75.15 / 75.28 / 95.05 - Base de Datos

El Modelo Lógico Relacional

Mariano Beiró

Dpto. de Computación - Facultad de Ingeniería (UBA)

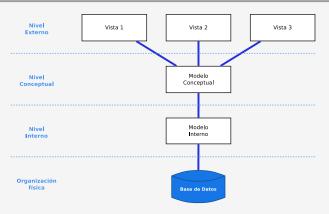
30 de agosto de 2022

Temas

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relaciona
 - Principios
 - Ejemplos
- 4 Bibliografía

Arquitectura de 3 capas ANSI/SPARC



Modelo lógico

Es un paso intermedio entre el modelo conceptual y el modelo interno, que formaliza las descripciones del modelo conceptual.

- 1 Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relaciona
 - Principios
 - Ejemplos
- 4 Bibliografía

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

Dominios y Relaciones

[ELM16 5.1]

- Las bases de datos relacionales se basan en un modelo lógico introducido por Codd en 1970, conocido como modelo relacional.
- El mismo es una formalización matemática basada en el concepto de relación.
- Algunas definiciones:
 - Dominio: Es un conjunto de valores homogéneos.
 - Ejemplo: $D_1 = \{Barcelona, Sevilla, Buenos Aires\}$
 - $D_2 = \{Argentina, España, Chile\}$
 - Producto cartesiano: $A \times B$ se define como el conjunto de pares (a,b) que cumplen que $a \in A$ y $b \in B$.
 - En nuestro ejemplo, $D_1 \times D_2 = \{(Barcelona, Argentina), (Barcelona, España), (Barcelona, Chile), (Sevilla, Argentina), (Sevilla, España), (Sevilla, Chile), (Buenos Aires, Argentina), (Buenos Aires, España), (Buenos Aires, Chile)}$
 - Relación: es un subconjunto de un producto cartesiano.
 - En el ejemplo anterior podría especialmente interesarnos la siguiente relación: R = {(Barcelona, España), (Sevilla, España), (Buenos Aires, Argentina)}

Formalización

- En el modelo relacional, un nombre de relación R junto con una lista de atributos asociados se denomina esquema de relación, y se denota de la siguiente manera:
 - \blacksquare $R(A_1, A_2, ..., A_n)$
 - Ejemplo: Películas(nombre_película, año, nombre_director, cant_oscars)
- Cada uno de los atributos A_i de un esquema de relación está asociado a un dominio particular, $dom(A_i)$. En nuestro ejemplo:
 - nombre_película → dom(nombre_pelicula) = string
 - lacksquare año o $dom(ilde{ano})=\mathbb{N}^+$
 - nombre_director → dom(nombre_director) = string
 - cant_oscars \rightarrow dom(cant_oscars) = \mathbb{N}^+

Relaciones y tuplas

- Una relación R con esquema de relación $R(A_1, A_2, ..., A_n)$ -estando los atributos A_i asociados a los dominios $D_i = dom(A_i)$ -es un subconjunto del producto cartesiano $D_1 \times D_2 \times ... \times D_n$.
 - Dado el esquema de relación Películas(nombre_película, año, nombre_director, cant_oscars), se cumple que:
 - Películas ⊂ dom(nombre_pelicula) × dom(año) × dom(nombre_director) × dom(cant_oscars)
 - Por ejemplo:
 - Películas = {(Kill Bill, 2003, Quentin Tarantino, 0), (Django Unchained, 2012, Quentin Tarantino, 2), (Star Wars III, 2005, George Lucas, 0), (El Cisne Negro, 2010, Darren Aronofsky, 1)}
- Una elemento de una relación se denomina tupla.
 - $t = (Kill \ Bill, 2003, Quentin \ Tarantino, 0)$ es una tupla de Películas.
 - Lo denotamos como $t \in Películas$.
 - u = (Kill Bill, 2003, Quentin Tarantino, 3) no es una tupla de Películas.
 - u ∉ Películas.

Relaciones y tuplas

Al margen...

Si pensamos a una relación como un predicado, podemos <u>predicar</u> que Películas($Kill\ Bill$, 2003, $Quentin\ Tarantino$, 3) = F.

