



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE CIENCIAS
FUNDAMENTOS DE BASES DE DATOS

Forma Normal de Boyce-Codd (BCNF)

Gerardo Avilés Rosas
gar@ciencias.unam.mx

Y ahora...
¿Cómo normalizo una
Base de Datos?



Una forma de acabar con anomalías como la redundancia es la **descomposición de relaciones**.

Dada una relación $R(A_1, A_2, \dots, A_n)$, se puede descomponer R en dos relaciones $S(B_1, B_2, \dots, B_i)$ y $T(C_1, C_2, \dots, C_j)$, tales que:

1. $\{A_1, A_2, \dots, A_n\} = \{B_1, B_2, \dots, B_i\} \cup \{C_1, C_2, \dots, C_j\}$
2. Las *tuplas* en la relación S son la proyección sobre $\{B_1, B_2, \dots, B_i\}$ de las *tuplas* en R .
3. De manera similar, las *tuplas* en T son la proyección de $\{C_1, C_2, \dots, C_j\}$ sobre R .

Veamos un ejemplo:

Sucursal	Alcaldía	Activo	Cliente	Préstamo	Importe
Centro	Cuauhtémoc	\$1,800 M	Santos	P-17	\$200,000
Copilco	Coyoacán	\$420 M	Gómez	P-23	\$400,000
Viveros	Coyoacán	\$340 M	López	P-15	\$300,000
Centro	Cuauhtémoc	\$1,800 M	Toledo	P-14	\$300,000
Eugenia	Benito Juárez	\$80 M	Santos	P-93	\$100,000
Zapata	Benito Juárez	\$1,600 M	Pérez	P-11	\$180,000
San Ángel	Álvaro Obregón	\$60 M	Vázquez	P-29	\$240,000
San Fernando	Tlalpan	\$740 M	López	P-16	\$260,000
Centro	Cuauhtémoc	\$1,800 M	González	P-23	\$400,000
Viveros	Coyoacán	\$340 M	Rodríguez	P-25	\$500,000

Si se descompone la relación en:

1. Una relación **Sucursal** (*Sucursal, Alcaldía, Activo, Cliente*) y
2. Una relación **Préstamo** (*Cliente, Préstamo, Importe*)

Sucursal	Alcaldía	Activo	Cliente
Centro	Cuauhtémoc	\$1,800 M	Santos
Copilco	Coyoacán	\$420 M	Gómez
Viveros	Coyoacán	\$340 M	López
Centro	Cuauhtémoc	\$1,800 M	Toledo
Eugenia	Benito Juárez	\$80 M	Santos
Zapata	Benito Juárez	\$1,600 M	Pérez
San Ángel	Álvaro Obregón	\$60 M	Vázquez
San Fernando	Tlalpan	\$740 M	López
Centro	Cuauhtémoc	\$1,800 M	González
Viveros	Coyoacán	\$340 M	Rodríguez

Cliente	Préstamo	Importe
Santos	P-17	\$200,000
Gómez	P-23	\$400,000
López	P-15	\$300,000
Toledo	P-14	\$300,000
Santos	P-93	\$100,000
Pérez	P-11	\$180,000
Vázquez	P-29	\$240,000
López	P-16	\$260,000
González	P-23	\$400,000
Rodríguez	P-25	\$500,000

1. Encontrar todas las sucursales que tienen préstamos con importe menor a **\$200,000**:

Con el esquema original:

Sucursal	Alcaldía	Activo	Cliente	Préstamo	Importe
Centro	Cuauhtémoc	\$1,800 M	Santos	P-17	\$200,000
Copilco	Coyoacán	\$420 M	Gómez	P-23	\$400,000
Viveros	Coyoacán	\$340 M	López	P-15	\$300,000
Centro	Cuauhtémoc	\$1,800 M	Toledo	P-14	\$300,000
Eugenia	Benito Juárez	\$80 M	Santos	P-93	\$100,000
Zapata	Benito Juárez	\$1,600 M	Pérez	P-11	\$180,000
San Ángel	Álvaro Obregón	\$60 M	Vázquez	P-29	\$240,000
San Fernando	Tlalpan	\$740 M	López	P-16	\$260,000
Centro	Cuauhtémoc	\$1,800 M	González	P-23	\$400,000
Viveros	Coyoacán	\$340 M	Rodríguez	P-25	\$500,000

