no hay violación de FNBC, ya que:
  - No hay dependencias funcionales no triviales
  - La clave constituye todos los atributos del esquema

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

6

## Problemas presentados

- Agregar un nuevo profesor, significa que deberá agregarse MAS de una tupla: anomalía de inserción

Materia	Profesor	Bibliografía
Base de Datos	Sánchez	Conceptos de Base de Datos
Base de Datos	Sánchez	Fundamentos de Base de Datos
Base de Datos	Hidalgo	Conceptos de Base de Datos
Base de Datos	Hidalgo	Fundamentos de Base de Datos
Base de Datos Avanzadas	Sánchez	Diseño Avanzado de Bases de Datos
Base de Datos Avanzadas	Martín	Diseño Avanzado de Bases de Datos

## Problemas presentados

- La solución es presentarlo en dos subesquemas, con los siguientes datos:

Materia	Profesor	Bibliografía
Base de Datos	Sánchez	Conceptos de Base de Datos
Base de Datos	Sánchez	Fundamentos de Base de Datos
Base de Datos	Hidalgo	Modelo Relacional
Base de Datos	López	Diseño Avanzado de Bases de Datos
Base de Datos Avanzadas	Sánchez	Diseño Avanzado de Bases de Datos
Base de Datos Avanzadas	Martín	Diseño Avanzado de Bases de Datos

- Dos esquemas que se encuentran en una forma normal más restrictiva que FNBC: 4ta. Forma Normal

## Problemas presentados

- Agregar un nuevo profesor, significa que deberá agregarse MAS de una tupla.

Materia	Profesor	Bibliografía
Base de Datos	Sánchez	Conceptos de Base de Datos
Base de Datos	Sánchez	Fundamentos de Base de Datos
Base de Datos	Hidalgo	Conceptos de Base de Datos
Base de Datos	Hidalgo	Fundamentos de Base de Datos
Base de Datos	López	Conceptos de Base de Datos
Base de Datos	López	Fundamentos de Base de Datos
Base de Datos Avanzadas	Sánchez	Diseño Avanzado de Bases de Datos
Base de Datos Avanzadas	Martín	Diseño Avanzado de Bases de Datos

## Definición de Dep. Multivaluada

- Dada la relación  $R(A)$ , se cumple la dependencia multivaluada  $X \twoheadrightarrow Y$  (leer  $X$  "multidetermina" a  $Y$ ) si, para cada valor de  $X$ , hay un conjunto de cero o más valores de  $Y$ , independientemente de los valores del resto de atributos  $\{A - X - Y\}$ .
- Las dependencias multivaluadas siempre se producen por parejas:
  - Si en el esquema  $R(A)$  existe la dependencia  $X \twoheadrightarrow Y$ , al mismo tiempo habrá de cumplirse  $X \twoheadrightarrow A - X - Y$ .
  - Se representa como  $X \twoheadrightarrow Y \mid \{A - X - Y\}$

## Problemas presentados

- Agregar un nuevo libro en la primer materia, significa que deberá agregarse ahora TRES tuplas!!!

Materia	Profesor	Bibliografía
Base de Datos	Sánchez	Conceptos de Base de Datos
Base de Datos	Sánchez	Fundamentos de Base de Datos
Base de Datos	Sánchez	Modelo Relacional
Base de Datos	Hidalgo	Conceptos de Base de Datos
Base de Datos	Hidalgo	Fundamentos de Base de Datos
Base de Datos	Hidalgo	Modelo Relacional
Base de Datos	López	Conceptos de Base de Datos
Base de Datos	López	Fundamentos de Base de Datos
Base de Datos	López	Modelo Relacional
Base de Datos Avanzadas	Sánchez	Diseño Avanzado de Bases de Datos

## Definición de Dep. Multivaluada

- Formalmente: Sea  $R$  una esquema de Relación y  $X \twoheadrightarrow Y$  y  $X \twoheadrightarrow R$ , la dependencia multivaluada  $X \twoheadrightarrow Y$  vale en cualquier relación  $r(R)$ , si para todo par de tuplas  $t_1, t_2 \in r$  tales que  $t_1[X] = t_2[X]$ , existe otro par de tuplas  $t_3, t_4 \in r$  tal que
 
$$t_1[X] = t_2[X] = t_3[X] = t_4[X]$$

$$t_1[Y] = t_3[Y]$$

$$t_2[Y] = t_4[Y]$$

$$t_1[R - X - Y] = t_3[R - X - Y]$$

$$t_2[Y] = t_4[Y]$$

$$t_1[R - X - Y] = t_4[R - X - Y]$$

## Gráficamente

$t_1$	A1 ... Ai	Ai+1 ... Aj	Aj+1 ... An
$t_2$	A1 ... Ai	Bi+1 ... Bj	Bj+1 ... Bn
$t_3$	A1 ... Ai	Ai+1 ... Aj	Bj+1 ... Bn
$t_4$	A1 ... Ai	Bi+1 ... Bj	Aj+1 ... An

