

### 3. MODELO ENTIDAD – RELACIÓN

#### 3.1 Conceptos básicos

#### 3.2 Representaciones gráficas

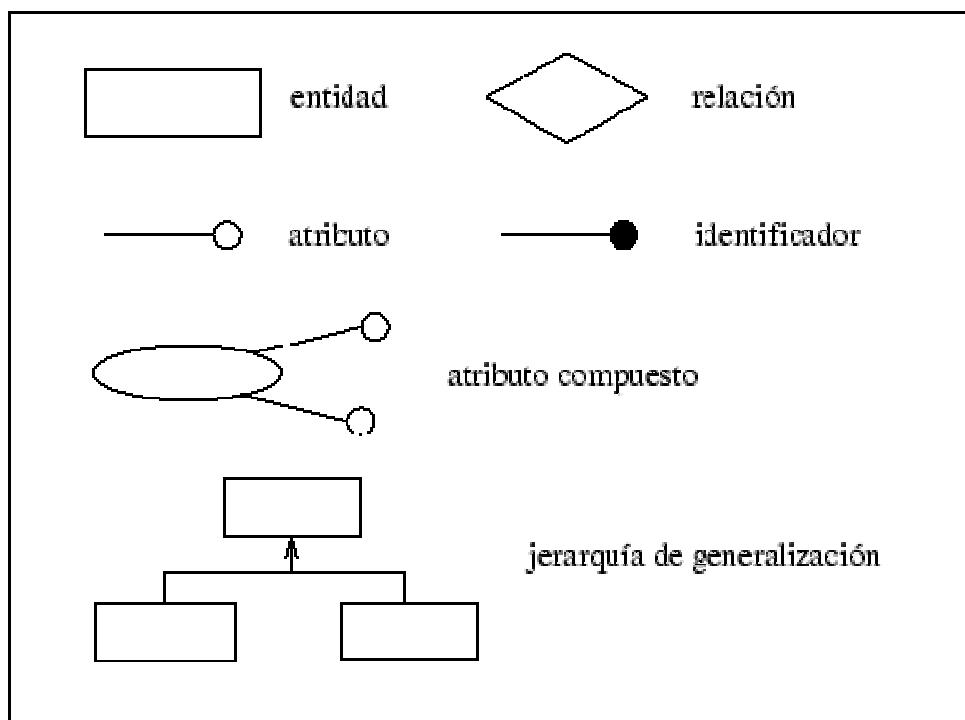
#### 3.3 Aplicaciones

#### 3.4 Modelo relacional

#### 3.1 Conceptos básicos

El modelo entidad-relación es el modelo conceptual más utilizado para el diseño conceptual de bases de datos. Fue introducido por Peter Chan en 1976. El modelo entidad-relación está formado por un conjunto de conceptos que permiten describir la realidad mediante un conjunto de representaciones gráficas y lingüísticas.

Originalmente, el modelo entidad-relación sólo incluía los conceptos de entidad, relación y atributo. Más tarde, se añadieron otros conceptos, como los atributos compuestos y las jerarquías de generalización, en lo que se ha denominado *modelo entidad-relación extendido*.



Conceptos del modelo entidad-relación extendido.

## Entidad

Cualquier tipo de objeto o concepto sobre el que se recoge información: cosa, persona, concepto abstracto o suceso. Por ejemplo: coches, casas, empleados, clientes, empresas, oficios, diseños de productos, conciertos, excursiones, etc. Las entidades se representan gráficamente mediante rectángulos y su nombre aparece en el interior. Un nombre de entidad sólo puede aparecer una vez en el esquema conceptual.

Hay dos tipos de entidades: fuertes y débiles. Una *entidad débil* es una entidad cuya existencia depende de la existencia de otra entidad. Una *entidad fuerte* es una entidad que no es débil.

## Conjunto de entidades

Es una colección de entidades que comparten los mismos atributos o características.

Ejemplos:

- Todos los atletas que participan en los Juegos Olímpicos, comparten sus atributos: nombre, número de identificación, edad, peso, categoría...
- Todos los países del mundo, comparten las características: nombre, continente, área, lengua principal, lengua secundaria, moneda, etc.

## Relación (interrelación)

Es una correspondencia o asociación entre dos o más entidades. Cada relación tiene un nombre que describe su función. Las relaciones se representan gráficamente mediante rombos y su nombre aparece en el interior.

Las entidades que están involucradas en una determinada relación se denominan *entidades participantes*. El número de participantes en una relación es lo que se denomina *grado* de la relación. Por lo tanto, una relación en la que participan dos entidades es una relación *binaria*; si son tres las entidades participantes, la relación es *ternaria*; etc.

Una *relación recursiva* es una relación donde la misma entidad participa más de una vez en la relación con distintos papeles. El nombre de estos papeles es importante para determinar la función de cada participación.

La *cordialidad* con la que una entidad participa en una relación especifica el número mínimo y el número máximo de correspondencias en las que puede tomar parte cada ocurrencia de dicha entidad. La participación de una entidad en una relación es *obligatoria (total)* si la existencia de cada

una de sus ocurrencias requiere la existencia de, al menos, una ocurrencia de la otra entidad participante. Si no, la participación es *opcional (parcial)*. Las reglas que definen la cordialidad de las relaciones son las *reglas de negocio*.

A veces, surgen problemas cuando se está diseñado un esquema conceptual. Estos problemas, denominados *trampas*, suelen producirse a causa de una mala interpretación en el significado de alguna relación, por lo que es importante comprobar que el esquema conceptual carece de dichas trampas. En general, para encontrar las trampas, hay que asegurarse de que se entiende completamente el significado de cada relación. Si no se entienden las relaciones, se puede crear un esquema que no represente fielmente la realidad.

