

BASES-DE-DATOS.pdf



Anónimo



Bases de Datos



2º Grado en Ingeniería Informática



Escuela Politécnica Superior de Córdoba
Universidad de Córdoba

quieres trabajar en Wuolah??

tú puedes ayudarnos a llevar **WUOLAH**
al siguiente nivel (o alguien que conozcas)

TE BUSCAMOS



sin ánimo de lucro, chequea esto:

- ☐ Todos los apuntes que necesitas están aquí
- ☐ Al mejor precio del mercado, desde **2 cent.**
- ☐ Recoge los apuntes en tu copistería más cercana o recíbelos en tu casa
- ☒ Todas las anteriores son correctas

BASES DE DATOS

1. INTRODUCCIÓN A LA BASE DE DATOS

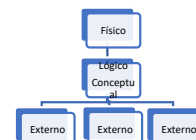
Este término apareció en los años sesenta, haciendo uso de un conjunto de ficheros, generalmente planos. Se caracterizaban por la gran redundancia y la poca garantía en la integridad de la información. Con los dispositivos de almacenamiento que permitían el acceso directo (*generalmente aleatorio*) los sistemas se hicieron más independientes de hardware.

Una base de datos garantiza la independencia de datos con respecto a los procedimientos. Esta independencia se realiza a dos niveles: lógica (*la modificación de la representación lógica general del dominio del problema no afecta a los programas de aplicación que la manipulan*) y física (*la distribución de datos en las unidades de almacenamiento es independiente de los cambios de la estructura lógica y, por tanto, de los procedimientos que manejan la misma*).

Las principales características de las bases de datos son:

- **Versatilidad para la representación de la información:** la información que existe en un dominio de un problema es única, pero pueden existir diferentes versiones de la información. *Adaptabilidad a nuevas características.*
- **Desempeño:** se debe asegurar un tiempo de respuesta adecuado en la comunicación hombre-máquina, permitiendo el acceso simultáneo a los ítems de datos por el mismo o distinto procedimiento. *Necesidad de una mayor capacidad de personas en línea en menor tiempo.*
- **Capacidad de acceso:** los usuarios reclaman a la base de datos constantemente información sobre los datos almacenados.
- **Mínima redundancia:** evitar la repetición es una de las características por las que surgió la tecnología de las bases de datos.
- **Simplicidad:** deben basarse en representaciones lógicas simples que permitan la verificación, en la representación de problemas que se plantean.
- **Integridad:** hace referencia a la veracidad de los datos almacenados con respecto a la información existente en el dominio del problema que trata la misma.
- **Seguridad y privacidad:** se encarga de proteger los datos contra su pérdida total o parcial por fallos del sistema o por accesos accidentales o intencionados a los mismos.
- **Afinación:** organización física de la base que determina el tiempo de respuesta de los procedimientos que operan sobre la misma.
- **Interfaz pasado-futuro:** es aceptado que el dominio de un problema cambia evolucionando con el tiempo.

Para que una base de datos pueda satisfacer las características anteriores, es necesario que los usuarios tengan una visión abstracta de los datos almacenados, aunque sin la necesidad de conocer cómo se organizan los



datos físicamente en la base. Existen tres visiones de los datos en una base de datos:

- **Visión Externa:** visión de datos que tienen los usuarios finales de una base de datos.
- **Visión Conceptual:** visión o representación del problema tal y como se representa en el mundo real.
- **Visión Física:** visión o representación de cómo la información es almacenada en los dispositivos de almacenamiento.

El nivel de descripción conceptual es el más importante. Se podrá hablar de un cuarto nivel, conocido como nivel lógico o canónico el cuál se verá afectado en función del nivel conceptual.

La granularidad es el nivel de detalle en que pueden ser descritas las representaciones externas derivadas de la representación lógica. A mayor granularidad de una representación externa (menor información a considerar) mayor será la independencia y viceversa. Un mayor granularidad proporciona una mayor complejidad en el software utilizado para realizar representaciones.

La integridad necesita que las representaciones a un determinado nivel de abstracción tengan en cuenta cómo se representa la información en los otros niveles. En ese instante, las diferentes representaciones de los datos se vinculan entre sí (ligadura). Existirán dos:

- **Ligadura lógica:** correspondiente al proceso de vinculación que se produce entre las representaciones externas y la lógica.
- **Ligadura física:** correspondiente al proceso de vinculación entre la representación lógica y la física.

Una base de datos es una colección de archivos relacionados que almacenan tanto una representación abstracta del dominio de un problema del mundo real, como los datos correspondientes a la información acerca del mismo. Las restricciones son innatas al problema o para que garanticen la integridad.

Un SGBD (*sistema general de bases de datos*) es una colección de programas de aplicación que proporcionan al usuario de la bases de datos los medios necesarios para realizar las tareas de la definición de los datos (*a los distintos niveles de abstracción*), manipulación de datos (*inserción, borrado, acceso o consulta a los mismos*), mantenimiento de la integridad de la base de datos (*integridad en cuanto a los datos en sí, sus valores y las relaciones entre ellos*), control de la privacidad y seguridad (*de los datos en la base de datos*) y el resto de características exigibles a una base de datos.

Los principales componentes de los SGBD son:

- **El lenguaje de definición de datos (DDL)** es un lenguaje artificial basado en un determinado modelo de datos que permite la representación lógica de los datos. La representación de los datos obtenida en este proceso de compilación obtenida



yo elijo cerveza SIN

Sea cual sea
el vehículo que
conduces, elige
cerveza SIN.

WWW.CONDUCCIONRESPONSABLECERVEZASIN.COM



**UNA GRAN CERVEZA.
UNA GRAN RESPONSABILIDAD.**

© CONDUCCIÓN RESPONSABLE, CERVEZA SIN es una iniciativa de la Asociación de Cerveceros de España con el apoyo de la Dirección General de Tráfico.



AERIVE



asfeme

asfeme

AOP



asapaym

CEA

ONBB

FEDRAL

ASOCIACIÓN DE BODEGAS

PMS

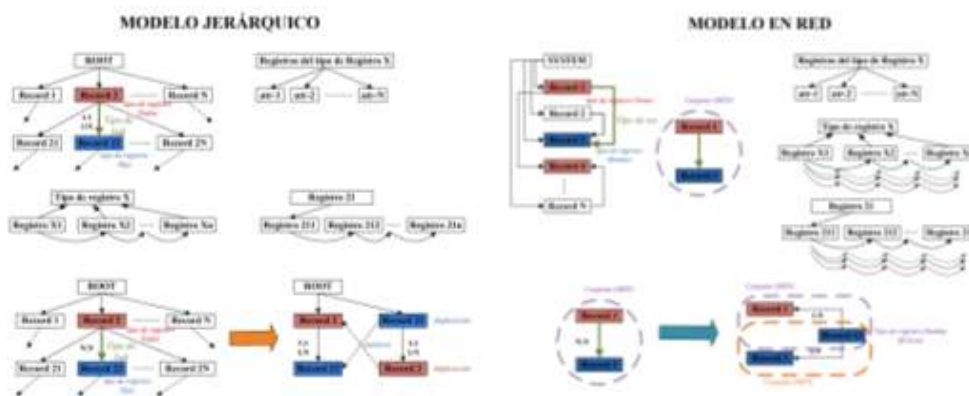
PONS

Rotary

STOP

en este proceso de compilación es almacenada en otro componente del SGBD conocido como Diccionario de Datos (*localiza información dentro de la base*).

- **El lenguaje de definición del almacenamiento de los datos (DSDL)** permite la definición de los datos en el nivel de representación físico.
- **El lenguaje de manipulación de datos (DML)** es un lenguaje artificial mediante el cual se realizan dos funciones bien diferentes en la gestión de datos: la definición del nivel externo o de usuario y la manipulación de los datos (*inserción, borrado, modificación y recuperación de los datos*). Dependiendo del modelo de datos en el cual se soportan existen dos tipos de DML, procedimentales, requieren que en las sentencias del lenguaje se especifique qué acciones/operaciones deben realizar, y no procedimentales, el DML es el encargado de determinar los procedimientos más efectivos.
- **El diccionario de datos** se trata de una meta-base de datos que contiene información sobre otra base de datos. En este diccionario se encontrará el esquema lógico, el esquema físico, subesquemas de la base de datos, restricciones de privacidad y acceso a los datos almacenados en la base y otra serie de información para garantizar la integridad.
- **El monitor responsable de la base de datos**, el cual se encarga de garantizar la privacidad de datos, la seguridad, la integridad, el acceso concurrente de los datos y la interacción con el sistema operativo y con el gestor de archivo de este.
- **El administrador de la base de datos** cuyas tareas serían la definición del esquema lógico, al igual que del físico y subesquemas, el control de la privacidad de los datos, la especificación de los procedimientos necesarios para el mantenimiento de la seguridad de los datos almacenados en la base de datos.
- **Los usuarios de la base de datos**, donde se diferenciarán entre usuarios terminales, técnicos, especializados y críticos.



2. REPRESENTACIÓN DE LOS PROBLEMAS DEL MUNDO REAL

El modelo de datos es una unidad de abstracción mediante el cual puede describirse un fenómeno abstracto o real. Mediante el uso de un modelo de datos se describen las propiedades que caracterizan al fenómeno y que lo diferencian de otros fenómenos *que se puedan o no describir* y las relaciones entre estas propiedades, y como estas evolucionan a lo largo del tiempo. Mediante un modelo de datos se describirán las características estáticas y dinámicas de un fenómeno.

Un modelo de datos es, por tanto, un conjunto de reglas de acuerdo con las cuales puede ser descrito un fenómeno.

Mediante el uso de un modelo de datos puede ser representado cualquier problema sobre el que se desea obtener información para su conocimiento y/o solución. El primer paso es la determinación de los límites del problema. Este proceso de simplificación es innato al proceso mental del ser humano y está basado en la capacidad de abstracción.

La abstracción es la capacidad mediante la cual una serie de objetos se categorizan en un nuevo objeto mediante una función de pertenencia. Al nuevo objeto se le llama clase o tipo de objeto y todos los elementos categorizados en esta clase tienen propiedades comunes, las cuales caracterizan la clase. En la definición de datos, la abstracción se utiliza tanto para la generalización o para la agregación.