- La dependencia multivaluada asegura que  $t_3$  existe
- La existencia de  $t_4$  se implica del carácter simétrico de la definición

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

13

## Uso de las DMV

- Las dependencias multivaluadas se utilizan de dos maneras:
  - Para controlar las relaciones para determinar si alcanzaron una forma “deseable” respecto a un conjunto de dependencias funcionales o multivaluadas
  - Para especificar las restricciones de un conjunto de relacionales buscado.
- Sólo nos interesan las relaciones que satisfagan un conjunto de dependencias funcionales y multivaluadas.

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

16

## Definición de Dep. Multivaluada

- En nuestro ejemplo del esquema  
*Clases (materia, profesor, bibliografía)*
- Se definen las siguientes dependencias  
*Materia ? ? profesor*  
*Materia ? ? bibliografía*
- Esta definición supone la formalización de la noción de que dado un valor particular de Y (Materia), tiene asociado un conjunto de valores Z (profesor) y un conjunto de valores W (bibliografía), tales conjuntos de valores independiente uno con el otro

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

14

## Uso de las DMV

- Si el esquema de relación tratado no satisface las dependencias multivaluadas, podríamos construir una relación  $r'$  agregando las correspondientes tuplas.
- O podríamos descomponer el esquema de relación  $r$  en nuevos subesquemas que respeten las restricciones deseables.
  - Discutir conveniencia de las aproximaciones

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

17

## Más de Dependencias MV

- Una dependencia multivaluada  $X ? ? Y$  es trivial en  $R$  si

$$Y ? ? X \\ \text{ó } X ? ? Y = R.$$

- Las dependencias no triviales necesitan al menos de 3 atributos
- Notar además:
  - Si  $Y ? ? Z$  entonces  $Y ? ? Z$
  - Podríamos decir que los valores de  $z$  son iguales

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

15

## Propiedades de las DMV

- Se aplica el símbolo  $?$  para indicar la implicación de dependencias (funcionales o multivaluadas)
- Estudiando las reglas de inferencia, valen  
 $X ? X ? X ? ? X$   
 $X ? ? X, X ? ? Y ? X ? ? (Y - X)$
- Si  $X ? Y = ?$ , afirmamos que  
 $X ? ? Y$  ssi  $?_{Y(?_{X=X}(r))} = ?_{Y(?_{X=X} ? Z=Z(r))}$
- Para todo valor  $xz$  de  $XZ$ , en toda ocurrencia  $r$  de  $R$ .

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

18

## Axiomas para DMV

- Si  $D$  nota un conjunto de dependencias funcionales y multivaluadas, entonces  $D^+$  es la clausura de  $D$ , constituida por el conjunto de dependencias funcionales que lógicamente pueden deducirse de  $D$ .
- Para ello, utilizamos un conjunto de axiomas, que similarmente a los axiomas utilizados para las dependencias funcionales, nos permiten construir secuencias de derivación para cada uno de los componentes de  $D^+$

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

19

## Axiomas para DMV

- Hay otro conjunto de axiomas que pueden deducirse utilizando los anteriores, tales como, entre otros:
- Unión para dependencias multivaluadas:  

$$X? ? Y? X? ? Z? X? ? ZY$$
- Intersección para dependencias multivaluadas:  

$$X? ? Y? X? ? Z? X? ? Z? Y$$
- Diferencia para dependencias multivaluadas:  

$$X? ? Y? X? ? Z? \\ X? ? Z - Y? X? ? Y - Z$$

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

22

## Axiomas para DMV

- Reflexividad para dependencias funcionales:  

$$Y? X? R? X? Y$$
- Aumentación para dependencias funcionales:  

$$X? Y? U? R? XU? YU$$
- Transitividad para dependencias funcionales:  

$$X? Y? Y? Z? X? Z$$
- Complementación para dep. multivaluadas:  

$$X? Y? X? ? (R - X - Y)$$

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

20

## Uso de Axiomas

- $R = (A, B, C, G, H, I)$
- $D = \{ A? ? B, B? ? HI, CG? ? H \}$
- Algunos miembros de  $D^+$ :
- $A? ? CGHI$ .  
 – Pues  $A? ? B$ , por complemento  $A? ? R - B - A$   
 – y  $R - B - A = CGHI$
- $A? ? HI$   
 – Pues  $A? ? B$  y  $B? ? HI$ , transitiva  
 – Se deduce que  $A? ? HI - B$   
 – Pues  $HI - B = HI$ , entonces  $A? ? HI$ .

