

Guía: Modelo Relacional:Normalización

Prof. Andrea Rodríguez

1. Defina *relación, dependencia funcional, clausura de un conjunto de atributos bajo un conjunto de dependencias funcionales, equivalencia de conjuntos de dependencias funcionales, dependencias funcionales triviales*
2. Demuestre las reglas de Armstrong (reglas 1, 2 y 3 de inferencias de dependencias funcionales). Demuestre además que son correctas.
3. Considere las siguientes tuplas de un esquema $R(A, B, C, D)$: $(a, b, c, d), (e, b, c, d), (f, c, c, d), (f, c, d, d)$. Indique las dependencias funcionales en forma canónica y dependencias multivaluadas *potenciales* que pueden existir asumiendo que estas tuplas satisfacen dichas dependencias.

Resp.

Las dependencias posibles son: $A \rightarrow B, A \rightarrow D, C \rightarrow D, B \rightarrow D, A \twoheadrightarrow B, A \twoheadrightarrow D, C \twoheadrightarrow D, A \twoheadrightarrow C, B \twoheadrightarrow D, B \twoheadrightarrow C, B \twoheadrightarrow A$.

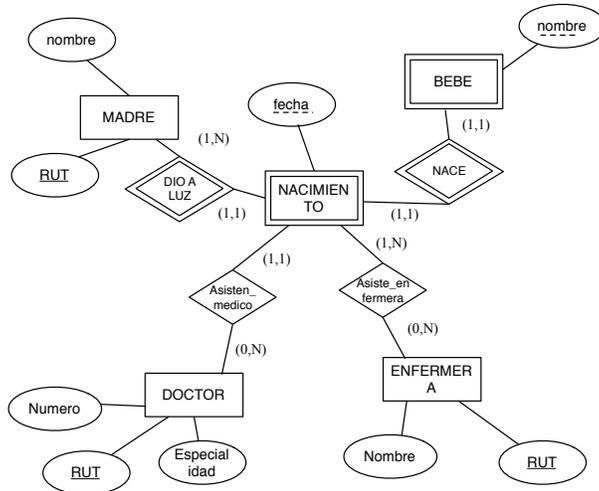
4. Considere de nuevo el esquema $R(A, B, C, D)$ con clave primaria AB e indique un conjunto de dependencias funcionales de manera que este esquema estaría en 2 FN pero no en 3FN.
5. Considere el esquema $R(A, B, C, D, E)$ con dependencias funcionales $B \rightarrow C, CA \rightarrow D, DE \rightarrow B$. Indique las claves candidatas, la forma normal en que se encuentra y si fuera necesario normalice para conseguir un esquema en BNFC. Indique si la descomposición posible conserva dependencias funcionales y/o no genera tuplas espurias. Explique sus respuestas.

Resp.

Las claves candidatas son: ADE o AEB o ACE . El esquema está en 2FN ya que los atributos que son dependientes parcialmente de las claves son también primos. Por la misma razón está en 3 FN. Sin embargo no está en BNFC porque en las 2 dependencias funcionales tenemos que el antecedente no es una superclave. Una descomposición que normaliza es la siguiente: $\underline{BC}, \underline{ACD}$ y \underline{BDE} y \underline{ACE} (La última para dejar una tabla con alguna clave candidata). Esta descomposición preserva dependencias y no genera tuplas espurias.

6. Para el siguiente MER, escriba su esquema relacional normalizado en 3era forma normal.
7. Para las siguientes preguntas considere el esquema $R(A, B, C, D, E, F, G)$, donde

Atributo	Descripción
A	ID Proyecto
B	Nombre Proyecto
C	Fecha de inicio
D	Fecha de término
E	Jefe de Proyecto
F	Departamento
G	Monto



Su conjunto de dependencias funcionales F es:

$$\begin{aligned}
 A &\rightarrow B, C, D, E, G \\
 B, C &\rightarrow A \\
 E &\rightarrow F
 \end{aligned}$$

Considere las siguientes preguntas para la definición del esquema relacional.

- Derive F^+ , excluyendo las triviales.
- Deduzca las claves candidatas y atributos primos
- Indique la forma normal en la que se encuentra el esquema
- Descomponga en el caso que no se encuentre normalizado el esquema, logrando una descomposición sin pérdida de dependencias funcionales y sin pérdida aditiva. En esta descomposición defina el esquema con claves primarias y claves foráneas

Resp.

Las soluciones son las siguientes:

- A continuación se usan las reglas de Armstrong para la derivación de F^+ .

$A \rightarrow B, C, D, E$	Por reflexibilidad y transitividad	$A \rightarrow B$
		$A \rightarrow C$
		$A \rightarrow D$
		$A \rightarrow E$
$A \rightarrow E$	Por transitividad	$A \rightarrow F$
$E \rightarrow F$		
$B, C \rightarrow A$	Por transitividad	$B, C \rightarrow D$
$A \rightarrow D$		$B, C \rightarrow E$
$A \rightarrow E$		$B, C \rightarrow F$
$A \rightarrow F$		$B, C \rightarrow G$
$A \rightarrow G$		

- b) Las claves candidatas son entonces: A o B,C. Entonces los atributos primos son A,B,C.
- c) La relación está en 2FN ya que todos los atributos no primo dependen funcionalmente en forma total de las claves candidatas. Sin embargo, no está en 3FN ya que F depende funcionalmente de E y E no es superclave y F no es un atributo primo.
- d) Se descompone en dos esquemas: $R_1(\underline{A}, B, C, D, G, E)$ FK E REFERENCES $R_2(E)$ y $R_2(\underline{E}, F)$

8. Usando el esquema de la pregunta anterior, escriba las siguientes preguntas en SQL:

- a) Encuentre el ID y nombre del primer proyecto que haya sido dirigido por el Jefe de Proyecto 'X'.
- b) Determine el número de proyectos y su monto total ejecutados por departamento 'Y' bajo la supervisión del Jefe de Proyecto 'X'.

Resp.

Las preguntas en SQL son:

- a) Encuentre el ID y nombre del primer proyecto que haya sido dirigido por el Jefe de Proyecto 'X'.

```
SELECT A,B
FROM R1
WHERE E = 'X' AND C IN
(SELECT E as E,MIN(C) AS C
FROM R1
WHERE E = 'X'
GROUP BY E) )
```

- b) Determine el número de proyectos y su monto total ejecutados por departamento 'Y' bajo la supervisión del Jefe de Proyecto 'X'.

```
SELECT count(*),sum(r1.G)
FROM R1 AS r1, R2 As r2
WHERE r1.E = r2.E AND r2.F = 'Y' AND r1.E = 'X' IN
```

9. Indique cómo implementaría la restricción de que no pueden existir dos proyectos simultáneos bajo la supervisión del mismo Jefe de Proyecto.

Resp.

Se puede aplicar a través de una ASSERTION

```
CREATE ASSERTION INCOMPATIBLE
CHECK(NOT EXISTS (SELECT *
FROM R1 AS r1, R2 As r2
WHERE r1.A <> r2.A AND r1.E = r2.E AND
((r1.C >= r2.C AND r1.C <= r2.D)
OR (r2.C >= r1.C AND r2.C <= r1.D))))
```

10. Considere la siguiente relación:

```
RELACION(NUM_PEDIDO,FECHA_PED,NUMCLIENTE,
NUM_ARTICULO,NOMBRE_ARTICULO,PRECIO_UNI,CANTIDAD)
```

Para esta relación, estos son las DF:

- NUM_PEDIDO \rightarrow FECHA_PED, NUMCLIENTE
- NUM_PEDIDO, NUM_ARTICULO \rightarrow CANTIDAD
- NUM_ARTICULO \rightarrow NOMBRE_ARTICULO, PRECIO_UNI

Determine las claves candidatas. Normalice (hasta donde pueda) la relación manteniendo las dependencias funcionales y sin generación de tuplas espurias.

11. Dado el siguiente conjunto de $DF = \{A \rightarrow B, BC \rightarrow E, DE \rightarrow A\}$ y el esquema $R(A,B,C,D,E)$:

- a) indique la cerradura DF^+ usando las propiedades de dependencias funcionales,
- b) indique la(s) clave(s) candidata(s), (c) indique en qué forma normal está, y
- c) si no estuviera en FNBC indique una descomposición, si es que existe, que preserve la dependencias funcionales y no genere tuplas espurias.

