



TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

- Hasta ahora, nuestras relaciones fueron agrupaciones de atributos en esquemas siguiendo el sentido común.
- Aunque hablamos de entidades y relaciones no hemos visto un método **formal** para determinar que un esquema es mejor que otro



TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

Normalmente, al seguir la metodología de diseño que hemos estudiado, realizando un adecuado diseño conceptual en el modelo entidad-relación, seguido con una cuidadosa transformación al modelo relacional, se pueden evitar estos problemas y obtener un esquema relacional exento de errores.



TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

Sin embargo, ante las posibles dudas respecto a si un determinado esquema relacional es correcto, es preferible aplicar a dicho esquema un método formal de análisis que determine lo que pueda estar equivocado y nos permita llegar a otro que nos asegure el cumplimiento de ciertos requisitos; este método formal es la teoría de la normalización.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

- La normalización nos provee pruebas para verificar si una relación pertenece a cierta “**forma normal**”.
- Veremos 3 formas normales (1FN, 2FN, 3FN y BC)

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

¿Cómo surgieron?

- En el proceso de normalización, según la propuesta original de Codd (1972), se somete un esquema de relación a una serie de pruebas para "certificar" si pertenece o no a una cierta forma normal.
- En un principio Codd propuso tres formas normales, a las cuales llamó primera (**1FN**), segunda (**2FN**) y tercera (**3FN**) forma normal.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN



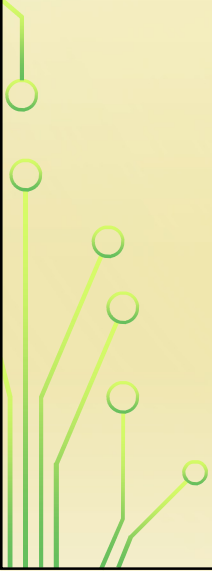
Posteriormente, Boyce y Codd propusieron una definición más estricta de 3FN, a la que se conoce como forma normal de BoyceCodd. Todas estas formas normales se basan en las dependencias funcionales entre los atributos de una relación.



TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

Más adelante se propusieron una cuarta forma normal (4FN) y una quinta (5FN), con fundamento en los conceptos de dependencias multivaluadas y dependencias de reunión, respectivamente.



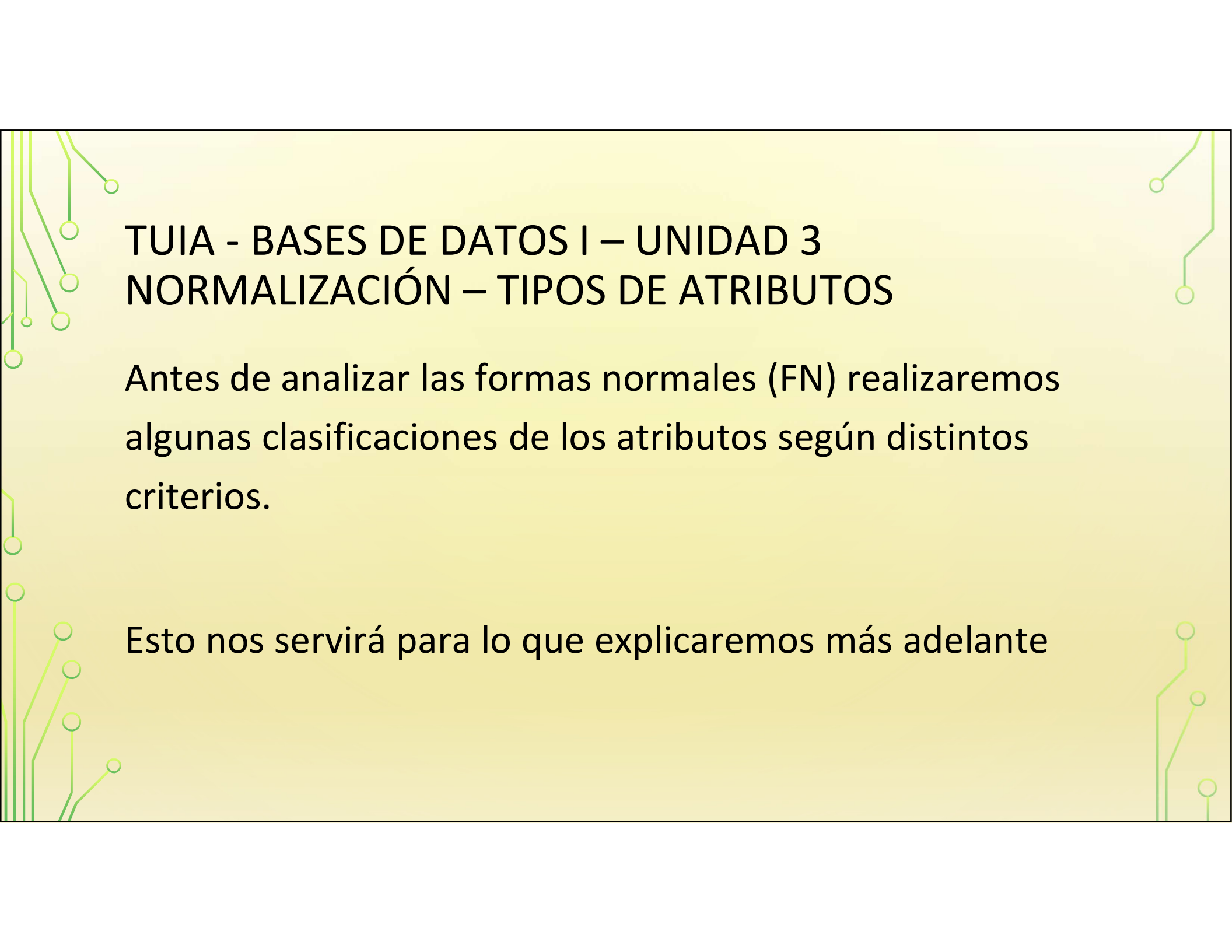
TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN

“El proceso de normalización permite que un conjunto de relaciones cumpla con una serie de propiedades deseables”

Este proceso descompone relaciones insatisfactorias en relaciones mas pequeñas, donde se busca evitar:

- Perdida de integridad de los datos
- Redundancia de datos
- Ocurrencia de anomalías



TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – TIPOS DE ATRIBUTOS

Antes de analizar las formas normales (FN) realizaremos algunas clasificaciones de los atributos según distintos criterios.

Esto nos servirá para lo que explicaremos más adelante

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – TIPOS DE ATRIBUTOS

a) Según la cantidad de elementos que lo componen

- **Simple o atómico:** no se puede descomponer en componentes más pequeños (Ej: nro. de documento, genero de un alumno).
- **Compuesto:** Se puede dividir en componentes más pequeños (Ej: Nombre completo, Domicilio).

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – TIPOS DE ATRIBUTOS

b) Según la cantidad de valores para una entidad

- **Monovaluado:** un atributo que tiene solo un valor para una entidad (Ej: Fecha de nacimiento).
- **Multivaluado:** cuando el atributo puede tener un conjunto de valores (Ej: Teléfonos de una persona).

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – TIPOS DE ATRIBUTOS

c) Según sea o no parte de una clave candidata (CC)

- **Primo:** si forma parte de una clave candidata
- **No primo:** si NO forma parte de una clave candidata

ESTUDIANTE	<u>Legajo</u>	Nombre	NroDoc	FechaNac	Carrera	Sexo
------------	---------------	--------	--------	----------	---------	------

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – EJEMPLO

Supongamos que queremos almacenar en una base de datos la siguiente información sobre los pacientes atendidos en una clínica médica:

- Nro. de historia clínica del paciente
- Nombre del paciente
- Código de la obra social del paciente
- Descripción de la obra social
- Para cada consulta médica: código y nombre del médico que lo atendió, fecha de la consulta y diagnóstico.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – EJEMPLO

PAC_CONS	<u>NroHist</u>	NomPac	CodOS	Descrip	Consulta			
					CodMed	NomMed	Fecha	Diagnostico
	1912	Gomez, A.	2	Osecac	101 103 104	Perez, J. Sosa, T. Juárez, I.	23/01/2001 24/02/2001 01/04/2002	Angina Otitis Hepatitis
	8123	Estévez, H.	3	Osde	103	Sosa, T.	30/08/2001	Angina
	5407	Alvarez, J.	1	OsUNR	103 103	Sosa, T. Sosa, T.	04/07/2002 09/09/2002	Gripe Bronquitis
	7191	Benitez, A.	2	Osecac	102 104	Tapia, D. Juárez, I.	12/06/2001 31/01/2002	Gripe Angina

