

Escrito por **Adrián Quiroga Linares**.

8.1 Introducción

La normalización del modelo relacional de una base de datos tiene una fuerte **base matemática** que se apoya en el concepto de dependencia funcional de atributos. El descubrimiento de dependencias funcionales entre atributos supone la representación de la semántica que se incorpora a la base de datos, se podría decir que es equivalente a descubrir conjuntos de entidades y relaciones en el MER.

Una vez representada la semántica mediante las dependencias funcionales el proceso de normalización consiste en el cumplimientos de determinadas condiciones denominadas **formas normales**.

A veces es necesario descomponer los esquemas en tablas más pequeñas siendo válidas, únicamente las descomposiciones reversibles, sin pérdidas, de forma que si se unen las tablas de vuelve a tener la original.

Conclusión: Un buen diseño de base de datos relacional evita la redundancia, asegura la consistencia, y facilita la representación precisa de datos independientemente.

La **Dependencia Funcional** es la base de la normalización. Un conjunto de atributos determina funcionalmente a otro si, para todas las tuplas de una relación, cuando los valores del primero se repiten, entonces los valores del segundo se repiten.

Las diversas formas normales van imponiendo sucesivas condiciones a las relaciones que forman la base de datos.

8.2 Primera Forma Normal

Cuando una tabla contiene únicamente atributos **atómicos**, se dice que está en la **Primera Forma Normal**. En una base de datos del MR esta primera forma es la única **obligatoria**, el resto de formas normales se limitan a mejorar el diseño evitando los problemas que podrían aparecer.

Está prohibido tener **atributos compuestos y multivalorados**.

8.3 Segunda Forma Normal (2FN)

Impide la existencia de **dependencias funcionales parciales** (*en las dependencias funcionales parciales el lado izquierdo es una parte de una superclave pero no la superclave completa*).

8.4 Tercera Forma Normal (3FN)

Impide la existencia de **dependencias funcionales transitivas** (*en las dependencias funcionales transitivas el lado izquierdo no forma parte de una superclave*).

