



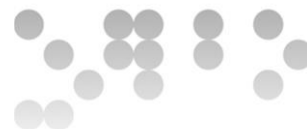
Nota: Propiedad intelectual

A menudo es inevitable, al producir una obra multimedia, hacer uso de recursos creados por terceras personas. Es por lo tanto comprensible hacerlo en el marco de una práctica de los estudios del Grado Multimedia, siempre y esto se documente claramente y no suponga plagio en la práctica.

Por lo tanto, al presentar una práctica que haga uso de recursos ajenos, se tiene que presentar junto con ella un documento en que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y su estatus legal: si la obra está protegida por el copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL ...). El estudiante tendrá que asegurarse que la licencia que sea no impide específicamente su uso en el marco de la práctica. En caso de no encontrar la información correspondiente tendrá que asumir que la obra está protegida por el copyright.

Habrán, además, adjuntar los ficheros originales cuando las obras utilizadas sean digitales, y su código fuente si corresponde.

Otro punto a considerar es que cualquier práctica que haga uso de recursos protegidos por el copyright no podrá en ningún caso publicarse en Mosaico, la revista del Graduado en Multimedia a la UOC, a no ser que los propietarios de los derechos intelectuales den su autorización explícita.



PEC2: NORMALIZACIÓN DE UN MODELO LÓGICO RELACIONAL Y USO DE SQL BÁSICO

ENUNCIADO

A.- Normalizar hasta BCNF el modelo lógico relacional dado.

Enunciado del ejercicio

Disponemos de una base de datos con datos de actuaciones de actores en películas. Interesa almacenar esta información en un conjunto de relaciones relacionales de forma que se minimicen las redundancias y se evite la pérdida de datos. Para conseguir este objetivo usad los criterios de la teoría de la normalización comentando, **a paso (1FN, 2FN, 3FN, BCNF)**. Normalizad hasta que sea posible. Tened en cuenta que puede ser necesario fusionar o crear nuevas tablas.

NOTA de CONTENIDO: Los datos existentes en las tablas siguientes son ficticios y no pretenden reflejar la realidad de personas existentes y localizaciones.

Tenemos una relación que almacena información sobre las actuaciones de los actores a diferentes películas con la siguiente intensión:

ACTUACIÓN (nombreActor, tituloPeli, datosPeli, datosDirector, codiPeli, papel)

Para dar contexto a la información que contiene la relación, diremos que:

Un actor se identifica por su nombre. Un determinado actor puede actuar en muchas películas. Cada película se identifica por un título. Además, cada película tiene asignado un código diferente (codiPeli).

De los datos de las películas, interesará tener acceso tanto a la duración como a la fecha de estreno de forma separada para realizar diferentes estadísticas.

Los directores de las películas se identifican por su nombre. Interesa también la nacionalidad de los directores de las películas para poder realizar un listado de las nacionalidades de los diferentes directores de películas.



La extensión de la relación en este momento es:

nombreActor	tituloPeli	datosPeli	datosDirector	codiPeli	papel
Leonardo DiCaprio	TITANIC	135min, 1998	James Cameron, Canadá	PL001	Protagonista
Bernard Hill	TITANIC	135min, 1998	James Cameron, Canadá	PL001	Secundario
Leonardo DiCaprio	DJANGO	170min, 2012	Quentin Tarantino, EEUU	PL002	Secundario

Normalizad hasta BCNF esta relación, indicando en qué forma normal se encuentra, y detallando el proceso de normalización. Hay que indicar las dependencias funcionales y las dependencias funcionales plenas en cada paso del proceso de normalización.

En el caso de BCNF hay que detallar claramente los determinantes y las claves candidatas.

Solución

Está en 1FN?

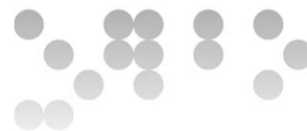
Cómo puede verse a partir de los datos informados, algunos atributos no son atómicos (datosPeli, datosDirector). Además, según dice el enunciado, tenemos que:

“Interesará tener acceso tanto a la duración como la fecha de estreno de forma independiente...”

|

“también hay que conocer la nacionalidad de los directores de las películas para poder realizar un listado de las nacionalidades de los directores de películas”

Hay que separar la información contenida en los atributos “datosPeli” y “datosDirector” para poder acceder a su contenido individualmente. Para ello atomizamos sus atributos y sus valores obteniendo:



ACTUACIÓN (nombreActor, tituloPeli, duración, anoEstreno, nombreDirector, paisDirector, codiPeli, papel)

<u>nombreActor</u>	<u>tituloPeli</u>	<u>duración</u>	<u>anoEstreno</u>	<u>nombreDirector</u>	<u>paísDirector</u>	<u>codiPeli</u>	<u>papel</u>
Leonardo DiCaprio	TITANIC	135min	1998	James Cameron	Canadá	PL001	Protagonista
Bernard Hill	TITANIC	135min	1998	James Cameron	Canadá	PL001	Secundario
Leonardo DiCaprio	DJANGO	170min	2012	Quentin Tarantino	EEUU	PL002	Secundario

Está en 2FN?

Para saber si la relación está en segunda forma normal hay que comprobar que todos los atributos que no forman parte de la clave primaria tienen una dependencia funcional plena con toda la clave primaria, excepto las claves alternativas, que ya se consideran normalizadas en segunda forma normal.

Veamos cuales son las dependencias:

ACTUACIÓN (nombreActor, tituloPeli, duración, anoEstreno, nombreDirector, paisDirector, codiPeli, papel)

- **duración** tiene dependencia funcional con tituloPeli
- **anoEstreno** tiene dependencia funcional con tituloPeli
- **nombreDirector** tiene dependencia funcional con tituloPeli
- **paísDirector** tiene dependencia funcional con nombreDirector
- **codiPeli** es clave alternativa (en el enunciado dice que el atributo codiPeli tiene “un código diferente para cada una de las películas”, o sea, que se podría sustituir la clave primaria el tituloPeli por el codiPeli y la clave primaria resultante, también identificaría correctamente a todas las ocurrencias.
- **papel** tiene dependencia funcional plena con toda la clave primaria puesto que, como se ve en los datos informados, un actor en una película puede desempeñar papel de protagonista, y en cambio en otra película, actuar como actor secundario.

Por lo tanto, tenemos estas dependencias:

- tituloPeli -> duración
- tituloPeli -> anoEstreno
- tituloPeli -> nombreDirector -> paisDirector
- tituloPeli <-> codiPeli
- {nombreActor, tituloPeli} -> papel



Así, rápidamente podemos ver que esta relación NO está en 2FN porque existen atributos, que no son clave alternativa, que no dependen de toda la clave, como por ejemplo duración, anoEstreno y nombreDirector .

Normalización a 2FN:

Para normalizar a 2FN, separaremos relaciones de forma que tengamos dependencias funcionales plenas en cada tabla. Así pues, por un lado separaremos los datos relativos a las actuaciones que dependen de toda la clave y crearemos nuevas tablas para el resto. En este caso tendremos:

ACTUACIÓN (nombreActor, tituloPeli, codiPeli, papel)

Donde {tituloPeli} referencia PELÍCULA {tituloPeli}

PELÍCULA (tituloPeli, duración, anoEstreno, nombreDirector, paisDirector)

Está en 3FN?

Para saber si la relación está en tercera forma normal, hay que comprobar que no existe ningún atributo no clave que dependa de otro atributo no clave o de un conjunto de atributos no clave. Se aplica también la excepción de las claves alternativas, como la segunda forma normal.

ACTUACIÓN (nombreActor, tituloPeli, codiPeli, papel)

- codiPeli es clave alternativa (se podría sustituir a la clave primaria el tituloPeli por codiPeli, y la clave primaria de la relación ACTUACIÓN seguiría identificando correctamente a todas las ocurrencias). Además, depende de un atributo que es clave primaria.
- papel depende funcionalmente de toda la clave primaria, puesto que un actor puede actuar en diferentes películas.

Por lo tanto, tenemos estas dependencias:

- tituloPeli <-> codiPeli
- {nombreActor, tituloPeli} -> papel

PELÍCULA (tituloPeli, duración, anoEstreno, nombreDirector, paisDirector)

Como ya se ha visto en la segunda forma normal, la duración, el año de estreno y el nombre del director tienen dependencia funcional con tituloPeli.

Por otro lado, el país del director tiene dependencia funcional con el nombre del Director.

Así que tenemos las siguientes dependencias:

- tituloPeli -> duración



- tituloPeli -> anoEstreno
- tituloPeli -> nombreDirector
- nombreDirector -> paisDirector.

Vemos pues, que esta relación NO está en tercera forma normal, dado que paisDirector tiene dependencia funcional con un atributo que no forma parte de la clave primaria, nombreDirector.

Normalización a 3FN

Se crea una relación para guardar la información relativa a los directores de películas:

ACTUACIÓN (nombreActor, tituloPeli, codiPeli, papel)

Donde {tituloPeli} referencia PELÍCULA {tituloPeli}

PELÍCULA (tituloPeli, duración, anoEstreno, nombreDirector)

Donde {nombreDirector} referencia DIRECTOR {nombreDirector}

DIRECTOR (nombreDirector, paisDirector)

Está en BCNF?

Cómo la relación está en 1FN, habrá que comprobar que todos los determinantes también son claves candidatas de la relación.

Tal como se indica en la teoría, un determinante es el atributo origen de una dependencia funcional. La única relación sospechosa de no estar en BCNF es la de ACTUACIÓN. Hagamos dos listas, una de determinantes y otra de claves candidatas. Si las listas coinciden, la relación está en BCNF. Si no, habrá que normalizar:

ACTUACIÓN (nombreActor, tituloPeli, codiPeli, papel)

- **tituloPeli** determina el valor de codiPeli.
- **codiPeli** determina el valor de tituloPeli.
- **{nombreActor, tituloPeli}** determinan el valor de papel.

Tenemos estas dependencias:

- tituloPeli -> codiPeli
- codiPeli -> tituloPeli
- {nombreActor, tituloPeli} -> papel
- {nombreActor, codiPeli} -> papel

Y por lo tanto, los determinantes son los conjuntos de atributos que hay a la izquierda de la dependencia y las claves candidatas son: {nombreActor, tituloPeli} y {nombreActor, codiPeli}.



Haciendo dos listas tenemos:

Determinantes	Claves candidatas
{nombreActor, tituloPeli}	{nombreActor, tituloPeli}
{nombreActor, codiPeli}	{nombreActor, codiPeli}
tituloPeli	
codiPeli	

Así, podemos observar que la lista de determinantes NO se igual a la lista de claves candidatas, con el que se deduce que la relación ACTUACIÓN NO está en BCNF

Normalización a BCNF

Para normalizar hasta BCNF hay que eliminar de la relación ACTUACIÓN el atributo codiPeli, que impide que la relación esté en BCNF.

Cómo ya se dispone de una relación para almacenar los datos de las películas, sólo hay que mover de lugar el atributo codiPeli, de ACTUACIÓN hacia PELÍCULA.

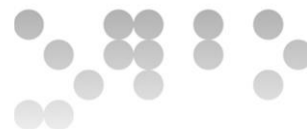
ACTUACIÓN (nombreActor, tituloPeli, papel)

Donde {tituloPeli} referencia PELÍCULA {tituloPeli}

PELÍCULA (tituloPeli, duración, anoEstreno, nombreDirector, codiPeli)

Donde {nombreDirector} referencia DIRECTOR {nombreDirector}

DIRECTOR (nombreDirector, paisDirector)

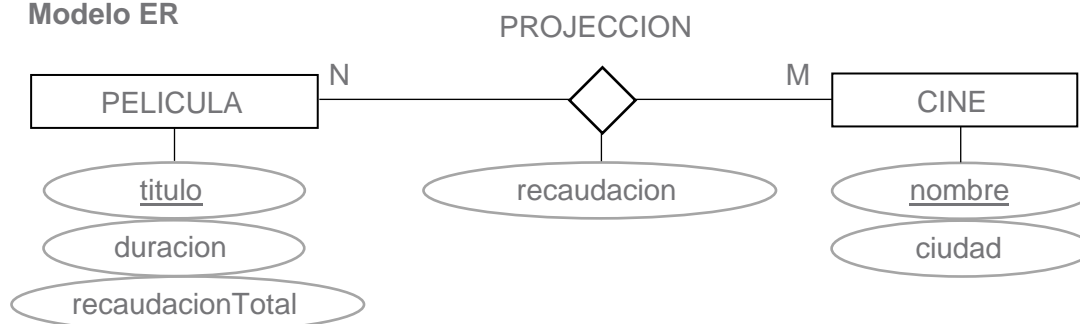


B.- Creación de tablas con instrucciones del DDL.

Dado un modelo el ER, el modelo lógico relacional que le corresponde y un conjunto de datos de una base de datos normalizada se pide:

1. Escribid las instrucciones SQL necesarias para **crear las tablas** (DDL).
2. Escribid también las instrucciones SQL necesarias **para insertar las filas en las tablas** creadas en el punto anterior (DML).

Modelo ER



Modelo Lógico Relacional

PELICULA (titulo, duracion, recaudacionTotal)

CINE (nombre, ciudad)

PROYECCION (tituloPeli, nombreCine, recaudacion)

Donde {tituloPeli} referencia a PELICULA {titulo}

Donde {nombreCine} referencia a CINE {nombre}

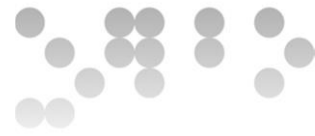
Tablas informadas

PELICULA

<u>titulo</u>	Duracion**	recaudacionTotal*
Ad Astra	122	500.000
El rey león	118	307.000
Los Goonies	114	390.000
Toy Story 4	89	410.000

* La recaudación máxima de una película en este ámbito puede ser de 999,999.999€

** La duración está expresada en minutos



CINE

<u>nombre</u>	<u>ciudad</u>
Arenas Multicines	Barcelona
Cine Maldà	Barcelona
Cinesa La Maquinista	Barcelona
JCA Cines Rambla	Lleida
Albèniz Centro	Girona

PROYECCION

<u>tituloPeli</u>	<u>nombreCine</u>	<u>recaudacion*</u>
Ad Astra	Arenas Multicines	180.000
Ad Astra	Cinesa La Maquinista	225.000
Ad Astra	JCA Cines Rambla	53.000
Ad Astra	Albèniz Centro	42.000
El rey león	Arenas Multicines	210.000
El rey león	Cine Maldà	97.000
Los Goonies	Arenas Multicines	210.000
Los Goonies	Cinesa La Maquinista	180.000

* La recaudación máxima de una película en este ámbito puede ser de 999.999€

SOLUCION

```
create table PELICULA
(
  titulo varchar(50),
  duracion integer,
  recaudacionTotal numeric (9,0),
  primary key (titulo)
```



```
)engine = innodb;
```

```
create table CINE  
(  
  nombre varchar(50),  
  ciudad varchar (50),  
  primary key (nombre)  
)engine = innodb;
```

```
create table PROYECCION  
(  
  tituloPeli varchar(50),  
  nombreCine varchar (50),  
  recaudacion numeric (6,0),  
  primary key (tituloPeli, nombreCine),  
  foreign key (tituloPeli) references PELICULA(titulo),  
  foreign key (nombreCine) references CINE(nombre)  
)engine = innodb;
```

```
insert into PELICULA values("Ad Astra",122,500000);  
insert into PELICULA values("El rey león",118,307000);  
insert into PELICULA values("Los Goonies",114,390000);  
insert into PELICULA values("Toy Story 4",89,410000);
```

```
insert into CINE values ("Arenas Multicines", "Barcelona");  
insert into CINE values ("Cine Maldà", "Barcelona");  
insert into CINE values ("Cinesa La Maquinista", "Barcelona");  
insert into CINE values ("JCA Cines Rambla", "Lleida");  
insert into CINE values ("Albèniz Centre", "Girona");
```

```
insert into PROYECCION values ("Ad Astra", "Arenas Multicines", 180000);  
insert into PROYECCION values ("Ad Astra", "Cinesa La Maquinista", 225000);  
insert into PROYECCION values ("Ad Astra", "JCA Cines Rambla", 53000);  
insert into PROYECCION values ("Ad Astra", "Albèniz Centro", 42000);  
insert into PROYECCION values ("El rey león", "Arenas Multicines", 210000);  
insert into PROYECCION values ("El rey león", "Cine Maldà", 97000);  
insert into PROYECCION values ("Los Goonies", "Arenas Multicines", 210000);
```



insert into PROYECCION values ("Los Goonies", "Cinesa La Maquinista", 180000);

C.- Consultas SQL sobre la base a datos

Dada la BD del ejercicio anterior (B) se piden las instrucciones SQL (DML) para dar respuesta a las siguientes cuestiones:

(Nota: Las instrucciones del DML se tienen que hacer con independencia de los datos que contengan las tablas. Sólo se puede usar en las instrucciones de la solución, los datos aportados por el enunciado. No se pueden usar datos extraídos directamente de las tablas si estos datos NO figuran en el enunciado de la pregunta)

1.- Mostrar el nombre y la ciudad de los cines donde se proyectan 2 o más películas. Hay que mostrar también cuántas películas se proyectan en cada uno de estos cines

```
select c.nombre, c.ciudad, count(*)
from cine c inner join proyeccion p donde c.nombre = p.nombreCine
group by c.nombre, c.ciudad
having count(*) >=2
```

2.- Mostrar el título y la duración de las películas que no se están proyectando actualmente en ningún cine.

```
select titulo, duración
from pelicula
where titulo not in (
    select tituloPeli
    from proyeccion
)
```

3.- Mostrar el título, la duración y la recaudación total de la película con mayor y menor importe de recaudación.

```
select *
from pelicula
where recaudacionTotal = (
    select max(recaudacionTotal)
    from pelicula
)
union
select *
from pelicula
where recaudacionTotal = (
    select min(recaudacionTotal)
    from pelicula
)
```



O BIEN

```
select *
from pelicula
where recaudacionTotal = (
select max(recaudacionTotal)
from pelicula
)
or recaudacionTotal = (
select min(recaudacionTotal)
from pelicula
)
```

4.- Qué instrucciones SQL harían falta por eliminar de la base de datos la película con menos recaudación total?

```
delete
from proyeccion
where titolPeli in(
select titol
from pelicula
where recaudacionTotal =(
select min(recaudacionTotal)
from pelicula
)
);
```

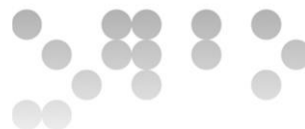
```
delete
from pelicula
where recaudacionTotal = (
select min(recaudacionTotal)
from pelicula
);
```

Criterios de corrección

Cada parte aportará el mismo valor a la nota final de la PEC.

El apartado A se dividirá en cuatro partes, correspondientes a cada una de las formas normales de la normalización a realizar.

Se tendrá en cuenta la corrección del proceso de normalización, y las explicaciones dadas para realizar tanto el análisis de la forma normal en la que se encuentre la relación en cada momento, como su normalización.



El apartado B se dividirá en dos partes. En la primera parte se tendrá en cuenta la correcta creación de las tablas pedidas, corrección sintáctica según SQL, creación de atributos, claves primarias, foráneas y elección de los tipos de datos más adecuados según los datos informados de cada atributo de la tabla.

En la segunda parte, se tendrá en cuenta la correcta inserción de los datos pedidos, respetando las reglas del modelo relacional (identidad e integridad referencial)

El apartado C se dividirá en cuatro partes, una por cada instrucción del DML a realizar.

Se tendrá en cuenta la corrección sintáctica de las instrucciones, según el lenguaje SQL, la no utilización de datos en el código de las instrucciones (a menos que el enunciado diga el contrario), y la adecuación del resultado de las instrucciones al que se pide en el enunciado.

Total de puntuación de la PEC2: 10 puntos.

La traducción de puntos a la nota es, para esta PEC:

- 10-9 puntos = "A"
- 8-7 puntos = "B"
- 6-5 puntos = "C +"
- 4-3 puntos = "C -"
- 2-1 punto = "D"

Formato de entrega

Tendréis que entregar este mismo documento, en formato PDF con las respuestas debidamente cumplimentadas en el buzón de entrega de actividades.

En el aula de teoría hay la información sobre las competencias y objetivos de esta PEC, así como las fechas importantes (fecha de inicio, de entrega, de publicación de notas y de solución) y los materiales complementarios necesarios para la realización de la PEC.