- **Generalización:** abstracción por la cual un conjunto de clases de objetos puede ser visto como una nueva clase de objetos más general. La generalización de objetos simples en una clase es denominada **clasificación** y se denomina **especialización e instanciación** a los procesos inversos a la generación y clasificación. La generalización puede asociarse con el concepto *es un...*
- **Agregación:** la capacidad de considerar un objeto basándose en los elementos que lo constituyen. El proceso inverso se conoce como **refinamiento**, mediante el cual se puede representar a aquellos objetos simples o propiedades que caracterizan a una clase de objetos. La agregación puede asociarse con el concepto *parte de...*

La representación de un problema puede llevarse a cabo haciendo uso de la abstracción de forma ascendente o descendente en complejidad. Es necesario, además, representar las interdependencias entre los elementos de este. De forma general, la representación de problemas requiere el seguimiento de los siguientes pasos.

- La definición del problema mediante una descripción simple.
- La definición de la arquitectura del problema, mediante una descripción de las partes importantes del sistema.
- La definición de la estructura del problema mediante la descripción de los elementos del sistema (*definición del objeto, medida del objeto, relaciones entre los objetos, restricciones*).

quieres trabajar
en Wuolah??

TE BUSCAMOS

sin ánimo
de lucro,
chequea esto:



tú puedes
ayudarnos a
llevar
WUOLAH
al siguiente
nivel
(o alguien que
conozcas)

El modelo **entidad-interrelación (E-R)** propuesto por Peter Chen para la representación conceptual de los problemas y como un medio para representar la visión de un sistema de forma global. Este permite la representación de cualquier tipo de sistema y a cualquier nivel de abstracción o refinamiento, está soportado en la representación de los datos, haciendo uso de grafos y de tablas y son representados los elementos que forman parte del sistema y las relaciones existentes entre ellos.

CONCEPTOS BÁSICOS

El **conjunto** es la agregación de una serie de objetos elementales mediante una función de pertenencia. No es importante el orden de los elementos dentro del conjunto, ni duplicación de este.

La **relación** es un conjunto que representa una correspondencia entre dos o más conjuntos. El orden de la relación es importante, *por lo que el par $\langle c1, c2 \rangle$ no tiene por qué ser igual al par $\langle c2, c1 \rangle$* . Las relaciones pueden ser binarias, ternarias o n-arias, y pueden ser definidas como el producto cartesiano de los conjuntos que intervienen en la relación ($R \subseteq C1 * C2 * ... Cn$).

Tanto los conjuntos como las relaciones pueden ser descritos en términos de intención y extensión. La **intención** es la descripción del tipo o clase de objeto (conjunto o relación). La **extensión** es la descripción de los elementos individuales o instancias de objetos (elementos del conjunto).

Se denomina **dominio** al conjunto homogéneo, es decir, a aquellos conjuntos cuyo elementos son homogéneos. *El dominio corresponde a los tipos de datos que serán almacenados mediante las restricciones en los valores que el atributo puede tomar.*

El **atributo** de un dominio es la intención de ese dominio, cuyo valor será la extensión del dominio. *Los atributos son las características que definen o identifica una entidad. Una entidad puede tener dos valores iguales para un mismo atributo.* Los atributos identificativos son aquellos que permiten diferenciar a una instancia de la entidad de otra (-).

Una **entidad** es un tipo de objeto (un conjunto) definido en base a la agregación de una serie de atributos. *Representa un objeto, cosa o concepto del mundo real el cual se diferencia de otro objeto, pudiendo ser física o conceptual.* Las entidades tienen, como los conjuntos, intenciones y extensiones.

La interpretación dada para las entidades puede ser igualmente propuesto para **interrelaciones**. Una interrelación representa la relación existente entre entidades, denominándose tipo de interrelación a la intención de la relación existente entre dos tipos de entidad. La extensión de un tipo de interrelación es denominada conjunto interrelaciones y representa a cada una de las posibles correspondencias entre los conjuntos de entidades que intervienen en el tipo de interrelación.

En el modelo E-R se considera que una entidad es un objeto real o abstracto que cumple las siguientes propiedades: tiene existencia propia, es distinguible del resto de las

WUOLAH

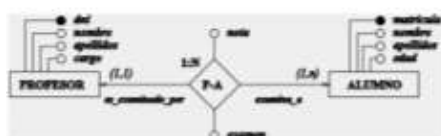
entidades (objeto) que intervienen en el sistema, las entidades de un mismo tipo están definidas en base a un mismo conjunto de atributos, cada uno de ellos definido en un mismo dominio.

Un tipo de interrelación es definido como una relación matemática entre n tipos de entidades. Así, el tipo de interrelación R puede ser definida como $R = \{[e_1, e_2, \dots, e_n] | e_1 \subseteq E_1, e_2 \subseteq E_2, \dots, e_n \subseteq E_n, \text{ donde } E_1, E_2, \dots, E_n, \text{ son tipos de entidades que intervienen en la interrelación. Diferenciaremos entre tipo de entidades fuertes, cuya existencia no depende de la existencia de ningún otro tipo de entidad en la consideración del problema, y, por otro lado, tipo de entidades débiles, cuya existencia depende de la existencia de un tipo de entidad fuerte. Diferenciaremos entre debilidad por identificación donde una entidad débil no puede ser identificada a no ser que se identifique un entidad fuerte por cuya existencia esté presente la debilidad y debilidad por existencia donde una entidad débil no puede ser identificada sin necesidad de identificar la entidad fuerte por la cual existe.}$

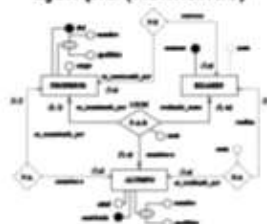


A los atributos que pueden tomar un conjunto de valores se les denomina atributos múltiples, poseen un nombre único. Se denomina identificador de un tipo de entidad al conjunto de atributos que no toma el mismo valor para dos entidades diferentes del mismo tipo. Un atributo asociado a un tipo de interrelación tiene la función de caracterizar la relación entre dos entidades, aportando información a la correspondencia entre los tipos de entidad.

Ejemplo



Ejemplo (Refinando)



3. EL MODELO ENTIDAD-INTERRELACIONAL EXTENDIDO (EE-R)

El modelo de **Entidad Relacional Extendido (EE-R)** permite la representación de cualquier tipo de relaciones existentes entre clases de objetos que considera los principios de abstracción. Dentro de la interrelaciones destacaremos las **reflexivas**, son relaciones unarias (*se involucra un único tipo de entidad*), las cuales no suelen verse en los modelos de datos, aunque pueden aparecer y las **exclusivas**, relaciones entre entidades diferentes, aunque no siempre son independientes.

La **generalización** es una abstracción que identifica una relación jerárquica que representa un tipo de entidad *es_un* subtipo de otro tipo de entidad representada a un nivel de abstracción mayor. Un subtipo de entidad es un tipo de entidad que mantiene un tipo de interrelación jerárquica con otro tipo de entidad. El comportamiento general es considerado por el tipo de entidad con la que mantiene el tipo de interrelación (**herencia**). La relación jerárquica puede ser n-aria entre un tipo de entidad y un conjunto de subtipos de ese tipo de entidad. Los subtipos deben tener un conjunto de propiedades que permitan esta discriminación.

Un tipo de interrelación jerárquica representa una **especialización** de un tipo de entidad en otros tipos de entidad. Esta especialización es debida a una diferencia en cuanto al número de propiedades que definen los subtipos de entidad y los diferentes valores que pueden ser medidos para una propiedad.

Existirán distintos tipos de especialización:

- **Exclusiva o sin solapamiento**
- **Inclusiva o con solapamiento**
- **Total**
- **Parcial**

Se considera **restricción** en el dominio del atributo a los valores de las cardinalidades máximas y mínimas de la interrelaciones. Estas se producen en las entidades débiles o las que participan en un modelo de interrelaciones jerárquicas.

Mediante una **descripción simple** se deben describir todos los elementos del problema y todas aquellas características de estos que permiten su identificación en el mundo real, así como dentro del conjunto de los objetos representados.

LINKS EJERCICIOS:

<https://www.studocu.com/es/document/universidad-de-malaga/base-de-datos/ejercicios-resueltos-base-de-datos-modelo-entidad-relacion/10015858>

<https://bustamanteg.files.wordpress.com/2010/06/ejercicios-mer-resueltos-para-publicar.pdf>

4. EL MODELO DE DATOS RELACIONAL

La complejidad del esquema va a depender tanto de la complejidad del problema en el mundo real, de la calidad de análisis de ese problema y de la evolución del problema y de los requisitos funcionales.

Una **tabla** es una matriz rectangular que puede ser descrita de forma simple matemáticamente y que posee las siguientes propiedades:

- i. Cada elemento de la matriz representa a un ítem de datos elemental.
- ii. Una tabla es homogénea por columnas y tienen asignado un nombre único.
- iii. Para una tabla todas las filas son diferentes, debido que no se admiten filas duplicadas.
- iv. Tanto las filas como las columnas pueden ser consideradas en cualquier secuencia.

La **relación** R es una relación entre n conjuntos si está formado por conjuntos de n tuplas no ordenadas $(d_1, d_2, d_3...)$ tales que $d_1 \in D_1, d_2 \in D_2... d_n \in D_n$.

Un **atributo** representa el uso de un dominio para una determinada relación (*el atributo aporta un significado semántico al dominio*). Un **dominio** es un conjunto homogéneo definido mediante el uso de la abstracción en base a otro conjunto.

La **intención** de una relación define todas las posibles extensiones admisibles para la misma. La intención de una relación define dos aspectos, una estructura de datos nominada en estructura como en ítem de datos que componen el objeto, y un conjunto de restricciones de integridad. La **extensión** de una relación representa a cada uno de los objetos (*tuplas*), pertenecientes a un mismo tipo (*relación*), existentes en el dominio del problema en el momento dado.

Las **claves de las relaciones** son el conjunto de atributos que forman parte de la intención de una relación específica, pudiendo tomar valores únicos en el dominio del problema para cualquier extensión de esa relación. Dentro de los esquemas relacionales, podremos destacar la **integridad**:

- Integridad de la clave: ningún atributo que forme parte de la clave candidata en una relación podrá tomar valores nulos para ninguna tupla de esa relación.
- Integridad de referencia: sea D , un dominio, y R_1 , una relación con un atributo $R_{1.a}$ definido en dicho dominio D , entonces en cualquier instante dado, cada valor de $R_{1.a}$ en R_1 debe ser nulo o bien igual a algún valor V , el cual existe en dicho instante, para un atributo $R_{2.b}$ definido en el mismo dominio sobre la relación R_2 y en la cual está definido como clave primaria.
- Otras restricciones: corresponderán con los valores permitidos para los atributos que forman parte de las relaciones existentes en el esquema.

Utilizando la teoría de la **normalización de relaciones**, existe un conjunto de reglas de normalización para satisfacer la no existencia de redundancias superfluas, aumentar el desempeño de las operaciones de actualización de la base de datos, representar de

quieres trabajar
en Wuolah??

TE BUSCAMOS

sin ánimo
de lucro,
chequea esto:



tú puedes
ayudarnos a
llevar
WUOLAH
al siguiente
nivel
(o alguien que
conozcas)

forma coherente los objetos y relaciones y garantizar la fiabilidad sobre la información mantenida en la base de datos.

La normalización de relaciones está basada en otra teoría conocida como **teoría de las dependencias**, donde dada una relación R , se dice que el atributo $R.y \in R$ es funcionalmente dependiente de otro atributo $R.x \in R$, y se expresa de la forma $R.x \rightarrow R.y$, y solo si, para cada valor de $R.x$ tiene asociado a él exactamente un valor de $R.y$ para cualquier extensión de la relación R . En el concepto de dependencia funcional se tiene en cuenta atributos de una misma relación (*no de diferentes relaciones*), pudiendo existir entre parejas de atributos de una relación, es intrínseca a la intención de una relación, pudiendo ser simples o estar formados por la agregación de varios atributos.

Hablaremos de **dependencia funcional completa** donde el atributo $R.y \in R$ es funcionalmente dependiente de forma completa de otro atributo $R.x \in R$ sí, y sólo si, depende funcionalmente de $R.x$ y no de ningún subconjunto de los atributos que formen parte del atributo $R.x$. *Es decir, que si el atributo $R.x$ es un agregado formado por la concatenación de varios atributos pertenecientes a la relación, entonces la dependencia funcional $R.x \rightarrow R.y$, puede no ser completa, ya que para existir una dependencia funcional de la forma $R.z \rightarrow R.y$, se pide $R.z \subset R.x$. Y, por otro lado, que el atributo de $R.y$ ya sea simple o complejo no tendrá relevancia para que la dependencia funcional $R.x \rightarrow R.y$ sea completa o no.*

Las principales propiedades de las dependencias funcionales serán:

- **Reflexiva:** dados los atributos a y b de una relación R , para los que se cumplen $R.b \subseteq R.a$ entonces, en la relación R , está presente una dependencia funcional de la forma $R.a \rightarrow R.b$.
- **Aumento:** dados los atributos a y b de una relación R en la que está presente la dependencia funcional $R.a \rightarrow R.b$, entonces también estará presente la dependencia funcional $R.(a + c) \rightarrow R.(b + c)$, siendo c cualquier otro atributo que forme parte de la intención de la relación R .
- **Transitiva:** dados los atributos a , b y c de una relación R en la que están presentes las dependencias funcionales $R.a \rightarrow R.b$ y $R.b \rightarrow R.c$, entonces también estará presente la dependencia funcional $R.a \rightarrow R.c$.
- **Unión:** dados los atributos a , b y c de una relación R en la que están presentes las dependencias funcionales $R.a \rightarrow R.b$ y $R.a \rightarrow R.c$, entonces también estará presente la dependencia funcional $R.a \rightarrow R.(b + c)$.
- **Pseudo-transitiva:** dados los atributos a , b , c y d de una relación R en la que están presentes las dependencias funcionales $R.a \rightarrow R.b$ y $R.(b + c) \rightarrow R.d$ entonces también estará presente $R.(a + c) \rightarrow R.d$.
- **Descomposición:** dados los atributos a , b y c de una relación R en la que está presente la dependencia funcional $R.a \rightarrow R.b$, y se cumple que $R.c \subseteq R.b$, entonces también estará presente $R.a \rightarrow R.c$.

La teoría de la normalización está basada en la aplicación de una serie de reglas a las que se les denomina **reglas de normalización**. La aplicación de una regla de

WUOLAH

normalización toma una relación como argumento de entrada y da como resultado dos o más relaciones, donde no se introducen nuevos atributos en el esquema relacional resultante de la normalización, la relación objeto de la aplicación es desestimada, los atributos de la relación objeto de la normalización pasan a formar parte de la intención de una o más de las relaciones resultantes y en la aplicación de normalización se ha debido eliminar al menos una dependencia existente entre los atributos de la relación.

- Primera forma normal – FN1: una relación R satisface FN1 si, y sólo si, todos los dominios subyacentes de la relación R contienen valores atómicos.
- Segunda forma normal – FN2: una relación satisface FN2 si, y sólo si, satisface FN1 y cada atributo de la relación depende funcionalmente de forma completa de la clave primaria de esa relación.
- Tercera forma normal – FN3: una relación satisface FN3 si, y sólo si, satisface FN2 y cada atributo no primo de la relación depende funcionalmente de forma transitiva de la clave primaria de esa relación, es decir, no pueden existir dependencias entre los atributos que no forman parte de la clave primaria de la relación R.
- Se denomina determinante principal a uno o un conjunto de atributos de una relación R de la cual depende funcionalmente de forma completa algún otro atributo de la misma relación. Una relación R satisface la forma normal de Boyce-Codd (FNBC) si, y sólo si, se encuentra en FN1 y cada determinante funcional es una clave candidata de la relación R.
- Cuarta forma normal – FN4: una relación satisface FN4 si, y sólo si, siempre que exista una dependencia multivaluada en R de la forma $R.x \twoheadrightarrow R.y$, todos los demás atributos de R son funcionalmente dependientes de $R.x$ ($R.x \rightarrow R.z, \forall R.z \in R$).
- *En una relación R puede estar presenta un caso especial de dependencias llamadas dependencias de reunión, las cuales son la base de la aplicación de la quinta forma normal – FN5. Dada una relación R de esquema $R(a1, a2...)$ se dice que existe una dependencia de reunión si, y sólo si, la relación R puede ser construida a partir de la reunión natural de las relaciones $R1, R2...$ obtenidas por la proyección de R sobre los atributos $a1, a2...$ respectivamente.*

5. ÁLGEBRA RELACIONAL SQL.