- El valor tomado por un atributo A en la tupla t se denota t[A] o t.A.
 - lacktriangledown $t[nombre_director] = t.nombre_director = Quentin Tarantino.$
 - t[nombre_pelicula, año] = t.(nombre_pelicula, año) = (Kill Bill, 2003).
- La cardinalidad de una relación R es la cantidad de tuplas que posee.
 - La simbolizaremos n(R).
 - Ejemplo: n(Películas) = 4.

Tablas

Una forma útil de representar una relación es a través de una tabla en la que las columnas representan los atributos y las filas representan las tuplas. En el caso de la relación Películas:

nombre_película	año	nombre_director	cant_oscars
Kill Bill	2003	Quentin Tarantino	0
Django Unchained	2012	Quentin Tarantino	2
Star Wars III	2005	George Lucas	0
El Cisne Negro	2010	Darren Aronofsky	1

Otra nomenclatura comúnmente utilizada –y más vinculada al nivel físico– habla de archivos en lugar de tablas, registros en lugar de filas, y campos en lugar de columnas.

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

Restricciones de dominio

[ELM16 5.1 5.2.1]

- Las relaciones del modelo relacional representan generalmente entidades o interrelaciones de nuestro modelo de datos.
- Deben cumplir una serie de restricciones de distintos tipos.

Restricciones de dominio

- Las restricciones de dominio especifican que dado un atributo *A* de una relación *R*, el valor del atributo en una tupla *t* debe pertenecer al dominio *dom*(*A*).
- Algunos dominios posibles:
 - \blacksquare N, N⁺, \mathbb{R} , \mathbb{R}^+ .
 - caracter, string, valor booleano, fecha, conjunto finito (categorías).
- En el modelo relacional se puede permitir que algunos de los atributos tomen un valor nulo (NULL).
- Los atributos deben ser atómicos (no se permiten atributos compuestos o multivaluados).

Restricciones de unicidad

[ELM16 5.2.2]

- No pueden existir dos tuplas distintas que coincidan en los valores de todos sus atributos.
 - I.e., una tupla no puede estar dos veces!
- Sin embargo, generalmente existe un subconjunto SK del conjunto de atributos $(A_1, A_2, ..., A_n)$ de R que cumple la condición de que dadas dos tuplas $s, t \in R$, las mismas difieren en al menos uno de los atributos de SK.
- Cuando un subconjunto SK cumple esta propiedad, diremos que SK es una superclave de R.
- Nos interesan aquellas superclaves que son <u>minimales</u>, es decir que no admiten ningún subconjunto propio con la misma propiedad. A estas superclaves las llamaremos claves candidatas o simplemente claves.
- De entre todas las claves candidatas elegiremos una como clave primaria de la relación. La indicaremos subrayada en el esquema.

Restricciones de unicidad

Ejemplo

Películas(nombre_película, año, nombre_director, cant_oscars)

- Si suponemos que no puede haber dos películas con el mismo nombre
 - "nombre_película" es clave candidata.
 - Es la única. La designaremos como clave primaria.
 - Lo representamos como: Películas(<u>nombre_película</u>, año, nombre_director, cant_oscars).
 - {nombre_película, cant_oscars} es superclave, pero no es clave candidata porque no es minimal.
- Si admitimos que existen películas distintas con el mismo nombre
 - Deberíamos crear un atributo "id" que nos permita identificar a cada película.
 - El esquema de relación sería ahora:
 Películas(<u>id</u>, nombre_película, año, nombre_director, cant_oscars)
 - "id" será la clave primaria.

Restricciones de unicidad

Veamos otro ejemplo

ActoresPelículas(nombre_película, año, nombre_director, cant_oscars, nombre_actor)

nombre_película	año	nombre_director	cant_oscars	nombre_actor
Kill Bill	2003	Quentin Tarantino	0	Uma Thurman
Django Unchained	2012	Quentin Tarantino	2	Leonardo DiCaprio
Django Unchained	2012	Quentin Tarantino	2	Samuel L. Jackson
Star Wars III	2005	George Lucas	0	Samuel L. Jackson
Star Wars III	2005	George Lucas	0	Natalie Portman
El Cisne Negro	2010	Darren Aronofsky	1	Natalie Portman

- Ya no alcanza "nombre película" para identificar una tupla.
- La clave de la relación ActoresPelículas es {nombre_película, nombre_actor}!

Nota: Desde ya, este es un diseño desprolijo, por contener información redundante sobre las películas en distintas tuplas.