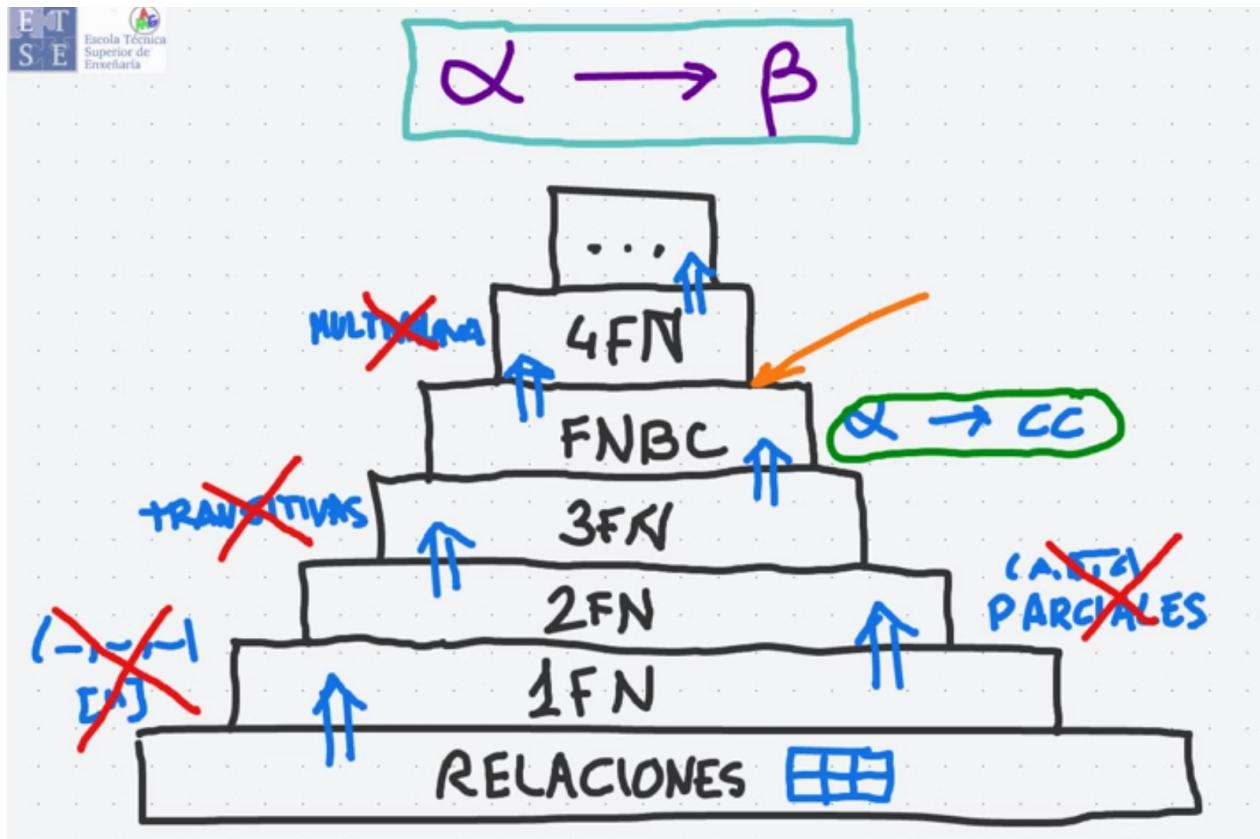


Figure 1: Archivos/imagenes/Pasted image 20241106123439.png

8.5 Forma Normal de Boyce-Codd (FNBC)

Obliga a que en todas las **dependencias funcionales del lado izquierdo** sea una superclave de la relación. Si en el **MER** tenemos un entidad débil nunca puede estar en FNBC.

Cuando se encuentra alguna relación con alguna dependencia funcional que no cumple las condiciones se **procede a descomponerla** (*habitualmente en 2*). Hay que escoger descomposiciones de reunión sin pérdidas que preserven las dependencias.

Por encima de la **FNBC** existen otras formas normales que van imponiendo condiciones cada vez más restrictivas, pero normalmente no se usan en la práctica. El objetivo habitual de la normalización es conseguir que todas las **relaciones estén en FNBC**.

8.6 Teoría de las Dependencias Funcionales

Una **dependencia funcional** entre dos conjuntos de atributos X y Y en una relación R se denota como:

$$X \rightarrow Y$$

Esto significa que si dos tuplas (*filas*) tienen los mismos valores para el conjunto de atributos X , entonces también deben tener los mismos valores para el conjunto Y . Es decir, el valor de X determina el valor de Y .

8.6.1 Cierre de un conjunto de dependencias funcionales

Cuando hablamos del **cierre de un conjunto de dependencias funcionales F** , denotado como F^+ , nos referimos al conjunto de todas las dependencias funcionales que se pueden derivar lógicamente a partir de las dependencias en F . El cierre es importante porque, cuando estamos comprobando si un esquema está en una forma normal (como la FNBC o la 3FN), no basta con mirar únicamente las dependencias que se nos dan. Debemos considerar todas las dependencias funcionales que pueden deducirse de ellas.

Ejemplo: Supongamos que tenemos el siguiente conjunto de dependencias funcionales para un esquema de relación $R(A, B, C, G, H, I)$:

$$F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$$

Queremos determinar si $A \rightarrow H$ se puede deducir a partir de F .

1. Sabemos que $A \rightarrow B$, lo que implica que si dos tuplas tienen el mismo valor para A , entonces también deben tener el mismo valor para B .
2. También sabemos que $B \rightarrow H$, lo que implica que si dos tuplas tienen el mismo valor para B , entonces deben tener el mismo valor para H .
3. Si $t_1[A] = t_2[A]$, entonces $t_1[B] = t_2[B]$ por $A \rightarrow B$.
4. Luego, $t_1[B] = t_2[B]$ implica que $t_1[H] = t_2[H]$ por $B \rightarrow H$.

Por lo tanto, hemos demostrado que $A \rightarrow H$ se cumple, lo que significa que $A \rightarrow H$ está implicado lógicamente por F .

