

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS



Tarea 4:

Normalización

Angel Christian Pimentel Noriega - 316157995

Mauricio Riva Palacio Orozco - 316666343

Alex Gerardo Fernandez Aguilar - 314338097

Martin Felipe Espinal Cruces - 316155362

1. Preguntas de repaso:

- ¿Qué son las anomalías de **inserción, borrado y actualización**?
 - Anomalia de insercion: Imposibilidad de dar de alta una tupla por no disponer del valor de un atributo principal. Cuando se inserta una nueva fila sin respetar las dependencias funcionales. La imposibilidad de añadir nuevos datos para el consecuente de la dependencia funcional sin que exista un antecedente para ella.
 - Anomalia de borrado o eliminación: Si un conjunto de valores queda vacío, podemos perder información adicional como efecto secundario. Se produce cuando se eliminan todas las filas en las que aparecen los datos redundantes por lo que se pierde los datos de la dependencia funcional. Pérdida de información por dar de baja una tupla.
 - Anomalia de actualización: Podemos cambiar información en una tupla y dejarla inalterada en otra. Se produce cuando se modifican las columnas con datos redundantes de sólo un subconjunto de las filas con el mismo dato.
- ¿Qué es una **dependencia funcional** y cómo se define?
 - Decimos que un atributo Y de una relación «depende funcionalmente» de otro atributo X de la relación si a todo valor de X le corresponde siempre el mismo valor de Y.
 - Las dependencias funcionales ayudan a especificar formalmente cuándo un diseño es correcto. Se trata de una relación unidireccional entre 2 atributos de tal forma que en un momento dado, para cada valor único de A, sólo un valor de B se asocia con él a través de la relación.
 - Sea el esquema de relación R definido sobre el conjunto de atributos A y sean X e Y subconjuntos de A llamados descriptores. Se dice que Y depende funcionalmente de X o que X determina o implica a Y, que se representa por $X \rightarrow Y$, si y solo si, cada valor de X tiene asociado en todo momento un único valor de Y.
- Sea **A** la llave de **R (A, B, C)**. Indica **todas** las dependencias funcionales que implica A. Como A es la llave de la relación R, A define funcionalmente a B y C:
 - $A \rightarrow B$
 - $A \rightarrow C$Aplicando la regla de la unión:
 - $A \rightarrow BC$
- ¿Qué es una **forma normal**? ¿Cuál es el objetivo de normalizar un modelo de datos?
 - Las formas normales (NF) proporcionan los criterios para determinar el grado de vulnerabilidad de una tabla a inconsistencias y anomalías lógicas. Cuanto más alta sea la forma normal aplicable a una tabla, menos vulnerable será a inconsistencias y anomalías. Cada tabla tiene una "forma normal más alta" (HNF): por definición, una tabla siempre satisface los requisitos de su HNF y de todas las formas normales más bajas que su HNF; también por definición, una tabla no puede satisfacer los requisitos de ninguna forma normal más arriba que su HNF.
 - 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.
 - Organizar los datos en grupos lógicos, de tal manera que cada grupo describa una pequeña parte del todo.
 - Minimizar la cantidad de datos duplicados almacenados en una base de datos.
 - Perfeccionar la organización de los datos de tal manera que, cuando se necesite introducir modificaciones, el cambio sólo deba aplicarse en un lugar.

- ¿En qué casos es preferible lograr **3NF** en vez de **BCNF**?
 - BCNF actúa de manera diferente a 3NF solo cuando hay varias claves candidatas superpuestas.
 - La razón es que la dependencia funcional $X \rightarrow Y$ es, por supuesto, verdadera si Y es un subconjunto de X . Entonces, en cualquier tabla que tenga solo una clave candidata y esté en 3NF, ya está en BCNF porque no hay ninguna columna (tampoco clave o no clave) que dependa funcionalmente de cualquier cosa además de esa clave.
 - BCNF es una forma normal en la que para cada una de las dependencias funcionales no triviales de una tabla, hay una superclave; 3NF es una forma normal en la que la tabla está en 2NF y cada atributo no principal no depende transitivamente de cada clave de la tabla.

2. Proporciona **algunos ejemplos** que demuestren que las siguientes reglas **no** son válidas:

a Si $\mathbf{A} \rightarrow \mathbf{B}$, entonces $\mathbf{B} \rightarrow \mathbf{A}$

■	<table> <tr><th>A</th><th>B</th></tr> <tr><td>a₁</td><td>b₁</td></tr> <tr><td>a₂</td><td>b₁</td></tr> </table>	A	B	a ₁	b ₁	a ₂	b ₁	En este caso $A \rightarrow B$ pero $B \not\rightarrow A$, debido a que hay 2 tuplas con b_1 .			
	A	B									
	a ₁	b ₁									
a ₂	b ₁										
<table> <tr><th>A</th><th>B</th></tr> <tr><td>a₁</td><td>b₁</td></tr> <tr><td>a₂</td><td>b₂</td></tr> <tr><td>a₃</td><td>b₁</td></tr> <tr><td>a₄</td><td>b₂</td></tr> </table>	A	B	a ₁	b ₁	a ₂	b ₂	a ₃	b ₁	a ₄	b ₂	En este caso $A \rightarrow B$ pero $B \not\rightarrow A$, debido a que hay 2 tuplas con b_1 y 2 tuplas con b_2 .
A	B										
a ₁	b ₁										
a ₂	b ₂										
a ₃	b ₁										
a ₄	b ₂										

b Si $\mathbf{AB} \rightarrow \mathbf{C}$, entonces $\mathbf{A} \rightarrow \mathbf{C}$ y $\mathbf{B} \rightarrow \mathbf{C}$

A	B	C
a ₁	b ₁	c ₁
a ₁	b ₂	c ₂

Cuadro 1: En este caso $AB \rightarrow C$ pero $A \not\rightarrow C$ y $B \rightarrow C$, esto pasa debido a que tenemos dos tuplas con a_1 .

A	B	C
a ₁	b ₁	c ₁
a ₂	b ₁	c ₂
a ₁	b ₂	c ₃

Cuadro 2: En este caso $AB \rightarrow C$ pero $A \not\rightarrow C$ y $B \not\rightarrow C$, esto pasa debido a que tenemos 2 tuplas con a_1 y 2 tuplas con b_1 .

c Si $\mathbf{AB} \rightarrow \mathbf{C}$ y $\mathbf{A} \rightarrow \mathbf{C}$, entonces $\mathbf{B} \rightarrow \mathbf{C}$

A	B	C
a ₁	b ₁	c ₁
a ₂	b ₁	c ₂

Cuadro 3: En este caso $AB \rightarrow C$ y $A \rightarrow C$ pero $B \not\rightarrow C$, esto pasa debido a que tenemos 2 tuplas con b_1 .

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_1	c_3
a_4	b_2	c_4

Cuadro 4: En este caso $AB \rightarrow C$ y $A \rightarrow C$ pero $B \not\rightarrow C$, esto pasa debido a que tenemos 2 tuplas con b_1 y 2 tuplas con b_2 .

3. Dada una relación $R(A, B, C, D, E, G)$ y el siguiente conjunto de dependencias funcionales F :

$$F = AB \rightarrow C, BC \rightarrow D, D \rightarrow EG, CG \rightarrow BD, C \rightarrow A, ACD \rightarrow B, BE \rightarrow C, CE \rightarrow AG$$

Para las siguientes sentencias, determina si son verdaderas o falsas. Para aquellas sentencias que resulten falsas,deberás explicar por qué consideras que no se cumplen:

No.	Sentencia	Verdadera	Falsa	Justificación
1.-	La cerradura de BC es {A, D, E, G}	-	Falso	La cerradura de $\{BC\}^+ = \{BCDEGA\}$
2.-	Todos los atributos de R están en $\{BC\}$	Verdadero	-	La cerradura de $\{BC\}^+ = \{BCDEGA\} = R$
3.-	La cerradura de AC es {A, C}	Verdadero	-	La única DF que cumple con AC es $C \rightarrow A$
4.-	ABC es una superllave de R	Verdadero	-	El subconjunto $\{BC\}$ es una llave candidata.
5.-	ABC es una llave candidata de R	-	Falso	Una llave candidata debe ser mínima.
6.-	BC es la única llave candidata de R	-	Falso	$\{AB\}$ también es una llave candidata

4. Considera la siguiente tabla, donde cada proyecto tiene asignados muchos empleados y cada empleado trabaja en muchos proyectos. Se muestra a continuación un extracto de la tabla **Proyectos**:

NumProy	NombreProy	Presupuesto	idEmp	NombreEmp	HrTrabajadas
P22	Cyclone	50000	E1001	Carlos	12
P22	Cyclone	50000	E2002	Juan	50
P21	IBM	20000	E3003	Patricia	40
P21	IBM	20000	E2002	Juan	30
P21	IBM	20000	E1001	Carlos	70

- ¿Qué problemas consideras que puede haber al almacenar los datos en esta tabla? Describe los problemas en términos de las anomalías que se pueden presentar.
Se pueden presentar las siguientes anomalías:
 - Redundancia:
La redundancia se presenta en el almacenamiento de las tuplas NombreProy y Presupuesto.
 - Eliminación:
Si se llega a eliminar a todos los empleados de un proyecto los datos de éste mismo proyecto serán eliminados.
- ¿Cuáles son las dependencias funcionales que cumplen en la relación **Proyectos**?
 - a) $\text{NumProy} \rightarrow \text{NombreProy}, \text{Presupuesto}.$
 - b) $\text{idEmp} \rightarrow \text{NombreEmp}.$
 - c) $\text{NumProy} + \text{idEmp} \rightarrow \text{HrTrabajadas}.$
- ¿Cuál sería alguna llave para la relación **Proyectos**?
Proponemos la llave compuesta **NumProy+idEmp**

- ¿La relación **Proyectos** cumple con **BCNF**? Justifica tu respuesta.
No, no cumple ya que en las dependencias funcionales a y b no son superllaves de la relación **Proyectos**.
5. Dado el conjunto de dependencias funcionales $F = \{A \rightarrow DC, BC \rightarrow F, F \rightarrow CH, AF \rightarrow B\}$
- Indica alguna llave candidata
AF
 - Encontrar las cerraduras: $\{AC\}^+$, $\{BA\}^+$, $\{AF\}^+$ y $\{A\}^+$
 $\{AC\}^+ = \{ACDBFH\}$
 $\{BA\}^+ = \{BADCFH\}$
 $\{AF\}^+ = \{AFDCBH\}$
 $\{A\}^+ = \{ADC\}$
 - Indicar si las siguientes dependencias funcionales son deducidas o no, justificando tu respuesta:
 $AC \rightarrow H$ No se puede porque solo una empieza con A, despues daria DC y necesitariamos transitividad con alguna de estas pero no hay ninguna que empieze con D, ya que el aumento agregaria una letra que no nos serviria ya que agregar C no serviria, por esto cualquier aplicacion de pseudo transividad nos incluiria D al inicio AD.
 $BA \rightarrow FC$ No se puede, si comenzando con $A \rightarrow DC$ y aunque se descomponga no se podria enlazar con ninguno sin agregar algo que no querramos a la parte izquierda
 $AF \rightarrow HB$ No, empezando por $A \rightarrow DC$ si se descompone o no para agregar algo se agregaria algo al lado izquierdo con pseudo-transitividad, si empezamos con AF, tampoco se puede utilizar nada que no agregase algo al lado izquierdo , y como falta un H no es Suficiente
 $A \rightarrow F$ No porque aunque se descomponga $A \rightarrow DC$ no podria conectarse con ningun otro sin agregar al al lado izquierdo
6. Para cada uno de los esquemas que se muestran a continuación:
- a $R(A, B, C, D, E, F, G)$ con $F = \{AB \rightarrow C, AB \rightarrow F, A \rightarrow D, A \rightarrow E, B \rightarrow G\}$
- b $R(A, B, C, D, E, F)$ con $F = \{AB \rightarrow C, BC \rightarrow AD, D \rightarrow E, CF \rightarrow B\}$
- Especifica de ser posible **dos DF no triviales** que se pueden derivar de las dependencias funcionales dadas
a- $AB \rightarrow CF, ABD \rightarrow E$
b- $CF \rightarrow AD, ABC \rightarrow E$
 - Indica **alguna llave candidata** para **R**
a- AB
b- ABC
 - Especifica **todas las violaciones** a la **BCNF**
a- $\{AB\}^+ = \{ABCFDEG\}$, $\{A\}^+ = \{ADE\}$ esta seria una violacion, $\{B\}^+ = \{BG\}$ entonces es violacion
Entonces las violaciones son : $\{A \rightarrow D, A \rightarrow E, B \rightarrow G\}$
b- $\{AB\}^+ = \{ABCDE\}$ violacion , $\{BC\}^+ = \{BCADE\}$ violacion , $\{D\}^+ = \{DE\}$ violacion, $\{CF\}^+ = \{CFB\}$ Violacion.
Entonces las violaciones son : $\{AB \rightarrow C, BC \rightarrow AD, D \rightarrow E, CF \rightarrow B\}$
 - **Normaliza** de acuerdo a **BCNF**, asegúrate de indicar cuáles son las relaciones resultantes con sus respectivas dependencias funcionales.
7. Para cada una de las siguientes relaciones con su respectivo conjunto de dependencias funcionales:
- a (A, B, C, D, E, F) con $F = \{AB \rightarrow C, BC \rightarrow AD, D \rightarrow E, CF \rightarrow B\}$
- b (A, B, C, D, E, F, G) con $F = \{A \rightarrow B, CD \rightarrow FG, G \rightarrow E, B \rightarrow D, A \rightarrow C, E \rightarrow A\}$
- Indica **alguna llave** para la relación R e indica las **violaciones** a **3NF** que encuentres en F

- a **ABDF** es una llave en **F**, las violaciones de **F** a **3FN** son: $\{BC \rightarrow AD, CF \rightarrow B\}$
- b **AD** es una llave en **F**, las violaciones de **F** a **3FN** son: $\{B \rightarrow D, E \rightarrow A, D \rightarrow FG\}$
- Encuentra el **conjunto mínimo** de dependencias funcionales equivalente a **F**.
 - a Para encontrar el conjunto mínimo de dependencias funcionales debemos primero encontrar atributos superfluos, entonces:
atributos superfluos por la izquierda:
 - $AB \rightarrow C$ entonces
 ¿A es superfluo? entonces $B \rightarrow C$ entonces $\{B\}^+ = \{B, C, A, D\}$
 ¿B es superfluo? entonces $A \rightarrow C$ entonces $\{A\}^+ = \{A, D, E\}$
 \therefore A es superfluo
 - $CF \rightarrow B$ entonces
 ¿C es superfluo? entonces $F \rightarrow B$ entonces $\{F\}^+ = \{F\}$
 ¿F es superfluo? entonces $C \rightarrow B$ entonces $\{C\}^+ = \{C, B, A, D, E\}$
 \therefore F es superfluo
 - \therefore el conjunto mínimo de **F** es: $\{B \rightarrow C, C \rightarrow B\}$
 - b \therefore el conjunto mínimo de **F** es: $\{A \rightarrow B, G \rightarrow E, A \rightarrow C\}$
- **Normaliza** de acuerdo a la **3NF**. Indica claramente las relaciones resultantes y en cada esquema, las dependencias funcionales que se cumplen.
 - a $\mathbf{F}_{\min} = \{B \rightarrow C, C \rightarrow B\}$
 $\mathbf{R}_1 = (\mathbf{B}, \mathbf{C})$ con $B \rightarrow C$
 $\mathbf{R}_2 = (\mathbf{B}, \mathbf{C})$ con $C \rightarrow B$
 $\mathbf{R}_3 = (\mathbf{A}, \mathbf{B}, \mathbf{D}, \mathbf{F})$ con $ABDF \rightarrow ABDF$
 - b $\mathbf{F}_{\min} = \{A \rightarrow B, G \rightarrow E, A \rightarrow C\}$ $\mathbf{R}_1 = (\mathbf{A}, \mathbf{B})$ con $A \rightarrow B$
 $\mathbf{R}_2 = (\mathbf{G}, \mathbf{E})$ con $G \rightarrow E$
 $\mathbf{R}_3 = (\mathbf{A}, \mathbf{C})$ con $A \rightarrow C$
 $\mathbf{R}_4 = (\mathbf{A}, \mathbf{D})$ con $AD \rightarrow AD$

8. Sea el esquema **R** (**A, B, C, D, E, F**). Se sabe que **F** es una llave candidata y se cumplen las dependencias funcionales **F** = $\{BD \rightarrow E, CD \rightarrow A, E \rightarrow C, B \rightarrow D\}$. Responde lo siguiente:

- ¿Qué puedes decir de $\{A\}^+$ y $\{F\}^+$?
 Ninguna es llave de **R** ya que a $\{A\}^+$ el resto de los atributos al igual que a $\{F\}^+$ (**B, C, D, E**)
- Calcula $\{B\}^+$, ¿qué puedes decir de esta cerradura?
 $\{B\}^+ = \{B, D, R, C, A, D\}$
 Es llave de **F**
- Obtén todas las **llaves candidatas**.
 - $\{B\}^+$
 - $\{D\}^+$
 - $\{EDB\}^+$
 - $\{CB\}^+$
 - $\{ABCDE\}^+$
- ¿**R** cumple con **BCNF**? ¿Cumple con **3NF**? (en caso contrario normaliza)
 Sea $\phi = \{BDCE\}$, ahora calcularemos la $\{BDCE\}^+ = \{BCDEA\}$, $\therefore \phi$ es superllave, ahora dado que ϕ contiene todos los atributos de lado izquierdo para toda **DF** entonces cumple la **BCNF**.
 Por otro lado una de las condiciones para que cumpla la **3NF** es que todos los atributos de lado izquierdo para todas las **DF** sea una superllave entonces también cumple la **3NF**
- Se ha decidido dividir **R** en las siguientes relaciones **S** (**A, B, C, D, F**) y **T** (**C, E**), ¿se puede recuperar la información de **R**?
 Si ya existe el atributo **C** en ambas relaciones y podrías obtener **R** si usas llamadas con join en **S** y **T**