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

23

## Axioma para DMV

- Aumentación para dep. multivaluadas:  

$$X? ? Y? V? W? XW? ? YV$$
- Transitividad para dep. multivaluadas:  

$$X? ? Y? Y? ? Z? X? ? (Z - Y)$$
- Replicación:  

$$X? Y? X? ? Y$$
- Pasaje (Coalescencia)  

$$X? ? Y? Z? Y? (?W) W? Y = ? \\ ? W? Z? X? Z$$

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

21

## Uso de Axiomas

- $R = (A, B, C, G, H, I)$
- $D = \{ A? ? B, B? ? HI, CG? ? H \}$
- $B? ? H$   
 – Utilizando la regla de coalescencia,  $B? ? HI$  dada y  
 – Como  $H? ? HI$  y  $CG? ? H$  y  $CG? ? HI = ?$ ,  
 – luego se satisface la regla y concluimos que  $B? ? H$ .
- $A? ? CG$   
 –  $A? ? CGHI$  y  $A? ? HI$   
 – Por la regla de la diferencia  $A? ? CGHI - HI$   
 – Como  $CGHI - HI = CG$ ,  $A? ? CG$ .

T&D BD – DI&E – FAEA - UNCOMA

Formas Normales Avanzadas

24

## Bases de Dependencia

- Si  $R$  es el conjunto de atributos de un esquema, es posible efectuarse una partición del conjunto  $(R-X)$  en descriptores  $Y_1, Y_2, \dots, Y_k$ , tales que si  $Z \rightarrow (R-X)$ , entonces  $X \rightarrow Z$  ssi  $Z$  puede expresarse como la unión de algunos de los  $Y_i$
- $R(A_1; \dots; A_n)$  esquema de relación  
D un conjunto de dependencias multivaluadas  
 $X \rightarrow \{A_1; \dots; A_n\}$  es  $DEP_D(X) = \{Y_1; \dots; Y_n\}$  si  $\{Y_k\}$  es una partición de  $U - X$   
 $Z \rightarrow (U - X)$  vale  $X \rightarrow Y \rightarrow Z \rightarrow Y_i$  para algún  $Y_i$

## Ejemplo: Base de Dependencias II

- $DEP_M(C) = \{HA, EG, P\}$
- Se deducen las siguientes dependencias multivaluadas:
  - $C \twoheadrightarrow HA$
  - $C \twoheadrightarrow EG$
  - $C \twoheadrightarrow P$
  - $C \twoheadrightarrow HAEG$
  - $C \twoheadrightarrow HAP$
  - $C \twoheadrightarrow EGP$
  - $C \twoheadrightarrow HAEGP$  (trivial)

## Bases de Dependencia

- $R(T, M)$  y sea  $U$  conjunto de descriptores  $Z \rightarrow T$ , tal que  $W \rightarrow Y \rightarrow M, y$   
 $W \rightarrow X, y Z = Y - X$  o  $Z = T - (X \rightarrow Y)$
- Para cada  $Z_i, Z_j \rightarrow U, Z_i \rightarrow Z_j \rightarrow ?$   
Reemplazar ambos con  $Z_i \rightarrow Z_j, Z_i - Z_j$  y  $Z_j - Z_i$
- Para cada  $V \rightarrow W \rightarrow M$ ,  
si  $Y \rightarrow U, Y \rightarrow W \rightarrow ? \rightarrow Y \rightarrow V = ?$   
Reemplazar  $Y$  por  $Y \rightarrow W$  y  $Y - W$
- $U$  es base de dependencias de  $X$  en  $M$ .

## Cuarta Forma Normal

- Forma normal más estricta que FNBC ya que elimina todas las anomalías y redundancia presentes indicadas por las dependencias multivaluadas
- Un esquema de relación está en 4 Forma Normal si para toda dependencia no trivial válida en el esquema su implicante siempre es clave.
- Es decir,  $R$  está en 4FN, si cuando existe  $X \rightarrow Y$  no trivial en  $F^+$  para  $R$ ,  $X$  es una superclave de  $R$ .

## Ejemplo: Base de Dependencias I

- Sea el esquema de relación CTHRSRG,
  - $C$  = Curso,  $P$  = Profesor,  $H$  = Hora,  $A$  = Aula,
  - $E$  = Estudiante,  $G$  = Grado.
- $F$  conjunto de dependencias  
 $F = \{C \twoheadrightarrow P, HA \twoheadrightarrow C, HP \twoheadrightarrow A, CE \twoheadrightarrow G, HE \twoheadrightarrow A, C \twoheadrightarrow HA\}$
- $U^{(0)} = \{HA, P, EGP, EGHA\}$
- $U^{(1)} = \{HA, EG, P, EGP\}$
- $U^{(2)} = \{HA, EG, P, P\}$
- $U^{(3)} = \{HA, EG, P\}$  conjunto base de dependencia

## Algoritmo de Descomposición

1. Encontrar la clausura del conjunto de dependencias
2. Repetir hasta que no haya subesquema que no estén en 4FN
  - Si  $X \rightarrow Y$  viola 4FN en  $R$ ,
    - Descomponer a  $R$  como  $R_1(X, Y)$  y  $R_2(X, R - Y)$
    - Proyectar las dependencias en cada subesquema

## Algoritmo 4FN



T&D BD - DI&E - FAEA - UNCOMA

Formas Normales Avanzadas

31

## Preservación de Dependencias

- La descomposición  $R_1, R_2, \dots, R_n$  preserva dependencias con respecto a  $D$ , si para todas las relaciones
- $r(R_1), r(R_2), \dots, r(R_n)$ , tales que  $\exists r_i$  satisface  $D_i$ ,
- existe una relación  $r(R)$  que satisface  $D$  y para lo cual  $\exists r_i, r_i = \pi_{R_i}(r)$
- Las descomposiciones que están en 4FN pueden no preservar dependencias (entre las dependencias funcionales o multivaluadas)

T&D BD - DI&E - FAEA - UNCOMA

Formas Normales Avanzadas

34

## Descomposición LJ en 4FN

- Teorema:**
- Sea  $R$  un esquema de relación,  $(R_1; R_2)$  una descomposición de  $R$  y  $D$  un conjunto de dependencias funcionales y multivaluadas.
- Entonces la descomposición tiene la propiedad join sin pérdida (LJ) con respecto a  $D$  si y solo si
 
$$(R_1 \bowtie R_2) \bowtie \dots \bowtie R_n = R$$

$$(R_1 \bowtie R_2) \bowtie \dots \bowtie R_n = R$$

T&D BD - DI&E - FAEA - UNCOMA

Formas Normales Avanzadas

32

## Problemas del Algoritmo

- Encontrar un cubrimiento
- Proyectar las dependencias en cada subesquema
- Encontrar las dependencias embebidas
  - Tipo especial de DM que no existen en un cierto esquema de relación pero que aparecen en una proyección de dicha relación. También se llaman jerárquicas.
  - Una relación  $R(A)$  satisface una DM embebida  $X \twoheadrightarrow Y|Z$ , con  $\{X \twoheadrightarrow Y \twoheadrightarrow Z\} \subseteq A$ , si dicha DM no existe en  $R$  y se cumple para cualquier extensión de un esquema  $R'(X \twoheadrightarrow Y \twoheadrightarrow Z)$ , con  $R'$  una proyección de  $R$  sobre los atributos  $\{X \twoheadrightarrow Y \twoheadrightarrow Z\}$  involucrados en la dependencia embebida.

T&D BD - DI&E - FAEA - UNCOMA

Formas Normales Avanzadas

35

## Preservación de Dependencias

- Sea  $R_1, R_2, \dots, R_n$  una descomposición de  $R$ , y  $D$  un conjunto de dependencias válidas en  $R$  (multivaluadas y funcionales)
- La restricciones de  $D$  en  $R_i$  es el conjunto  $D_i$ , que consiste de:
  - Todas las dependencias funcionales de  $D^+$  que incluyen los atributos de  $R_i$
  - Todas las dependencias multivaluadas de la forma  $W \twoheadrightarrow V \twoheadrightarrow R_i$ , con  $W$  en  $R_i$  y  $W \twoheadrightarrow V$  perteneciente a  $D^+$ .

T&D BD - DI&E - FAEA - UNCOMA

Formas Normales Avanzadas

33