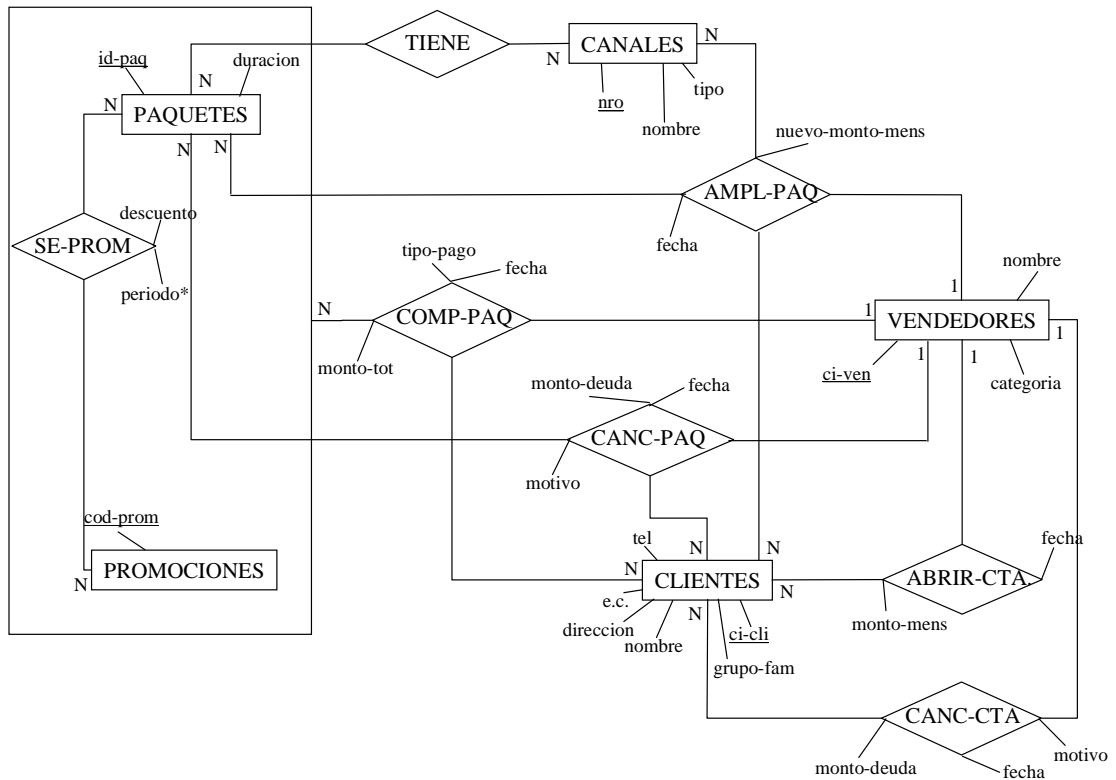


BASES DE DATOS – feb. 2001 – Soluciones

Ejercicio 1)



RNE:

- Un cliente puede cancelar una cuenta si previamente la abrió y no fue cancelada
- Un cliente puede cancelar un paquete si previamente lo compró y no fue cancelado
- Un cliente puede agregar canales si solo si tiene un paquete

Ejercicio 2

Si se dice que se cumple la dependencia funcional $X \rightarrow Y$ sobre una tabla R , se está diciendo que se cumple que:

a) Dada una instancia cualquiera r de R , no hay dos tuplas diferentes que tengan el mismo valor en X .

b) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que $t_1[Y]=t_2[Y]$.

c) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que existen en r dos tuplas t_3 y t_4 tales que $t_1[X]=t_3[X]$ y $t_2[X]=t_4[X]$ y $t_1[Y]=t_4[Y]$ y $t_2[Y]=t_3[Y]$.

d) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que existen en r dos tuplas t_3 y t_4 tales que $t_1[X]=t_3[X]$ y $t_2[X]=t_4[X]$ y $t_1[Y]=t_4[Y]$ y $t_2[Y]=t_3[Y]$ y $t_1[R-XY]=t_3[R-XY]$ y $t_2[R-XY]=t_4[R-XY]$.