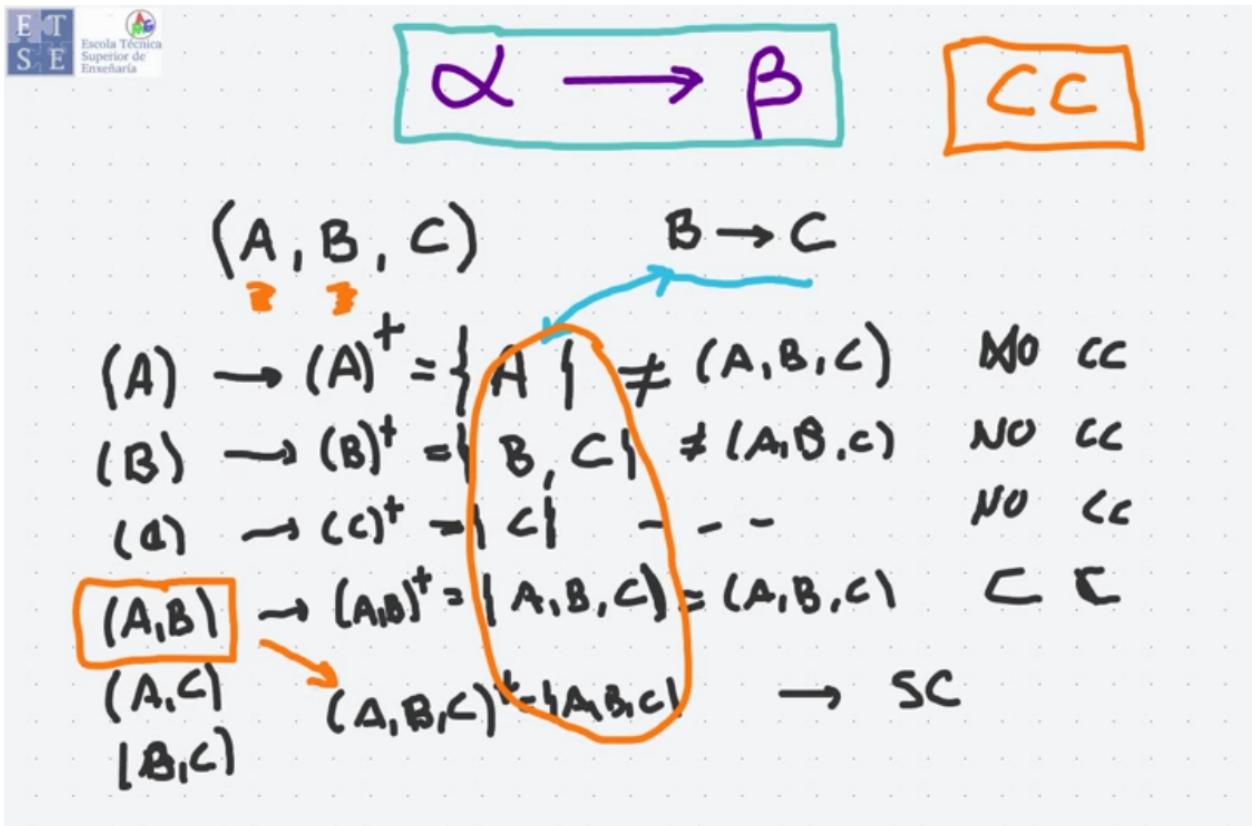


Figure 2: Archivos/imagenes/Pasted image 20241106162630.png

8.6.2 Axiomas de Armstrong

Los **axiomas de Armstrong** y sus reglas adicionales permiten descubrir nuevas dependencias funcionales a partir de un conjunto de dependencias dados.

- Reflexividad:** Si $\alpha \subseteq \beta$, entonces $\beta \rightarrow \alpha$.
- Aumentatividad:** Si $\alpha \rightarrow \beta$ y γ es un conjunto de atributos, entonces $\gamma\alpha \rightarrow \gamma\beta$.
- Transitividad:** Si $\alpha \rightarrow \beta$ y $\beta \rightarrow \gamma$, entonces $\alpha \rightarrow \gamma$.

Otras reglas útiles

Además de los axiomas de Armstrong, existen reglas adicionales que se usan para derivar dependencias funcionales más fácilmente:

- Unión:** Si $\alpha \rightarrow \beta$ y $\alpha \rightarrow \gamma$, entonces $\alpha \rightarrow \beta\gamma$.
- Descomposición:** Si $\alpha \rightarrow \beta\gamma$, entonces $\alpha \rightarrow \beta$ y $\alpha \rightarrow \gamma$.
- Pseudotransitividad:** Si $\alpha \rightarrow \beta$ y $\gamma\beta \rightarrow \delta$, entonces $\alpha\gamma \rightarrow \delta$.

8.7 Recubrimiento Canónico

El **recubrimiento canónico** es el conjunto mínimo equivalente de dependencias funcionales. Dos conjuntos son equivalentes si sus cierres son iguales.

8.8 Pasos para la normalización

Paso 1: Análisis semántico para obtener la relación universal U de atributos y el conjunto de dependencias funcionales del problema

Paso 2: Primera Forma Normal

Paso 3: Obtención del recubrimiento canónico del conjunto de dependencias funcionales

Paso 3.1: Aplicar la regla de la descomposición de Armstrong

Paso 3.2: Eliminar dependencias funcionales implicadas lógicamente

Paso 3.3: Aplicar la regla de la unión de Armstrong

Cálculo de Claves Candidatas de una relación

¿Está U en Forma Normal de Boyce-Codd?

Paso 4: Descomposición por localización de atributos equivalentes

Paso 5: Descomposición por eliminación de dependencias funcionales que no permiten cumplir las condiciones de Forma Normal de Boyce-Codd

Figure 3: Archivos/imagenes/Pasted image 20241224121236.png

TIPs: Está en Boyce-Codd si todos los determinantes del recubrimiento canónico ($a \rightarrow B$, a es el determinante) forman parte de la clave candidata calculada con el cierre. **Si no están en Boyce-Codd**, localizamos atributos equivalentes ($a \rightarrow B$ y $B \rightarrow a$), si pasa esto se llevan todos los atributos equivalentes a una nueva relación y en la relación universal (U) se deja solo uno de esos atributos. **Después**, si sigue sin cumplir alguna relación $a \rightarrow B$ se crea una nueva tabla con los atributos a, B y la tabla original se elimina B .

NOR CASO 03

Universidad

En esta Universidad [U], un profesor [P] se identifica por un código de profesor [CP] y, además, ocurre la curiosa circunstancia de que todos los profesores [PS] tienen nombres [NP] distintos.

Cada asignatura que se imparte, que tiene un nombre [A], tiene un único profesor como responsable [PR], si bien un mismo profesor puede ser responsable de más de una asignatura [PA]. Las asignaturas se dividen en uno o más grupos [G]. Todo estudiante [E], en cada asignatura, pertenece a un único grupo [GU].

Cada profesor [PD] depende siempre y únicamente de un departamento [D]. Así mismo, toda asignatura está ligada a un único departamento [AD], el del profesor responsable de la misma.

Si ponen este tremendo cigarro, para sacar las dependencias funcionales, pensar las multiplicidades que se darían. Si son del tipo 1 a 1 o N a N se determinan mutuamente, si son 1..N o N..1 el lado N determina al lado 1.

[!Importante] **Normalización:** Técnica para producir un conjunto de relaciones con una serie de propiedades deseables a partir de unos requisitos de datos.

1FN: Relación en la que la intersección de toda fila y columna contiene un valor atómico

Dependencia Funcional: Describe la relación existente entre atributos de una relación. Si A y B son conjuntos de atributos de la relación R, B será funcionalmente dependiente de A ($A \rightarrow B$) si cada valor de A está asociado con exactamente un valor de B.

Determinante: Grupo de atributos del lado izquierdo de una Dependencia Funcional.

Dependencia Funcional Completa: Si A y B son conjuntos de atributos de la relación R, B depende funcionalmente de manera completa de A si B depende funcionalmente de A pero no de ningún subconjunto de A.

Dependencia Parcial: Dependencia funcional en la cual existe algún atributo que puede eliminarse del determinante y la dependencia continúa verificándose

Dependencia Transitiva: Si A, B y C son conjuntos de atributos de la relación R, C depende transitivamente de A a través de B si C depende funcionalmente de B y B depende funcionalmente de A.

2FN: Relación que está en 1FN y en la que todo atributo que no sea de una clave candidata depende funcionalmente de forma completa de cualquier clave candidata.

3FN: Relación que está en 2FN y en la que ningún atributo que no sea de una clave candidata depende transitivamente de cualquier clave candidata

FNBC: Relación en la que el determinante de todas las dependencias funcionales es una clave candidata.

Dependencia Multivaluada: Dependencia entre conjuntos de atributos A, B y C con valores independientes de modo que para cada valor de A hay un conjunto de valores de B y un conjunto de valores de C.

4FN: Relación que está en FNBC y no contiene dependencias multivaluadas.

5FN: Relación que no presenta dependencias de combinación (propiedad de la descomposición de relaciones que garantiza que no se generan tuplas incorrectas al volver a combinar relaciones de una descomposición mediante una operación de reunión natural)