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – EJEMPLO

PAC_CONS	<u>NroHist</u>	NomPac	CodOS	Descrip	<u>CodMed</u>	NomMed	<u>Fecha</u>	Diagnostico
	1912	Gomez, A.	2	Osecac	101	Perez, J.	23/01/2001	Angina
	1912	Gomez, A.	2	Osecac	103	Sosa, T.	24/02/2001	Otitis
	1912	Gomez, A.	2	Osecac	104	Juárez, I.	01/04/2002	Hepatitis
	8123	Estévez, H.	3	Osde	103	Sosa, T.	30/08/2001	Angina
	5407	Alvarez, J.	1	OsUNR	103	Sosa, T.	04/07/2002	Gripe
	5407	Alvarez, J.	1	OsUNR	103	Sosa, T.	09/09/2002	Bronquitis
	7191	Benitez, A.	2	Osecac	102	Tapia, D.	12/06/2001	Gripe
	7191	Benitez, A.	2	Osecac	104	Juárez, I.	31/01/2002	Angina

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – ANOMALÍAS

a) Anomalías de inserción:

Al insertar nuevas tuplas (una nueva consulta) puede generar valores distintos en aquellos atributos que se repiten.

Si agregamos una entidad nueva hay que resolver el tema de los atributos multivaluados. Puede violarse integridad de entidades.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – ANOMALÍAS

b) Anomalías de eliminación:

En caso de eliminar una columna, si esta es la última para un paciente dado, se eliminan también los datos del paciente.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – ANOMALÍAS

c) Anomalías de actualización:

Si cambiamos la descripción de una Obra Social, hay que cuidarse de cambiarlo en todas las repeticiones.

TUIA - BASES DE DATOS I – UNIDAD 3

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

- El concepto individual más importante en el diseño de esquemas relacionales es el de dependencia funcional.
- Una dependencia funcional es una restricción entre dos conjuntos de atributos de la base de datos.
- Suponiendo que una relación R tiene un conjunto de atributos A_1, A_2, \dots, A_n .
- $R = \{A_1, A_2, \dots, A_n\}$.

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

Una dependencia funcional, denotada por:

$$\mathbf{X} \rightarrow \mathbf{Y}$$

entre dos conjuntos de atributos **X** e **Y** que son subconjuntos de **R**, especifica una restricción sobre las posibles tuplas que podrían formar un ejemplar de relación **r** de **R**.

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

$R = \{A_1, A_2, \dots, A_n\}.$

$X \rightarrow Y$

La restricción dice que, para cualesquiera dos tuplas t_1 y t_2 , de **R** tales que $t_1[X] = t_2[X]$, debemos tener también $t_1[Y] = t_2[Y]$

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

$R = \{A1, A2, \dots, An\}.$

$X \rightarrow Y$

Establece que los valores de **Y** de una tupla **r** dependen de los valores de **X**.

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

$R = \{A1, A2, ..., An\}.$

$X \rightarrow Y$

O bien que los valores de **X** de una tupla *r* determinan de manera única los valores de **Y**.

Es decir que **Y** depende funcionalmente de **X**.



TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

Decimos que **X** determina funcionalmente a **Y** porque siempre que 2 tuplas coincidan en su valor de **X**, necesariamente deben coincidir en su valor de **Y**.

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

Si X es una **clave candidata** (CC) entonces $X \rightarrow Y$ para cualquier subconjunto de atributos Y de R .

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

Consideremos el esquema de relación PAC_CONS, algunas de las dependencias funcionales que se pueden establecer son las siguientes:

df1: { NroHist, CodMed, Fecha } \rightarrow Diagnostico

df2 : CodMed \rightarrow NomMed

df3 : NroHist \rightarrow { NomPac, CodOS, Descrip }

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

df1 : La combinación de los valores (NroHist, CodMed, Fecha) determina de manera única el diagnóstico del paciente en una consulta (Diagnostico).

df2 : El código del médico (CodMed) determina de manera única su nombre (NomMed).

df3 : El número de historia clínica (NroHist) determina de manera única el nombre de ese paciente (NomPac) y el código (CodOS) y la descripción (Descrip) de su obra social.

TUIA - BASES DE DATOS I

NORMALIZACIÓN – DEPENDENCIAS FUNCIONALES

Representación gráfica:

