



Preparador Informática

PRÁCTICA 8

BASES DE DATOS

(NORMALIZACIÓN)

(SOLUCIONES Ejercicios 17 al 19)

EJERCICIO 17. PELÍCULA

17. Dada la relación siguiente:

PELÍCULA (Título, Año, Duración, Tipo, Estudio, Dirección_Estudio, Actor)

Y teniendo en cuenta además las siguientes consideraciones:

- De un título se pueden haber realizado varias versiones en distintos años pero nunca con los mismo actores.
- Un determinado actor puede haber participado en varias películas durante un año.
- No existe ningún estudio que esté ubicado en varias ciudades.
- Un actor puede trabajar con distintos estudios.

Y los siguientes datos de ejemplo:

| Título | Año | Duración | Tipo | Estudio | Dirección Estudio | Actor |
|---------------------|------|----------|-------|---------|----------------------|-----------------|
| Star Wars | 1977 | 124 | Color | Fox | Hollywood | Carrie Fisher |
| Star Wars | 1977 | 124 | Color | Fox | Hollywood | Mark Hamill |
| Star Wars | 1977 | 124 | Color | Fox | Hollywood | Harrison Ford |
| Mighty Ducks | 1991 | 104 | Color | Disney | Buena Vista | Emilio Estevez |
| Ben Hur | 1959 | 212 | Color | MGM | Hollywood | Charlton Heston |
| Ben Hur | 1959 | 212 | Color | MGM | Hollywood | Martha Scott |
| El retorno del Jedi | 1983 | 124 | Color | Fox | Hollywood | Carrie Fisher |

Se pide:

- a) Explicar todos los tipos de anomalías que existen en la relación Película. (Debes razonar la respuesta de acuerdo a los datos contenidos en la relación)
- b) ¿Cuáles son las dependencias funcionales existentes en la relación Película? Para explicarlas puedes utilizar las siguientes abreviaturas:
 Título (T) Año (A) Duración (D) Tipo (Ti)
 Estudio (E) Dirección_Estudio (Di) Actor (Ac)
- c) ¿La relación Película se encuentra en 3FN? En caso negativo, debes hacer que esté en 3FN.

EJERCICIO 17. PELÍCULA. SOLUCIÓN PROPUESTA

- a) Explicar todos los tipos de anomalías que existen en la relación Película. (Debes razonar la respuesta de acuerdo a los datos contenidos en la relación)

En primer lugar, analizaremos las anomalías producidas por la Dependencia Funcional $E \rightarrow Di$. Esta dependencia provoca la redundancia de datos al almacenar repetidas veces la dirección de cada estudio. Esta redundancia existente también puede provocar inconsistencias.

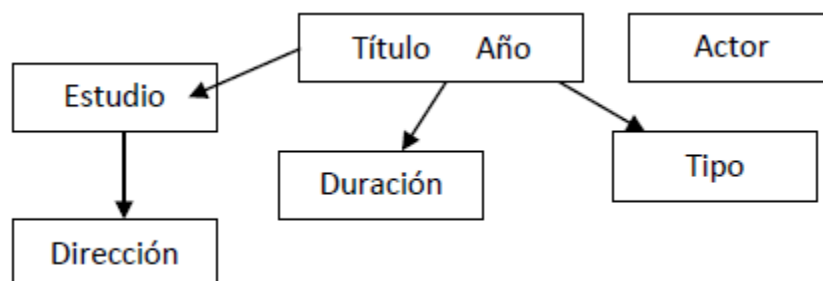
Por otro lado, al incluir el atributo 'Actor' en la relación, estamos incluyéndolo, por consecuencia, en la clave candidata de la relación, y al existir dependencias funcionales de parte de la clave candidata, provoca un almacenamiento redundante de todos los atributos dependientes. En este caso estaríamos redundando los atributos: 'Duración', 'Tipo' y 'Estudio'. Y como consecuencia de la redundancia de 'Estudio' también se redunda el atributo 'Dirección'.



b) ¿Cuáles son las dependencias funcionales existentes en la relación Película? Para explicarlas puedes utilizar las siguientes abreviaturas:

| | | | |
|-------------|------------------------|--------------|------------|
| Título (T) | Año (A) | Duración (D) | Tipo (Ti) |
| Estudio (E) | Dirección_Estudio (Di) | | Actor (Ac) |

Estas dependencias pueden verse muy claramente en el siguiente gráfico:



De Título y Año podemos saber el Estudio. $T A \rightarrow E$

De Título y Año podemos saber la Duración. $T A \rightarrow D$

De Título y Año podemos saber el Tipo. $T A \rightarrow Ti$

De Estudio podemos saber la Dirección. $E \rightarrow Di$

Y Actor no tiene dependencia funcional con nadie.

A la vista de estas dependencias funcionales la clave primaria de la relación PELÍCULA debe ser la concatenación de los atributos Título, Año y Actor.

c) ¿La relación Película se encuentra en 3FN? En caso negativo, debes hacer que esté en 3FN.

Vamos a realizar las observaciones pertinentes:

1ª Forma Normal:

Una relación está en 1FN si no existen atributos multivaluados ni relaciones anidadas. Dicho de otra forma, estará en 1FN si los atributos no clave, dependen funcionalmente de la clave.

Comprobando la tabla que tenemos se observa que ya se encuentra en 1ª Forma Normal ya que no tiene atributos multivaluados.

2ª Forma Normal

Una relación está en 2FN si está en 1FN y todos los atributos no clave tienen dependencia funcional completa de la clave primaria. De acuerdo con esta definición, cada tabla que tiene un atributo único como clave, está en segunda forma normal.

Esta tabla tiene varios atributos que actúan como clave principal. Para las relaciones en las que la clave primaria contiene múltiples atributos, como es el caso, ningún atributo no clave debería depender funcionalmente de una parte de la clave Primaria.

En este caso vemos que los atributos {E, D, Di, Ti} dependen de una parte de la clave primaria {A, T} pero no dependen de Ac.

Vamos a descomponer y crear una nueva relación para cada clave parcial con sus atributos dependientes. Nos debemos asegurar de mantener una relación con la clave primaria original y todos los atributos que dependen funcionalmente en forma total de ella.

Descomponemos **PELÍCULA** (Título, Año, Duración, Tipo, Estudio, Dirección_Estudio, Actor) en dos tablas:

P1 (T, A, Ac)

P2 (T, A, D, Ti, E, Di)

De esta forma ya las tendremos en 2FN y si queremos las podemos renombrar para que esté más claro:

PELICULA_ACTOR (T, A, Ac)

PELICULA (T, A, D, Ti, E, Di)

3ª forma normal

Una relación está en 3FN si y solo si está en 2FN y todos sus atributos no clave dependen no transitivamente de la clave primaria. Dicho de otro modo, si y sólo si los atributos no clave son mutuamente independientes y son dependientes por completo de la clave primaria.

En este caso vemos como en la tabla PELICULA tenemos Di que depende de un atributo no clave E, por lo que debemos descomponer y crear una tabla que incluya el atributo o atributos no clave que determinen funcionalmente a otros no clave quedando de la siguiente forma:

PELICULA1 (T, A, D, Ti, E)

PELICULA2 (E, Di)

De esta forma ya las tendremos en 3FN y si queremos las podemos renombrar para que esté más claro:

PELICULA (T, A, D, Ti, E)

ESTUDIO (E, Di)

Finalmente me han quedado tres tablas de la siguiente manera que ya sí están en 3FN:

PELICULA_ACTOR (T, A, Ac)

PELICULA (T, A, D, Ti, E)

ESTUDIO (E, Di)

EJERCICIO 18. BIBLIOTECA

18. Dado el siguiente enunciado:

Una biblioteca guarda, para cada uno de sus préstamos, los siguientes datos:

Número, nombre, apellido, dirección y teléfono del socio - ISBN del libro - Título del libro - Código del autor del libro - Nombre y apellido del autor del libro - Fecha de nacimiento y de muerte del autor del libro - Tipo de autor(*) - Número de páginas del libro - Cantidad de días por los que se presta el libro - Fecha en la que se presta el libro - Fecha esperada de devolución - Fecha real de devolución.

