

Segundo Parcial de Fundamentos de Bases de Datos

Noviembre 2021

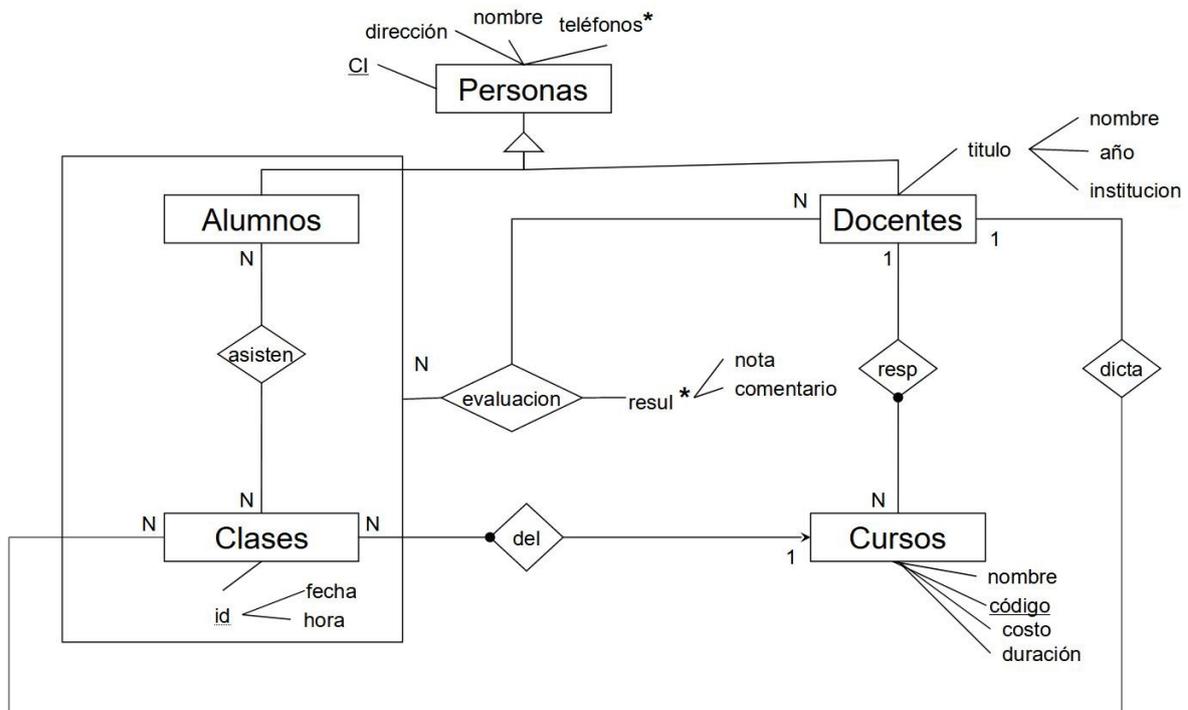
Duración: 3 horas

Presentar la resolución del parcial:

- Con las hojas numeradas y escritas de un solo lado. Comience cada ejercicio en una hoja nueva.
- Con la cantidad de hojas entregadas en la primer hoja.
- Con cédula de identidad y nombre en cada hoja.
- **Escrita a lápiz y en forma prolija.**

Ejercicio 1 (18 puntos)

Considere el siguiente Modelo de Entidad Relación (MER) con información sobre cursos y un conjunto de Restricciones No Estructurales (RNEs) sobre él.