La respuesta b coincide con la definición de dependencia funcional. La a es una definición para clave, la d es la definición de dependencia multivaluada y la c no es nada particular visto en el curso.

Ejercicio 3

Si se dice que se cumple la dependencia multivaluada $X \twoheadrightarrow Y$ sobre una tabla R , se está diciendo que se cumple que:

a) Dada una instancia cualquiera r de R , no hay dos tuplas diferentes que tengan el mismo valor en X .

b) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que $t_1[Y]=t_2[Y]$.

c) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que existen en r dos tuplas t_3 y t_4 tales que $t_1[X]=t_3[X]$ y $t_2[X]=t_4[X]$ y $t_1[Y]=t_4[Y]$ y $t_2[Y]=t_3[Y]$.

d) Dada una instancia cualquiera r de R y dada cualquier pareja de tuplas t_1 y t_2 de r tales que $t_1[X]=t_2[X]$, se cumple que existen en r dos tuplas t_3 y t_4 tales que $t_1[X]=t_3[X]$ y $t_2[X]=t_4[X]$ y $t_1[Y]=t_4[Y]$ y $t_2[Y]=t_3[Y]$ y $t_1[R-XY]=t_3[R-XY]$ y $t_2[R-XY]=t_4[R-XY]$.

La respuesta d es la correcta. La c es muy parecida pero no indica que sucede con las partes de la tupla que no aparecen en la dependencia. Esto permitiría que los esos valores pudieran ser diferentes en cada tupla, o sea que no se cumpliría el "cruzamiento".

Ejercicio 4

A partir de la siguiente instancia particular de una tabla R con esquema A,B,C , se puede decir que:

A	B	C
1	2	Juan
1	3	Lucia
1	2	Lucia
1	3	Juan

a) En la tabla R se cumple la dependencia funcional $A \rightarrow B$.

b) En la tabla R se cumple la dependencia multivaluada $A \twoheadrightarrow B$.

c) En la tabla R no se cumple la dependencia funcional $A \rightarrow B$.

d) Nada en particular con respecto a A y B .

Al observar la tabla se encuentra que las tuplas cumplen la relación adecuada para que exista una dependencia multivaluada. Sin embargo, esto es UNA INSTANCIA EN PARTICULAR y para que se cumpla la multivaluada se debe garantizar que EN TODA INSTANCIA SE CUMPLE ESA CONDICION. De esta forma, la única respuesta aceptable es la c.

Ejercicio 5

Dado un esquema $R(A,B,C,D,E,G,H)$ sobre el que se cumplen las siguientes dependencias: $\{A \rightarrow D, CD \rightarrow E, E \rightarrow GH, GH \rightarrow C, B \twoheadrightarrow C\}$

1) (2 pts).

Indique la opción que contenga solamente a todas las claves de R.

- a) AB
- b) ABC, ABE.
- c) **ABC, ABE, ABHG.**
- d) ABCD.

2) (2 pts).

Indicar el resultado de aplicar a R el algoritmo para llevar a 3NF con join sin pérdida visto en el curso.

- a) $R_1(AD), R_2(CDE), R_3(BC), R_4(EGHC)$
- b) $R_1(AD), R_2(CDE), R_3(EGH), R_4(GHC), R_5(BC)$
- c) $R_1(AD), R_2(CDE), R_3(EGH), R_4(GHC)$
- d) **$R_1(AD), R_2(CDE), R_3(EGH), R_4(GHC), R_5(BCA)$**

Para garantizar el resultado correcto, es necesario asegurar la inclusión de una clave en el esquema de una tabla. Si no es así, entonces hay que agregar una tabla cuyo esquema sea una clave. Además se resumen las tablas que surgen de dependencias con igual lado izquierdo. Por esto la respuesta correcta es la d. Hay que notar que la última dependencia es multivaluada y por lo tanto NO INTERVIENE EN EL ALGORITMO DE 3NF ¡!!.

3) (2 pts)

Indicar el resultado de aplicar el algoritmo para llevar a 4NF con join sin pérdida visto en el curso. Durante el algoritmo, las dependencias se deben considerar en el orden en que aparecen escritas.

- a) $R_1(AD), R_2(CDE), R_3(EGH), R_4(GHC), R_5(BC)$
- b) **$R_1(AD), R_2(EGH), R_3(BC), R_4(ABE)$**
- c) $R_1(AD), R_2(EGH), R_3(BC), R_4(ABEG)$
- d) $R_1(AD), R_2(ABC), R_3(ABE)$

Al considerar las dependencias en el orden escrito, entonces se debe particionar primero en base a la dependencia $A \rightarrow D$. De esta forma, D se separa del resto de los atributos obteniendo la partición $R_1(AD) R_i(ABCEGH)$. Al considerar la dependencia $CD \rightarrow E$ se observa que se perdió. Por lo tanto se sigue con la siguiente $E \rightarrow GH$. Como cualquier clave debe incluir a B es claro que viola y al partir se obtiene $R_2(EGH) R_j(ABCE)$. Al considerar la siguiente dependencia se observa que también se perdió por lo que sólo queda la dependencia multivaluada y es claro que B por sí sólo no es clave por lo que queda es la última partición en $R_3(BC)$ y $R_4(ABE)$.

4) (2 pts).

Indicar la opción que contiene más dependencias funcionales que se pierden en el proceso de llevar a 4NF considerando las dependencias en el orden en que están escritas. Si considera que no se pierde ninguna dependencia marque la opción d.

- a) $GH \rightarrow C$, $CD \rightarrow E$
- b) $EG \rightarrow H$, $GH \rightarrow C$
- c) $A \rightarrow D$, $EG \rightarrow H$
- d) No se pierde ninguna dependencia funcional.

Ejercicio 6)

a) Escribir una historia H1 donde se ejecuten 2 transacciones de Venta_entradas, que cumpla:

i) H1 serializable y no recuperable.

H1: r1(X) w1(X) r2(X) w2(X) r1(Y) w1(Y) r2(Y) w2(Y) c2 c1

ii) H1 no serializable y estricta.

H1: r1(X) r2(X) w2(X) r2(Y) w2(Y) c2 w1(X) r1(Y) w1(Y) c1

iii) H1 serial.

H1: r1(X) w1(X) r1(Y) w1(Y) c1 r2(X) w2(X) r2(Y) w2(Y) c2

iv) H1 serializable y que evita abortos en cascada.

H1: r1(X) w1(X) r1(Y) w1(Y) c1 r2(X) w2(X) r2(Y) w2(Y) c2

b) Escribir una historia H2 donde se ejecuten 2 transacciones de Venta_entradas y éstas sigan el protocolo 2PL básico. Decir si H2 es serializable y si es recuperable, justificando.

H2: r1(X) r1(X) w1(X) w1(X) r1(Y) r1(Y) w1(Y) u1(X) r2(X) r2(X) w2(X)
w2(X) u2(X) w1(Y) u1(Y) c1 r2(Y) r2(Y) w2(Y) w2(Y) u2(Y) c2

H2 es serializable, ya que las transacciones cumplen 2PL.

H2 es recuperable, ya que T2 lee de T1 y T1 hace commit antes que T2.

c) Escribir una historia H2 donde se ejecuten 2 transacciones de Venta_entradas y éstas sigan el protocolo 2PL estricto. Decir si H2 es serializable y si es recuperable, justificando.

H2: r1(X) r1(X) w1(X) w1(X) r1(Y) r1(Y) w1(Y) w1(Y) c1 u1(X) u1(Y) r2(X)
r2(X) w2(X) w2(X) r2(Y) r2(Y) w2(Y) w2(Y) c2 u2(X) u2(Y)

H2 es serializable, ya que las transacciones cumplen 2PL.

H2 es recuperable, ya que las transacciones cumplen 2PL estricto, lo cual garantiza historias estrictas, y una historia estricta es recuperable.

Ejercicio 7 (25 ptos).

Las siguiente base de datos contiene información sobre CD's , sus temas y sus intérpretes.

CD(id_cd, nom_cd, int_cd)

Cada tupla de la tabla representa un CD diferente. **id_cd** es una identificación que tiene cada CD. **Nom_cd** es un string con el nombre del CD e **int_cd** es un string que contiene el nombre del grupo o solista que interpreta todos los temas del CD. Un valor válido para este último atributo es "ensalada". En ese caso los temas pueden ser ejecutados por diferentes intérpretes.

Tema(id_tema, nom_tema, autor)

Cada tupla de esta tabla representa un tema diferente. El primer atributo es un identificador para cada tema. El segundo es un string con el nombre del tema y el tercero es un string con el nombre del autor.

Grabacion(id_cd,ubicacion, id_tema, minutos, int_tema)

Cada tupla de esta tabla representa cada aparición de un determinado tema en determinado CD. En cada tupla se tiene en que **ubicación** del CD **id_cd** aparecer el tema **id_tema** durando determinada cantidad de **minutos** e interpretado por **int_tema**.

Int_integracion(interperte, id_cd, ubicacion, instrumento, ejecutante)

Esta tabla detalla la integración de cada grupo intérprete en cada tema de cada CD. Cada tupla contiene el grupo **intérprete** del tema que está en determinada **ubicación** en determinado **id_cd**, indicando el instrumento (ej, guitarra, bajo, voz) y el nombre del ejecutante. Notar que en esta tabla pueden existir varias tuplas para el mismo intérprete en la misma grabación, ya que puede ser un grupo y aún varias tuplas para el mismo ejecutante, ya que un mismo ejecutante puede tocar varios instrumentos en la misma grabación.

Resolver en Cálculo Relacional.

Devolver los nombres de los CD's grabados por un único grupo, en los que George Harrison toca la guitarra en al menos un tema y John Lennon no aparece ejecutando ningún instrumento en ningún tema de esos CD's.

SOLUCIÓN:

$$\{ t / \exists c (CD(c) \wedge c[int_cd] \neq "ensalada" \wedge t[nom_cd]=c[nom_cd] \wedge \\ \exists i.(Int_Integración(i) \wedge i[id_cd]=c[id_cd] \wedge I[interperte]=c[int_cd] \wedge \\ I[instrumento]="guitarra" \wedge I[ejecutante]="George Harrison" \wedge \\ \neg \exists i_1(Int_Integración(i_1) \wedge i_1[id_cd]=c[id_cd] \wedge \\ i_1[interperte]="John Lennon")) \}$$

Resolver en Álgebra Relacional.

Devolver los nombres de los integrantes de Pink Floyd que están en todos los discos que son solamente de ese grupo.

SOLUCION:

$D_{pf} \equiv \pi_{id_cd}(\sigma_{int_cd='Pink Floyd'}(CD))$

$IntPFCD \equiv \pi_{ejecutante,id_cd}(\sigma_{interperte='Pink Floyd'}(Int_Integración) * DPF)$

$Resultado \equiv IntPFCD \% DPF$

Elegir el lenguaje, resolver y fundamentar por qué ese lenguaje.

Devolver los nombres de los CD de Pink Floyd que duran entre 55 y 74 minutos.

SOLUCIÓN:

Select nom_cd

From CD A, Grabacion B

Where A.int_cd='Pink Floyd'

Group by A.id_cd

Having sum(B.minutos) >= 55 and sum(B.minutos)<=74.