Una de las trampas que pueden encontrarse ocurre cuando el esquema representa una relación entre entidades, pero el camino entre algunas de sus ocurrencias es ambiguo. El modo de resolverla es reestructurando el esquema para representar la asociación entre las entidades correctamente.

Otra de las trampas sucede cuando un esquema sugiere la existencia de una relación entre entidades, pero el camino entre una y otra no existe para algunas de sus ocurrencias. En este caso, se produce una pérdida de información que se puede subsanar introduciendo la relación que sugería el esquema y que no estaba representada.

### Conjunto de relaciones

Consiste en una colección de relaciones de la misma naturaleza.

Ejemplo:

Dados los **conjuntos de entidades** "Habitación" y "Huésped", todas las relaciones de la forma habitación-huésped, permiten obtener la información de los huéspedes y sus respectivas habitaciones.

La dependencia o asociación entre los conjuntos de entidades es llamada **participación**. En el ejemplo anterior los conjuntos de entidades "Habitación" y "Huésped" **participan** en el conjunto de relaciones habitación-huésped.

Se llama **grado** del conjunto de relaciones a la cantidad de conjuntos de entidades participantes en la relación.

### Atributo

Es una característica de interés o un hecho sobre una entidad o sobre una relación. Los atributos representan las propiedades básicas de las

entidades y de las relaciones. Toda la información extensiva es portada por los atributos. Gráficamente, se representan mediante bolitas que cuelgan de las entidades o relaciones a las que pertenecen.

Cada atributo tiene un conjunto de valores asociados denominado *dominio*. El dominio define todos los valores posibles que puede tomar un atributo. Puede haber varios atributos definidos sobre un mismo dominio.

Los atributos pueden ser simples o compuestos. Un *atributo simple* es un atributo que tiene un solo componente, que no se puede dividir en partes más pequeñas que tengan un significado propio. Un *atributo compuesto* es un atributo con varios componentes, cada uno con un significado por sí mismo. Un grupo de atributos se representa mediante un atributo compuesto cuando tienen afinidad en cuanto a su significado, o en cuanto a su uso. Un atributo compuesto se representa gráficamente mediante un óvalo.

Los atributos también pueden clasificarse en monovalentes o polivalentes. Un *atributo monovalente* es aquel que tiene un solo valor para cada ocurrencia de la entidad o relación a la que pertenece. Un *atributo polivalente* es aquel que tiene varios valores para cada ocurrencia de la entidad o relación a la que pertenece. A estos atributos también se les denomina *multivaluados*, y pueden tener un número máximo y un número mínimo de valores. La *cardinalidad* de un atributo indica el número mínimo y el número máximo de valores que puede tomar para cada ocurrencia de la entidad o relación a la que pertenece. El valor por omisión es (1,1).

Por último, los atributos pueden ser derivados. Un *atributo derivado* es aquel que representa un valor que se puede obtener a partir del valor de uno o varios atributos, que no necesariamente deben pertenecer a la misma entidad o relación.

## Identificador

Un identificador de una entidad es un atributo o conjunto de atributos que determina de modo único cada ocurrencia de esa entidad. Un identificador de una entidad debe cumplir dos condiciones:

1. No pueden existir dos ocurrencias de la entidad con el mismo valor del identificador.
2. Si se omite cualquier atributo del identificador, la condición anterior deja de cumplirse.

Toda entidad tiene al menos un identificador y puede tener varios identificadores alternativos. Las relaciones no tienen identificadores.

## Jerarquía de generalización

Una entidad  $E$  es una generalización de un grupo de entidades  $E^1$ ,  $E^2$ , ...  $E^n$ , si cada ocurrencia de cada una de esas entidades es también una ocurrencia de  $E$ . Todas las propiedades de la entidad genérica  $E$  son heredadas por las subentidades.

Cada jerarquía es total o parcial, y exclusiva o superpuesta. Una jerarquía es *total* si cada ocurrencia de la entidad genérica corresponde al menos con una ocurrencia de alguna subentidad. Es *parcial* si existe alguna ocurrencia de la entidad genérica que no corresponde con ninguna ocurrencia de ninguna subentidad. Una jerarquía es *exclusiva* si cada ocurrencia de la entidad genérica corresponde, como mucho, con una ocurrencia de una sola de las subentidades. Es *superpuesta* si existe alguna ocurrencia de la entidad genérica que corresponde a ocurrencias de dos o más subentidades diferentes.

Un *subconjunto* es un caso particular de generalización con una sola entidad como subentidad. Un subconjunto siempre es una jerarquía parcial y exclusiva.

## Atributos en relaciones

Las relaciones también pueden tener atributos asociados. Se representan igual que los atributos de las entidades. Un ejemplo típico son las relaciones de tipo "histórico" donde debe constar una fecha o una hora. Por ejemplo, supongamos que es necesario hacer constar la fecha de emisión de una factura a un cliente, y que es posible emitir duplicados de la factura (con distinta fecha). En tal caso, el atributo "Fecha de emisión" de la factura debería colocarse en la relación "se emite".

## Herencia

La herencia es un intento de adaptación de estos diagramas al paradigma orientado a objetos. La herencia es un tipo de relación entre una entidad "padre" y una entidad "hijo". La entidad "hijo" hereda todos los atributos y relaciones de la entidad "padre". Por tanto, no necesitan ser representadas dos veces en el diagrama. La relación de herencia se representa mediante un triángulo interconectado por líneas a las entidades. La entidad conectada por el vértice superior del triángulo es la entidad "padre". Solamente puede existir una entidad "padre" (herencia simple). Las entidades "hijo" se conectan por la base del triángulo.

