

EJERCICIO: NORMALIZACIÓN DE BD

25 de agosto de 2017

Generar una BD en 3ª Forma Normal para el siguiente caso

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

Numero, 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

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 de a 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.

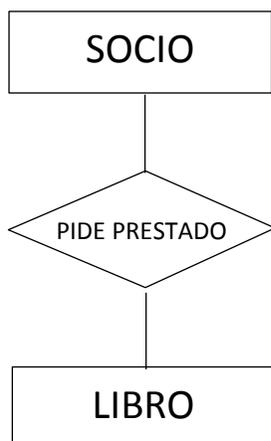
Si surgen dudas respecto a otras condiciones, consultar con el docente y registrar por escrito.

(*) 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.

RESOLUCION

Hay muchas maneras de empezar la resolución, ninguna está mal, y es un hecho que a medida que se van resolviendo las relaciones de los datos en función del enunciado y sus condiciones, tras el proceso se termina arribando al mismo modelo...

1 - INICIALMENTE, armamos el Diagrama Entidad-Relación (DER). Algunos empezaron con esta primera relación



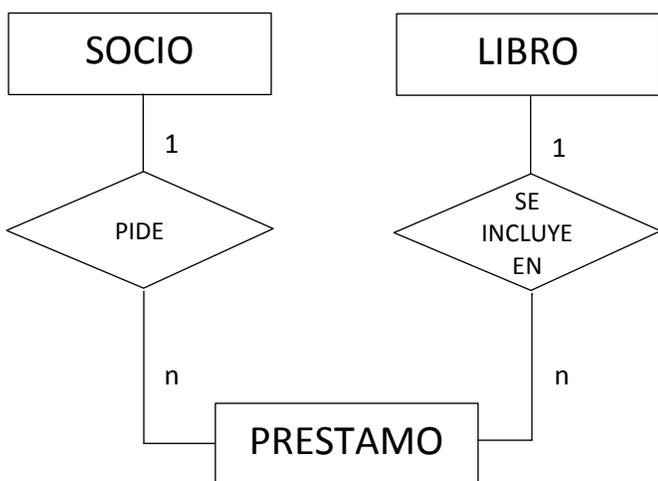
Donde:

- **SOCIO**, tiene los siguientes datos: Numero, nombre, apellido, dirección y teléfono del socio
- **LIBRO**: 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

PERO, OBSERVAMOS que hay datos que deberíamos de poner en la acción **PIDE PRESTADO**: 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

Esto nos lleva a entender que si **PIDE PRESTADO** debe almacenar datos o propiedades, debemos cambiarlo a una entidad, por lo tanto, arribamos al siguiente diagrama:

2 – “PIDE PRESTADO” PASA A SER LA ENTIDAD “PRESTAMO”, y lo aplicamos al DER. (Algunos empezaron habiendo resuelto desde el principio esta relación)



Donde:

- **SOCIO**, tiene los siguientes datos: Numero, nombre, apellido, dirección y teléfono del socio
- **LIBRO**: 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
- **PRESTAMO**: 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

3 – ARMADO DE TABLAS

Con esta primer forma, vamos a presentar las tablas que se conformarían, para comenzar con la normalización:

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

OBSERVACION: REFERENCIAS Y CRITERIOS DE LOS NOMBRES: *Por una cuestión de compatibilidad con cualquier sistema, sea el lenguaje en el que se trabaja, sea el sistema operativo con el que se trabaje, tanto en el servidor como en el usuario, etc., se recomienda que a la hora de darle nombre a los atributos o propiedades:*