Esquemas de base de datos

[ELM16 5.2.3]

- Las bases de datos almacenan múltiples esquemas de relación, muchas veces relacionados entre ellos.
- En el modelo relacional, una base de datos se representa a través de un esquema de base de datos relacional.
- Un esquema de base de datos relacional S es un conjunto de esquemas de relación $S = \{R_1, R_2, ..., R_m\}$ junto con una serie de restricciones de integridad.
- Ejemplo: Cine = {Películas, Actores, Actuaciones}, en donde:
 - Películas(nombre_película, año, nombre_director, cant_oscars)
 - Actores(nombre_actor, país)
 - Actuaciones(nombre_película, nombre_actor)

Esquemas de base de datos

La siguiente es una posible instancia de nuestra base de datos Cine:

PELÍCULAS					
nombre_película	año	nombre_director	cant_oscars		
Kill Bill	2003	Quentin Tarantino	0		
Django Unchained	2012	Quentin Tarantino	2		
Star Wars III	2005	George Lucas	0		
El Cisne Negro	2010	Darren Aronofsky	1		

ACTORES					
nombre_actor	país				
Uma Thurman	Estados Unidos				
Leonardo DiCaprio	Estados Unidos				
Samuel L. Jackson	Estados Unidos				
Natalie Portman	Israel				

ACTUACIONES					
r					
n					
Caprio					
ckson					
ckson					
an					
an					
ck ck ia					

Restricciones de integridad

[ELM16 5.2.4]

- Restricción de integridad de entidad: La clave primaria de una relación no puede tomar el valor nulo.
- Restricción de integridad referencial: Cuando un conjunto de atributos FK de una relación R hace referencia a la clave primaria de otra relación S, entonces para toda tupla de R debe existir una tupla de S cuya clave primaria sea igual al valor de FK, a menos que todos los atributos de FK sean nulos.

Cine = {Películas, Actores, Actuaciones}

- Películas(<u>nombre_película</u>, año, nombre_director, cant_oscars)
- Actores(<u>nombre_actor</u>, país)
- Actuaciones(nombre_película, nombre_actor)
- Ejemplo: Si una tupla en Actuaciones hace referencia "Star Wars III", entonces debe existir "Star Wars III" en la relación Películas.

Restricciones de integridad

- Formalmente:
 - Sean $R(A_1, A_2, ..., A_r)$ y $S(B_1, B_2, ..., B_s)$ dos esquemas de relación.
 - $FK \subset (A_1, A_2, ..., A_r)$ hace referencia a S, cuya clave primaria es $PK \subset (B_1, B_2, ..., B_s)$.
 - Entonces:

$$\forall t \in R : t[FK] \neq NULL \rightarrow \exists s \in S : s[PK] = t[FK].$$

- FK se denomina clave foránea de S en R.
- Estas restricciones generalmente surgen de las interrelaciones entre entidades de nuestro modelo conceptual.

Restricciones de integridad

```
Cine = {Películas, Actores, Actuaciones}

Películas(nombre_película, año, nombre_director, cant_oscars)

Actores(nombre_actor, país)

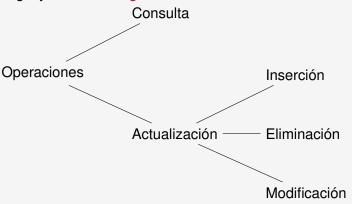
Actuaciones(nombre_película, nombre_actor)
```

- En nuestro ejemplo de la base de datos Cine, Actuaciones.nombre_película es clave foránea y hace referencia a la relación Películas. Asímismo, Actuaciones.nombre_actor es clave foránea y hace referencia a la relación Actores.
- Indicaremos a las claves foráneas con un subrayado punteado.

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

Operaciones

Las operaciones del modelo relacional se especifican a través de lenguajes como el álgebra relacional o el cálculo relacional.



Operaciones e integridad

[ELM16 5.3]

- Operaciones de consulta
 - No modifican ninguna relación existente.
 - Por lo tanto no violan ningún tipo de restricción.
- Operaciones de inserción de tuplas
 - Pueden violar restricciones de dominio, de unicidad y de integridad de entidad o referencial.
 - El SGBD debería rechazar una inserción que viola algún tipo de restricción.
- Operaciones de eliminación
 - Sólo pueden violar restricciones de integridad referencial.
 - Cuando R referencia a S, y se intenta eliminar una tupla de S que es referenciada por alguna/s tupla/s en R.
 - Hay tres estrategias posibles: rechazar la eliminación, eliminar en cascada, o poner en NULL los atributos referenciales de las tuplas de R.

Operaciones e integridad

- Operaciones de modificación
 - Si se modifica una clave foránea, se debe verificar que sus nuevos valores referencien a una tupla existente de la relación referenciada, o bien sean todos nulos. De lo contrario se debería rechazar la operación.
 - Si se modifica una clave primaria, puede violarse cualquiera de las restricciones de integridad, y se combinan las situaciones indicadas para inserción y eliminación.
- A veces es necesario realizar una serie de operaciones por completo, o bien no realizarlas
 - Surge el concepto de transacción, como conjunto ordenado de operaciones que, o se ejecutan por completo, o no se ejecutan.
 - La ejecución de una transacción es a todo o nada.
 - Si una transacción no puede terminar de realizarse porque una de sus operaciones viola alguna restricción de integridad, entonces debe dejarse la base de datos en el estado anterior al inicio de la misma.

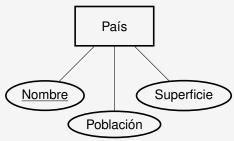
- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

[ELM16 9]

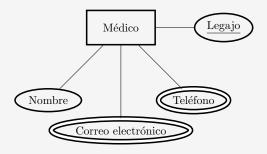
- Cada entidad del modelo ER producirá generalmente una relación del modelo relacional.
 - Hay excepciones!



Países(nombre_país, población, superficie)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

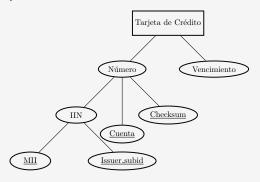
Atributos multivaluados:



Médicos(<u>legajo_médico</u>, nombre_médico) Telefonos(<u>legajo_médico</u>, teléfono) Mails(<u>legajo_médico</u>, mail)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Atributos compuestos:

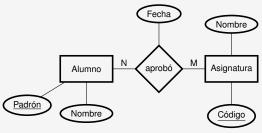


TarjetasCrédito(MII, issuer_subid, cuenta, checksum, fecha_venc)

Se representan a través de sus sub-atributos simples.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

 Cada interrelación N:M del modelo ER producirá una relación del modelo relacional.



Alumnos(padrón, nombre_alumno)

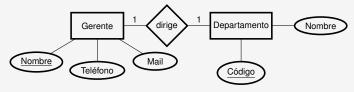
Asignaturas(código_asignatura, nombre_asignatura)

Aprobaciones (padrón, código_asignatura, fecha_aprobación)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Cuando la cardinalidad es 1:1 tenemos varias posibilidades.

Forma 1: "Relación asociada a la interrelación"

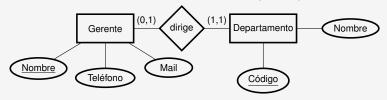


Gerentes(nombre_gerente, teléfono, mail)
Departamentos(código_dpto, nombre_dpto)
Dirige(nombre_gerente, código_dpto)

Debemos elegir sólo una de las claves foráneas como clave primaria de Dirige.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Forma 2a: Recomendada si Gerente tiene participación total

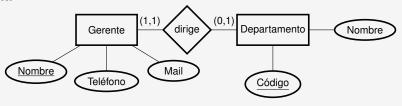


Gerentes(nombre_gerente, teléfono, mail, codigo_dpto)
Departamentos(codigo_dpto, nombre_dpto)

Una interrelación con cardinalidad 1:1 puede representarse incluyendo la clave primaria de una de las entidades participantes como clave foránea en la relación correspondiente a la otra entidad participante, siempre que esta última tenga participación total.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

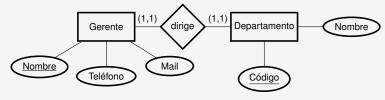
 Forma 2b: Recomendada si Departamento tiene participación total.



Gerentes(nombre_gerente, teléfono, mail)
Departamentos(código_dpto, nombre_dpto, nombre_gerente)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Forma 3: Ambas tienen participación total

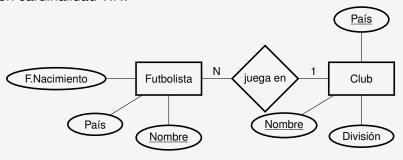


GerentesDepartamentos(<u>nombre_gerente</u>, teléfono, mail, código_dpto, nombre_dpto)

La clave puede ser o bien "nombre_gerente", o bien "código_dpto". Ambas son *claves candidatas*. En este caso hemos elegido la primera como clave primaria.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

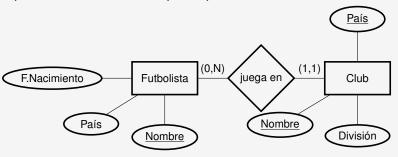
Con cardinalidad 1:N:



Futbolistas(<u>nombre_futbolista</u>, f_nac, país) Clubes(nombre_club, país, división) JuegaEn(<u>nombre_futbolista</u>, <u>nombre_club</u>, <u>país</u>)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

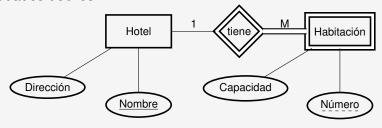
Aunque si Futbolista tuviera participación total sería conveniente:



Futbolistas(nombre_futbolista, f_nac, país_futbolista, nombre_club, país_club) Clubes(nombre_club, país_club, división)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Entidades débiles:

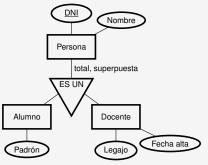


Hoteles(<u>nombre_hotel</u>, dirección) Habitaciones(<u>número_habitación</u>, <u>nombre_hotel</u>, capacidad)

Nota: No tiene sentido agregar una relación que represente la interrelación "tiene".

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

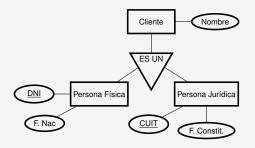
Generalización/Especialización:



Personas(<u>DNI</u>, nombre_persona) Alumnos(<u>DNI</u>, padrón) Docentes(<u>DNI</u>, legajo, fecha_alta)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

Unión:



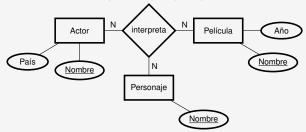
Clientes(id_cliente, nombre_cliente)
PersonasFísicas(<u>DNI</u>, fecha_nacimiento, id_cliente)
PersonasJurídicas(<u>CUIT</u>, fecha_constitución, id_cliente)

En este ejemplo debemos crear una clave subrogada para identificar a los clientes.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

- Interrelaciones ternarias
 - Con cardinalidad N:N:N

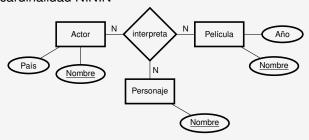
En cada película (Star Wars, ...), distintos actores (Harrison Ford, Carrie Fisher, ...) interpretan distintos personajes (Han Solo, Princesa Leia, ...). A veces un mismo actor puede interpretar más de un personaje en una misma película, e inclusive un mismo personaje en una película puede ser interpretado por más de un actor. Por último, existen personajes que aparecen en muchas películas.



Se crea una relación que representa a la interrelación.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

- Interrelaciones ternarias
 - Con cardinalidad N:N:N



Actores(nombre_actor, país)

 $Películas(\underline{nombre_película},\,a\tilde{n}o)$

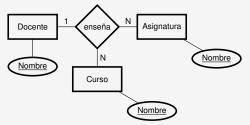
Personajes(nombre_personaje)

Interpreta(nombre_actor, nombre_película, nombre_personaje)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

- Interrelaciones ternarias
 - Con cardinalidad 1:N:N

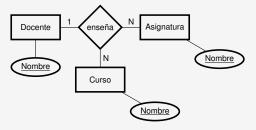
En una escuela, los docentes enseñan distintas asignaturas ("Matemática", "Literatura", ...) en distintos cursos (3°A, 3°B, ...). Cada asignatura en cada curso es enseñada por un único docente.



Recordemos que la cardinalidad de una entidad determina la cantidad de instancias de relación en que puede aparecer, fijadas las instancias de los otros tipos de entidades.