A pesar de que **SQL** utiliza operadores algebraicos matemáticos, éstos operan de forma diferente a como lo hacían en la teoría de conjuntos ya que se aplican sobre tablas relacionales en lugar de sobre conjunto o relaciones. SQL es el lenguaje estandarizado y de uso universal utilizado por los SGBD basados en el álgebra relacional. Como cualquier lenguaje de un SGBD, SQL se encarga de la descripción, manipulación, control y seguridad de los datos.

Definiremos como **compatibles** a dos relaciones R_1 y R_2 , si ambas tienen el mismo grado y el atributo enésimo de R_1 está definido en el mismo dominio que el atributo enésimo de la relación R_2 , si bien el nombre de los atributos puede ser diferente.

La **unión** de dos relaciones compatibles R_1 y R_2 , es una nueva relación R_3 , también compatible, cuyo esquema es igual al esquema de R_1 y R_2 , y cuya extensión está formada por la agrupación, sin repetición, de las extensiones de R_1 y R_2 .

La **diferencia (minus)** de dos relaciones compatibles R_1 y R_2 es una nueva relación R_3 , también compatible, cuyo esquema es igual al esquema de R_1 y R_2 y cuya extensión está formada por aquellas tuplas de la relación de R_1 que no se encuentran en la relación R_2 .

La **selección (select)** sobre una relación R_1 mediante una cualificación Q es una nueva relación R_2 , cuyo esquema es igual a R_1 , y cuya extensión está formada por todas aquellas tuplas de R_1 que satisfacen la cualificación Q .

La **proyección (project)** sobre una relación R_1 con esquema $R_1.a_i, R_1.a_j \dots R_1.a_z$ mediante un subesquema $S(R_1) \equiv R_1.a_m \dots R_1.a_p$, donde $a_m \geq a_i$ y $a_p \leq a_z$, es una nueva relación R_2 , cuyo esquema es igual al subesquema $S(R_1)$, y cuya extensión es igual a todas las tuplas de R_1 sin repetición sobre el subesquema $S(R_1)$.

El producto cartesiano de dos relaciones R_1 y R_2 no necesariamente compatibles es una nueva relación R_3 , cuyo esquema es igual a la concatenación de los esquemas de R_1 y R_2 , y cuya extensión está formado por el conjunto de tuplas que se obtiene de concatenar cada una de las tuplas de R_1 con todas y cada una de las tuplas de R_2 .

La **intersección (intersect)** de dos relaciones compatibles R_1 y R_2 es una nueva relación R_3 , también compatible, cuyo esquema es igual al esquema de R_1 y R_2 , y cuya extensión está formada por el conjunto de tuplas que son comunes a R_1 y R_2 .

La **reunión (join)** de dos relaciones (*compatibles o no*) con al menos un atributo con el dominio común y sobre una cualificación da lugar a una nueva relación cuya intención es la concatenación de R_1 y R_2 y cuya extensión está formada por las tuplas que resultan del producto cartesiano de $R_1 * R_2$ que satisfacen la cualificación Q . Existirá distintos tipos de reunión como:

- Equireunión: cualificación q es igual.
- Thetareunión: cualificación q no es igual.
- Reunión natural: se representa por el signo $|x|$, y es una equireunión sobre todos los atributos comunes existentes entre las relaciones R_1 y R_2 .

- **Autoreunión:** es una reunión natural sobre una misma relación, es decir $R1=R2$.
- **Semireunión:** esta operación se realiza entre dos relaciones $R1$ y $R2$ sobre una cualificación multiatributo Q en la que puede aparecer cualquier operador de comparación y el resultado es una relación $R3$ de esquema igual a $R1$, cuya extensión está formada por todas aquellas tuplas de $R1$ que intervienen en la reunión de $R1$ y $R2$.

La **división** de una relación $R1$ con esquema $R1.ai, R1.aj \dots R1.az$ entre una relación $R2$ con esquema $R1.am, R1.an \dots R1.ap$, donde $am \geq ai$ y $ap \leq az$, es una relación $R3$ cuyo esquema es igual a la diferencia del esquema de $R1$ menos el esquema de $R2$, y cuya extensión es igual a todas las tuplas de $R1$ sin repetición para las cuales está presente toda la extensión de la relación $R2$ (*operación compleja que permite expresar de forma sencilla el operador \forall*).

