

Tema 6: Diseño de Bases de Datos Relacionales (Parte 2)

Bases de Datos



Indice

- 6. Forma Normal de Boyce-Codd
- 7. Tercera forma normal
- 8. Cuarta Forma Normal
- 9. Otras Formas Normales

Forma Normal de Boyce-Codd

- A través de las dependencias funcionales se pueden definir varias formas normales que definan diseños adecuados de bases de datos.
- Una de las formas normales más deseables es la **Forma Normal de Boyce-Codd (FNBC)**.

Forma Normal de Boyce-Codd

- Se dice que un esquema R está en FNBC respecto a un conjunto de dependencias funcionales F si, para todas las dependencias funcionales F^+ de la forma $\alpha \rightarrow \beta$, donde $\alpha \subseteq R$ y $\beta \subseteq R$ se cumple al menos una de las siguientes condiciones:
 - $\alpha \rightarrow \beta$ es una dependencia funcional trivial (es decir, $\beta \subseteq \alpha$)
 - α es una superclave del esquema R
- Un diseño de BBDD está en FNBC si cada relación lo está.

Forma Normal de Boyce-Codd

- Ejemplo: dados los siguientes esquemas y sus correspondientes dependencias funcionales

Esquema-cliente = (nombre-cliente, calle-cliente, ciudad-cliente)
Nombre-cliente → calle-cliente ciudad-cliente

Esquema-sucursal = (nombre-sucursal, activo, ciudad-sucursal)
Nombre-sucursal → activo ciudad-sucursal

Esquema-info-préstamo = (nombre-sucursal, nombre-cliente, número-préstamo, importe)
Numero-prestamo → importe nombre-sucursal

Forma Normal de Boyce-Codd

- Se puede comprobar que *Esquema-cliente* está en FNBC.
 - Obsérvese que *nombre-cliente* es una clave candidata
 - Obsérvese que todas las dependencias funcionales no triviales tienen a *nombre-cliente* a la izquierda de la flecha
- Igualmente *Esquema-sucursal* está en FNBC.
- Sin embargo *Esquema-info-préstamo* no está en FNBC
 - obsérvese que *Número-préstamo* no es una superclave ya que puede haber un préstamo concedido a dos personas
 - Sin embargo la dependencia *Numero-préstamo* → *importe* es no trivial.

Forma Normal de Boyce-Codd

- Por otra parte *Esquema-info-préstamo* sufre el problema de la repetición de información, si hay varios clientes asociados a un préstamo hay que repetir el nombre de la sucursal y el importe por cada cliente.
- Se puede eliminar esta redundancia rediseñando la base de datos de forma que todos los esquemas esten en FNBC.
- Esto lo podemos hacer descomponiendo las relaciones que no están en FNBC.

Forma Normal de Boyce-Codd

- Podríamos descomponer Esquema-info-préstamo en dos esquemas:

Esquema-préstamo=(número-préstamo, nombre-sucursal, importe)

Esquema-prestatario=(nombre-cliente, número-préstamo)

- Esta es una descomposición de reunión sin pérdida.

Forma Normal de Boyce-Codd

- Las dependencias funcionales de cada una de estas relaciones son:

- Para *Esquema-préstamo* tenemos:

número-préstamo → *importe nombre-sucursal*

que es una superclave de *Esquema-préstamo*

- Para *Esquema-prestatario* sólo hay dependencias funcionales triviales

- Así pues ambas están en FNBC.

Forma Normal de Boyce-Codd

- Además se elimina la redundancia:
 - En la relación *Esquema-préstamo* hay exactamente una tupla para cada préstamo
 - En la relación *Esquema-prestatario* hay una sólo tupla para cada cliente de cada préstamo
- Por tanto no hay que repetir el nombre de la sucursal y el importe una vez por cada importe asociado a un préstamo

Algoritmo de descomposición

- Si R no está en FNBC se puede descomponer en un conjunto de esquemas en FNBC R_1, R_2, \dots, R_n utilizando el siguiente algoritmo que utiliza las dependencias (“testigos”) que demuestran la violación de la FNBC para llevar a cabo la descomposición.

```
resultado := {R};  
hecho := falso;  
calcular  $F^+$ ;  
while (not hecho) do  
  if (hay un esquema  $R_i$  de resultado que no esté en FNBC)  
  then begin  
    sea  $\alpha \rightarrow \beta$  una dependencia funcional no trivial  
    que se cumple en  $R_i$  tal que  $\alpha \rightarrow \beta$  no esté en  
     $F^+$  y  $\alpha \cap \beta = \emptyset$ ;  
    resultado := (resultado -  $R_i$ )  $\cup$  ( $R_i - \beta$ )  $\cup$  ( $\alpha, \beta$ );  
  end  
else hecho := cierto;
```

Forma Normal de Boyce-Codd

- No todas las descomposiciones FNBC conservan las dependencias.
- Puesto que los objetivos de diseño son tres:
 1. Reunión sin pérdida
 2. FNBC
 3. Conservación de las dependenciasy puesto que 1. es una condición esencial para la descomposición, para evitar pérdidas de información, entonces a veces habrá que abandonar la FNBC o el requisito de conservación de las dependencias.
- La **tercera forma normal** es una relajación de la FNBC y permite obtener siempre descomposiciones que conservan las dependencias.