## Entidades fuertes y débiles

Cuando una entidad participa en una relación puede adquirir un papel *fuerte* o *débil*. Una entidad débil es aquella que no puede existir sin participar en la relación, es decir, aquella que no puede ser únicamente identificada solamente por sus atributos. Una entidad fuerte (también conocida como entidad regular) es aquella que sí puede ser identificada únicamente. En los casos en que se requiera, se puede dar que una entidad fuerte "preste" algunos de sus atributos a una entidad débil para que, esta última, se pueda identificar.

Las entidades débiles se representan mediante un **doble rectángulo**, es decir, un rectángulo con doble línea.

## Cardinalidad de las relaciones

El tipo de cardinalidad se representa mediante una etiqueta en el exterior de la relación, respectivamente: "1:1", "1:N" y "N:M", aunque la notación depende del lenguaje utilizado, la que más se usa actualmente es el unificado. Otra forma de expresar la cardinalidad es situando un símbolo cerca de la línea que conecta una entidad con una relación:

- "**0**" si cada instancia de la entidad no está obligada a participar en la relación.
- "**1**" si toda instancia de la entidad está obligada a participar en la relación y, además, solamente participa una vez.
- "**N**", "**M**", ó "**\*\***" si cada instancia de la entidad no está obligada a participar en la relación y puede hacerlo cualquier número de veces.

Ejemplos de relaciones que expresan cardinalidad:

- Cada esposo (entidad) está casado (relación) con una única esposa (entidad) y viceversa. Es una relación 1:1.
- Una factura (entidad) se emite (relación) a una persona (entidad) y sólo una, pero una persona puede tener varias facturas emitidas a su nombre. Todas las facturas se emiten a nombre de alguien. Es una relación 1:N.
- Un cliente (entidad) puede comprar (relación) varios artículos (entidad) y un artículo puede ser comprado por varios clientes distintos. Es una relación N:M.

## Restricciones

Son reglas que deben mantener los datos almacenados en la base de datos.

## Correspondencia de cardinalidades

Dado un conjunto de relaciones en el que participan dos o más conjuntos de entidades, la correspondencia de cardinalidad indica el número de entidades con las que puede estar relacionada una entidad dada.

Dado un conjunto de relaciones binarias y los conjuntos de entidades A y B, la correspondencia de cardinalidades puede ser:

- **Uno a uno:** Una entidad de A se relaciona únicamente con una entidad en B y viceversa.
- **Uno a varios:** Una entidad en A se relaciona con cero o muchas entidades en B. Pero una entidad en B se relaciona con una única entidad en A.
- **Varios a uno:** Una entidad en A se relaciona exclusivamente con una entidad en B. Pero una entidad en B se puede relacionar con 0 o muchas entidades en A.
- **Varios a varios:** Una entidad en A se puede relacionar con 0 o muchas entidades en B y viceversa.

## Restricciones de participación

Dado un conjunto de relaciones R en el cual participa un conjunto de entidades A, dicha participación puede ser de dos tipos:

- **Total:** Cuando cada entidad en A participa en al menos una relación de R.
- **Parcial:** Cuando al menos una entidad en A NO participa en alguna relación de R.

## Claves

Es un subconjunto del conjunto de atributos comunes en una colección de entidades, que permite identificar únicamente cada una de las entidades pertenecientes a dicha colección. Asimismo, permiten distinguir entre sí las relaciones de un conjunto de relaciones.

Dentro de los conjuntos de entidades existen los siguientes tipos de claves:

- **Superclave:** Es un subconjunto de atributos que permite distinguir únicamente cada una de las entidades de un conjunto de entidades. Si otro atributo unido al anterior subconjunto, el resultado seguirá siendo una superclave.

- **Clave candidata:** Dada una superclave, si ésta deja de serlo removiendo únicamente uno de los atributos que la componen, entonces ésta es una clave candidata.
- **Clave primaria:** Es una clave candidata, elegida por el diseñador de la base de datos, para identificar únicamente las entidades en un conjunto de entidades.

Los valores de los atributos de una clave, no pueden ser todos iguales para dos o más entidades.

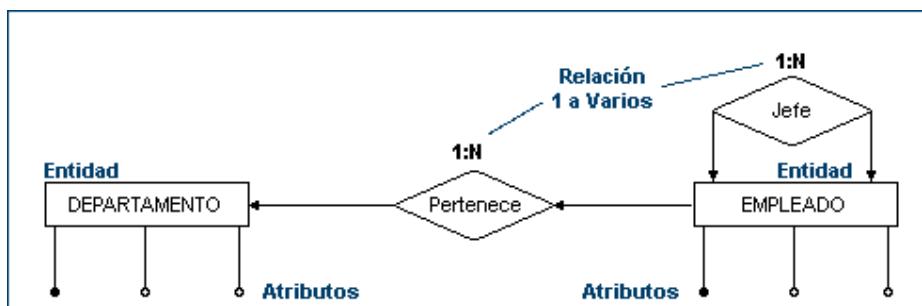
Para poder distinguir únicamente las relaciones en un conjunto de relaciones R, se deben considerar dos casos:

- **R NO tiene atributos asociados:** En este caso, se usa como clave primaria de R la unión de las claves primarias de todos los conjuntos de entidades participantes.
- **R tiene atributos asociados:** En este caso, se usa como clave primaria de R la unión de los atributos asociados y las claves primarias de todos los conjuntos de entidades participantes.

Si el conjunto de relaciones, R, sobre las que se pretende determinar la clave primaria está compuesto de relaciones binarias, con los conjuntos de entidades participantes A y B, se consideran los siguientes casos, según sus cardinalidades:

- **R es de muchos a uno de A a B** entonces sólo se toma la clave primaria de A, como clave primaria de R.
- **R es de uno a muchos de A a B** entonces se toma sólo la clave primaria de B, como clave primaria de R.
- **R es de uno a uno de A a B** entonces se toma cualquiera de las dos claves primarias, como clave primaria de R.

Ejemplo sencillo de diagrama Entidad/Relación:



- Tenemos un empleado que puede ser jefe de otros empleados, pero

esos empleados no pueden tener más de un jefe.

- Un empleado pertenece a un solo departamento, pero un departamento puede tener varios empleados.

### **Técnica para el modelado de datos utilizando diagramas entidad relación.**

No es la única técnica pero sí la más utilizada. Brevemente consiste en los siguientes pasos:

1. Se parte de una descripción textual del problema o sistema de información a automatizar (los requisitos).
2. Se hace una lista de los sustantivos y verbos que aparecen.
3. Los sustantivos son posibles entidades o atributos.
4. Los verbos son posibles relaciones.
5. Analizando las frases se determina la cardinalidad de las relaciones y otros detalles.
6. Se elabora el diagrama (o diagramas) entidad-relación.
7. Se completa el modelo con listas de atributos y una descripción de otras restricciones que no se pueden reflejar en el diagrama.

Dado lo rudimentario de esta técnica se necesita cierto entrenamiento y experiencia para lograr buenos modelos de datos.

El modelado de datos no acaba con el uso de esta técnica. Son necesarias otras técnicas para lograr un modelo directamente implementable en una base de datos. Brevemente:

- Transformación de relaciones múltiples en binarias.
- Normalización de una base de datos de relaciones (algunas relaciones pueden transformarse en atributos y viceversa).
- Conversión en tablas (en caso de utilizar una base de datos relacional).
- Etc.

## **3.2 Representaciones gráficas**

El primer paso en el diseño de una base de datos es la producción del esquema conceptual. Normalmente, se construyen varios esquemas conceptuales, cada uno para representar las distintas visiones que los

usuarios tienen de la información. Cada una de estas visiones suelen corresponder a las diferentes áreas funcionales de la empresa como, por ejemplo, producción, ventas, recursos humanos, etc.

Estas visiones de la información, denominadas *vistas*, se pueden identificar de varias formas. Una opción consiste en examinar los diagramas de flujo de datos, que se pueden haber producido previamente, para identificar cada una de las áreas funcionales. La otra opción consiste en entrevistar a los usuarios, examinar los procedimientos, los informes y los formularios, y también observar el funcionamiento de la empresa.

A los esquemas conceptuales correspondientes a cada vista de usuario se les denomina *esquemas conceptuales locales*. Cada uno de estos esquemas se compone de entidades, relaciones, atributos, dominios de atributos e identificadores. El esquema conceptual también tendrá una documentación, que se irá produciendo durante su desarrollo. Las tareas a realizar en el diseño conceptual son las siguientes:

1. Identificar las entidades.
2. Identificar las relaciones.
3. Identificar los atributos y asociarlos a entidades y relaciones.
4. Determinar los dominios de los atributos.
5. Determinar los identificadores.
6. Determinar las jerarquías de generalización (si las hay).
7. Dibujar el diagrama entidad-relación.
8. Revisar el esquema conceptual local con el usuario.

## **1. Identificar las entidades**

En primer lugar hay que definir los principales objetos que interesan al usuario. Estos objetos serán las entidades. Una forma de identificar las entidades es examinar las especificaciones de requisitos de usuario. En estas especificaciones se buscan los nombres o los sintagmas nominales que se mencionan (por ejemplo: número de empleado, nombre de empleado, número de inmueble, dirección del inmueble, alquiler, número de habitaciones). También se buscan objetos importantes como personas, lugares o conceptos de interés, excluyendo aquellos nombres que sólo son propiedades de otros objetos. Por ejemplo, se pueden agrupar el número de empleado y el nombre de empleado en una entidad denominada *empleado*, y agrupar número de inmueble, dirección del inmueble, alquiler y número de habitaciones en otra entidad denominada *inmueble*.

Otra forma de identificar las entidades es buscar aquellos objetos que existen por sí mismos. Por ejemplo, *empleado* es una entidad porque los empleados existen, sepamos o no sus nombres, direcciones y teléfonos. Siempre que sea posible, el usuario debe colaborar en la identificación de las entidades.

A veces, es difícil identificar las entidades por la forma en que aparecen en las especificaciones de requisitos. Los usuarios, a veces, hablan utilizando ejemplos o analogías. En lugar de hablar de empleados en general, hablan de personas concretas, o bien, hablan de los puestos que ocupan esas personas.

Para liarlo aún más, los usuarios usan, muchas veces, sinónimos y homónimos. Dos palabras son sinónimos cuando tienen el mismo significado. Los homónimos ocurren cuando la misma palabra puede tener distintos significados dependiendo del contexto.

No siempre es obvio saber si un objeto es una entidad, una relación o un atributo. Por ejemplo ¿cómo se podría clasificar *matrimonio*? Pues de cualquiera de las tres formas. El análisis es subjetivo, por lo que distintos diseñadores pueden hacer distintas interpretaciones, aunque todas igualmente válidas. Todo depende de la opinión y la experiencia de cada uno. Los diseñadores de bases de datos deben tener una visión selectiva y clasificar las cosas que observan dentro del contexto de la empresa u organización. A partir de unas especificaciones de usuario es posible que no se pueda deducir un conjunto único de entidades, pero después de varias iteraciones del proceso de análisis, se llegará a obtener un conjunto de entidades que sean adecuadas para el sistema que se ha de construir.

Conforme se van identificando las entidades, se les dan nombres que tengan un significado y que sean obvias para el usuario. Los nombres de las entidades y sus descripciones se anotan en el diccionario de datos. Cuando sea posible, se debe anotar también el número aproximado de ocurrencias de cada entidad. Si una entidad se conoce por varios nombres, éstos se deben anotar en el diccionario de datos como alias o sinónimos.

## 2. Identificar las relaciones

Una vez definidas las entidades, se deben definir las relaciones existentes entre ellas. Del mismo modo que para identificar las entidades se buscaban nombres en las especificaciones de requisitos, para identificar las relaciones se suelen buscar las expresiones verbales (por ejemplo: oficina tiene empleados, empleado gestiona inmueble, cliente visita inmueble). Si las especificaciones de requisitos reflejan estas relaciones es porque son importantes para la empresa y, por lo tanto, se deben reflejar en el esquema conceptual.

Pero sólo interesan las relaciones que son necesarias. En el ejemplo anterior, se han identificado las relaciones *empleado gestiona inmueble* y *cliente visita inmueble*. Se podría pensar en incluir una relación entre empleado y cliente: *empleado atiende a cliente*, pero observando las especificaciones de requisitos no parece que haya interés en modelar tal relación.

La mayoría de las relaciones son binarias (entre dos entidades), pero no hay que olvidar que también puede haber relaciones en las que participen más de dos entidades, así como relaciones recursivas.

Es muy importante repasar las especificaciones para comprobar que todas las relaciones, explícitas o implícitas, se han encontrado. Si se tienen pocas entidades, se puede comprobar por parejas si hay alguna relación entre ellas. De todos modos, las relaciones que no se identifican ahora se suelen encontrar cuando se valida el esquema con las transacciones que debe soportar.

Una vez identificadas todas las relaciones, hay que determinar la cardinalidad mínima y máxima con la que participa cada entidad en cada una de ellas. De este modo, el esquema representa de un modo más explícito la semántica de las relaciones. La cardinalidad es un tipo de restricción que se utiliza para comprobar y mantener la calidad de los datos. Estas restricciones son aserciones sobre las entidades que se pueden aplicar cuando se actualiza la base de datos para determinar si las actualizaciones violan o no las reglas establecidas sobre la semántica de los datos.

Conforme se van identificando las relaciones, se les van asignando nombres que tengan significado para el usuario. En el diccionario de datos se anotan los nombres de las relaciones, su descripción y las cardinalidades con las que participan las entidades en ellas.

### **3. Identificar los atributos y asociarlos a entidades y relaciones**

Al igual que con las entidades, se buscan nombres en las especificaciones de requisitos. Son atributos los nombres que identifican propiedades, cualidades, identificadores o características de entidades o relaciones.

Lo más sencillo es preguntarse, para cada entidad y cada relación, ¿qué información se quiere saber de ...? La respuesta a esta pregunta se debe encontrar en las especificaciones de requisitos. Pero, en ocasiones, será necesario preguntar a los usuarios para que aclaren los requisitos. Desgraciadamente, los usuarios pueden dar respuestas a esta pregunta que también contengan otros conceptos, por lo que hay que considerar sus respuestas con mucho cuidado.

Al identificar los atributos, hay que tener en cuenta si son simples o compuestos. Por ejemplo, el atributo *dirección* puede ser simple, teniendo la dirección completa como un solo valor: 'San Rafael 45, Almazora'; o puede ser un atributo compuesto, formado por la *calle* ('San Rafael'), el *número* ('45') y la *población* ('Almazora'). El escoger entre atributo simple o compuesto depende de los requisitos del usuario. Si el usuario no necesita acceder a cada uno de los componentes de la dirección por separado, se puede representar como un atributo simple. Pero si el usuario quiere acceder a los componentes de forma individual, entonces se debe representar como un atributo compuesto.

También se deben identificar los atributos derivados o calculados, que son aquellos cuyo valor se puede calcular a partir de los valores de otros atributos. Por ejemplo, el número de empleados de cada oficina, la edad de los empleados o el número de inmuebles que gestiona cada empleado. Algunos diseñadores no representan los atributos derivados en los esquemas conceptuales. Si se hace, se debe indicar claramente que el atributo es derivado y a partir de qué atributos se obtiene su valor. Donde hay que considerar los atributos derivados es en el diseño físico.

Cuando se están identificando los atributos, se puede descubrir alguna entidad que no se ha identificado previamente, por lo que hay que volver al principio introduciendo esta entidad y viendo si se relaciona con otras entidades.

Es muy útil elaborar una lista de atributos e ir eliminándolos de la lista conforme se vayan asociando a una entidad o relación. De este modo, uno se puede asegurar de que cada atributo se asocia a una sola entidad o relación, y que cuando la lista se ha acabado, se han asociado todos los atributos.

Hay que tener mucho cuidado cuando parece que un mismo atributo se debe asociar a varias entidades. Esto puede ser por una de las siguientes causas:

- Se han identificado varias entidades, como *director*, *supervisor* y *administrativo*, cuando, de hecho, pueden representarse como una sola entidad denominada *empleado*. En este caso, se puede escoger entre introducir una jerarquía de generalización, o dejar las entidades que representan cada uno de los puestos de empleado.
- Se ha identificado una relación entre entidades. En este caso, se debe asociar el atributo a una sola de las entidades y hay que asegurarse de que la relación ya se había identificado previamente. Si no es así, se debe actualizar la documentación para recoger la nueva relación.

Conforme se van identificando los atributos, se les asignan nombres que tengan significado para el usuario. De cada atributo se debe anotar la siguiente información:

- Nombre y descripción del atributo.
- Alias o sinónimos por los que se conoce al atributo.
- Tipo de dato y longitud.
- Valores por defecto del atributo (si se especifican).
- Si el atributo siempre va a tener un valor (si admite o no nulos).
- Si el atributo es compuesto y, en su caso, qué atributos simples lo forman.
- Si el atributo es derivado y, en su caso, cómo se calcula su valor.
- Si el atributo es multievaluado.

#### **4. Determinar los dominios de los atributos**

El dominio de un atributo es el conjunto de valores que puede tomar el atributo. Por ejemplo el dominio de los números de oficina son las tiras de hasta tres caracteres en donde el primero es una letra y el siguiente o los dos siguientes son dígitos en el rango de 1 a 99; el dominio de los números de teléfono y los números de fax son las tiras de 9 dígitos.

Un esquema conceptual está completo si incluye los dominios de cada atributo: los valores permitidos para cada atributo, su tamaño y su formato. También se puede incluir información adicional sobre los dominios como, por ejemplo, las operaciones que se pueden realizar sobre cada atributo, qué atributos pueden compararse entre sí o qué atributos pueden combinarse con otros. Aunque sería muy interesante que el sistema final respetara todas estas indicaciones sobre los dominios, esto es todavía una línea abierta de investigación.

Toda la información sobre los dominios se debe anotar también en el diccionario de datos.

#### **5. Determinar los identificadores**

Cada entidad tiene al menos un identificador. En este paso, se trata de encontrar todos los identificadores de cada una de las entidades. Los identificadores pueden ser simples o compuestos. De cada entidad se escogerá uno de los identificadores como clave primaria en la fase del diseño lógico.

Cuando se determinan los identificadores es fácil darse cuenta de si una entidad es fuerte o débil. Si una entidad tiene al menos un identificador,

es *fuerte* (otras denominaciones son *padre*, *propietaria* o *dominante*). Si una entidad no tiene atributos que le sirvan de identificador, es *débil* (otras denominaciones son *hijo*, *dependiente* o *subordinada*).

Todos los identificadores de las entidades se deben anotar en el diccionario de datos.

## 6. Determinar las jerarquías de generalización

En este paso hay que observar las entidades que se han identificado hasta el momento. Hay que ver si es necesario reflejar las diferencias entre distintas ocurrencias de una entidad, con lo que surgirán nuevas subentidades de esta entidad genérica; o bien, si hay entidades que tienen características en común y que realmente son subentidades de una nueva entidad genérica.

En cada jerarquía hay que determinar si es total o parcial y exclusiva o superpuesta.

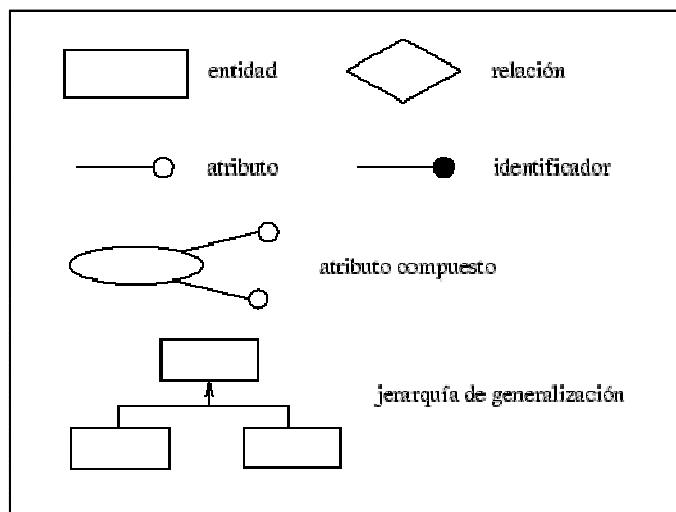
## 7. Dibujar el diagrama entidad-relación

Una vez identificados todos los conceptos, se puede dibujar el diagrama entidad-relación correspondiente a una de las vistas de los usuarios. Se obtiene así un esquema conceptual local.

## 8. Revisar el esquema conceptual local con el usuario

Antes de dar por finalizada la fase del diseño conceptual, se debe revisar el esquema conceptual local con el usuario. Este esquema está formado por el diagrama entidad-relación y toda la documentación que describe el esquema. Si se encuentra alguna anomalía, hay que corregirla haciendo los cambios oportunos, por lo que posiblemente haya que repetir alguno de los pasos anteriores. Este proceso debe repetirse hasta que se esté seguro de que el esquema conceptual es una fiel representación de la parte de la empresa que se está tratando de modelar.

Originalmente, el modelo entidad-relación sólo incluía los conceptos de entidad, relación y atributo. Más tarde, se añadieron otros conceptos, como los atributos compuestos y las jerarquías de generalización, en lo que se ha denominado *modelo entidad-relación extendido*.



*Conceptos del modelo entidad-relación extendido.*

### Representación gráfica de: Artículos y Pedidos,

Una base de datos para una pequeña empresa debe contener información acerca de clientes, artículos y pedidos.

Hasta el momento se registran los siguientes datos en documentos varios

- Para cada cliente: Número de cliente (único), Direcciones de envío (varias por cliente), Saldo, Límite de crédito (depende del cliente, pero en ningún caso debe superar los 3.000.000 pts), Descuento.
- Para cada artículo: Número de artículo (único), Fábricas que lo distribuyen, Existencias de ese artículo en cada fábrica,

Descripción del artículo.

- Para cada pedido: Cada pedido tiene una cabecera y el cuerpo del pedido.

La cabecera está formada por el número de cliente, dirección de envío y fecha del pedido. El cuerpo del pedido son varias líneas, en cada línea se especifican el número del artículo pedido y la cantidad.

Además, se ha determinado que se debe almacenar la información de las fábricas.

Sin embargo, dado el uso de distribuidores, se usará: Número de la fábrica (único) y Teléfono de contacto.

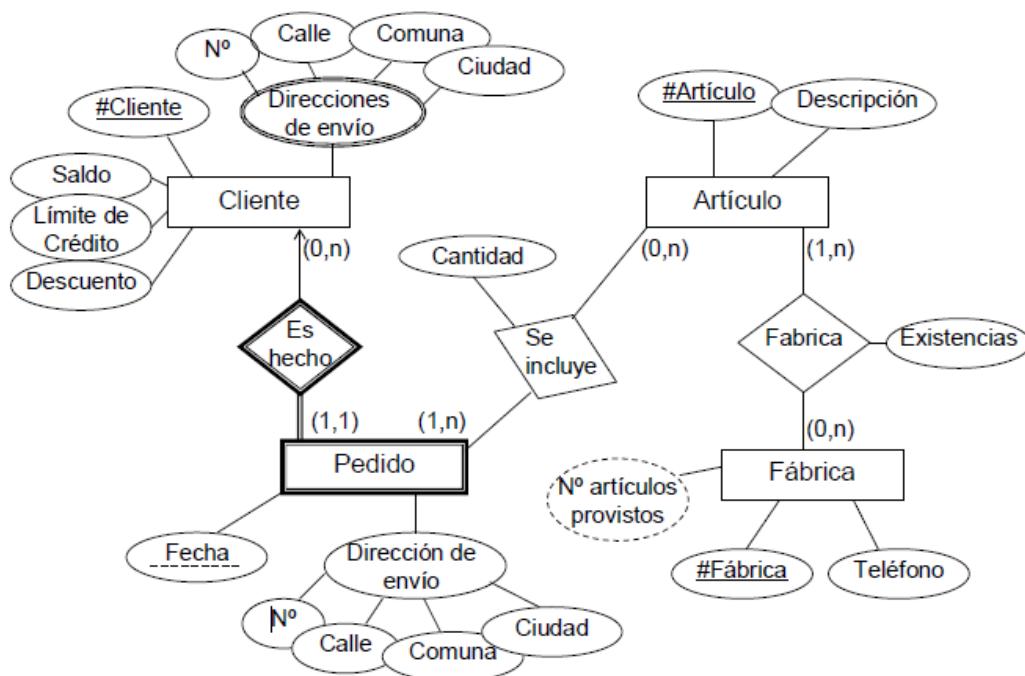
Y se desean ver cuántos artículos (en total) provee la fábrica.

También, por información estratégica, se podría incluir información de fábricas alternativas respecto de las que ya fabrican artículos para esta empresa.

*Nota: Una dirección se entenderá como Nº, Calle, Comuna y Ciudad.*

*Una fecha incluye hora.*

Se pide hacer el diagrama ER para la base de datos que represente esta información.



- Notas:
- El Nº de artículos provistos es la suma de las existencias de cada artículo
  - Se podría almacenar una fábrica de la cual no se tengan artículos

## Representación gráfica de: Sistema de ventas

Le contratan para hacer una BD que permita apoyar la gestión de un sistema de ventas.

La empresa necesita llevar un control de proveedores, clientes, productos y ventas.

Un proveedor tiene un RUT, nombre, dirección, teléfono y página web.

Un cliente también tiene RUT, nombre, dirección, pero puede tener varios teléfonos de contacto.

La dirección se entiende por calle, número, comuna y ciudad.

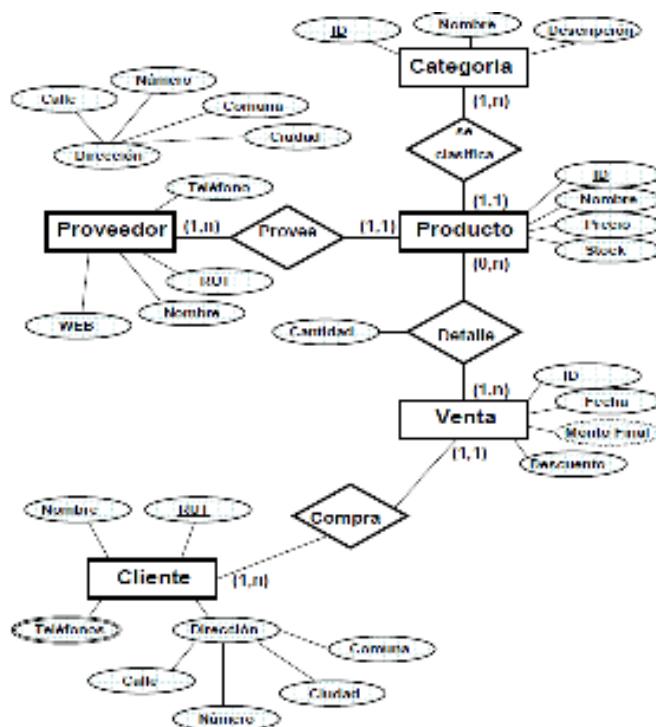
Un producto tiene un id único, nombre, precio actual, stock y nombre del proveedor.

