



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

FUNDAMENTOS DE BASES DE DATOS - 7094

T A R E A 5

EQUIPO:

DEL MONTE ORTEGA MARYAM MICHELLE - 320083527

SOSA ROMO JUAN MARIO - 320051926

CASTILLO HERNÁNDEZ ANTONIO - 320017438

ERIK EDUARDO GÓMEZ LÓPEZ - 320258211

FECHA DE ENTREGA:

5 DE NOVIEMBRE DE 2024

PROFESOR:

M. EN I.Z GERARDO AVILÉS ROSAS

AYUDANTES:

LUIS ENRIQUE GARCÍA GÓMEZ

KEVIN JAIR TORRES VALENCIA

RICARDO BADILLO MACÍAS

ROCÍO AYLIN HUERTA GONZÁLEZ



Tarea 5

1. 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 $\{A, D, E, G\}$		✓	$\{BC\}^+ = \{BCDEGA\}$
2	Todos los atributos de R están en la cerradura de BC	✓		
3	La cerradura de AC es $\{A, C\}$	✓		
4	ABC es una superllave de R	✓		Como BC es llave pues tiene todos los atributos de R , agregar A significa que es superllave.
5	ABC es una llave candidata de R		✓	Contiene redundancia podemos eliminar A sin destruir la propiedad de identificación única .
6	BC es la única llave candidata de R		✓	Porque $\{AB\}^+ = \{ABCDEG\}$ cumple con identificación única y no redundancia .

2. 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. 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.

Algunos de los principales problemas que vemos en la tabla de este esquema donde cada proyecto puede tener múltiples empleados y cada empleado puede trabajar en múltiples proyectos son:

- **Redundancia de datos:** Esto es debido a que los números de proyectos (*NumProy*), nombres de proyectos (*NombreProy*), presupuestos (*Presupuesto*), (*idEmp*) y nombres de empleados (*NombreEmp*) se repiten.
- **Anomalía de inserción:** No podemos agregar un proyecto sin asociarlo a al menos un empleado y viceversa.
- **Anomalía de actualización:** Si un empleado cambia de nombre, o si se actualiza el nombre de un proyecto o si se cambia de presupuesto, (*en general, si actualiza alguna información*) se debe actualizar en todas las filas donde aparezca información relacionada lo que puede llevar a inconsistencias.
- **Anomalía de eliminación:** Si se elimina una fila, se puede perder información sobre el proyecto o el empleado si no hay otras filas que contengan esa información.

- ¿Cuáles son las dependencias funcionales que cumplen en la relación **Proyectos**?

Las dependencias funcionales en la tabla **Proyectos** se identifican observando cómo ciertos atributos determinan de forma única a otros. Analizamos la relación entre los atributos en cada fila, lo que nos permite inferir las siguientes dependencias:

- **NumProy \rightarrow NombreProy, Presupuesto:** Cada número de proyecto (*NumProy*) está asociado de manera única a un nombre de proyecto (*NombreProy*) y un presupuesto. Esto significa que, si conocemos el número de proyecto, podemos deducir su nombre y presupuesto. **Por Ejemplo:** El proyecto con *NumProy* = P22 siempre tiene el nombre Cyclone y el presupuesto de 50000 en todas las filas donde aparece.
- **idEmp \rightarrow NombreEmp:** Cada empleado tiene un identificador único (*idEmp*) que determina su nombre. Esto significa que, con el identificador del empleado, es posible determinar un solo valor para el nombre del empleado. *Veamos el siguiente ejemplo:* El empleado con *idEmp* = E1001 tiene siempre el nombre Carlos en las filas en las que aparece.
- **NumProy, idEmp \rightarrow HrTrabajadas:** Las (*HrTrabajadas*) dependen tanto del proyecto como del empleado, ya que cada combinación de proyecto-empleado define una cantidad única de horas trabajadas. Por lo tanto, se requiere conocer ambos valores (*NumProy* e *idEmp*) para determinar las horas trabajadas. Por ejemplo: La combinación de (P22, E1001) da *HrTrabajadas* = 12, mientras que (P22, E2002) da *HrTrabajadas* = 50.

- ¿Cuál sería alguna llave para la relación **Proyectos**?

Una llave candidata para la relación **Proyectos** es el conjunto de atributos **{NumProy, idEmp}** ya que la combinación de **NumProy** (Número de Proyecto) y **idEmp** (Identificación del Empleado) es única para cada registro en la tabla. En otras palabras, cada par de **NumProy** e **idEmp** determina de manera única las **HrTrabajadas** (Horas Trabajadas) por un empleado en un proyecto específico. Por lo tanto, **{NumProy, idEmp}** cumple con los requisitos de una llave candidata, ya que no hay dos filas en la tabla que tengan la misma combinación de estos valores.

- ¿La relación **Proyectos** cumple con **BCNF**? Justifica tu respuesta.

Para determinar si la relación **Proyectos** cumple con la (BCNF), necesitamos verificar si todas las dependencias funcionales no triviales en la relación tienen como lado izquierdo una superllave. Recordemos que una relación está en BCNF si, para toda dependencia funcional $X \rightarrow Y$, X es una superllave de la relación.

En esta relación **Proyectos**, identificamos las siguientes dependencias funcionales:

- **NumProy \rightarrow NombreProy, Presupuesto**: Esto indica que el número de proyecto (*NumProy*) determina de manera única tanto el nombre del proyecto como su presupuesto. Sin embargo, **NumProy** no es una superllave de la relación, ya que no determina de manera única todos los atributos de la tabla (por ejemplo, no determina *idEmp* ni *HrTrabajadas*). Esta dependencia funcional hace que la relación **Proyectos** no esté en BCNF.
- **idEmp \rightarrow NombreEmp**: Aquí, el identificador del empleado (*idEmp*) determina de forma única el nombre del empleado (*NombreEmp*). Similar al caso anterior, **idEmp** no es una superllave de la relación, pues no determina atributos como *NumProy* o *HrTrabajadas*. Esto también viola la condición de BCNF.
- **NumProy, idEmp \rightarrow HrTrabajadas**: La combinación de **NumProy** e **idEmp** determina de manera única las horas trabajadas (*HrTrabajadas*), y esta combinación es una llave candidata de la relación, ya que identifica de forma única todas las tuplas en la tabla. Esta dependencia sí cumple con la condición de BCNF.

Aunque **NumProy, idEmp \rightarrow HrTrabajadas** cumple con BCNF, las dependencias **NumProy \rightarrow NombreProy, Presupuesto** y **idEmp \rightarrow NombreEmp** no lo hacen, ya que sus lados izquierdos no son superllaves. \implies **la relación Proyectos no está en BCNF** debido a que existen dependencias funcionales donde el determinante no es una superllave. Para cumplir con BCNF, sería necesario descomponer la tabla en relaciones que eliminen estas dependencias.

3. Para cada uno de los **esquemas** que se muestran a continuación, con su respectivo **conjunto de dependencias funcionales**:

- 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 $F = \{AB \rightarrow C, BC \rightarrow AD, D \rightarrow E, CF \rightarrow B\}$

■ Indica **alguna llave candidata** para la relación **R**.

- a. $\{AB\}^+ = \{ABCFDEG\}$ es llave candidata pues cumple con **identificación única** por tener a todos los atributos de **R** y **no redundancia**. pues si eliminamos a cualquiera de sus atributos, no se cumple con la identificación única.
- b. $\{CF\}^+ = \{CFABDE\}$ notemos que aquí tenemos que incluir a F pues la única manera de agregarlo es directamente (no existe DF con F a la derecha) y no podemos quitar a ninguno de los atributos de la llave candidata pues no se cumple con la identificación única.

■ Especifica **todas las violaciones** a la **BCNF**.

- a. Calculamos la cerraduras de los lados izquierdos: (ninguna es trivial)

$$\begin{aligned}\{AB\}^+ &= \{ABCFDEG\} \text{ Es llave} \\ \{AB\}^+ &= \{ABCFDEG\} \text{ Es llave} \\ \{A\}^+ &= \{ADE\} \text{ Violación a BCNF, la elegimos para normalizar} \\ \{A\}^+ &= \{ADE\} \text{ Violación a BCNF} \\ \{B\}^+ &= \{BG\} \text{ Violación a BCNF}\end{aligned}$$

- b. Calculamos la cerraduras de los lados izquierdos: (ninguna es trivial)

$$\begin{aligned}\{AB\}^+ &= \{ABCDE\} \text{ Violación a BCNF, la elegimos para normalizar} \\ \{BC\}^+ &= \{BCADE\} \text{ Violación a BCNF} \\ \{D\}^+ &= \{DE\} \text{ Violación a BCNF} \\ \{CF\}^+ &= \{CFBADE\} \text{ Es llave}\end{aligned}$$

■ **Normaliza** de acuerdo con **BCNF**, asegúrate de indicar cuáles son las **relaciones resultantes** con sus respectivas **dependencias funcionales**.

- a. Elegimos la primera violación y dividimos **R**:

$$\begin{aligned}R_1(A, D, E) \text{ con } F &= \{A \rightarrow D, A \rightarrow E\}; \{A\}^+ = \{ADE\} \rightarrow A \text{ es llave para } R_1 \\ R_2(A, B, C, F, G) \text{ con } F &= \{AB \rightarrow C, AB \rightarrow F, B \rightarrow G\}\end{aligned}$$

En R_1 no hay violación a BCNF, revisamos R_2 :

$$\begin{aligned}\{AB\}^+ &= \{ABCFG\} \text{ Es llave para } R_2 \\ \{B\}^+ &= \{BG\} \text{ Violación a BCNF, la elegimos para normalizar}\end{aligned}$$

Divido R_2 :

$$\begin{aligned}R_3 &= (B, G) \text{ con } F = \{B \rightarrow G\}; \{B\}^+ = \{BG\} \rightarrow B \text{ es llave para } R_3 \\ R_4 &= (B, A, C, F) \text{ con } F = \{AB \rightarrow C, AB \rightarrow F\}; \{AB\}^+ = \{ABCF\} \rightarrow A \text{ es llave para } R_4\end{aligned}$$

Finalmente:

$$\begin{aligned} R_1(A, D, E) \text{ con } F &= \{A \rightarrow D, A \rightarrow E\} \\ R_3(B, G) \text{ con } F &= \{B \rightarrow G\} \\ R_4(B, A, C, F) \text{ con } F &= \{AB \rightarrow C, AB \rightarrow F\} \end{aligned}$$

b. Elegimos la primera violación y dividimos **R**:

$$\begin{aligned} R_1(A, B, C, D, E) \text{ con } F &= \{AB \rightarrow C, BC \rightarrow AD, D \rightarrow E\}; \\ R_2(A, B, F) \text{ con } F &= \{ABF \rightarrow ABF\}; \rightarrow ABF \text{ es llave para } R_2 \end{aligned}$$

Aqui observamos 2 cosas, la primera es que R_2 al no cumplir ninguna dependencia funcional, solo se tiene la trivial y trivialmente es llave, ademas de esto, vemos que perdemos la DF $CF \rightarrow B$ por lo que deberiamos parar la normalización en R_2 , pero como el profesor hizo vamos a seguir por fines didacticos. Revisamos R_1 :

$$\begin{aligned} \{AB\}^+ &= \{ABCDE\} \text{ Es llave para } R_1 \\ \{BC\}^+ &= \{BCADE\} \text{ Es llave para } R_1 \\ \{D\}^+ &= \{DE\} \text{ Violación a BCNF, la elegimos para normalizar} \end{aligned}$$

Dividimos R_1 :

$$\begin{aligned} R_3(D, E) \text{ con } F &= \{D \rightarrow E\}; \{D\}^+ = \{DE\} \rightarrow D \text{ es llave para } R_3 \\ R_4(D, A, B, C) \text{ con } F &= \{AB \rightarrow C, BC \rightarrow AD\}; \end{aligned}$$

Revisamos R_4 :

$$\begin{aligned} \{AB\}^+ &= \{ABCD\} \text{ Es llave para } R_4 \\ \{BC\}^+ &= \{BCAD\} \text{ Es llave para } R_4 \end{aligned}$$

Finalmente:

$$\begin{aligned} R_2(A, B, F) \text{ con } F &= \{ABF \rightarrow ABF\} \\ R_3(D, E) \text{ con } F &= \{D \rightarrow E\} \\ R_4(D, A, B, C) \text{ con } F &= \{AB \rightarrow C, BC \rightarrow AD\} \end{aligned}$$

Como nota importante, se presenta join con perdida, DF perdidas: $CF \rightarrow B$

- 4.
- 5.
- 6.