12. Dado el siguiente esquema:

CLIENTE(cid: integer, enombre: string, phone: integer)
 VUELOS(vid: string, hora: time, pid: string, aerolinia: string) FK pid REFERNCES PILOTO(pid)
 INSTANCIA_VUELO(vid: string, dia: date, estado: string)
 FK vid REFERENCES VUELOS
 TICKET(cid: integer, vid: string, dia: date, costo: integer)
 FK cid REFERENCES CLIENTE(cid)
 FK vid, dia REFERENCES INSTANCIA_VUELO(vid, dia)
 PILOTO(pid: integer, enombre: string, phone: integer)

En este esquema, el estado de una instancia de vuelo puede tomar los siguientes valores: atrasado, a la hora, cancelado.

- a) Escriba la consulta: “Recupere el código de vuelo que tenga el mayor promedio de número de clientes en las dos últimas semanas, considerando sólo aerolíneas con más de dos vuelos diarios”. Utilice el concepto de vista para escribir la consulta.
 - b) Indique si este tipo de restricciones son posibles de especificar, y si fuera así, con qué mecanismo (triggers assertions) se implementarían y cuándo se activaría el mecanismo:
 - Los vuelos sólo se realizan si es que tienen más de 10 pasajeros.
 - Un vuelo que ha sido cancelado más de dos veces es eliminado del sistema.
13. Asuma un esquema de relación $R(A,B,C)$ y dos descomposiciones: $R_1(A,B)$ y $R_2(B,C)$. Asuma las siguientes instancias $R_1 = \{(5, 1), (6, 1)\}$ y $R_2 = \{(1, 8), (1, 9)\}$. Indique las tuplas que están definitivamente en R y cuáles posiblemente están. Según eso, ¿ Es B una clave de R ?
14. Considere el siguiente esquema de base de base de datos
- PROVEEDOR(sid: integer, nombre: string, direccion, string)
 ITEM(pid: integer, nombre: string, color: string)
 CATALOGO(sid: integer, pid: integer, costo: real)
 FK (sid) REFERENCES PROVEEDOR(sid)
 FK (pid) REFERENCES ITEM(pid)

Basado en este esquema pequeño, conteste lo siguiente:

- a) Escriba en SQL: Encuentre los items que tienen al menos dos proveedores.

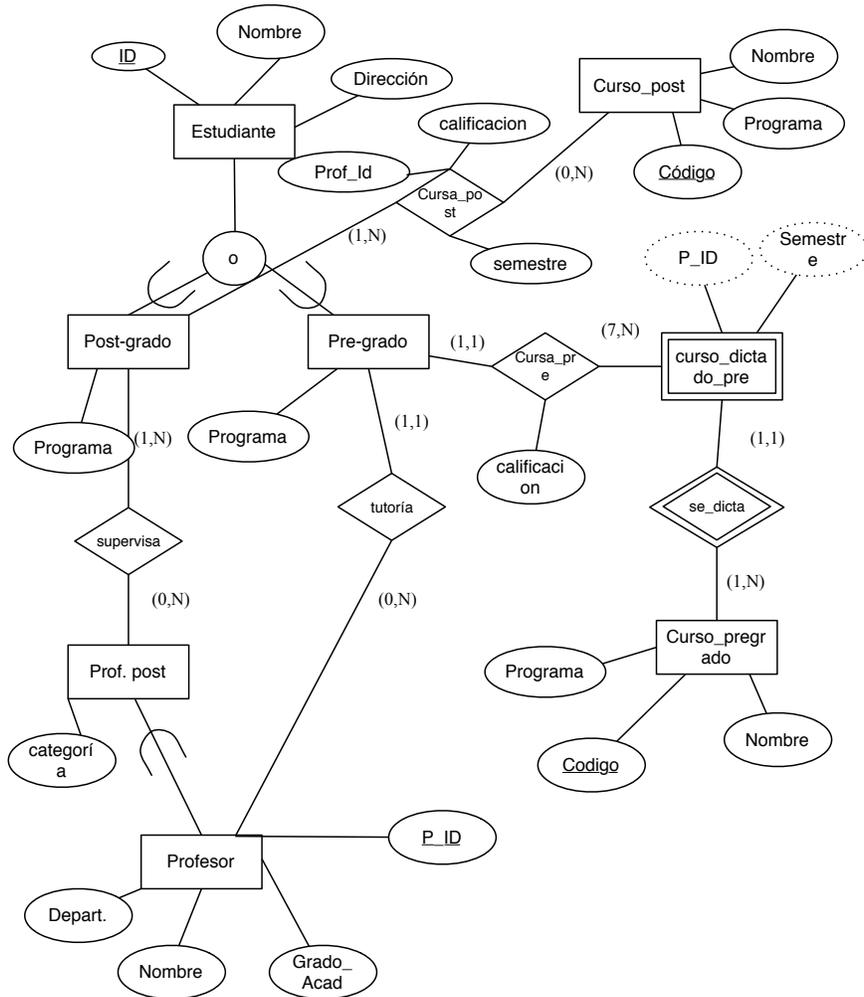
b) Dada la siguiente consulta en SQL, indique su correspondiente consulta en lenguaje natural, en álgebra relacional y cálculo de tuplas y de dominio.

```

SELECT g.pid
FROM CATALOGO g, PROVEEDORES p
WHERE p.nombre = 'FIRESTONE' AND g.sid = p.sid
AND g.costo ≥ ALL (SELECT g2.cost
FROM CATALOGO g2, PROVEEDORES p2
WHERE p2.nombre = 'FIRESTONE'
AND g2.sid = p2.sid)

```

15. Diseñe el esquema de una base de datos relacional normalizada en tercera forma normal para el modelo conceptual siguiente:



16. Dado el esquema de base de datos anterior, escriba en álgebra relacional y en cálculo relacional de tuplas y de dominio las siguientes consultas:

- Liste el nombre de los alumnos que son alumnos de pre y post grado.
- Liste el nombre los profesores que supervisen alumnos de postgrado, siendo estos alumnos también de oregrado

- Liste el nombre los alumnos que no hayan repetido ningun curso de pregrado
- Liste el nombre los alumnos que hayan repetido un curso de post-grado
- List el nombre de los cursos de pregrado cuyos profesores hagan tambien clases en el postgrado en el mismo semestre.

17. Considere la siguiente instancia de corredores:

Rut	Nombre	Edad	Rank
11111111	Ulloa	24	1
22222222	Salas	27	1
33333333	Roa	30	2
44444444	Fuentes	21	null

Considere las siguientes consultas (1 p):

(1) `SELECT C.nombre`
`FROM Corredores C`
`WHERE NOT EXISTS (SELECT *`
`FROM Corredores C2`
`WHERE C2.edad < 22 AND C.rank <= S2.rank)`

(2) `SELECT *`
`FROM Corredores C.nombre`
`WHERE C.rank > ANY (SELECT C2.rank`
`FROM Corredores C2`
`WHERE C2.edad < 22)`

Indique la respuesta de ambas consultas. ¿Ambas entregan el mismo resultado siempre?

Resp. Se hizo en clases

18. Dado el siguiente conjunto de $DF = \{A \rightarrow B, BC \rightarrow E, DE \rightarrow A\}$ y el esquema $R(A,B,C,D,E)$:

- indique la cerradura DF^+ usando las propiedades de dependencias funcionales,
- indique la(s) clave(s) candidata(s), (c) indique en qué forma normal está, y
- si no estuviera en FNBC indique una descomposición, si es que existe, que preserve la dependencias funcionales y no genere tuplas espurias.

Resp:

a) Sin considerar DF no triviales, tenemos lo siguiente:

- Por pseudotransitividad: $BC \rightarrow E, DE \rightarrow A \models BCD \rightarrow A$
- Por aumento y descomposición: $BC \rightarrow E \models BCD \rightarrow E$
- Por transitividad: $DE \rightarrow A, A \rightarrow B \models DE \rightarrow B$
- Por aumento y descomposicion: $DE \rightarrow A \models CDE \rightarrow A, DE \rightarrow B \models CDE \rightarrow B$
- Por pseudotransitividad: $A \rightarrow B, BC \rightarrow E \models AC \rightarrow E$
- Por aumento y descomposición: $A \rightarrow B \models ACD \rightarrow B, AC \rightarrow E \models ACD \rightarrow E$

b) las claves candidatas son: BCD, ACD, EDC.

- c) La relación está en 3FN. Se encuentra en 2FN porque todos los atributos son primos (participan en alguna de las claves primarias) y por lo tanto no importan que dependan parcialmente de una clave candidata. No está en FNBC porque siempre deben depender de una superclave.

La decompesición es: $R_1(A, B)$, $R_2(B, C, E)$, $R_3(D, E, A)$, $R_4(B, C, D)$ Ojo, deben conservar en alguna relación alguna de las claves candidatas.

19. Dado el siguiente esquema:

CLIENTE(cid: integer, enombre: string, phone: integer) PK cid
VUELOS(vid: string, hora: time, pid: string, aerolinia: string) PK vid, FK pid de PILOTO
INSTANCIA_VUELOS(vid: string, dia: date, estado: string) PK vid, dia, FK vid de VUELOS
TICKET(cid: integer, vid: string, dia: date, costo: integer) PK cid, vid,
FK cid de CLIENTE, FK vid, dia de INSTANCIA_VUELO
PILOTO(pid: integer, enombre: string, phone: integer) PK pid

En este esquema, el estado de una instancia de vuelo puede tomar los siguientes valores: atrasado, a la hora, cancelado.

- a) Escriba la consulta: “Recupere el código de vuelo que tenga el mayor promedio de número de clientes en las dos últimas semanas, considerando sólo aerolíneas con más de dos vuelos diarios”. Utilice el concepto de vista para escribir la consulta.
- b) Indique si este tipo de restricciones son posibles de especificar, y si fuera así, con qué mecanismo (triggers assertions) se implementarían y cuándo se activaría el mecanismo:
- Los vuelos sólo se realizan si es que tienen más de 10 pasajeros.
 - Un vuelo que ha sido cancelado más de dos veces es eliminado del sistema.

Resp.

- a) SQL:

```

CREATE VIEW LINEAS(aerolinea,numero)
AS SELECT vid,COUNT(*)
FROM VUELOS
GROUP BY VUELOS
HAVING COUNT(*) > 2;

```

```

CREATE VIEW PASAJES(vid,dia,cuenta)
AS SELECT t.vid,t.dia,COUNT(*)
FROM VUELOS as V, LINEAS as l, TICKET as t
WHERE v.estado <> 'CANCELADO' AND l.aerolina = v.aerolinea AND v.vid = t.vid
        AND t.dia > '16/09/2008'
GROUP BY t.vid, t.dia;

```

```

CREATE VIEW PASAJES_AVG(vid,promedio)
AS SELECT vid,AVG(cuenta)
FROM PASAJES
GROUP BY vid;

```

```

SELECT p.vid
FROM PASAJES_AVG AS p
WHERE p.promedio IN (SELECT MAX(promedio) FROM PASAJES_AVG);

```

b) Restricciones:

- 1) Se puede con un trigger, en el cual el sistema verifique el numero de pasajeros al momento de comenzar los preparativos del vuelo (1 o 2 hrs antes) y se cancela en caso que no hayan suficientes pasajeros
- 2) Se puede hacer que cuando se modifique el estado de un vuelo a cancelado, se verifique si ha sido cancelado más de dos veces, y si es así, se elimina en cascada.

20. Asuma un esquema de relación $R(A,B,C)$ y dos descomposiciones: $R_1(A,B)$ y $R_2(B,C)$. Asuma las siguientes instancias $R_1 = \{(5, 1), (6, 1)\}$ y $R_2 = \{(1, 8), (1, 9)\}$. Indique las tuplas que están definitivamente en R y cuáles posiblemente están. Según eso, ¿ Es B una clave de R ?

Resp.

El atributo B no es una clave candidata porque aparece en más de una tupla. Las tuplas que son necesarias son $\{(5, 1, 8), (6, 1, 9)\}$ or $\{(5, 1, 9), (6, 1, 8)\}$. Como máximo podrían existir: $\{(5, 1, 8), (5, 1, 9), (6, 1, 8), (6, 1, 9)\}$ (el producto cartesiano).