(*) Un autor puede ser: autor principal, coautor, prologuista, traductor o compilador. Un mismo autor puede ser, por ejemplo, prologuista de un libro y compilador de otro.

Condiciones:

- 1) Cada socio tiene un número único que lo identifica.
- 2) Se registra un solo domicilio y un solo teléfono por cada socio.
- 3) Cada libro tiene un ISBN único que lo identifica. Suponer que la biblioteca solo tiene un ejemplar de cada libro.
- 4) Cada autor se identifica con un número único.
- 5) Un libro tiene siempre al menos un autor. Puede tener más de uno. No hay libros de autores desconocidos.
- 6) Un socio puede sacar muchos libros, pero solo uno por vez (no se le presta un nuevo libro hasta que no haya devuelto el anterior).
- 7) Para cada libro, se decide por cuántos días se prestará. Esta cantidad de días no depende del socio, sino del libro.
- 8) Un mismo lector puede llevarse varias veces el mismo libro.

Y después de hacer el diagrama Entidad-Relación y el paso al modelo relacional nos dan estas tablas que supuestamente sirven para modelar todo el enunciado anterior.

| SOCIO | PRESTAMO | LIBRO |
|------------------|---------------------------|--------------------------|
| <u>nro_socio</u> | <u>id_prestamo</u> | <u>isbn_libro</u> |
| nombre_socio | días_prestamo | titulo_libro |
| apellido_socio | fecha_prestamo | codigo_autor |
| direccion_socio | fecha_esperada_devolucion | nombre_autor |
| telefono_socio | fecha_devolucion | fecha_nac_y_muerte_autor |
| | | tipo_autor |
| | | paginas_libro |

El ejercicio consiste en buscar todos los fallos que tienen estas tablas y solucionarlos para que cumpla con todos los requerimientos del enunciado y además estén en 3FN.

EJERCICIO 18. BIBLIOTECA. SOLUCIÓN PROPUESTA

- 1) La tabla SOCIO tiene como clave nro_socio que corresponde al número de socio y lo identifica en forma única, por tanto, es correcto.
- 2) La tabla PRESTAMO no tenía sugeridos en el enunciado una clave, pero sabemos que necesitamos algún dato que identifique de forma única un préstamo puntual. Por lo tanto, se genera un identificador, por ejemplo, id_prestamo
- 3) La tabla libro, tiene como clave el ISBN, que sabemos es un numero identificador único para cualquier libro, con el nombre isbn_libro, por tanto, es correcto.
- 4) En la tabla PRESTAMO tenemos un atributo que no es necesario almacenar porque se puede calcular dependiendo de otros. Sabiendo la fecha de préstamo, y la cantidad de días del préstamo, se puede autocalcular la fecha esperada de devolución, por lo tanto, fecha_esperada_devolucion no es necesario así que ese atributo lo eliminaremos de la tabla.
- 5) La tabla PRESTAMO relaciona al socio que saca el préstamo del libro, y al libro que pide prestado, por lo tanto, nro_socio y isbn_libro deben estar en esta tabla, y son claves foráneas así que habrá que incluirlas quedando las tablas de la siguiente forma:

| SOCIO | PRESTAMO | LIBRO |
|------------------|--------------------------------------|-------------------|
| <u>nro_socio</u> | <u>id_prestamo</u> | <u>isbn_libro</u> |
| nombre_socio | dias_prestamo | titulo_libro |
| apellido_socio | fecha_prestamo | código_autor |
| direccion_socio | fecha_esperada_devolucion | nombre_autor |
| telefono_socio | fecha_devolucion | fechanac_autor |
| | nro_socio | fechamu_autor |
| | isbn_libro | tipo_autor |
| | | paginas_libro |

ANALICEMOS AHORA LAS TABLAS POR SEPARADO A VER SI CUMPLEN LAS FORMAS NORMALES:

- **TABLA SOCIO:**

Si leemos la condición 1: "Cada socio tiene un número único que lo identifica.", y la condición 2: "Se registra un solo domicilio y un solo teléfono por cada socio." Podemos afirmar que estas 2 condiciones confirman que la tabla SOCIO va a estar en:

- 1FN porque no hay atributos multivaluados.
- También está 2FN porque cualquier atributo que no es clave es posible encontrarla con un único valor de la clave primaria
- Y también en 3FN porque no hay dependencias funcionales transitivas.

- **TABLA LIBRO:**

Según la condición 3: "Cada libro tiene un ISBN único que lo identifica. Suponer que la biblioteca solo tiene un ejemplar de cada libro." Esto verifica que con solo tener el ISBN, isbn_libro en la tabla LIBRO, se identifica un único libro (si hubieran más de un ejemplar de un mismo libro, deberíamos prever la una tabla auxiliar para identificar cada uno de los ejemplares para un mismo isbn_libro), y confirma la condición de isbn_libro como clave primaria.



La tabla libros está en 1FN y también lo está en 2FN ya que cualquier tabla donde haya un solo atributo como clave lo está. Pero no está en 3FN ya que existe dependencia transitiva entre los atributos de la siguiente forma:

- Los atributos titulo_libro y paginas_libro sí dependen de isbn_libro.
- El atributo código_autor sí depende de isbn_libro pero los atributo nombre_autor, fechanac_autor, fechamu_autor, y tipo_autor dependen de codigo_autor por lo que aquí encontramos una dependencia transitiva que hay que evitar para que esté en 3FN.

Por lo tanto, hay que separarlas en 2 tablas: los datos de los autores estarán en una tabla independiente que llamaremos AUTOR. En consecuencia, en la tabla libro, codigo_autor era un atributo más, y ahora será una clave foránea para relacionarse con la tabla AUTOR. La tabla LIBRO está así en 3FN.

Según la condición 7: “Para cada libro, se decide por cuántos días se prestará. Esta cantidad de días no depende del socio, sino del libro.” Por lo tanto, dias_prestamo debería estar en la tabla LIBRO.

Por el momento omitiremos el último dato, que son los TIPOS DE AUTOR, el cual analizaremos en el siguiente paso. quedando de la siguiente manera:

| LIBRO | AUTOR |
|-------------------|---------------------|
| <u>isbn_libro</u> | <u>código_autor</u> |
| titulo_libro | nombre_autor |
| paginas_libro | fechanac_autor |
| dias_prestamo | fechamu_autor |
| código_autor | tipo_autor |

- TABLA PRESTAMO:

Según la condición 6: “Un socio puede sacar muchos libros, pero solo de a uno por vez (no se le presta un nuevo libro hasta que no haya devuelto el anterior).” Esto verifica que con solo tener el id_prestamo, en la tabla PRESTAMO, se identifica un único socio con un único libro (si se prestaran más de un libro a un mismo socio, deberíamos prever la una tabla auxiliar para identificar cada uno de los isbn_libro para un mismo id_prestamo). De esta manera, la tabla PRESTAMO también está en 2FN, y como no hay dependencias funcionales transitivas, también está en 3FN.

Según la condición 8: “Un mismo lector puede llevarse varias veces el mismo libro. Esto junto a lo que evaluamos de la condición 6, ya queda cubierto, ya que los prestamos quedan identificados en forma única con el id_prestamo, e identifican cada vez una fecha, un socio, y un libro... y no generaría ninguna condición a prever si el socio volviera a sacar prestado el mismo libro en otro momento... y no cambia la condición de la tabla PRESTAMO, que está en 3FN. Por tanto, hasta el momento tendríamos las tablas como nuestro a continuación:

| SOCIO | PRESTAMO | LIBRO | AUTOR |
|---|---|---|--|
| <u>nro_socio</u> nombre_socio apellido_socio direccion_socio telefono_socio | <u>id_prestamo</u> fecha_prestamo fecha_devolucion nro_socio isbn_libro | <u>isbn_libro</u> titulo_libro paginas_libro dias_prestamo código_autor | <u>código_autor</u> nombre_autor fechanac_autor fechamu_autor tipo_autor |

Todavía no hemos hablado de la condición número 5 que dice: “Un libro tiene siempre al menos un autor. Puede tener más de uno. No hay libros de autores desconocidos.”.

