

# Informe 2

## Creación de Base de Datos para Películas

CC3201 - Base de Datos

**Nombre:** Franco Curotto

**Profesor:** Jorge Perez

**Auxiliar:**

**Fecha:** 21 de Octubre del 2013

### 1. Introducción

En este informe se continuará con la creación de la base de datos de películas de la entrega 1. El procedimiento fue básicamente ampliar el modelo de la entrega 1, agregándole los elementos que no se habían alcanzado a incorporar. Además se asignó las restricciones de llaves primarias y foráneas, se agregaron más datos a las tabla, y se detectaron y eliminaron las dependencias funcionales que existían en las tablas de la primera entrega.

### 2. Cambios Realizados

En la primera entrega se tenían las siguientes tablas:

1. directores(nombre, sexo, nacionalidad, fecha\_de\_nacimiento)
2. actores(nombre, sexo, nacionalidad, fecha\_de\_nacimiento)
3. peliculas(titulo, año, genero, director, productora, duracion\_min, pais, presupuesto, recaudacion, animacion, color, resumen)
4. premios\_peliculas(pelicula, premio)
5. casting(pelicula, actor)

Para la tablas de *directores* y *actores* mantuvo el esquema, asignando de llave el atributo *nombre* a ambas tablas. Se escogió usar los nombres de llave en vez de inventar un ID, debido a que se espera manejar una base de datos pequeñas, donde la probabilidad de que coincidan los nombres es muy baja (en el nombre se incluye el primer nombre y el apellido). Es trivial ver que ambas tablas están en BCNF.

Para la tabla de películas se hicieron dos cambios. Primero, se eliminó el atributo *genero* para dejarlo en una tabla aparte llamada *genero*. Esto es debido a que una película puede ser de muchos géneros, la nueva tabla permite esta

posibilidad. El segundo cambio consiste en agregarle un atributo *pelicula\_ID* a las películas y hacer éste la llave de las películas. Esto fue debido a que se consideró importante manejar el caso de que dos películas tengan el mismo nombre, lo que se da en la vida real. Agregarle otro atributo a la llave aparte del título no soluciona el problema, ya que existe la probabilidad, aunque baja, de que existan dos películas con el mismo nombre en el mismo año, o con el mismo nombre y director, y así sucesivamente para todas las combinaciones (no sé mucho de películas así que no sé qué tan cierto sea esto, y usar una ID soluciona inmediatamente el problema), por esto mismo la tabla está en BCNF. Como llave foránea, se crea una restricción para *directores* hacia el atributo *nombre* de la tabla *directores*, ya que el director debe existir para que dirija una película.

Para la tabla de casting se agregó el atributo *rol*, el cual designa si el actor desempeñó un rol de personaje principal, secundario o extra. Claramente todos los atributos en esta tabla son llave, un actor puede actuar en varias películas, una película puede tener varios actores, y pueden haber varios personajes con el mismo rol en cada película, por lo que no se necesita crear una restricción de llave primaria. El mismo argumento anterior se puede usar indicando que la tabla está en BCNF. Como llave foránea de definen las evidentes, de *pelicula\_ID* en la tabla *peliculas*, y de *actor* a *nombre* en la tabla *actores*.

Para la tabla *premios.peliculas* se agregó el atributo *festival*, para indicar el festival de cine donde la película ganó el premio. Claramente la llave tiene que ser todos los atributos, ya que una película puede ganar muchos premios, un premio puede ser ganado por muchas películas, y pueden existir premios iguales de distintos festivales. Por lo mismo la tabla está en BCNF. Como llave foránea se creó la restricción de *peliculas* a *pelicula\_ID* de la tabla *peliculas*. Una dependencia que (supongo) debería existir y no está es que dos películas no puedan ganar el mismo premio en el mismo festival, el mismo año, por lo que se debe tener cuidado cuando se agrega una tupla a esta tabla.

Además de estas tablas, se agregaron 5 nuevas tablas: *cines*, *genero*, *palabras\_claves*, *estrenos*, y *casting-ganador*.

La tabla *cines* tiene los atributos *nombre\_cine* y *ciudad*. Esta tabla contiene los cines que ofrecen las películas de la base de datos, y la ciudad donde se encuentran esos cines, para que el usuario tenga información para ir a ver las películas. En esta tabla se considera el nombre del cine único, por lo que se le crea una restricción de llave primaria.

La tabla de género tiene los atributos *pelicula* y *genero*. Es evidente que la única llave son los dos atributos, por lo que no es necesario aplicar la restricción de llave primaria. Como restricción de llave foránea se aplica la evidente para *pelicula* hacia *pelicula\_ID*.

La tabla *palabras\_claves* tiene los atributos *pelicula* y *palabra*, con las palabras claves de cada película. Aquí también es evidente que la única llave son los dos atributos, por lo que la tabla está en BCNF. La única restricción de llave foránea es la evidente de *pelicula* a *pelicula\_ID*.

La tabla *estrenos* tiene los atributos *pelicula*, *cine*, *fecha\_estreno*, *fecha\_retiro*. La tabla indica qué cine está proyectando qué películas y el intervalo de días en que está en cartelera. Como llave primaria se asigna la restricción a la tupla (*pelicula*, *cine*), por lo que se asume que una vez que una película está en cartelera, no puede volver a estarlo en el futuro (a menos que elimine la tupla), que es lo usual en los cines. Claramente la tabla está en BCNF. También se

tienen las restricciones de llave foránea de *pelicula* a *pelicula\_ID*, y de *cine* a *nombre\_cine* en la tabla *cine*.

La tabla *casting\_ganador*, tiene de atributos *pelicula*, *actor*, *premio*, *festival*. La idea de esta tabla es indicar los actores que han ganado un premio específico para actores, como "mejor actor principal", "mejor actriz de reparto", etc. Un premio indicado en esta tabla debe estar también indicado en la tabla *premios\_peliculas*, para así poder contar fácilmente los premios ganados por las películas y por los actores por separado. Por lo anterior, se crearon las restricciones de llave foránea de (*pelicula*, *premio*, *festival*) a (*pelicula*, *premio*, *festival*) en la tabla *premios\_peliculas*. Como llave primaria se tienen todos los atributos, por lo que el tabla está en BCNF.

La base de datos final quedó con las siguientes tablas:

1. *directores*(nombre, sexo, nacionalidad, fecha\_de\_nacimiento)
2. *actores*(nombre, sexo, nacionalidad, fecha\_de\_nacimiento)
3. *cines*(nombre\_cine, ciudad)
4. *peliculas*(pelicula\_ID, titulo, año, director, productora, duracion\_min, pais, presupuesto, recaudacion, animacion, color, resumen)
5. *premios\_peliculas*(*pelicula*, *premio*, *festival*)
6. *casting*(*pelicula*, actor, rol)
7. *genero*(*pelicula*, genero)
8. *palabras\_claves*(*pelicula*, palabra)
9. *estrenos*(pelicula, cine, fecha\_estreno, fecha\_retiro)
10. *casting\_ganador*(*pelicula*, *actor*, *premio*, *festival*)

### 3. Consultas de Prueba

Como consultas de prueba se hicieron dos pruebas de la entrega 1. Ambas pruebas resultaron ser exitosa.

- Consulta: Lista de los actores que han participado en más de N películas del género drama (con N un input entregado por el usuario), en este caso, como no se ha hecho la interfaz se considera  $N = 1$ .

Código en SQL:

```
select a.nombre as 'nombre actor', count(a.nombre) as 'n peliculas'
from actores as a, peliculas as p, casting as c, genero as g
where a.nombre = c.actor and
p.pelicula_ID = c.pelicula and
p.pelicula_ID = g.pelicula and
g.genero = 'drama'
group by a.nombre
having(count(a.nombre) >1);
```

Resultado:

nombre actor	n películas
Eduardo Pachuca	2
Fernandich Moraliz	2
Sebastian Needles	3
Valentina Soto	2

- Consulta: Director con más premios en sus películas dirigidas.

Código SQL

```
select e.nombre as 'nombre director', max(e.c) as 'nro premios'
from(
select d.nombre, count(d.nombre) as c
from directores as d, peliculas as p, premios_peliculas as pp
where d.nombre = p.director and p.pelicula_ID = pp.pelicula
group by d.nombre ) e;
```

Resultado:

nombre director	nro premios
Franco Coppola	5