Tercera Forma Normal

- En los casos en que la descomposición en FNBC no conserva las dependencias se pueden hacer dos cosas:
 - Para cada actualización soportar el coste computacional del cálculo de las reuniones para buscar violaciones
 - Emplear una descomposición alternativa (la 3FN) que, a diferencia de la FNBC, puede tener alguna redundancia.
- La FNBC exige que todas las dependencias funcionales no triviales sean de la forma $\alpha \rightarrow \beta$ donde α es una superclave, la 3FN relaja la restricción permitiendo dependencias funcionales no triviales cuyo lado de la izquierda no sea una superclave.

Tercera Forma Normal

- Un esquema de una relación R está en 3FN respecto de un conjunto de dependencias funcionales F si, para todas las dependencias funcionales F^+ de la forma $\alpha \rightarrow \beta$ donde $\alpha \subseteq R$ y $\beta \subseteq R$ se cumple una de las tres condiciones:
 - $\alpha \rightarrow \beta$ es una dependencia funcional trivial
 - α es una superclave de R
 - Cada atributo A de $\beta - \alpha$ está contenido en alguna clave candidata de R
- Obsérvese que la última condición no indica que una sola clave candidata deba contener todos los atributos de $\alpha \rightarrow \beta$, cada atributo puede estar contenido en una clave candidata diferente.

Tercera Forma Normal

- Obsérvese que si un esquema satisface FNBC también satisface 3FN, pero no al revés.
- Al igual que en el caso de FNBC existe un algoritmo (**algoritmo de síntesis de 3FN**) que permite descomponer, con reuniones sin pérdida y conservando las dependencias funcionales

Algoritmo de síntesis de 3FN

```
sea  $F_c$  un recubrimiento canónico de  $F$ ;  
 $i := 0$ ;  
for each dependencia funcional  $\alpha \rightarrow \beta$  de  $F_c$  do  
  if ninguno de los esquemas  $R_j, j = 1, 2, \dots, i$  contiene  $\alpha\beta$   
  then begin  
     $i := i + 1$ ;  
     $R_i := \alpha\beta$ ;  
  end  
if ninguno de los esquemas  $R_j, j = 1, 2, \dots, i$  contiene una clave  
candidata de  $R$   
then begin  
   $i := i + 1$ ;  
   $R_i :=$  cualquier clave candidata de  $R$ ;  
end  
return  $(R_1, R_2, \dots, R_i)$ 
```


Comparación entre 3FN y FNBC

- Ventajas 3FN:

- conserva las dependencias funcionales

- Inconvenientes 3FN:

- si no se eliminan las dependencias transitivas de las relaciones puede que se tengan que emplear valores nulos para representar algunas de las relaciones
- Repetición de la información

Cuarta Forma Normal

- Como hemos visto, el problema de repetición de información se produce en las anteriores normalizaciones.
- Para atacar este problema hay que introducir el concepto de dependencias multivaloradas.
- Hemos visto que las dependencias funcionales impiden que ciertas tuplas estén en una relación, si $\alpha \rightarrow \beta$ entonces no puede haber dos tuplas con el mismo valor de α y distinto valor de β .
- Las dependencias multivaloradas no impiden esto sino que exigen que estén presentes otras tuplas de una cierta forma.

Cuarta Forma Normal

• Sea R un esquema de una relación y sean $\alpha \subseteq R$ y $\beta \subseteq R$. La **dependencia multivalorada** $\alpha \twoheadrightarrow \beta$ se cumple en R si en toda relación legal $r(R)$, para todo par de tuplas t_1 y t_2 tales que $t_1(\alpha) = t_2(\alpha)$, existen unas tuplas t_3 y t_4 tales que:

$$t_1[\alpha] = t_2[\alpha] = t_3[\alpha] = t_4[\alpha]$$

$$t_3[\beta] = t_1[\beta]$$

$$t_3[R - \beta] = t_2[R - \beta]$$

$$t_4[\beta] = t_2[\beta]$$

$$t_4[R - \beta] = t_1[R - \beta]$$

Cuarta Forma Normal

- Intuitivamente, la dependencia multivalorada $\alpha \twoheadrightarrow \beta$ indica que la relación entre α y β es independiente de la relación entre α y $R - \beta$.
- Se dice que la dependencia $\alpha \twoheadrightarrow \beta$ es **trivial** en el esquema R si todas las relaciones del esquema la satisfacen.

Dependencias Multivaloradas

- Ejemplos:

Cod_Curso	Texto	Idioma
C1	Introducción a las BBDD	Español
C1	Introducción a las BBDD	Inglés
C1	Modelo Relacional	Español
C2	Diseño de BBDD	Francés
C2	Modelo Relacional	Español

Asignatura	Profesor	Texto
Ficheros y BD	López	Diseño de BBDD
Ficheros y BD	López	Fundamentos BD
Ficheros y BD	Pérez	Diseño de BBDD
Ficheros y BD	Pérez	Fundamentos BD
BD Avanzadas	López	Diseño Avanzado
BD Avanzadas	Pérez	Diseño Avanzado

CodCurso \twoheadrightarrow Texto
CodCurso \twoheadrightarrow Idioma

Asignatura \twoheadrightarrow Texto
Asignatura \twoheadrightarrow Profesor

Dependencias Multivaloradas

- Comprobación Asignatura $\rightarrow \rightarrow$ Profesor:

$$t_1[\alpha] = t_2[\alpha] = t_3[\alpha] = t_4[\alpha]$$

$$t_3[\beta] = t_1[\beta]$$

$$t_3[R-\beta] = t_2[R-\beta]$$

$$t_4[\beta] = t_2[\beta]$$

$$t_4[R-\beta] = t_1[R-\beta]$$

Tupla	(α)Asignatura	(β)Profesor	Texto
t_1	Ficheros y BD	López	Diseño de BBDD
t_3	Ficheros y BD	López	Fundamentos BD
t_4	Ficheros y BD	Pérez	Diseño de BBDD
t_2	Ficheros y BD	Pérez	Fundamentos BD
	BD Avanzadas	López	Diseño Avanzado
	BD Avanzadas	Pérez	Diseño Avanzado

Resumen de las Formas Normales

- Las Formas Normales vistas hasta ahora son generadas a partir de las Dependencias Funcionales
- Con estas Formas Normales se llega a un grado suficiente en el diseño de las bases de datos:
 - La 1FN es obligada en toda base de datos.
 - La FNBC es la deseada en el diseño de base de datos, pero puede no ser aconsejable en determinadas ocasiones, por su pérdida de dependencias funcionales.
 - La 3FN se considera como satisfactoria en el diseño de bases de datos, ya que conserva la información y las dependencias funcionales.

4ª Forma Normal

- La 4FN es una especialización de la forma normal Boyce-Codd.
- Una relación R está en 4FN con respecto a un conjunto de dependencias funcionales multivaloradas F si para toda dependencia multivalorada en F^+ de la forma $\alpha \twoheadrightarrow \beta$ con $\alpha \subseteq R$ y $\beta \subseteq R$, se cumple, como mínimo, una de las condiciones siguientes:
 - $\alpha \twoheadrightarrow \beta$ es una dependencia multivalorada trivial
 - α es una superclave en el esquema R

4ª Forma Normal

- Obsérvese que la definición es análoga a definición de FNBC excepto que en 4FN se emplean dependencias multivaloradas.
- Por tanto un esquema que esté en 4FN está en FNBC.
- Igual que antes, existe un algoritmo de descomposición idéntico al de la FNBC:

Algoritmo de descomposición

```
resultado := {R};  
hecho := falso;  
calcular  $F^+$ ; Dado el esquema  $R_i$ , supongamos que  $F_i$  denota la  
restricción de  $F^+$  a  $R_i$   
while (not hecho) do  
    if (hay un esquema  $R_i$  en resultado que no se halla en 4FN  
    con respecto a  $F_i$ )  
        then begin  
            supongamos que  $\alpha \twoheadrightarrow \beta$  es una dependencia  
            multivalorada no trivial que se cumple en  $R_i$  tal  
            que  $\alpha \rightarrow R_i$  no se halla en  $F_i$  y  $\alpha \cap \beta = \emptyset$ ;  
            resultado := (resultado -  $R_i$ )  $\cup$  ( $R_i - \beta$ )  $\cup$  ( $\alpha, \beta$ );  
        and  
        else hecho := verdadero;
```

Otras Formas Normales

- Hay otras formas normales, además de las anteriores, que tampoco se pueden considerar definitivas.
- Entre ellas está la 2FN se basa en el concepto de dependencia plena.
- Se dice que una relación está en 2FN si:
 - Está en 1FN
 - Cada atributo no principal tiene dependencia funcional completa respecto de una de las claves de la relación.
- La 2FN no se cumple cuando algún atributo no principal depende funcionalmente de algún subconjunto de la clave.

Otras Formas Normales

- Otra también muy conocida es la 5FN también llamada Forma Normal de Reunión por Proyección (FNRP).
- Este tipo de normalizaciones se emplea raramente, ya que no hay un conjunto de reglas de inferencia seguras y completas que permitan razonar con las restricciones.