- *No se deben usar símbolos, letras acentuadas, ni caracteres especiales*
- *En lo posible, se utilizaran nombres que remitan de forma única y directa al atributo o propiedad que estamos identificando. Esto es algo que solo será de nuestra elección en el caso de trabajos que diseñemos nosotros es del principio. Esto no es así, cuando se trabaja en grupo y los datos con los que se trabajan y el diseño es compartido, o en el caso de que estemos trabajando con un diseño ya establecido (reforma o adecuación de una base de datos)*
- *En lo posible, se simplificaran los nombres sin perder su significado. Generalmente se usan abreviaturas y/o se eliminan los nexos de los enunciados:
Ejemplo: "isbn del libro" escribirlo como "isbn libro"
Ejemplo: "numero de socio" escribirlo como "nro socio"*
- *No se deben usar espacios en blanco: por lo tanto, hay dos alternativas que se suelen aplicar:
 - *Reemplazar los espacios en blanco por guiones bajos, como aplique yo en las tablas presentadas arriba
Ejemplo: "isbn libro" escribirlo como "isbn_libro"*
 - *Eliminar los espacios en blanco y diferenciar las palabras escribiendo su carácter inicial en mayúsculas:
Ejemplo: "isbn libro" escribirlo como "IsbnLibro"**
- *Se debe aplicar **un único criterio** para todos los nombres:
Por ejemplo: en el ejercicio actual, todos los nombre están en minúscula y los espacios en blanco fueron reemplazados por guiones bajos*
- *Es importante no cambiar los nombres de los atributos durante todo el desarrollo del trabajo. Principalmente, suele ser una error común cambiar por mayúsculas y minúsculas:
Ejemplo: "ID_socio" es diferente a "id_socio", y diferente a "Id_Socio"
La razón principal de esta precaución es que muchos de los lenguajes de programación y sistemas operativos son **Case Sensitive**, o sea que diferencian los nombres si se los escribe con alguna variación de minusulas y mayúsculas.
En el ejemplo anterior, "ID_socio", "id_socio", y "Id_Socio" se consideran como 3 nombres diferentes*

PRIMER ANALISIS DE LAS TABLAS:

- Claves primarias:
 - La tabla SOCIO tiene como clave **nro_socio** que corresponde al número de socio y lo identifica en forma única
 - 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, genero un identificador **id_prestamo**
 - La tabla libro, tiene como clave el ISBN, que sabemos es un numero identificador único para cualquier libro, con el nombre **isbn_libro**
- Eliminamos los datos que se pueden autocalcular:
 - 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**
- Normalizacion a 1FN de las tablas:
 - La tabla SOCIO, teniendo como clave primaria **nro_socio**, esta en 1FN
 - La tabla PRESTAMO, teniendo como **clave primaria id_prestamo**, esta en 1FN. Esta tabla 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**.
 - En la tabla LIBRO, la clave **isbn_libro** es la clave primaria.
Pero **no cumple con la atomicidad de los atributos**, porque **fecha_nac_y_muerte_autor**, esto se resuelve separándolos en 2 atributos simples: **fechanac_autor** y **fechamu_autor**. Con este cambio la tabla está en 1FN.

Por lo tanto, volcamos todas estas modificaciones en las tablas:

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

Ahora continuando con este análisis, evaluamos las condiciones e intentaremos llevar un orden lógico, para hacer más claro el análisis:

- Según 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.”, Estas 2 condiciones confirman que **la tabla SOCIO**:
 - **no solo esta en 1FN,**
 - **sino también en 2FN**, cualquier atributo que no es clave es posible encontrarla con un único valor de la clave primaria
 - y como no hay dependencias funcionales transitivas, **también esta en 3FN.**
- 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 IBSN, **isbn_libro** en la tabla LIBRO, se identifica un único libro (si hubieran mas

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.**

- Según la **condición 4:** “Cada autor se identifica con un número único.”. Esto, en principio, no nos aporta ningún dato adicional, ya que hasta este punto verificamos que un libro se identifica de manera única a un libro, y esto también identificaría al autor.
Lo que no está escrito, pero sabemos de nuestra vida real, es que **un autor puede haber escrito varios libros**, esta situación nos indicaría que:
 - que la tabla libros está en 1FN, si consideramos una clave doble, conjugando *isbn_libro* y *código_autor*,

<i>isbn_libro</i>	<i>titulo_libro</i>	<i>codigo_autor</i>	<i>nombre_autor</i>	<i>fechanac_autor</i>	<i>fechamu_autor</i>	<i>tipo_autor</i>	<i>paginas_libro</i>
11111	LA LUNA	XXX3	JUAN PEREZ	1928	2003	AUTOR PRINCIPAL	358
11112	EL SOL	XXX4	LALO LANDA	1956	0	TRADUCTOR	425
11113	LA TIERRA	XXX3	JUAN PEREZ	1928	2003	AUTOR PRINCIPAL	268

- pero no está en 2FN, ya que:
 - los atributos *titulo_libro* y *paginas_libro* dependen de *isbn_libro*
 - pero los atributos *nombre_autor*, *fechanac_autor*, *fechamu_autor*, y *tipo_autor* dependen de la clave completa (*codigo_autor* + *isbn_libro*)
 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 esta así en 2FN.
 - En la tabla AUTOR, *codigo_autor* es una clave primaria**, y esta tabla, así conformada, **está en 2FN**

- OMITIREMOS por el momento la condición 5, hasta que este creada la tabla AUTOR**

- 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 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**
- Según la **condición 8:** “Un mismo lector puede llevarse varias veces el mismo libro.
Esto, de acuerdo 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

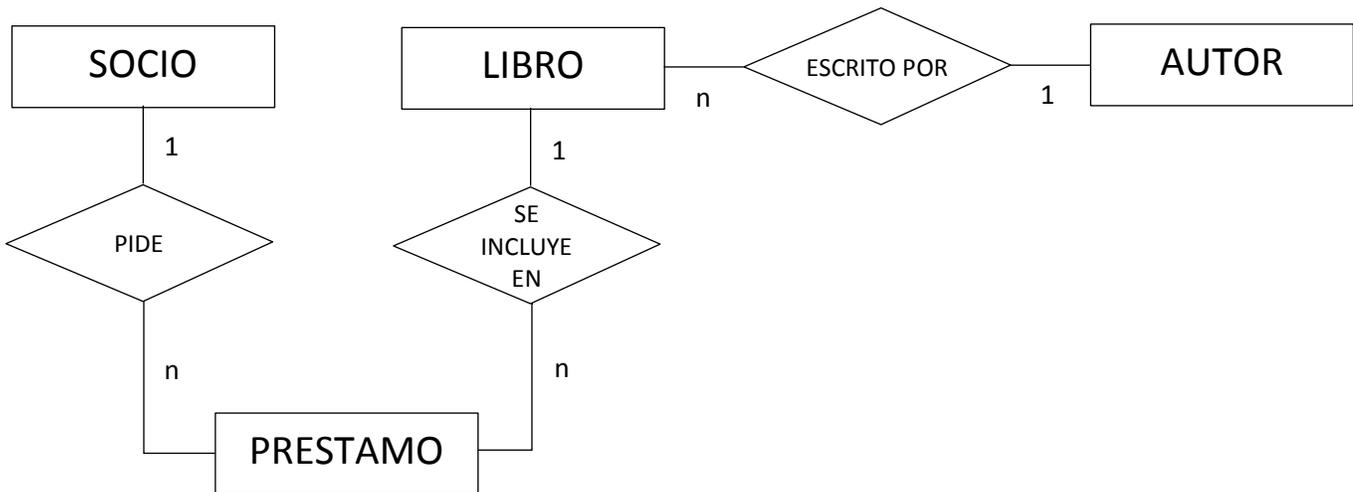
Hasta acá, el análisis de las condiciones, en primera instancia, por el momento **omitiremos también el último dato, que son los TIPOS DE AUTOR**, el cual analizaremos en el siguiente paso.

SEGUNDO ANALISIS DE LAS TABLAS:

Haciendo todas estas modificaciones en las tablas quedarían de la siguiente manera:

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

Nuestro DER, se vería así:



Las tablas SOCIO y PRESTAMO ya están en 3FN, así que terminaremos de analizar la tabla AUTOR

Ahora continuando con este análisis, evaluamos las condiciones las condiciones restantes:

- Según la **condición 5**: “Un libro tiene siempre al menos un autor. Puede tener más de uno. No hay libros de autores desconocidos.”.

<u>isbn_libro</u>	titulo_libro	paginas_libro	dias_prestamo	<u>codigo_autor</u>
11111	LA LUNA	325	10	XXX3
11111	LA LUNA	325	10	XXX8
11112	EL SOL	623	7	XX12
11113	LA TIERRA	268	12	XXX5

Con esta condición, ahora la tabla LIBRO:

- nuevamente deja de estar en 2FN (y en consecuencia tampoco en 3FN), ya que para un mismo *isbn_libro*, debería de asociarse con varios *codigo_autor*:
 - Si consideramos la clave doble (*isbn_libro* + *codigo_autor*), esta en 1FN
 - Los atributos *titulo_libro*, *paginas_libro*, y *dias_prestamo*, dependen únicamente de *isbn_libro*
 - El *codigo_autor* es el único atributo que obliga a repetir la clave principal y los otros atributos para poder referir a autores diferentes.

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.”.

<u>codigo_autor</u>	nombre_autor	fechanac_autor	fechamu_autor	tipo_autor
XXX3	JUAN PEREZ	1928	2003	AUTOR PRINCIPAL
XXX3	JUAN PEREZ	1928	2003	PROLOGUISTA
XXX4	LALO LANDA	1956	0	TRADUCTOR
XXX5	LALI GERA	1917	1987	COMPILADOR

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 realizo un autor específico, indentificado con el atributo *código_autor*, en un libro en particular, indentificado con el atributo *isbn_libro*, **DEBERIA estar en la tabla AUTORES,** que recién creamos.

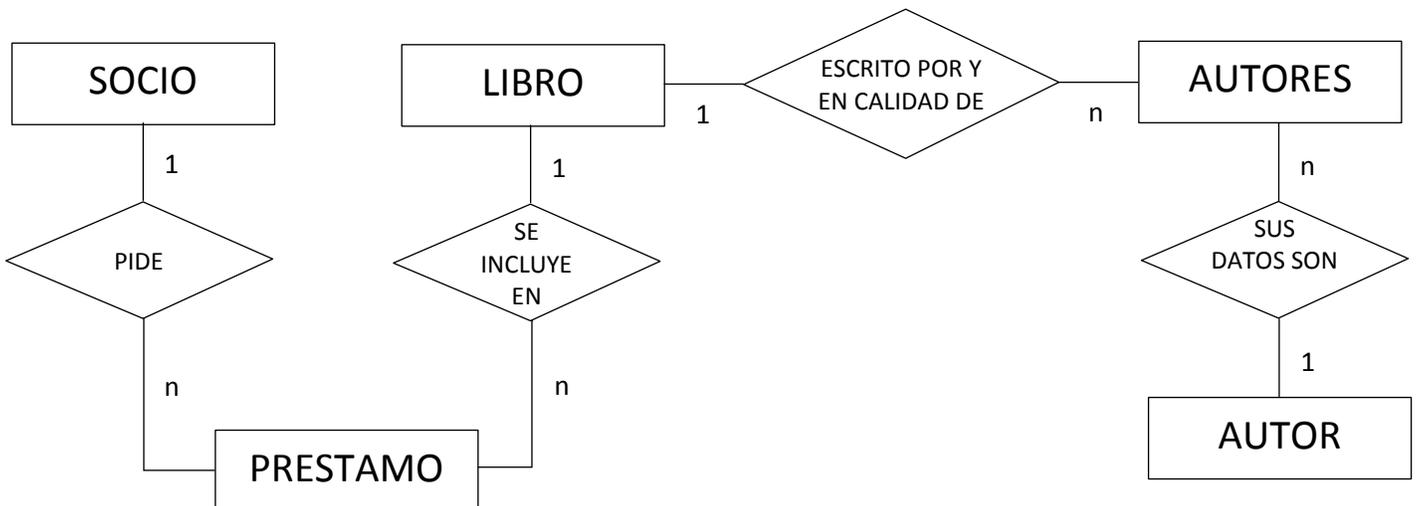
isbn_libro	codigo_autor	tipo_autor
11111	XXX3	AUTOR PRINCIPAL
11112	XXX3	PROLOGUISTA
11112	XXX4	TRADUCTOR
11113	XXX5	COMPILADOR

DISPOSICION FINAL, TABLAS NORMALIZADAS:

Haciendo todas estas modificaciones en las tablas, quedarían de la siguiente manera:



EL DER FINALMENTE SE VERIA ASI:



BASE DE DATOS NORMALIZADA, TABLAS FINALES CON LOS CAMPOS RELACIONADOS:

