

# CC42A - AUXILIAR 5

## NORMALIZACIÓN

Prof. Claudio Gutierrez  
Prof. aux. Karina Figueroa

5 de mayo de 2003

### 1. Introducción

El objetivo de la normalización consiste en eliminar anomalías, como redundancia, anomalías de actualización y anomalías de eliminación.

### 2. Forma Normal de Boyce Codd (FNBC)

Esta normalización garantiza que no se presenten anomalías.

*Una relación  $R$  esta en forma normal de este tipo si: siempre que haya una dependencia no trivial  $A_1A_2 \dots A_n \rightarrow B$  de  $R$ , sucede que  $\{A_1, A_2, \dots A_n\}$  es una super llave de  $R$ .*

En otras palabras, el lado izquierdo de toda dependencia funcional no trivial ha de ser una superllave (la superllave no necesariamente debe ser mínima).

#### 2.1. Estrategia de descomposición

La estrategia de descomposición basada en una violación de la FNBC que aplicaremos consiste en:

- Buscar la(s) dependencia funcional que viole la FNBC, es decir, que  $\{A_1, A_2, \dots, A_n\}$  no sea una superllave
- Agregar al lado derecho tantos atributos como están funcionalmente determinados por  $\{A_1, A_2, \dots, A_n\}$

El algoritmo para realizar la descomposición a partir de las dependencias que violan la FNBC, será entonces:

- Sea  $R$  una relación
- Sea  $C \rightarrow D_1 D_2 \dots D_n$  la dependencia funcional que viola la FNBC
- Sea  $XY$  la llave de la relación  $R$
- paso 1.- generar una relación  $R_1$  con los atributos de  $C, D_1, D_2, \dots, D_n$
- paso 2.- generar una relación  $R_2$  con los atributos  $R - R_1$
- paso 3.- Revisar las nuevas dependencias funcionales formadas y si alguna viola la FNBC regresar al paso 1.

IMPORTANTE.- No siempre es posible realizar una normalización de este tipo.

## 2.2. Ejemplo 1

Considere la relación mostrada en la tabla 2.2. Las dependencias funcionales de esta relación son:

- título, año  $\rightarrow$  tamaño, tipo, Estudio
- Estudio  $\rightarrow$  Dirección

título	año	duración	Tipo	Estudio	Dirección
Star Wars	1977	124	color	Fox	Hollywood
Mighty Ducks	1991	104	color	Disney	Buena Vista
Wayne's World	1992	95	color	Paramount	Hollywood

Cuadro 1: Relación Película

### 2.2.1. Solución

De las dependencias funcionales dadas, es claro que título y año son las llaves de esta relación y la dependencia  $Estudio \rightarrow Direccion$  viola la FNBC, puesto que Estudio no es una superllave.

Siguiendo el algoritmo descrito previamente, generamos una relación  $R_1$  (Estudio, Dirección) y generamos una segunda relación con  $R_2$  (título, año, duración, tipo, estudio).

Como no existen nuevas dependencias funcionales, ni hay dependencias funcionales que aún se sigan violando, las relaciones  $R_1$  y  $R_2$  están en FNBC.

## 2.3. Ejemplo 2

Considere una relación  $R(\text{titulo}, \text{año}, \text{estudio}, \text{presidente}, \text{DirecPresidente})$ , las dependencias funcionales de esta relación son:

1.  $\text{titulo}, \text{año} \rightarrow \text{estudio}$
2.  $\text{estudio} \rightarrow \text{presidente}$
3.  $\text{presidente} \rightarrow \text{DirecPresidente}$

### 2.3.1. Solución

La única llave de esta relación es título y año, por lo tanto las dependencias funcionales 2 y 3 violan la FNBC.

Siguiendo la estrategia de descomposición, agregamos los atributos determinados funcionalmente por *estudio*, y con esto tenemos:  $\text{estudio} \rightarrow \text{presidente}$ ,  $\text{DirecPresidente}$ .

Siguiendo el algoritmo descrito, tenemos que  $R_1$ (estudio, presidente, DirecPresidente), y  $R_2$ (titulo, año, estudio).

En la relación  $R_1$  aún existe una dependencia funcional(3) que viola la FNBC, por lo tanto se debe repetir el procedimiento, por lo tanto obtenemos nuevas relaciones  $R_{1a}$ (estudio, presidente) y  $R_{1b}$ (presidente, DirecPresidente).

Estas nuevas relaciones  $R_1, R_{1a}$  y  $R_{1b}$  están en la FNBC.

### 3. Tercera Forma normal (3FN)

En ocasiones se tiene un esquema relacional que sus dependencias funcionales no están en FNBC, pero no se desea descomponerla mas. Esta relación puede estar en la 3FN.

Una relación  $R$  está en la tercera forma normal (3FN) si: siempre que  $A_1A_2 \dots A_n \rightarrow B$  es una dependencia no trivial,  $\{A_1A_2 \dots A_n\}$  será una superllave o  $B$  sea miembro de alguna llave

Note la diferencia entre la 3FN y la FNBC, es el enunciado “o B será un miembro de alguna llave”.

El algoritmo para la descomposición en esta forma, es similar al algoritmo de la FNBC.

#### 3.1. Ejemplo 1

En cada uno de los siguientes esquemas relacionales y conjuntos de dependencias funcionales:

1.  $R(A,B,C,D)$  con las dependencias funcionales  $AB \rightarrow C$ ,  $C \rightarrow D$  y  $D \rightarrow A$ .
2.  $R(A,B,C,D)$  con las dependencias funcionales  $B \rightarrow C$  y  $B \rightarrow D$ .

Indique todas la violaciones de la FNBC, considerando las dependencias que no estén incluidas en el conjunto, pero se derivan de ellas. Si es posible descomponga en FNBC y si no es posible descompongalas en 3FN.

### 3.1.1. Solución 1

Para poder conocer que dependencias funcionales violan la FNBC, es necesario encontrar las llaves, esto se puede hacer a partir del cálculo de las cerraduras de los atributos.

Tomando las cerraduras de todos los atributos, vemos cuales son llaves.

$$AB^+ = ABCD \quad BC^+ = ABCD \quad BD^+ = ABCD \quad ABC^+ = ABCD \quad ABD^+ = ABCD \quad BCD^+ = ABCD.$$

De estas dependencias funcionales, podemos eliminar aquellas cuyo lado izquierdo sea representado por otra cerradura si le sacamos alguno de sus atributos, en este caso, las dependencias ABC, ABD y BCD se descartan quedando entonces las llaves AB, BC, y BD.

Las dependencias funcionales para este problema son:  $C \rightarrow A$ ,  $C \rightarrow D$ ,  $D \rightarrow A$ ,  $AB \rightarrow D$ ,  $AB \rightarrow C$ ,  $AC \rightarrow D$ ,  $BC \rightarrow A$ ,  $BC \rightarrow D$ ,  $BD \rightarrow A$ ,  $BD \rightarrow C$ ,  $CD \rightarrow A$ ,  $ABC \rightarrow D$ ,  $ABD \rightarrow C$ ,  $BCD \rightarrow A$ .

Teniendo las dependencias funcionales y las llaves, veamos ahora las violaciones a la FNBC, si alguna dependencia no tiene alguno de los pares (llaves) en el lado izquierdo es una violación a la FNBC. Podemos identificar entonces las dependencias que cumplen lo anterior  $C \rightarrow A$ ,  $C \rightarrow D$ ,  $D \rightarrow A$ ,  $AC \rightarrow D$  y  $CD \rightarrow A$ .

Una alternativa es usar  $C \rightarrow D$ . Con esto obtenemos dos relaciones (C, D) y (ABC). La relación (C, D) esta en FNBC, sin embargo la relación (ABC) no esta en FNBC, por que AB y BC son llaves y la dependencia  $C \rightarrow A$  no cumple con la condición de Boyce Codd. Sin embargo no podemos descomponer la relación por que nos quedaríamos sin llaves :(

Ahora, si observamos los lados derechos de las dependencias funcionales, todo lado derecho es parte de al menos una de las llaves mínimas de ABCD, entonces la relación ya esta en 3FN y no es necesario decomponerla.

### 3.1.2. Solución 2

Para el segundo esquema la única llave es AB. Así,  $B \rightarrow C$  y  $B \rightarrow D$  son violaciones a la FNBC pues ninguna de ellas tiene la llave contenida en su lado izquierdo. Estas son violaciones no triviales a la FNBC. La razón es que las únicas dependencias derivadas no triviales deben tener A y B en el lado

izquierdo y por lo tanto contener una llave.

Una posible descomposición de la FNBC es AB y BCD, AB es la llave única para AB y B es la llave única para BCD.

A partir de que solo hay una llave para ABCD, las violaciones a la 3FN son las mismas, y la descomposición también.

## 4. Cuarta Forma normal (dependencias multivaloradas)

Una dependencia multivaluada es una proposición sobre alguna relación  $R$  que, cuando se corrigen los valores de un conjunto de atributos, los otros atributos son independientes de los valores de los atributos restantes de la relación.

$$A_1 A_2 \dots A_n \twoheadrightarrow B_1 B_2 \dots B_n$$

Afirmamos que esta dependencia multivaluada se satisface si:

En cada pareja de tuplas  $t$  y  $u$  de la relación  $R$  que concuerdan en todos los atributos  $A$ , podemos encontrar en  $R$  alguna tupla  $v$  que concuerden:

- Con  $t$  y  $u$  en los atributos  $A$
- Con  $t$  en los atributos  $B$
- Con  $u$  en los atributos de  $R$  que no figuran entre los atributos  $A$  ni  $B$ .

Nótese que se puede aplicar esta regla con  $t$  y  $u$  intercambiadas, para inferir la existencia de una cuarta tupla.

El algoritmo de descomposición para esta forma, es igual al algoritmo para la FNBC.

### 4.1. Ejemplo 1

Suponga que tiene una relación  $R(A, B, C)$  con una dependencia multivaluada  $A \twoheadrightarrow B$ . Si sabemos que las tuplas  $(a, b_1, c_1)$ ,  $(a, b_2, c_2)$  y  $(a, b_3, c_3)$  están en la instancia actual de  $R$ . Encuentre otras tuplas de esta relación.

#### 4.1.1. Solución

Sea  $t = (a, b_1, c_1)$ ,  $u = (a, b_2, c_2)$ , y  $v = (a, b_3, c_3)$ .

De la definición de dependencias funcionales tenemos que existen las tuplas

- de  $t$  y  $u$  tenemos  $(a, b_1, c_2)$  y  $(a, b_2, c_1)$
- de  $t$  y  $v$  tenemos  $(a, b_1, c_3)$  y  $(a, b_3, c_1)$
- de  $u$  y  $v$  tenemos  $(a, b_2, c_3)$  y  $(a, b_3, c_2)$

#### 4.2. Ejemplo 2

Sea  $R(A, B, C, D)$  con las dependencias multivaluadas  $A \twoheadrightarrow B$  y  $A \twoheadrightarrow C$ , Encuentre todas las violaciones de la 4FN y descomponga las relaciones en una colección de esquemas relacionales en la 4FN.

##### 4.2.1. Solución

Puesto que no hay dependencias funcionales, la única llave contiene los cuatro atributos ABCD. Así, cada una de las dependencias multivaluadas no triviales  $A \twoheadrightarrow B$  y  $A \twoheadrightarrow C$  violan la 4FN.

Podemos descomponer los atributos de estas dependencias, primero en AB y ACD. En este caso, AB está en 4FN, pero ACD no, porque  $A \twoheadrightarrow C$  es una violación a la 4FN. Luego descomponemos ACD en AC y AD, las cuales si están en 4FN.