Restricciones no estructurales:

1) $\text{Alumnos} \cap \text{Docentes} = \emptyset$

2) Un alumno que asiste a una clase que se dicta, debe tener una evaluación donde participa el docente que la dicta

$$(\forall d \in \text{docentes}) (\forall a \in \text{alumnos}) (\forall c \in \text{clases}) \\ (\langle a, c \rangle \in \text{asisten} \wedge \langle d, c \rangle \in \text{dicta} \rightarrow \langle d, \langle a, c \rangle \rangle \in \text{evaluacion})$$

3) Un alumno no puede asistir a más de una clase en a la misma fecha, hora.

$$(\forall a \in \text{alumnos}) (\forall c1, c2 \in \text{clases}) \\ (\langle a, c1 \rangle \in \text{asisten} \wedge \langle a, c2 \rangle \in \text{asisten} \rightarrow \text{id}(c1) \neq \text{id}(c2))$$

5) Toda persona tiene al menos un teléfono.

$(\forall p \in \text{Personas}) (\text{COUNT}(\text{telefonos}(p)) > 0)$

6) El costo de los cursos es mayor o igual 0, mientras que la duración es mayor a 0.

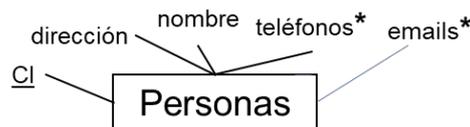
$(\forall c \in \text{Cursos}) (c.\text{costo} \geq 0 \wedge c.\text{duracion} > 0)$

Se pide:

1. Teniendo en cuenta el MER y las restricciones no estructurales presentadas:

- elabore un esquema relacional que incluya los esquemas relación así como las claves (tanto las primarias como las alternativas) y las dependencias de inclusión,
- indicar las dependencias funcionales que se cumplen y que no han sido ya expresadas sobre cada una de las relaciones del esquema relacional obtenido en la parte (a). **NO** incluir ni triviales ni redundantes.
- indicar las restricciones no estructurales que no fueron posibles expresar ni en (a) y ni en (b). **NO** se pide la elaboración de un método para controlarlas.

2. Considere que ahora se quiere incluir la información de las direcciones de correo (emails) de las personas. Para ello, se modifica la entidad Personas:



y se define el siguiente esquema relación:

PERSONAS(cí, dirección, nombre, teléfono, email)

- Indique las dependencias funcionales que se cumplen en este esquema relación dada la modificación de la entidad Personas.
- Se tiene información de una persona llamada Draco Malfoy, de la que se conoce que su cédula es 7.563.107-8, su dirección es Rivera 3456, sus teléfonos son 2467.879 y 4345.2222 y su emails son dracom@gmail.com y dmalfoy@hotmail.com. Dar una instancia del esquema Personas que represente a esta información.
- ¿Qué dificultades encuentra en este esquema? Usando conceptos de normalización vistos en el curso, ¿cómo lo mejoraría?

Ejercicio 2 (18 puntos)

Considere el esquema de relación $R(A,B,C,D,E,G,H)$, el conjunto de dependencias funcionales $F = \{C \rightarrow D, A \rightarrow BH, CA \rightarrow E, H \rightarrow G, HE \rightarrow CA\}$, y la siguiente descomposición:

$$\rho = \{R1(A,B,C,E), R2(A,D,E,G,H)\}$$

1) Para cada una de las afirmaciones, indique si es verdadera o falsa, **justificando**.

- C no pertenece a ninguna clave de R según F.
- (AEH) es clave de R según F.
- R está en 2NF respecto a F.
- La descomposición ρ tiene JSP.
- $\Pi_{R2}(F) = \{H \rightarrow G\}$

2)

- Encuentre todas las claves de R según F.
- Obtenga una descomposición en 3NF de R con JSP y preservación de dependencias, aplicando el algoritmo visto en el curso y justificando cada paso.

Ejercicio 3 (12 puntos)

Considere el esquema relacional de un subconjunto de la base de datos del sitio IMDB, que almacena, entre otras cosas, información sobre películas, y que fue utilizada en la tarea de laboratorio del curso.

PELICULAS(pel_i_d, nombre, fecha, duracion, presupuesto, ganancia, web, votos_prom)

Contiene información sobre las películas: su identificador, su nombre (en el idioma original), su fecha de estreno, su duración en minutos, el presupuesto que se tenía para realizar la película, los ingresos que generó, un sitio web (que puede no existir) y el promedio de votos que la película ha recibido en el sitio IMDB hasta el momento.

SAGAS(saga_id, nombre, fecha)

Contiene información sobre las sagas: su identificador, su nombre y la fecha de estreno de la última película perteneciente a la saga.

PELI_EN_SAGA(saga_id, pel_i_d)

Almacena las películas que pertenecen a cada saga.

$$\begin{aligned} \Pi_{\text{pel}_i\text{d}}(\text{PELI_EN_SAGA}) &\subseteq \Pi_{\text{pel}_i\text{d}}(\text{PELICULAS}) \\ \Pi_{\text{saga_id}}(\text{PELI_EN_SAGA}) &\subseteq \Pi_{\text{saga_id}}(\text{SAGAS}) \end{aligned}$$

Se sabe que todos los atributos tienen distribución uniforme (recuerde que $V(A,T)$ es la cantidad de valores distintos que tiene un atributo A en una tabla T). Además, de cada tabla se conoce la siguiente información:

Relación R	n_R	Atributos	Índices
PELICULAS P	5000	$V(\text{fecha}, P) = 1000$	Índice primario sobre pel_i_d
SAGAS S	80		Índice primario sobre saga_id
PELI_EN_SAGA PS	1000		Índice Primario sobre (saga_id, pel_i_d)

Considere la siguiente consulta sobre el esquema dado:

```

SELECT p.id, p.name
FROM PELI_EN_SAGA ps, SAGAS s, PELICULAS p
WHERE ps.saga_id = s.id AND p.id=ps.peli_id AND p.fecha= '1-1-2020';

```

Se pide:

1. Dar el árbol canónico para la consulta.
2. Aplicar las heurísticas para optimización llegando al plan lógico optimizado. Explique cada uno de los pasos ejecutados.
3. Dar un plan físico para el plan lógico obtenido en la **parte 2**, utilizando índices en los casos en que sea posible.
4. Considere ahora que se agrega un índice secundario en el atributo fecha de la tabla PELÍCULAS, ¿haría algún cambio al plan físico propuesto en la **parte 3**?
5. Se decide modificar el índice primario de la tabla PELI_EN_SAGA, de forma que el nuevo índice primario sea (peli_id, saga_id), ¿haría algún cambio al plan físico propuesto en la **parte 3**?
6. Finalmente, se desea agregar a la consulta una selección sobre la tabla SAGAS que no afecta el orden de las hojas del árbol obtenido en la **parte 2**. Además, esa selección se resuelve en el plan físico mediante la operación de búsqueda lineal. Dar el nuevo plan físico adecuado para esta consulta.

Ejercicio 4 (12 puntos)

Considere las siguientes transacciones:

$T_1: r_1(x), w_1(y), w_1(x), r_1(y), c_1$

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

a) Justificando su respuesta, escriba una historia no serializable con estas transacciones.

b) En esas transacciones se agregaron los bloqueos de read-write, sin seguir ninguno de los protocolos 2PL visto en el curso:

$T_{11}: rl_1(x), r_1(x), u_1(x), wl_1(y), w_1(y), u_1(y), wl_1(x), w_1(x), u_1(x), rl_1(y), r_1(y), u_1(y), c_1$

$T_{21}: rl_2(y), r_2(y), u_2(y), wl_2(x), w_2(x), u_2(x), wl_2(y), w_2(y), u_2(y), c_2$

Si es posible, escriba una historia no serializable con estas transacciones. Justifique su respuesta.

c) Escriba las transacciones T_{11}' y T_{21}' de forma de que sean iguales con respecto a los read y write, pero teniendo los bloqueos y desbloqueos de manera que sigan 2PL estricto. Justifique su respuesta.

d) Si es posible, escriba con las transacciones de la parte c, una historia no serializable. Justifique su respuesta.