En SQL, se podrá definir tablas del esquema relacional mediante **create table**, incluyendo los distintos atributos a partir de los cuales se forman. También podremos consultar la información mediante **select**. Por otro lado, podremos manipular los datos, mediante **insert**, actualizar la tabla mediante **update** o borrar elementos de las tablas mediante **delete**.

```
CREATE TABLE NombreTabla (
    nombre-atributo-1: DOMINIO (opción),
    nombre-atributo-2: DOMINIO (opción),
    .....
    nombre-atributo-n: DOMINIO (opción),
    PRIMARY KEY (lista-atributos),
    FOREIGN KEY (lista-atributos),
    REFERENCES NombreTabla (clave-principal)
    [ON DELETE CASCADE] );

SELECT [DISTINCT | ALL] (lista-atributos-1 | *)
[INTO cursor]
FROM lista-tablas
[WHERE condición-1]
[GROUP BY lista-atributos-2]
[HAVING condición-2]
[ORDER BY lista-atributos-3 [ASC | DESC]]

INSERT [INTO] (lista-atributos)
NombreTabla [lista-tablas]
VALUES (lista-valores)

UPDATE NombreTabla
SET atributo-i = valor
[WHERE Condición];

DELETE [FROM] NombreTabla
[WHERE Condición];
```

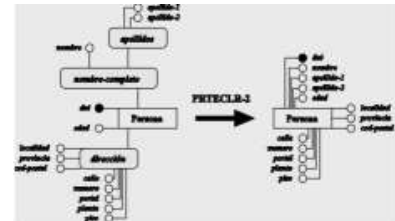
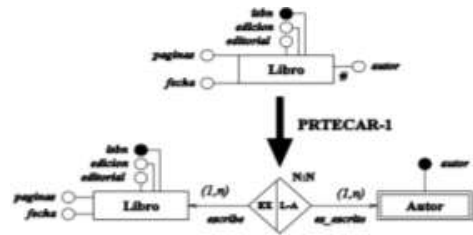
- ☐ Todos los apuntes que necesitas están aquí
- ☐ Al mejor precio del mercado, desde **2 cent.**
- ☐ Recoge los apuntes en tu copistería más cercana o recíbelos en tu casa
- ☒ Todas las anteriores son correctas

6. TRADUCCIÓN DE ESQUEMAS E-R A ESQUEMAS RELACIONALES.

6.1. TRADUCCIÓN DE ESQUEMAS E-R A ESQUEMAS RELACIONALES I.

Las PRTECAR se basan en la aplicación de FN1 a los objetos que forman parte de los esquemas conceptuales. El principio en el que se basan estas reglas es el eliminar de estos tipos de objetos de los atributos correspondientes a las entidades con múltiples valores y atributos correspondientes a los tipos de entidades o interrelaciones compuestas.

- **PRTECAR-1.** Todos los atributos múltiples (*atributos que pueden tomar más de un valor*) se transformarán en un tipo de entidad débil por existencia con relaciones uno a muchos o muchos a muchos. En caso de que el atributo del tipo de entidad débil creado no pudiera identificar a las entidades de este tipo: el tipo de entidad se considera que es débil por identificación a tipo de entidad con el que mantiene relación (*heredando sus atributos*), o bien, se añadirá un nuevo atributo que permita identificar sin ambigüedad a las entidades de este tipo de entidad débil.
- **PRTECAR-2.** Todos los atributos compuestos asociados con los tipos de entidad y los tipos de interrelación deben ser descompuestos en los atributos simples que formen parte o intervengan en los atributos compuestos correspondientes.



Estas reglas permiten la transformación de los esquemas sin pérdida de información, conservando el nivel de representación del problema. La aplicación de estas reglas dependerá del tipo de objeto del esquema conceptual que se debe transformar y de la cardinalidad de las relaciones que los objetos mantienen con otros objetos en el esquema conceptual.

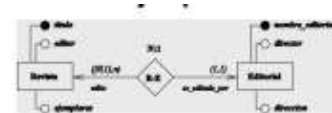
La transformación de tipos de interrelación uno a uno se centrará en el valor de la cardinalidad máxima, donde los dos tipos de entidad participan de forma completa, uno de los dos participa de forma parcial o los dos participan de forma parcial en tipo interrelación.

- **PRTECAR-2.1.** Si en un tipo de interrelación, los dos tipos de entidad participan de forma completa, dependerá del identificador:
 - Si los dos tienen el mismo identificador, los dos se transformarán en única tabla formada por la agregación de los atributos de las dos entidades.
 - Si tienen identificadores diferentes, cada tipo de entidad se transforma en tabla y, la clave principal será el identificador de cada

uno de los tipos de entidad y la clave foránea, el identificador de otro tipo de entidad.

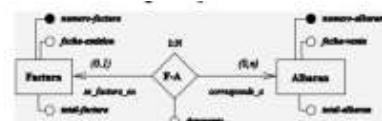
- Si tienen la misma clave, pero uno es entidad débil, entonces se procede a alguna de las dos formas expuesta anteriormente en función de los requisitos funcionales.
- **PRTECAR-2.2.** Si en un tipo de interrelación binaria, alguno de los tipos de entidad participa de forma parcial, no se crea una tabla para el tipo de interrelación (*el identificador será el atributo de la tabla del otro tipo de entidad, no pudiendo tomar valores nulos*) o se construirá una nueva tabla correspondiente al tipo de interrelación, formada por los identificadores de las dos entidades, considerando clave al identificador que participa de forma parcial.
- **PRTECAR-2.3.** Si en un tipo de interrelación binaria, los dos tipos de entidades participan de forma parcial, se construye una nueva tabla correspondiente al tipo de interrelación y cuyos atributos serán los identificadores de las dos entidades. La clave principal será el identificador de una de las entidades, definiendo al otro identificador como clave alterna.

