

Segundo Parcial de Fundamentos de Bases de Datos

Noviembre 2022

Indicaciones Generales:

- La duración de la prueba es de **tres (3)** horas.
- En la prueba **NO** se permite consultar material alguno.
- Empezar cada ejercicio en una hoja nueva.
- Escribir con lápiz y de un solo lado de las hojas.
- Numerar todas las hojas. Incluir en cada hoja la cédula y el nombre. En la primer hoja, incluir la cantidad de hojas que se entregan.

Ejercicio 1. (10 puntos)

Considere los siguientes esquemas relación con sus respectivos conjuntos de dependencias:

R1 (A,B,C,D) F1 = {CB → A, B → A, D → AC }

R2 (A,C,E,G) F2 = {E → C, G → AE}

R3 (A,B,E) F3 = {A → BE, B → A, E → BA }

- (a) ¿Cuál es la máxima forma normal en que se encuentra cada uno de los esquemas relación anteriores? Justificar la respuesta.

Solución.

R1 (A,B,C,D) F1 = {CB → A, B → A, D → AC }

Claves: B y D pertenecen a todas las claves ya que no pertenecen a los lados derechos de las dependencias de F1.

$(BD)^+ = B,D,A,C = R1$, por lo tanto BD es la única clave de R1.

$D \rightarrow A$, D es parte de una clave y A no es primo, la dependencia viola la condición de 2NF.

R1 esta en 1NF

R2 (A,C,E,G) F2 = {E → C, G → AE}

Claves: G pertenece a todas las claves ya que no pertenece a los lados derechos de las dependencias.

$G^+ = G,C,E,A = R2$, por lo tanto G es la única clave de R2.

$E \rightarrow C$ E no es superclave, por lo tanto la df viola las condiciones de BCNF. E no es superclave y C no es primo (no pertenece a la única clave), por lo tanto la df viola las condiciones de 3NF.

E no es parte de una clave por lo que se cumplen las condiciones de 2NF.

$G \rightarrow AE$

G es superclave por lo tanto $G \rightarrow AE$ cumple las condiciones de BCNF.

Por lo tanto R2 se encuentra en 2NF

R3 (A,B,E) F3 = {A → BE, B → A, E → BA }

$A^+ = A,B,E = R3$ $B^+ = B,A,E = R3$

Por lo tanto A y B son superclaves. $E \rightarrow BA$ es una dependencia multivaluada trivial ya que involucra a todos los atributos del esquema, por lo tanto **R3 esta en 4NF**

- (b) ¿Cuál es la máxima forma normal en que se encuentra el esquema de base de datos (R1, R2, R3)? Justificar la respuesta.

Solución.

La forma normal en que se encuentra el esquema de base de datos es la mínima forma normal en que se encuentran TODOS los esquemas relación que componen el esquema de bd. Como R1 se encuentra en 1NF por lo tanto el esquema de bd se encuentra en 1NF.

Ejercicio 2. (20 puntos)

Considere el esquema de relación R, R(A,B,C,D,E,G,H) y el conjunto de dependencias F que se cumplen sobre R.

$$F = \{CD \rightarrow AH, G \rightarrow BE, DH \rightarrow G, A \rightarrow B, CE \rightarrow AG\}$$

- (a) Determinar si la descomposición $\rho_1 = (R1, R2)$ con

R1(A,C,H,G)

R2(C,G,E,D,B)

es una descomposición de R con join sin pérdida respecto a F. Justificar la respuesta.

Solución.

Para aplicar el teorema de JSP se debe verificar si se cumple al menos una de las siguientes dependencias:

$$R1 \cap R2 \rightarrow R1 - R2$$

$$R1 \cap R2 \rightarrow R2 - R1$$

Para verificar esto, calculamos $R1 \cap R2$ y su clausura:

$$R1 \cap R2 = \{C, G\}$$

$$CG^+ = \{C, G, B, E, A\}$$

Las dependencias a chequear serían:

$$CG \rightarrow AH$$

$$CG \rightarrow EDB$$

La primera no se cumple dado que H no aparece en la clausura de CG y la segunda tampoco se cumple, dado que D no aparece en la clausura de CG.

Como no se cumple ninguna de esas dependencias, la descomposición tiene pérdida de información, o sea, NO TIENE JSP.

- (b) Dada la siguiente descomposición de R, $\rho_2 = (R1, R2, R3)$ con

R1 (C,D,G,B)

R2 (D,H,E,B)

R3 (C,G,A)

1. Calcular la proyección de F en cada uno de los subesquemas de ρ_2 .

Solución.

