

# TEORIA-Test-Tema-4-NORMALIZACION...



**NachoPiece** 



**Bases de Datos** 



2º Grado en Ingeniería Informática



Escuela Técnica Superior de Ingeniería Informática Universidad de Málaga



# Inteligencia Artificial & Data Management

MADRID









# Esto no son apuntes pero tiene un 10 asegurado (y lo vas a disfrutar igual).

Abre la **Cuenta NoCuenta** con el código <u>WUOLAH10</u>, haz tu primer pago y llévate  $10 \in$ .

Me interesa



Este número es indicativo del riesgo del producto, siendo 1/6 indicativo de menor riesgo y 6/6 de mayor riesgo.

ING BANK NV se encuentra adherid al Sistema de Garantía de Depósito Holandés con una garantía de hast 100.000 euros por depositante. Consulta más información en ing.e







Mezclar atributos de distintas entidades no es adecuado, provoca redundancia en el almacenamiento además de anomalías de inserción, modificación y borrado.

Supongamos la relación EMP\_PROJ (Empleado#, Proyecto#, ENombre, PNombre, N\_horas)

**Anomalía de inserción:** No podremos insertar un proyecto a menos que asignemos a este al menos a un empleado, tampoco podremos insertar un empleado a menos que exista un proyecto.

**Anomalía de modificación:** Cambiar el nombre de un proyecto, hará que cambien para todos los empleados que estén asignados a este proyecto.

**Anomalía de borrado:** Si borramos un proyecto se eliminarán todos los empleados asociados a este, a la inversa, si eliminamos todos los empleados de un proyecto, este se eliminará también.



En el diseño de una base de datos, cada tupla ("atributo") que se relacione con otra de la misma tabla debe representar a una entidad o instancia de esta relación.

- Los atributos de diferentes entidades no deben mezclarse.
- Utilizaremos claves foráneas para referenciar a otras entidades.
- La entidad y sus atributos deben formar una unidad lo más independiente posible.
- Debemos evitar la redundancia y las anomalías.
- La normalización proporciona un procedimiento sistemático para lograrlo.



Consulta condiciones aquí





NachoPiece



# **DEPENDENCIAS FUNCIONALES**

En la siguiente tabla podemos ver algunos ejemplos de dependencias funcionales como, por ejemplo:

País -> Indicativo.

País y número -> Abonado.

Indicativo	Numero	Pais	Abonado
34	952131000	España	Universidad de Málaga
34	951299316	España	Turismo Andaluz
34	951299300	España	Turismo Andaluz
1	800-275-2273	EEUU	Apple
1	800-426-9400	EEUU	Microsoft
1	951299300	Canadá	Peter Smith
1	800-555-0077	EEUU	Turismo Andaluz
	***		111

### Definición:

Entre los atributos A1, A2, . . ., An y B de una relación R hay una Dependencia Funcional (DF) si siempre que dos tuplas de R coincidan en sus valores de A1, A2, . . ., An, coinciden también en sus valores de B.

**Notación:** A1,A2,...,An  $\rightarrow$  B.

Generalizamos la definición: A1,A2,...,An  $\rightarrow$  B1,B2,...,Bm. Si  $A\rightarrow$ B, A se denomina el determinante y B el determinado.

### **CLAVES**

Dada una relación R, podemos decir que un conjunto de atributos (K) son clave si:

**Identificación**: Es posible identificar una única fila por el valor en la clave.

Indistinción: Es imposible que dos tuplas de R coincidan en la clave.

Determinación: La clave determinan funcionalmente a todos los atributos de la relación.

 $K \rightarrow A1, A2, \dots, An.$ 

### Algunas definiciones importantes:

**Clave minimal:** Si no existe un subconjunto de claves dentro de un conjunto de claves, esta se considera minimal.

Super-clave: Cuando una clave contiene una clave minimal como subconjunto propio.

**Clave primaria:** Es una clave seleccionada entre todas las claves de la relación, se considera la más importante.

NachoPiece -2-



ING BANK NV se encuentra adherido al Sistema de Garantía de Depósitos Holandés con una garantía de hasta 100.000 euros por depositante. Consulta más información en ing.es

# Que te den **10 € para gastar** es una fantasía. ING lo hace realidad.

Abre la **Cuenta NoCuenta** con el código **WUOLAH10**, haz tu primer pago y llévate 10 €.