- **RTECAR-3.1.** Si en un tipo de interrelación binaria 1:N, ambos tipos de entidades participan de forma total o el tipo de entidad que interviene con cardinalidad máxima muchos participa de forma parcial, el identificador de la entidad con cardinalidad máxima uno forma parte de la tabla del tipo de entidad con cardinalidad *muchos*. Este atributo será definido como clave foránea de esta tabla (*no toma valores nulos*). Si el tipo de interrelación tuviera atributos asociados, estos atributos pasan a formar parte de la tabla correspondiente al tipo de entidad que participa con cardinalidad máxima muchos.



Editorial (*nombre_editorial, direccion, director*)
Revista (*titulo, editor, ejemplares, nombre_editorial*)

- **PTECAR-3.2** Si en un tipo de interrelación binaria 1:N, ambos tipos de entidades participan de forma parcial o el tipo de entidad que interviene con cardinalidad máxima uno participa de forma parcial, entonces la interrelación forma una nueva tabla, donde se encontrarán los identificadores de los tipos de entidad y por todos los atributos asociados al tipo de interrelación, cuya clave principal será el identificador del tipo de entidad que interviene con cardinalidad muchos, definiendo el resto como claves foráneas.



Factura (*numero_factura, fecha_emision, total_factura*)
Albaran (*numero_albaran, fecha_venta, total_albaran*)
Fact_Alba (*numero_albaran, numero_factura, descuento*)

- **PTERCAR-4.** Si en un tipo de interrelación binaria N:N, se genera una nueva tabla para representar el tipo de interrelación, formada por los identificadores de los tipos de entidad que intervienen en el tipo de interrelación y por todos



Autobus (*matricula, licencia, categoria*)
Conductor (*carnet#, nombre, experiencia*)
Condu_Auto (*carnet#, matricula, trayecto*)

los atributos asociados al tipo de interrelación. La clave principal será la agregación de los atributos identificadores correspondientes a los tipos de entidad que intervienen en el tipo de interrelación.

6.2. TRADUCCIÓN DE ESQUEMAS E-R A ESQUEMAS RELACIONALES II.

Para el proceso de transformación de las relaciones en las que intervienen más de dos tipos se debe aplicar la regla RTECAR-4. Cada tipo de relación se transformará en una tabla y el tipo de interrelación se transforma en una nueva tabla con los atributos de identificadores en el tipo de entidad participantes en el tipo de interrelación y los atributos asociados al tipo de interrelación. La clase de esta tabla será la agregación de los atributos identificadores de los tipos de entidad.

Dentro de las interrelaciones reflexivas, son relaciones binarias en las que únicamente interviene un tipo de entidad. Dentro de las reflexivas diferenciaremos entre reflexivas N:N, donde se generará una tabla para el tipo de entidad y otra para el tipo de interrelación (PTECAR-4), siendo los atributos de esta tabla: los atributos asociados al tipo de interrelación, el identificador del tipo de entidad desempeñando uno de los papeles en el tipo de interrelación y el identificador del tipo de entidad desempeñando el otro papel. Por otro lado, las reflexivas 1:N, se genera una tabla para el tipo de entidad, añadiendo como clave foránea el identificador del tipo de entidad con cardinalidad 1 (PTECAR-3.1) y se generará una tabla para el tipo de entidad y otra para la interrelación con ambos identificadores tomando el de cardinalidad N como clave (RTECAR-3.1).

La aplicación de una regla depende de la magnitud de la especialización, tipo de especialización, otros tipos de interrelación y criterios de procesamiento.

- **PRTECAR-3.** Eliminación del supertipo de entidad. Transfiere todos los atributos e interrelaciones del supertipo a cada uno de los subtipos, manteniéndose los tipos de interrelación en los que interviene en los que intervenga cada uno de los subtipos de entidad. Además, el atributo calificador del tipo interrelación, si estuviera presenta, se puede desestimar. Si el tipo de interrelación jerárquica es exclusivo, los subtipos intervendrán de forma parcial en los tipos de interrelación transferidos desde el supertipo. Esta regla es conveniente aplicarla cuando las interrelaciones jerárquica son totales y exclusivas. Una transformación haciendo uso de esta regla, introduciendo una redundancia de información, el número de interrelación aumenta...
- **PRTECAR-4.** Eliminación de los subtipos de entidad, transfiriendo los atributos de los subtipos al supertipo y cada uno de los tipos de interrelación que mantuvieran los subtipos de entidad serán considerados para el supertipo. Se diferenciará entre interrelación jerárquica exclusiva (*participa con cardinalidad mínima cero en todas las interrelaciones*) e inclusiva (*participará con las cardinalidades que participaba cada subtipo de entidad*). El atributo calificador del tipo de interrelación, pasa a formar del supertipo de entidad, si el tipo de interrelación es exclusivo no formará parte de la clave, si el tipo de interrelación es inclusivo formará parte de la clave,

originado redundancia de los atributos del supertipo para cada instancia de los subtipos y si el tipo de interrelación es parcial, podrá tomar valores nulos para representar a entidades que no se especifican.

- **PRTECAR-5.** Eliminación de la jerarquía. El tipo de interrelación jerárquica se transformará en tantos tipos de interrelación uno a uno como subtipos de interrelación uno a uno como subtipos de entidad estén presentes. Los subtipos de entidad participarán si el tipo de interrelación jerárquica es exclusivo, participarán con cardinalidad mínima 0 y si el tipo de interrelación jerárquica es inclusivo, participarán con cardinalidad mínima cero a uno. El supertipo participa con cardinalidades mínima y máxima igual a 1.