La proyección de un conjunto F de dependencias sobre un esquema de relación R, es el conjunto de todas las dependencias de F^+ que tienen ambos lados incluidos en R. Es por esto que, al calcular las proyecciones, deben tenerse en cuenta las clausuras de los atributos o bien las reglas de inferencia de dependencias. Se asume que basta dar un cubrimiento de las dependencias que proyectan, que no tiene por qué ser minimal.

Las proyecciones son las siguientes:

$$\pi_{R1}(F) = \{CD \rightarrow G, G \rightarrow B\}$$

$$\pi_{R2}(F) = \{DH \rightarrow EB\}$$

$$\pi_{R3}(F) = \{CG \rightarrow A\}$$

- II. Determinar si las dependencias de $F : CD \rightarrow A, DH \rightarrow G$ se preservan en la descomposición ρ_2 . Justificar la respuesta.

Solución.

Para determinar si se preservan las dependencias indicadas, hay que verificar si se cumplen con respecto a la unión de los conjuntos de dependencias asociados a $R1, R2$ y $R3$.

Sea $F_1 = \pi_{R1}(F) \cup \pi_{R2}(F) \cup \pi_{R3}(F)$:

$$CD_{F_1}^+ = \{C, D, B, G, A\}$$

$$DH_{F_1}^+ = \{D, H, E, B\}$$

Dado el resultado anterior, la dependencias $CD \rightarrow A$ se preserva, pero la dependencia $DH \rightarrow G$ no se preserva.

- III. ¿ Es posible que en R2 además de las dependencias proyectadas de F se cumpla la dependencia multivaluada no funcional $E \twoheadrightarrow B$? Justificar la respuesta.

Solución.

Para chequear si es posible, se debe estudiar la validez de la dependencia teniendo en cuenta el resto de las dependencias involucradas. Si la dependencia se cumple, las instancias deben tener la siguiente estructura.

Nota: Tid es un identificador de tupla para poder referenciarlas en el resto de la explicación.

Tid	D	H	E	B
1	d_1	h_1	e_1	b_1
2	d_2	h_2	e_1	b_2
3	d_1	h_1	e_1	b_2
4	d_2	h_2	e_1	b_1

En este esquema de relación, se cumple la dependencia funcional $DH \rightarrow EB$. Esta dependencia, obliga a que se cumpla que $b_1 = b_2$ para las tuplas 1 y 3 o 2 y 4. Sin embargo, esto obligaría a que la multivaluada $E \twoheadrightarrow B$ fuera en realidad una dependencia funcional.

Por todo lo expuesto, no es posible que se cumpla esa multivaluada no funcional.

Ejercicio 3. (15 puntos)

Considere el MER de la Figura 1 que representa la base de datos de una academia de baile.

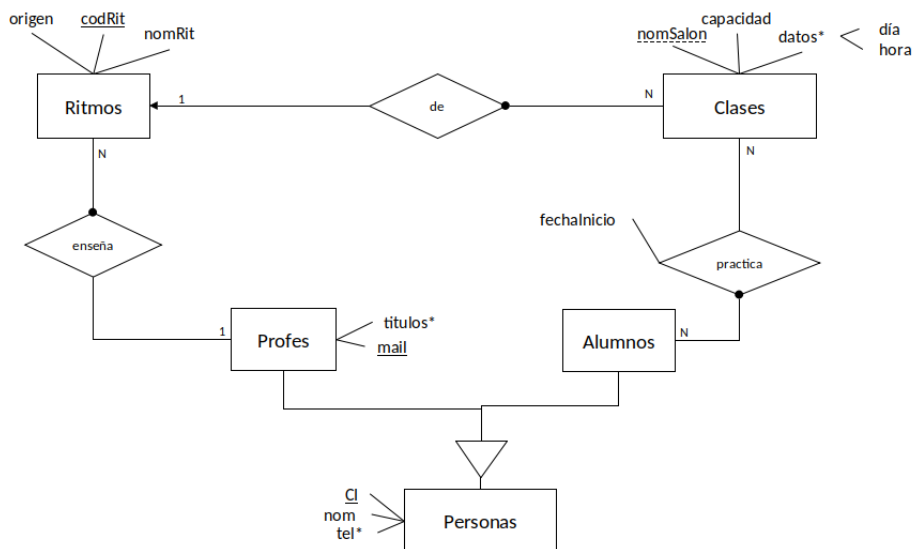


Figura 1: MER academia de baile

Se propone el siguiente esquema R como representación relacional del MER de la Figura 1.