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

- Interrelaciones ternarias
 - Con cardinalidad 1:N:N



Docentes(<u>nombre_docente</u>)

Cursos(nombre_curso)

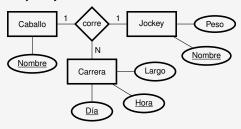
Asignaturas(nombre_asignatura)

Enseña(nombre_curso, nombre_asignatura, nombre_docente)

¿Cómo transformamos nuestro modelo ER en un modelo relacional?

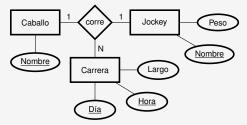
- Interrelaciones ternarias
 - Con cardinalidad 1:1:N

En un hipódromo se corren varias carreras a diario, en las cuales participan jockeys y caballos. En una carrera, cada jockey está asociado a un caballo, y el caballo sólo es montado por ese único jockey. Sin embargo, en distintas carreras un jockey puede variar de caballo, y un mismo caballo puede ser montado por distintos jockeys.



¿Cómo transformamos nuestro modelo ER en un modelo relacional?

- Interrelaciones ternarias
 - Con cardinalidad 1:1:N



Caballos(nombre_caballo)
Jockeys(nombre_jockey, peso)
Carreras(día_carrera, hora_carrera, largo)
Corre(día_carrera, hora_carrera, nombre_caballo, nombre_jockey)

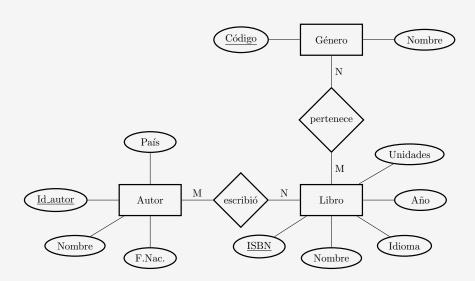
■ {día_carrera, hora_carrera, nombre_jockey} también es clave candidata.

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relacional
 - Principios
 - Ejemplos
- 4 Bibliografía

Librería "Jennifer"

Los dueños de esta librería desean crear una base de datos de libros que contenga información sobre los libros actualmente en venta, y que permita hacer búsquedas por nombre o país de origen del autor, género, idioma y año.

Librería "Jennifer"



Librería "Jennifer"

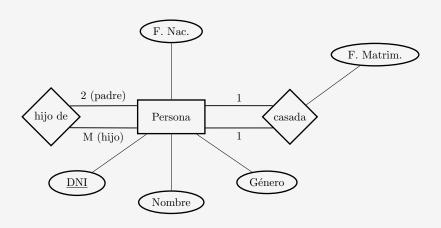
Autores(<u>id_autor</u>, nombre_autor, fecha_nacimiento, país) Libros(<u>ISBN</u>, nombre, idioma, año, unidades) Escribió(<u>id_autor, ISBN</u>) Géneros(<u>código_género</u>, nombre_género) Pertenece(<u>ISBN</u>, <u>código_género</u>)

Ejemplo 2 RENAPER

El Registro Nacional de las Personas quiere mantener una base de datos con el nombre, DNI, género y fecha de nacimiento de cada ciudadano argentino. Asimismo desea tener registrados todos los matrimonios en curso (no divorciados) incluyendo la fecha de matrimonio, y los nacimientos de personas indicando la identidad de los padres en caso que la misma sea conocida.

Hipótesis: Suponga que todas las personas son argentinas.

Ejemplo 2 RENAPER



Personas(<u>DNI</u>, nombre, género, fecha_nacimiento) HijoDe(<u>DNI_padre, DNI_hijo</u>) CasadaCon(<u>DNI1</u>, <u>DNI2</u>, fecha_matrimonio)

- Diseño lógico
- 2 Modelo Relacional
 - Estructura
 - Restricciones
 - Operaciones
- 3 Pasaje del modelo conceptual al modelo relaciona
 - Principios
 - Ejemplos
- 4 Bibliografía

Bibliografía

[ELM16] Fundamentals of Database Systems, 7th Edition.

R. Elmasri, S. Navathe, 2016.

Capítulo 5, Capítulo 9

[GM09] Database Systems, The Complete Book, 2nd Edition.

H. García-Molina, J. Ullman, J. Widom, 2009.

Capítulo 2.1, 2.2, Capítulo 4.5, 4.6

[SILB19] Database System Concepts, 7th Edition.

A. Silberschatz, H. Korth, S. Sudarshan, 2019.

Capítulo 2, Capítulo 6