Eugenia y Zapata

Con el esquema fraccionado:

Sucursal	Alcaldía	Activo	Cliente
<u>Centro</u>	Cuauhtémoc	\$1,800 M	Santos
Copilco	Coyoacán	\$420 M	Gómez
Viveros	Coyoacán	\$340 M	López
Centro	Cuauhtémoc	\$1,800 M	Toledo
<u>Eugenia</u>	Benito Juárez	\$80 M	Santos
<u>Zapata</u>	Benito Juárez	\$1,600 M	Pérez
San Ángel	Álvaro Obregón	\$60 M	Vázquez
San Fernando	Tlalpan	\$740 M	López
Centro	Cuauhtémoc	\$1,800 M	González
Viveros	Coyoacán	\$340 M	Rodríguez

N

Eugenia, Zapata y Centro

Cliente	Préstamo	Importe
Santos	P-17	\$200,000
Gómez	P-23	\$400,000
López	P-15	\$300,000
Toledo	P-14	\$300,000
Santos	P-93	\$100,000
Pérez	P-11	\$180,000
Vázquez	P-29	\$240,000
López	P-16	\$260,000
González	P-23	\$400,000
Rodríguez	P-25	\$500,000

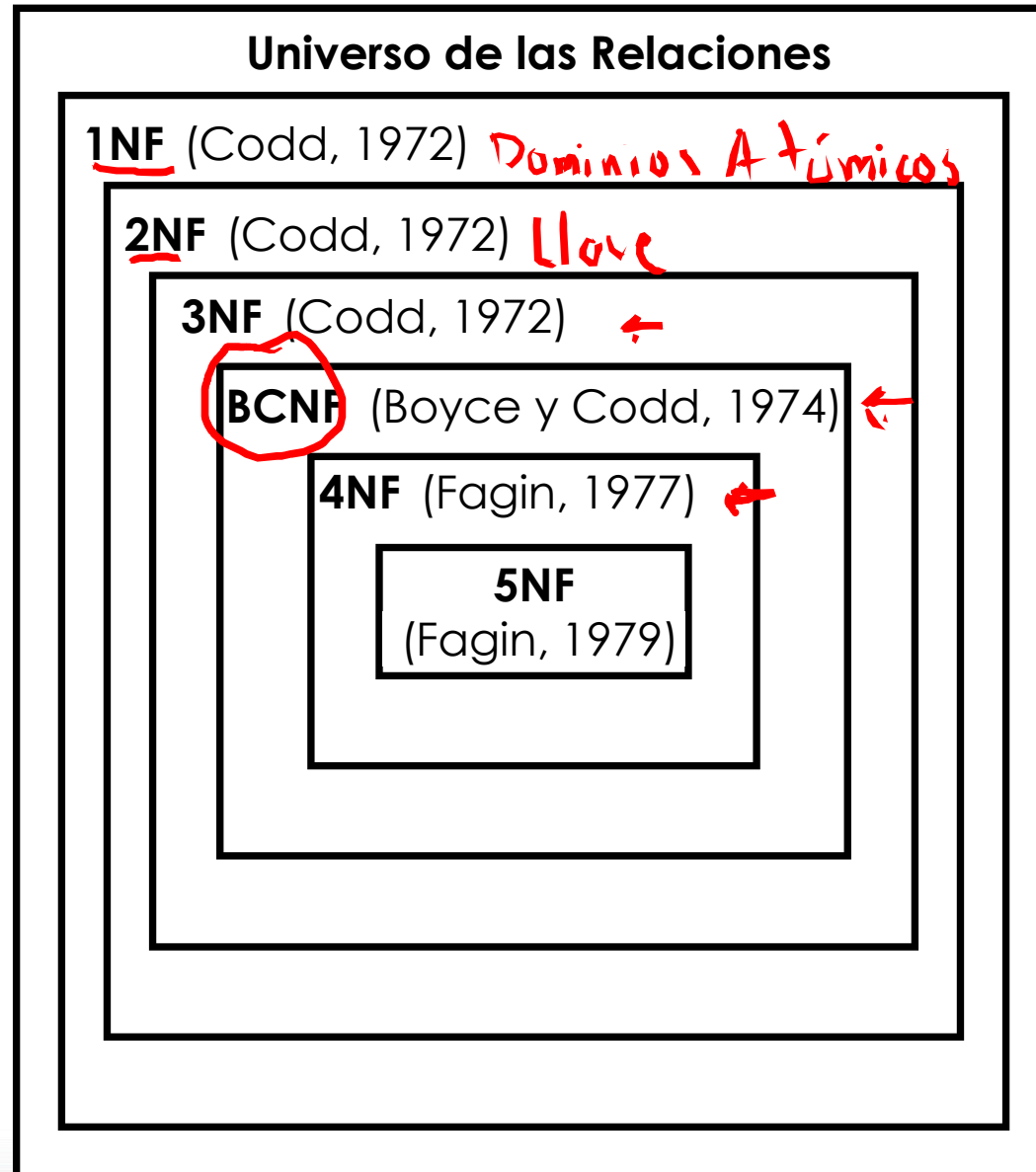
2. Indicar los préstamos que se tienen en cada sucursal el cliente **Santos**:

$\sigma_{\text{nom_cliente} = \text{'Santos'}} (\text{Sucursal} \bowtie \text{Prestamo})$

Sucursal	Alcaldía	Activo	Cliente	Préstamo	Importe
Centro	Cuauhtémoc	\$1,800 M	Santos	P-17	\$200,000
Centro	Cuauhtémoc	\$1,800 M	Santos	P-93	\$100,000
Eugenia	Benito Juárez	\$80 M	Santos	P-17	\$200,000
Zapata	Benito Juárez	\$1,600 M	Santos	P-93	\$180,000

Ambas relaciones tienen a **Cliente** en común, así que para reunir las se usa este atributo, que **no resulta adecuado** puesto que un cliente puede tener **varios préstamos** no necesariamente en la misma sucursal.

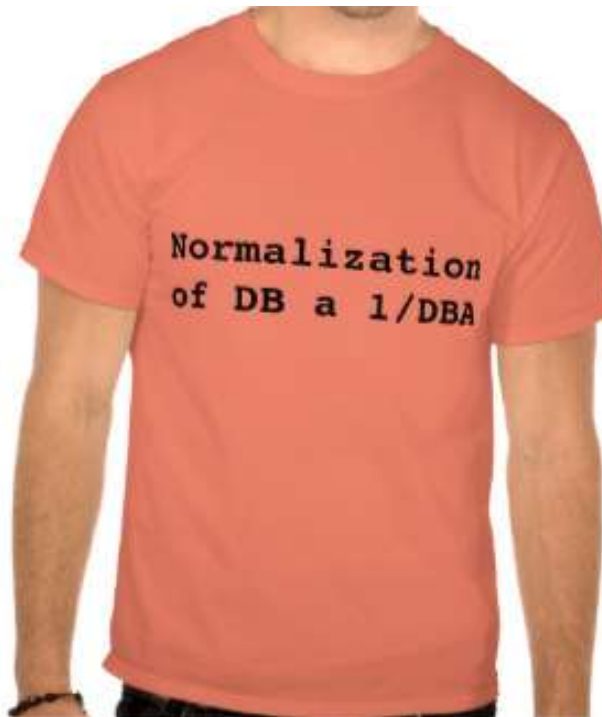
- La **Normalización** es una técnica desarrollada inicialmente por **E.F. Codd en 1972**, para diseñar la estructura lógica de una BD en el modelo relacional.
- Se trata de un proceso en el cual se van comprobando el cumplimiento de una serie de reglas (restricciones) por parte de un esquema de relación.
- Cada regla que se cumple, aumenta el grado de normalización del esquema.
- Cuando una regla no cumple, el esquema de relación se debe descomponer en varios esquemas que sí la cumplan por separado.



Objetivos de la normalización

Para normalizar una BD se desea:

- Eliminar la duplicidad de información.
- Que las relaciones fraccionadas tengan un **join** sin pérdida.
- Conservar las dependencias funcionales.



Forma Normal de Boyce-Codd

Una relación R está en BCNF si y sólo si en toda DF no trivial $A_1, A_2, \dots, A_n \rightarrow B$ para R , se tiene que $\{A_1, A_2, \dots, A_n\}$ es superllave de R .

Por ejemplo:

- La relación $C(\text{nombre}, \text{calle}, \text{ciudad})$ con DF = $\{\text{nombre} \rightarrow \text{calle ciudad}\}$

$$\{n\}^+ = \{n, c, cd\} \quad \therefore n \text{ ES UNA LLAVE}$$

$$\alpha \rightarrow \beta \text{ (DF)} : \text{LA DF ES TRIVIAL SI } \beta \subseteq \alpha$$

- La relación $S(\text{nombre}, \text{no_préstamo})$ con DF = $\{\text{nombre} \rightarrow \text{no_préstamo}\}$

$$\{n\}^+ = \{n, p\} \quad \therefore n \text{ ES LLAVE}$$

...Forma Normal de Boyce-Codd

Cualquier relación de dos atributos **A** y **B** está en **BCNF** si:

1. No hay **DF no triviales**, se mantiene entonces la condición BCNF, debido a que sólo una **DF no trivial** puede violar esta condición (notar que {A,B} es la única llave).
2. Si se tiene **A** → **B**, pero no **B** → **A**, entonces **A** es la única llave y cada dependencia no trivial contiene **A** en la izquierda, por tanto no hay violación a la condición BCNF.
3. Si **B** → **A** y no se tiene **A** → **B** es un caso simétrico al anterior.
4. Si se tiene **A** → **B** y **B** → **A**. Entonces tanto **A** como **B son llaves**, y cualquier dependencia tiene al menos uno de ellos en su lado izquierdo, por tanto **no puede haber violación** de la norma BCNF.

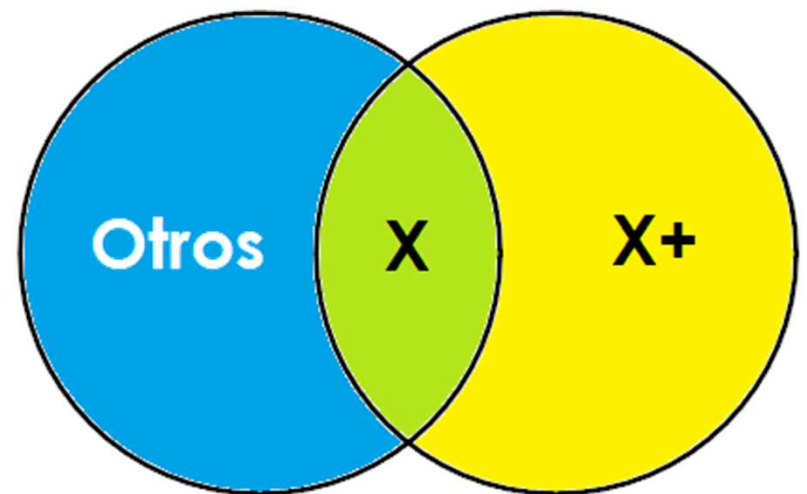
...Forma Normal de Boyce-Codd

Es posible dividir cualquier relación en otras con las siguientes propiedades:

- **Son esquemas de relaciones en BCNF.**
- **Los datos en la relación original se representan fielmente por las relaciones que son resultado de la descomposición.**

La estrategia a seguir es:

1. **Buscar una DF no trivial $X \rightarrow B$ que viole BCNF.**
2. **Calcular X^+ .**
3. **Fraccionar R en $R_1(X^+) \cup R_2((R-X^+) \cup X)$.**
4. **Encontrar las DF para las nuevas relaciones.**



Se debe aplicar la regla de descomposición tantas veces como sea necesario hasta que todas las relaciones estén en **BCNF**.

La relación

Prestamo (^SnombreSuc, ^Cnombre_cliente, ^Pno_prestamo, ⁱimporte)

DF = {no_prestamo → importe, nombreSuc}

$$P \rightarrow i \ S$$

$$\Rightarrow \{P\}^+ = \{P \ i \ S\} \quad \times \text{ BCNF}$$

se descompone en:

$$\therefore \{P \ C\}^+ = \{P \ C \ i \ S\} \quad \checkmark \text{ LLAVE}$$

$R(P, i, S) \quad F = \{P \rightarrow i \ S\} \quad \downarrow$ LA LLAVE ES P \checkmark

$S(P, C) \quad \text{LA LLAVE ES } PC \Rightarrow PC \rightarrow PC \quad \checkmark$
 $\underbrace{PC \rightarrow PC}_{\text{DF. TRIVIAL}}$

No estés muy orgulloso de haber comprendido estas notas.

La habilidad para manejar la Normalización de BD es insignificante comparado con el poder de la Fuerza.