# Quiero el cash

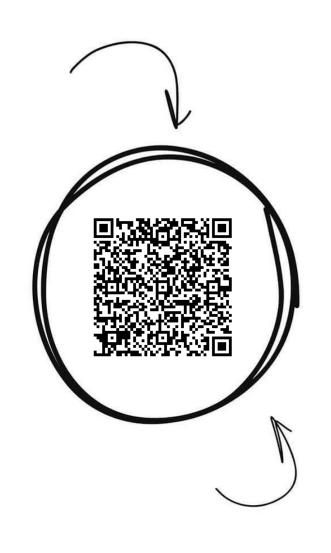
Consulta condiciones aquí







# **Bases de Datos**



Banco de apuntes de la



# Comparte estos flyers en tu clase y consigue más dinero y recompensas

- Imprime esta hoja
- Recorta por la mitad
- Coloca en un lugar visible para que tus compis puedan escanar y acceder a apuntes
- Llévate dinero por cada descarga de los documentos descargados a través de tu QR





# **NORMALIZACIÓN**

- Es el proceso de descomponer tablas o relaciones "mal" construidas, generalmente dividiéndola en relaciones más pequeñas.
- Una forma normal es una condición que cumple una relación de acuerdo a sus claves y dependencias.
- Las tres primeras formas normales (1FN,2FN,3FN) y la forma normal de Boyce-Codd (FNBC) están basadas en sus claves y dependencias funcionales.
- La cuarta forma normal (4FN) se basa en las claves y las dependencias multivaluadas (MVD) de una relación. La quinta forma normal (5FN) se basa en las claves y las dependencias de unión (JD) de una relación.
- No obstante, lo anterior, con la FNBC se obtiene un diseño de la BD correcto, por lo que en este tema no se estudian la 4FN ni la 5FN.
- Para asegurar el mejor diseño posible de una base de datos aparte de las formas normales anteriores es necesario tener en cuenta otras propiedades.

### **DEFINICIONES PREVIAS**

- Se define un **atributo primario** como aquel que es miembro de alguna clave candidata. Cualquier otro atributo es considerado un atributo no primario.
- Una dependencia funcional A  $\rightarrow$  B es **plena** si no existe otra dependencia funcional C  $\rightarrow$  B donde C es un subconjunto propio de A.
- Una dependencia funcional  $A \to B$  es **transitiva** si existe un atributo C con C ni está contenida ni es  $A \to B$  es transitiva si existe un atributo C con C ni está contenida ni es C tal que existen las dependencias funcionales  $A \to C$  y  $C \to B$ .

NachoPiece -3-



# Esto no son apuntes pero tiene un 10 asegurado (y lo vas a disfrutar igual).

Abre la **Cuenta NoCuenta** con el código <u>WUOLAH10</u>, haz tu primer pago y llévate 10 €.

Me interesa



5 Este número es indicativo del riesgo del producto, siendo 1/6 indicativo de menor riesgo y 6/6 de mayor riesgo.

ING BANK NV se encuentra adherid al Sistema de Garantía de Depósito Holandés con una garantía de hast 100.000 euros por depositante. Consulta más información en ing.e





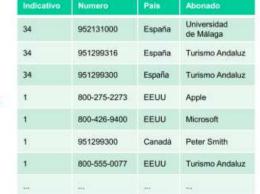
### (1FN) FORMA NORMAL Nº1

- La primera forma normal es aquella que no permite atributos compuestos o multivalor; es decir, atributos cuyos valores para una tupla individual no son atómicos.

(Se considera que cualquier tabla, dada las restricciones del DDL está en 1FN)







pormalización



# (2FN) FORMA NORMAL Nº2

- Un esquema de relación R se encuentra en 2FN si todo atributo no primario de R se encuentra en una dependencia funcional plena con la clave primaria.
- ${\sf -}$  Si R no cumple la 2FN tiene que descomponerse en varias relaciones 2FN, por cada clave parcial con sus atributos dependientes.

Consu	lta	





Indicativo	Numero	Pais	Abonado
34	952131000	España	Universidad de Málaga
34	951299316	España	Turismo Andaluz
34	951299300	España	Turismo Andaluz
1	800-275-2273	EEUU	Apple
1	800-426-9400	EEUU	Microsoft
1	951299300	Canadá	Peter Smith
1	800-555-0077	EEUU	Turismo Andaluz
++	24.5	1966	411

Numero	Pais	Abonado
952131000	España	Universidad de Málaga
951299316	España	Turismo Andaluz
951299300	España	Turismo Andaluz
800-275-2273	EEUU	Apple
800-426-9400	EEUU	Microsoft
951299300	Canadá	Peter Smith
800-555-0077	EEUU	Turismo Andaluz
	Care	1044

Indicativo	Pais
34	España
1	EEUU
1	Canadá
1440	3448





# (3FN) FORMA NORMAL N°3

- Un esquema de relación R se encuentra en 3FN si se está en 2FN y ningún atributo no primario de R es transitivamente dependiente de la clave primaria a través de atributos no primarios. Si la dependencia transitiva es a través de una clave, no hay problema.
- Si R no cumple la 3FN tiene que descomponerse en varias relaciones 3FN que incluyan los atributos no primarios que determinen funcionalmente otros atributos no primarios.



Nacimiento	Numero	Pais	Abonado		Numero	Pals	Abonado	Nacimiento	Abonado
1974	952131000	España	Francisca		952131000	España	Francisca	1974	Francisca
1974	951299316	España	Lucas		951299316	España	Lucas	1974	Lucas
1964	951299300	España	Lolo		951299300	España	Lolo	1964	Lolo
1998	800-275-2273	EEUU	Antonio	-	800-275-2273	EEUU	Antonio	1998	Antonio
1998	800-426-9400	EEUU	Miguel		800-426-9400	EEUU	Miguel	1998	Miguel
1989	951299300	Canadá	Peter Smith		951299300	Canadá	Peter Smith	1989	Peter Smith
1979	800-555-0077	EEUU	Tomás		800-555-0077	EEUU	Tomás	1979	Tomás
10.5	22	125	125			Sal I		7775	122

### **DEFINICIONES GENERALES**

- Las definiciones anteriores consideran sólo las claves primarias, aunque también pueden definirse en función de cualquier clave candidata.
- Un esquema de relación R se encuentra en 2FN si cualquier atributo no primario de R se encuentra en una DF plena con todas las claves de R.
- Un esquema de relación R se encuentra en 3FN si cuando existe una DF  $X \rightarrow A$  en R entonces se cumple una de las siguientes condiciones:
- a) X es una superclave de R.
- b) A es un atributo primario de R.
- La FNBC se puede definir como una 3FN en la que no se permite b).

NachoPiece



# FORMA NORMAL BOYCE-CODD (FNBC)

- Un esquema de relación R se encuentra en FNBC si cuando existe una DF  $X \rightarrow A$  en R entonces se cumple que X es una superclave de R.
- Dicho de otro modo, una relación está en FNBC si y solo si está en 3FN y cada determinante es una clave candidata
- Cada forma normal es estrictamente más fuerte que la anterior
- Existen relaciones que están en 3FN pero no en FBC. El objetivo es conseguir que todas las relaciones de nuestro diseño estén en FNBC.

# **EJEMPLOS FNBC**

# FNBC. Ejemplo. Direcciones Postales:

<b>CPost</b>	Dir	Ciudad
3000	C/ Las Flores Nº17	Merida
4858	Av. Bolívar este Nº72	Maracay

A Ciudades diferentes le corresponden códigos postales distintos.

En este caso hay dependencia entre el Código Postal → Ciudad, ya que, conocido el Código Postal se puede conocer la Ciudad, y además Dirección, Ciudad → Código Postal. Existe pues un determinante, Código Postal que no es clave candidata Para transformar la tabla en una tabla en FNBC se crea una tabla de Códigos Postales y Ciudades, eliminando de la tabla original la Ciudad, obteniéndose dos tablas, una con los atributos Dirección y Código Postal y otra con el Código Postal y la Ciudad

Tablas en FNBC:

CÓDIGOS_CIUDADES		CÓDIGOS_DIRECCIONES		
CPost	Ciudad	CPost	Dir	
3000	Merida	3000	C/ Las Flores Nº17	
4858	Maracay	4858	Av. Bolivar este Nº72	

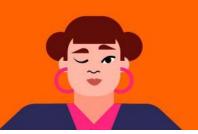
NachoPiece -



# Esto no son apuntes pero tiene un 10 asegurado (y lo vas a disfrutar igual).

Abre la **Cuenta NoCuenta** con el código <u>WUOLAH10</u>, haz tu primer pago y llévate 10 €.





Este número es indicativo del riesgo del producto, siendo 1/6 indicativo de menor riesgo y 6/6 de mayor riesgo.

ING BANK NV se encuentra adheria al Sistema de Garantía de Depósito Holandés con una garantía de hasi 100.000 euros por depositante. Consulta más información en ing.e











do your thing

NachoPiece

# FNBC. Otro Ejemplo. Consideremos una empresa donde

Consideremos una empresa donde un trabajador puede trabajar en varios departamentos. En cada departamento hay varios responsables, pero cada trabajador sólo tiene asignado uno. Tendríamos una tabla con las columnas:

### IDTrabajador, IDDepartamento, IDResponsable

La única clave candidata es IDTrabajador (que será por tanto la clave primaria). Si añadimos la limitación de que el responsable sólo puede serlo de un departamento, este detalle produce una dependencia funcional ya que: IDResponsable → IDDepartamento Por lo tanto hemos encontrado un determinante (IDResponsable) que sin embargo no es clave candidata. Por ello, esta tabla no está en FNBC. En este caso la redundancia ocurre por mala selección de clave. La repetición del par [IDDepartamento + IDResponsable] es innecesaria y evitable

Una forma sencilla de comprobar si una relación se encuentra en FNBC consiste en comprobar, además de que esté en 3FN, lo siguiente:

- (1) Si no existen claves candidatas compuestas (con varios atributos), está en FNBC.
- (2) Si existen varias claves candidatas compuestas y éstas tienen un elemento común, puede no estar en FNBC. Sólo si, para cada dependencia funcional en la relación, el determinante es una clave candidata, estará en FNBC.

### ALGUNOS VIDEOS QUE AYUDAN A ENTENDER LAS FORMAS NORMALES

1FN,2FN Y 3FN: https://www.youtube.com/watch?v=bO18omSzeR4

3FN: https://www.youtube.com/watch?v=-LrUJR0G\_6g

FNBC: <a href="https://www.youtube.com/watch?v=1C9dSd4ci0o">https://www.youtube.com/watch?v=1C9dSd4ci0o</a>