Con esta condición, ahora la tabla LIBRO deja de cumplir las condiciones de normalización. Para solucionar esto, la relación entre el ISBN de cada libro, atributo isbn_libro, y los autores, atributo codigo_autor, se establece a través de una tabla separada que llamaremos AUTORES. De esta manera, conformaríamos las tablas LIBRO y AUTORES, quedando en la 2FN, y como no hay dependencias funcionales transitivas, también está en 3FN.

Según la última condición citada en el enunciado: “Un autor puede ser: autor principal, coautor, prologuista, traductor o compilador. Un mismo autor puede ser, por ejemplo, prologuista de un libro y compilador de otro.”. De acuerdo al enunciado, el tipo_autor, es un atributo que puede variar de acuerdo al libro, por lo tanto, este atributo NO DEBERIA estar en la tabla AUTOR. Este atributo, como está relacionado estrechamente con la actividad que realiza un autor específico, identificado con el atributo código_autor, en un libro en particular, identificado con el atributo isbn_libro, DEBERIA estar en la tabla AUTORES, que acabamos de crear.

| SOCIO | PRESTAMO | LIBRO | AUTOR | AUTORES |
|---|---|---|--|--|
| <u>nro_socio</u> nombre_socio apellido_socio direccion_socio telefono_socio | <u>id_prestamo</u> fecha_prestamo fecha_devolucion nro_socio isbn_libro | <u>isbn_libro</u> titulo_libro paginas_libro dias_prestamo | <u>código_autor</u> nombre_autor fechanac_autor fechamu_autor | <u>isbn_libro</u> <u>código_autor</u> tipo_autor |

Y ahora ya sí tenemos todas las tablas normalizadas hasta la 3FN y cumpliendo todos los requerimientos del enunciado.

EJERCICIO 19. COMPAÑÍA DE SEGUROS

19. Dado el siguiente enunciado:

Una compañía de seguros desea que se haga un diseño de una base de datos para gestionar toda la información referente a los seguros que ofrece, los clientes a los que atiende y los agentes de seguros que trabajan para la compañía. Esta compañía ofrece tres tipos de seguros:

- Seguros de Hogar: los seguros de este tipo ofrecidos por la compañía están ofertados de forma fija (es decir se han hecho estudios previos), según el valor del continente (la casa), el contenido (muebles, electrodomésticos, joyas, etc.), riesgos auxiliares (responsabilidad civil, asalto y otros). Para cada oferta hay una prima asignada.
- Seguros de Vida: de la misma forma que los de hogar, existen varias ofertas fijas según la edad y profesión del cliente, y la cobertura económica del seguro. De la misma forma que en los seguros de Hogar, existe una prima fija para cada oferta.
- Seguros de Automóvil: también existen ofertas fijas, según la categoría de coche (utilitario, gama media, gama alta, gran turismo, lujo, etc.), años del vehículo, edad del conductor y cobertura (todo riesgo, franquicia, terceros, etc.). A cada una de estas ofertas le corresponde una prima.

Para llevar un control de las comisiones que se llevan los agentes y de sus carteras correspondientes, la compañía necesita tener almacenados los datos de los agentes, considerándose de interés el nombre, NIF, dirección y teléfono. Para el pago de comisiones y carteras (se entiende por "cartera" la comisión anual del agente mientras el seguro este vigente), será necesario saber qué agente ha realizado qué seguro y en qué fecha.

La compañía considera como datos de interés referentes al cliente (sea cual sea el seguro que contrate), los siguientes: Nombre, dirección, teléfono y NIF.

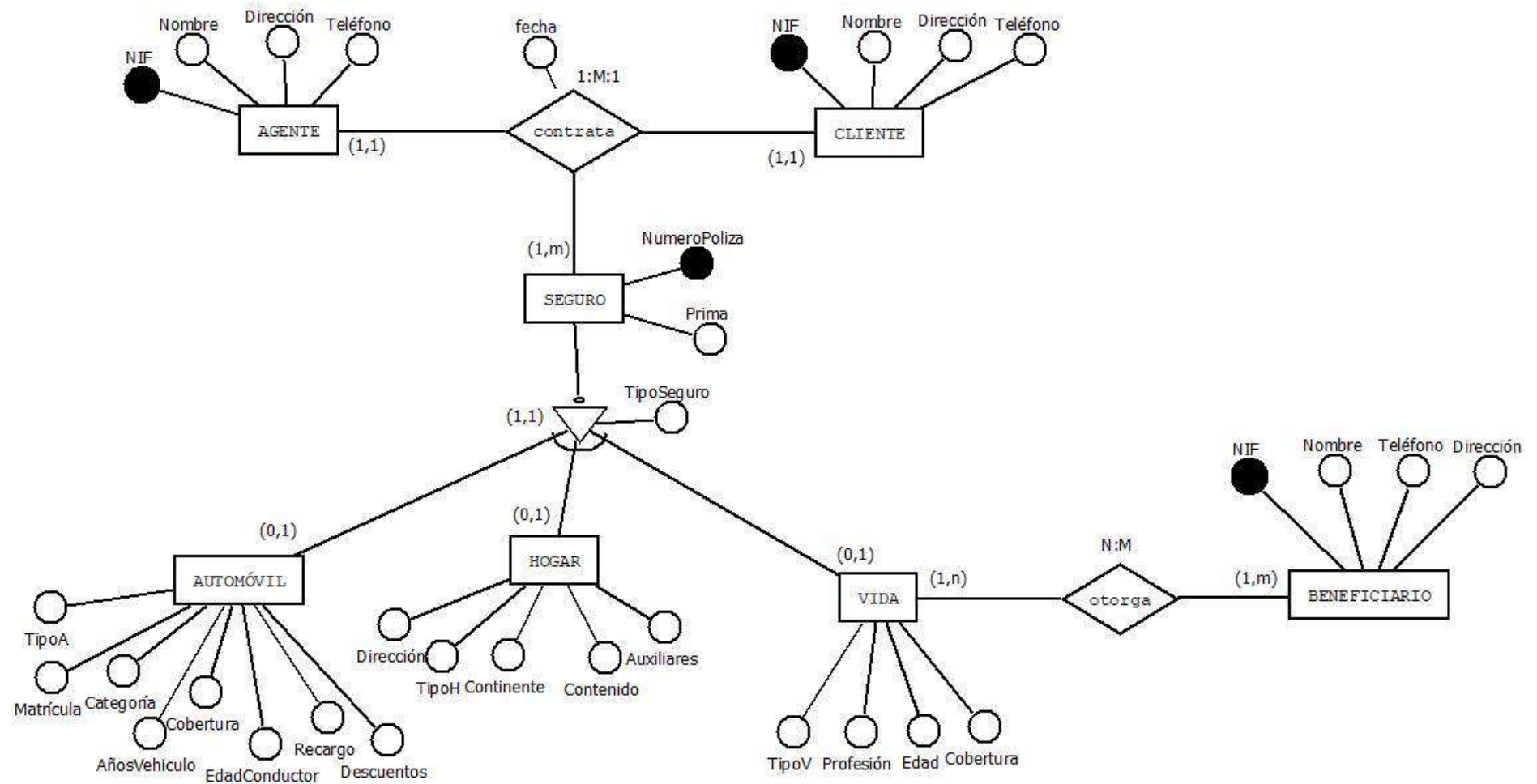
Otras consideraciones sobre la contratación de seguros por parte del cliente son:

- Seguros Hogar: fecha del contrato del seguro y dirección del inmueble asegurado.
- Seguros Automóvil: fecha contratación, matrícula del vehículo, recargos y descuentos.
- Otras consideraciones: Un cliente puede contratar más de un seguro de Vida, más de un seguro de Hogar y más de un seguro de Automóvil. Además, estos contratos pueden realizarse a través de distintos agentes. Los beneficiarios de seguros de vida pueden serlo de varios seguros, e incluso de varios clientes distintos. Por supuesto un cliente puede nombrar a varios beneficiarios de un mismo seguro de vida.

Se pide:

- a) Diseñar el diagrama Entidad-Relación correspondiente al enunciado anterior.
- b) Realizar el paso al modelo relacional (paso a tablas), asegurarnos de que las tablas estén normalizadas hasta la 3FN y realizar el grafo relacional.



EJERCICIO 19. COMPAÑÍA DE SEGUROS. SOLUCIÓN PROPUESTA

PASO A TABLAS

Para realizar el paso a tablas se siguen los siguientes criterios:

- Todas las entidades se convierten en tabla y todo atributo se transforma en columna dentro de la tabla. En nuestro ejercicio se crearán tablas de las entidades AGENTE, CLIENTE, SEGURO y BENEFICIARIO con sus correspondientes atributos.
- Las entidades AUTOMÓVIL, HOGAR y VIDA que forman parte de una especialización también tendrán su propia tabla con sus atributos correspondientes y donde su atributo clave o principal para todas ellas es el atributo clave de la entidad SEGURO.
- Respecto a las relaciones de este diagrama he seguido los siguientes criterios:
 - o Las relaciones N:M se convierten a tabla, teniendo como atributos los propios de la relación y los atributos claves de las entidades que relacionan. Tendrán como clave primaria la concatenación de los atributos principales de cada una de las entidades que relacionan que serán clave ajena respecto a cada una de las tablas donde ese atributo es clave primaria. Así por ejemplo en este caso tenemos la relación “otorga”.
 - o La relación contrata es de tipo 1:M:1. En tal caso debemos propagar las claves principales de las entidades que actúan con cardinalidad 1 a la tabla que actúa con cardinalidad M (Tabla SEGURO). El atributo “fecha” de la relación contrata igualmente lo trasladamos a la entidad seguro y de esta manera la relación “contrata” no sería necesario pasarla a tabla.

De esta forma, se obtienen las siguientes tablas donde aparecen subrayados los atributos clave:

AGENTE (NIF, Nombre, Dirección, Teléfono)

CLIENTE (NIF, Nombre, Dirección, Teléfono)

SEGURO (NumeroPoliza, Prima, TipoSeguro, Fecha, NIF_Cliente, NIF_Agente)

AUTOMOVIL (NumeroPolizaSeguro, TipoA, Matrícula, Categoría, AñosVehículo, Cobertura, EdadConductor, Recargo, Descuentos)

HOGAR (NumeroPolizaSeguro, TipoH, Dirección, Continente, Contenido, Auxiliares)

VIDA (NumeroPolizaSeguro, TipoV, Profesión, Edad, Cobertura)

OTORGA (NumeroPolizaSeguro, NIF_Beneficiario)

BENEFICIARIO (NIF, Nombre, Dirección, Teléfono)

NORMALIZACIÓN

El proceso de normalización es un proceso importante en nuestra base de datos. Los objetivos del mismo son los siguientes:

- Controlar la redundancia de la información.
- Evitar pérdidas de información.
- Capacidad para representar toda la información.
- Mantener la consistencia de los datos.

1ª Forma Normal:

Una relación está en 1FN si no existen atributos multivaluados ni relaciones anidadas. Dicho de otra forma, estará en 1FN si los atributos no clave, dependen funcionalmente de la clave.

Comprobando las tablas obtenidas en el paso a tablas se observa que ya se encuentran en 1ª Forma Normal.

2ª Forma Normal

Una relación está en 2FN si está en 1FN y todos los atributos no clave tienen dependencia funcional completa de la clave primaria. De acuerdo con esta definición, cada tabla que tiene un atributo único como clave, está en segunda forma normal.

Comprobando las tablas obtenidas en el paso a tablas se observa que ya se encuentran en 2ª Forma Normal.

3ª forma normal

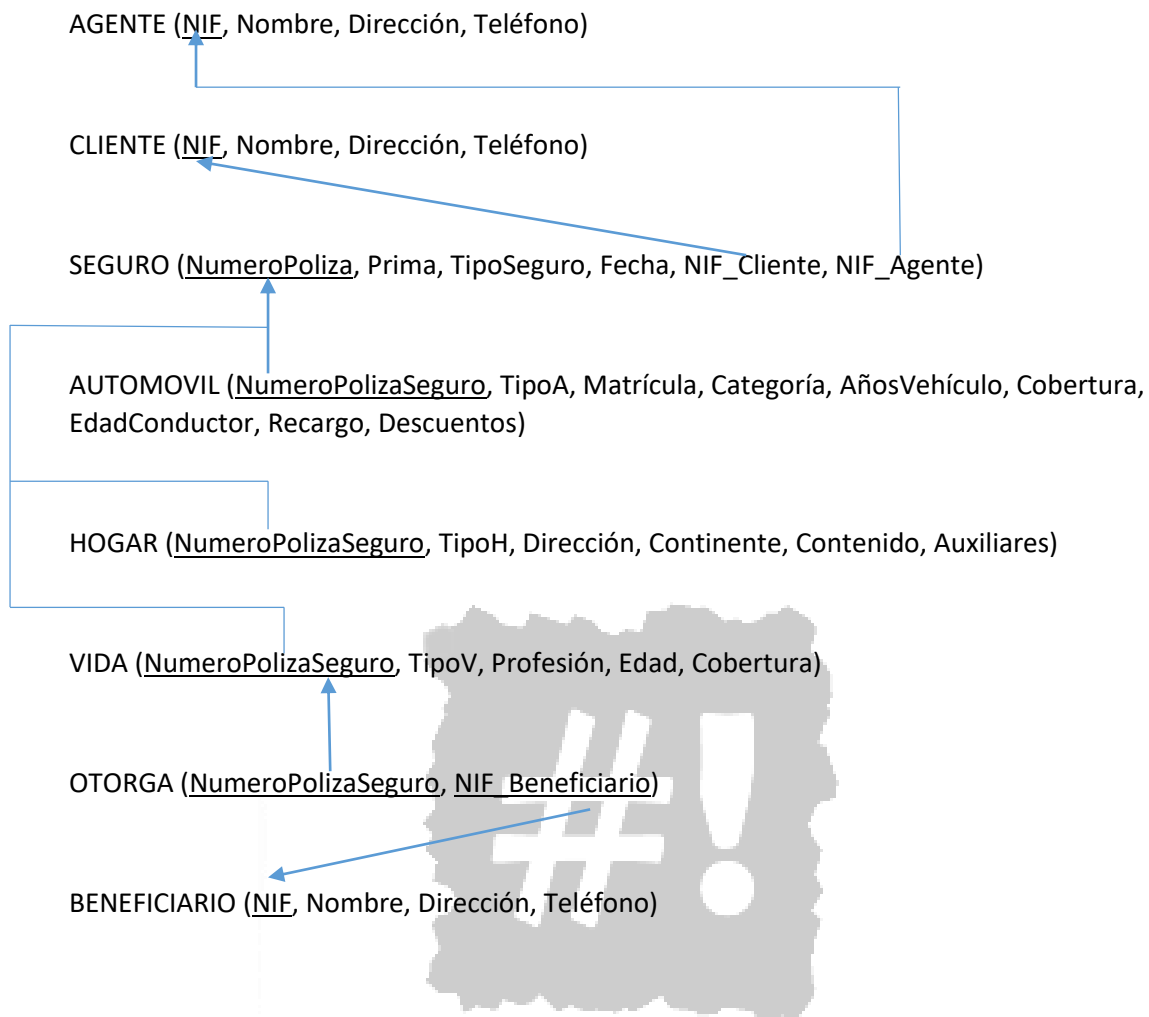
Una relación está en 3FN si y solo si está en 2FN y todos sus atributos no clave dependen no transitivamente de la clave primaria. Dicho de otro modo, si y sólo si los atributos no clave son mutuamente independientes y son dependientes por completo de la clave primaria.

Comprobando las tablas obtenidas en el paso a tablas se observa que ya se encuentran en 3ª Forma Normal.

Por tanto, las tablas obtenidas en el paso a tablas se encuentran normalizadas hasta la 3ªFN.

GRAFO RELACIONAL.

El grafo relacional es un grafo compuesto de un conjunto de nodos, donde cada nodo representa un esquema de relación, es decir, una tabla de la BD con su nombre y atributos. Se dibuja, además, un conjunto de arcos que conectan los atributos que constituyen la clave ajena de una tabla respecto a la clave primaria de la otra tabla referenciada. Los arcos están direccionados de modo que el arco parta de la clave ajena y la flecha señale a la tabla referenciada.



Preparador Informática