- Ritmos(codRit, origen, nomRit)
- Clases(codRit nomSalon, capacidad)
- ClasesDatos(codRit, nomSalon, dia, hora)
- Personas(ciPers, nomPers)
- Alumnos(ciAl)
- AlumnosTel(ciAl, telAl)
- Profes(ciProf, mail)
- ProfesTelTit(ciProf, telProf, tit)
- Practica(ciAl, nomSalon, codRit, fechaInicio)
- Enseña(ciProf, codRit)

Se sabe que en R se cumplen las siguientes dependencias de inclusión:

- $\Pi_{codRit}(Clases) \subseteq \Pi_{codRit}(Ritmos)$
- $\Pi_{ciAl}(Alumnos) \subseteq \Pi_{ciPers}(Personas)$
- $\Pi_{ciAl}(AlumnosTel) \subseteq \Pi_{ciAl}(Alumnos)$
- $\Pi_{ciProf}(Profes) \subseteq \Pi_{ciPers}(Personas)$
- $\Pi_{ciProf}(ProfesTelTit) \subseteq \Pi_{ciProf}(Profes)$
- $\Pi_{ciAl}(Practica) = \Pi_{ciAl}(Alumnos)$
- $\Pi_{nomSalon, codRit}(Practica) \subseteq \Pi_{nomSalon, codRit}(Clases)$
- $\Pi_{ciProf}(Enseña) \subseteq \Pi_{ciProf}(Profes)$
- $\Pi_{codRit}(Enseña) \subseteq \Pi_{codRit}(Ritmos)$

- (a) El esquema relacional propuesto, ¿representa adecuadamente el MER de la Figura 1? En caso de que no lo represente indique todos los problemas/errores detectados. Justifique usando la metodología de pasaje de MER al modelo Relacional vista en el curso.

Solución.

El esquema relacional propuesto no representa el MER de la figura. Se detectan los siguientes errores.

- La relación **Enseña** representa a una relación N:N y no a una relación N:1 como se indica en la Figura. Para representar a una relación N:1 la clave de la relación debería ser solo el atributo que representa a la entidad que está del lado N de la relación, en este caso **codRit**. Además, al tratarse de una relación N:1 con totalidad del lado N, la entidad con

cardinalidad 1 puede incluirse en la tabla que representa a la entidad con cardinalidad N. Para el MER de la figura, la tabla **Enseña** puede no considerarse e incluir el atributo **ciProf** en la tabla **Ritmos**.

- La relación **ProfesTelTit** representa tanto a los teléfonos como a los títulos de los profesores. Esto es incorrecto ya que ambos atributos son multivaluados y según la metodología vista en el curso, se debe crear una tabla por cada atributo multivaluado.
 - En relación a lo anterior, al no haber RNE para la categorización, no se puede suponer que todas las personas van a ser o bien profesores o bien alumnos. Esta suposición la podemos ver en la inclusión de los teléfonos exclusivamente en los alumnos y en los profesores, representados en las tablas **AlumnosTel** y **ProfesTelTit**.
 - Mirando las dependencias de inclusión, $\Pi_{codRit}(Enseña) \subseteq \Pi_{codRit}(Ritmos)$ debería llevar una igualdad ya que hay una totalidad del lado de **Ritmos** en el MER. En caso de que se eliminara la tabla **Enseña**, esta dependencia de inclusión no sería necesaria.
 - Falta la dependencia de inclusión para la tabla **ClasesDatos** ya que para que una instancia sea válida, debe realmente representar una clase. La dependencia que se debería agregar es $\Pi_{codRit, nomSalon}(ClasesDatos) \subseteq \Pi_{codRit, nomSalon}(ClasesDatos)$
- (b) Ahora, indique qué dependencias funcionales se cumplen sobre cada una de las relaciones de *R*, con respecto al MER de la Figura 1. Si hay dependencias multivaluadas no triviales sobre alguna de las tablas de *R*, indíquelas y justifique su respuesta.

Solución.

Las dependencias funcionales que se cumplen en cada relación son las siguientes:

Ritmos(codRit, origen, nomRit)

codRit → origen, nomRit

Clases(codRit nomSalon, capacidad)

codRit, nomSalon → capacidad

ClasesDatos(codRit, nomSalon, dia, hora)

- **Personas**(ciPers, nomPers)

ciPers → nomPers

Alumnos(ciAl)

-

AlumnosTel(ciAl, telAl)

-

Profes(ciProf, mail)

ciProf → mail

mail → ciProf

ProfesTelTit(ciProf, telProf, tit)

-

Practica(ciAl, nomSalon, codRit, fechaInicio)

ciAl, nomSalon, codRit → fechaInicio

Enseña(ciProf, codRit)

codRit → ciProf

Las dependencias multivaluadas no triviales son las siguientes:

ciProf → telProf

ciProf → tit

Estas dmv surgen de la tabla **ProfesTelTit**, que representa los teléfonos y los títulos de los profesores que son multivaluados e independientes entre sí.

- (c) Analice si las claves de las relaciones de *R* son correctas en función de la parte b. Justifique.

Solución.

De acuerdo a las dependencias funcionales identificadas en la parte b se puede observar que la única clave identificada que viola alguna de las dependencias funcionales detectadas es la tabla Enseña.

- (d) Finalmente, teniendo en cuenta las partes b y c, determine si el esquema R está en 4NF, justificando su respuesta.

Solución.

El esquema no está en 4NF porque la tabla ProfTelTit tiene dependencias multivaluadas no triviales, como fue indicado en la parte b.

Ejercicio 4. (15 puntos)

Considere el grafo de seriabilidad de la Figura 2:

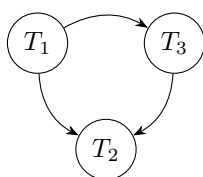


Figura 2: Grafo de Seriabilidad

Considere además las siguientes transacciones:

$$T_1 : r_1(x), w_1(z), c_1.$$

$$T_2 : r_2(y), w_2(x), c_2.$$

$$T_3 : r_3(z), w_3(y), c_3.$$

- (a) Indique si una historia con el grafo de la Figura 2 es o no serializable. Justifique su respuesta.
- (b) I. Copie la siguiente tabla en su hoja e indique en ella las operaciones en conflicto de las transacciones mencionadas, considerando en cada caso el orden indicado por los arcos del grafo.

Arco	Operaciones en Conflicto
$T_1 \rightarrow T_2$	
$T_1 \rightarrow T_3$	
$T_2 \rightarrow T_3$	

Solución.

Arco	Operaciones en Conflicto
$T_1 \rightarrow T_2$	$r_1(x), w_2(x)$
$T_1 \rightarrow T_3$	$w_1(z), r_3(z)$
$T_2 \rightarrow T_3$	$r_2(y), w_3(y)$

- II. Construya una historia H en donde al menos dos transacciones se entrelacen en operaciones de read y write, su grafo de seriabilidad sea el de la Figura 2 y que utilice las transacciones de la parte anterior.

Solución.

Una forma de construir la historia pedida, es copiar las tres transacciones en orden y luego mover “hacia atrás” las operaciones de forma que queden entrelazadas, respetando los conflictos de la parte anterior.

Un resultado posible es el siguiente:

$$H = r_1(x), r_2(y), w_1(z), c_1, w_2(x), r_3(z), c_2, w_3(y), c_3$$

- III. Considere las siguientes clases de historias vistas en el curso: no recuperable, recuperable, evita abortos en cascada, estricta. Indique a cuales de estas clases pertenece la historia H que construyó en la parte II, justificando su respuesta.

Solución.

Si se hace una revisión de la historia anterior, se puede ver que:

- T_1 no lee de ninguna otra transacción.
- T_2 no lee de ninguna otra transacción.
- T_3 lee de T_1 el ítem Z de ninguna otra transacción, y lo lee después del commit de T_1 .

Dado que no existen lecturas de transacciones sin confirmar, la historia es **estricta**.

- (c) Considere el log de la Figura 3 que usa las transacciones anteriores.

1.[start, T_1]	9.[read, T_3 , Z , 20]
2.[read, T_1 , X , 10]	10.[write, T_2 , X , 10, 21]
3.[write, T_1 , Z , 15, 20]	11.[commit, T_2]
4.[commit, T_1]	12.[write, T_3 , Y , 10, 5]
5.[checkpoint]	
6.[start, T_2]	
7.[read, T_2 , Y , 10]	
8.[start, T_3]	



Figura 3: Log de la Base

NOTA: En los registros 3, 10 y 12, el primer valor es el *before image* del ítem y el segundo el *after image*.

- I. Indique que transacciones se deben ignorar, rehacer o deshacer en el caso de una estrategia de *Actualización Inmediata (redo/undo)*. Justifique su respuesta.

Solución.

En las estrategias *redo/undo* no hay nada que garantice que los datos quedaron grabados en la base si el commit se realiza después del checkpoint. Por esto es que se asume que las transacciones que confirmaron luego del checkpoint se deben rehacer, las que no llegaron a confirmar se deben deshacer y las que confirmaron antes del checkpoint se ignoran.

Conclusión:

- T_1 : se debe ignorar.
- T_2 : se debe rehacer.
- T_3 : se debe deshacer y relanzar al final.

- II. Indique que transacciones se deben ignorar, rehacer o deshacer en el caso de una estrategia de *Actualización Diferida (redo/no undo)*. Justifique su respuesta.

Solución.

En las estrategias *redo/no undo* existe la garantía que las transacciones que no confirmaron, no dejaron datos en la base. Por esto, no es necesario deshacer ninguna transacción. Sólo se deben rehacer las transacciones que confirmaron luego del checkpoint.

Conclusión:

- T_1 : se debe ignorar.
- T_2 : se debe rehacer.
- T_3 : se relanza al final.

Preguntas del laboratorio (10 puntos)

- (a) Considere el siguiente diagrama entidad relación sobre sagas y películas extraído de la tarea 3 del laboratorio.

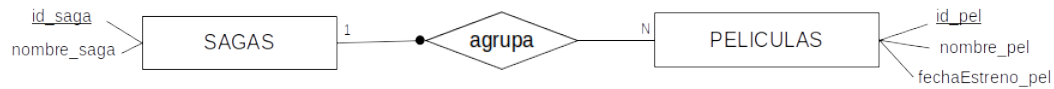


Figura 4: MiniMovies simplificado

Para simplificar y mejorar una consulta sobre sagas que se sabe es muy frecuente, quienes diseñan la base de datos decidieron agregar el atributo `cant_pels_saga` en el esquema relación `peliculas`. Este nuevo atributo representa la cantidad de películas de la saga. A continuación se presenta la tabla `peliculas` resultante. Como en la Tarea 3, no se definieron claves primarias en este esquema.

```
peliculas(id_pel, nombre_pel, fechaEstreno_pel, id_saga, nombre_saga, cant_pels_saga)
```

- I. Explique brevemente una anomalía de inserción en la que el valor del atributo `cant_pels_saga` puede quedar inconsistente.
- II. Se ha determinado hasta ahora el siguiente conjunto de dependencias funcionales

$$F_{peliculas} = \{ id_pel \rightarrow nombre_pel \text{ fechaEstreno_pel}, id_saga \rightarrow nombre_saga \}$$
 Complete las dependencias funcionales que se cumplen en este esquema tomando en cuenta el MER y el nuevo atributo. Evitar redundancia tanto de atributos como de dependencias funcionales.

(b) Considere la siguiente consulta, propuesta en la letra de la Tarea 4.

```

SELECT j.name, m.name
FROM jobs j, casts c, movies m, movie_genres mg
WHERE j.id = c.job_id
AND c.movie_id = m.id AND m.id = mg.movie_id
AND j.name = 'Director' AND mg.genre_id = 18
    
```

Se propone el siguiente plan lógico luego de aplicar algunas de las heurísticas vistas en el curso.

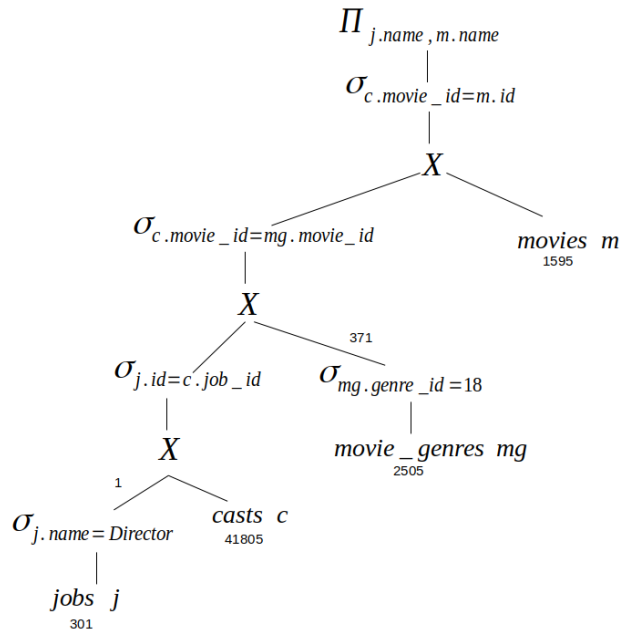


Figura 5: Plan lógico para la Consulta

Indique si el plan lógico de la Figura 5 corresponde a aplicar correctamente las heurísticas vistas en el curso, y de lo contrario indique cuáles son los problemas que detecta.