Además se organizan en categorías, y cada producto va sólo en una categoría.

Una categoría tiene id, nombre y descripción.

Por razones de contabilidad, se debe registrar la información de cada venta con un id, fecha, cliente, descuento y monto final.

Además se debe guardar el precio al momento de la venta, la cantidad vendida y el monto total por el producto.



### 3.3 Aplicaciones

#### Definición de Aplicación

(Application). Programa informático que permite a un usuario utilizar una computadora con un fin específico. Las aplicaciones son parte del

software de una computadora, y suelen ejecutarse sobre el sistema operativo.

Una aplicación de software suele tener un único objetivo: navegar en la web, revisar correo, explorar el disco duro, editar textos, jugar (un juego es un tipo de aplicación), etc. Una aplicación que posee múltiples programas se considera un paquete.

Son ejemplos de aplicaciones Internet Explorer, Outlook, Word, Excel, WinAmp, etc.

### **Características de las aplicaciones**

En general, una aplicación es un programa compilado (aunque a veces interpretado), escrito en cualquier lenguaje de programación.

Las aplicaciones pueden tener distintas licencias de distribución como ser freeware, shareware, trialware, etc. Para más información ver: Licencias de software.

Las aplicaciones tienen algún tipo de interfaz, que puede ser una interfaz de texto o una interfaz gráfica (o ambas).

También hay que destacar que la distinción entre aplicaciones y sistemas operativos muchas veces no es clara. De hecho, en algunos sistemas integrados no existe una clara distinción para el usuario entre el sistema y sus aplicaciones.

Debido a la creciente aceptación de las bases de datos por parte de la industria y el gobierno en el plano comercial, y a una variedad de aplicaciones científicas y técnicas, el diseño de bases de datos desempeña un papel central en el empleo de los recursos de información en la mayoría de las organizaciones.

### **3.4 Modelo relacional**

#### **Principios básicos de diseño**

- Todos los datos se representan en tablas.
- Incluso los resultados de cualquier consulta son otra tabla.
- Las tablas están compuestas por filas y columnas.

- Las filas y las columnas, en principio, carecen de orden (p.ej., el orden en el que se muestren las filas y las columnas no importa).
- Las filas sólo se ordenan si se le indica a la base de datos que lo haga, mediante el correspondiente comando.
- De no ser así, el orden será arbitrario, y puede cambiar en caso de tratarse de una base datos dinámica.
- El orden de las columnas lo determina cada consulta.
- Cada tabla tiene una **clave primaria**, un identificador único, compuesto por una o más columnas.
- La mayoría de las claves primarias están formadas por una única columna (p.ej., CIUDAD\_ID).
- Para establecer una relación entre dos tablas es necesario incluir, en forma de columna, en una de ellas la clave primaria de la otra. A esta columna se le llama **clave secundaria**.
- **Estos dos conceptos --clave primaria y secundaria-- son los más importantes en el diseño de bases de datos. Es importante dedicarles tiempo, para entender bien en qué consisten y cómo funcionan.**

Por Ejemplo:

Tabla del empleado

\* Cada una las columnas representa a los atributos de la entidad empleado

Tabla de empleado

Nombre	Puesto	Salario	R.F.C
Juan Pérez Cota	Vendedor	5,000	PECJ500922XYZ
Nora Méndez Angel	Vendedor	5,000	MEAN761014ABC

\* Registros que contienen la información de la entidad empleado

Tabla artículo

Clave	Descripción	Costo
C001	Colcha matrimonial	200

## Cualidades de un buen diseño de base de datos

- Reflejar la estructura del problema en el mundo real.
- Ser capaz de representar todos los datos esperados, incluso con el paso del tiempo.
- Evitar el almacenamiento de información redundante.
- Proporcionar un acceso eficaz a los datos.
- Mantener la integridad de los datos a lo largo del tiempo.
- Ser claro, coherente y de fácil comprensión.
- *Nota: A veces, estos objetivos pueden ser contradictorios.*

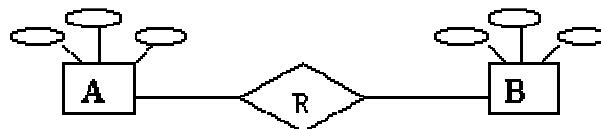
## Definir modelo Entidad/Interrelación (E/R)

- El modelo Entidad/Interrelación (E/R): un método de diseño de bases de datos.
- Muestra de una versión simplificada.
- Representa los datos mediante una serie de **entidades** que disponen de **atributos**.
- Una entidad es una clase de objetos o conceptos claramente identificable.
- Las entidades establecen **interrelaciones** con otras entidades.
- El resultado de este proceso es una base de datos **normalizada** que facilita el acceso a los datos y evita su duplicado.  
*Nota: en su mayor parte, el diseño formal de una base de datos se centra en la normalización de la base y en asegurar que el diseño se ajuste a un nivel de normalización (p.ej., first normal form, second normal form, etc.). Este nivel de formalidad va mucho más allá, pero es importante saber que existen tales formalidades.*

## Proceso de diseño en el modelo E-R

- Identificar las entidades que debe presentar la base de datos.
- Determinar las **cardinalidades** de las interrelaciones establecidas entre las distintas entidades y clasificar estas interrelaciones entre los siguientes tipos:

- **Uno a uno** (p.ej., una parcela sólo tiene una dirección).
- **Uno a muchos** (p.ej., en una parcela pueden ocurrir varios incendios).
- **Muchos a muchos** (p.ej., la venta de parcelas: una misma parcela la pueden vender varios propietarios y cada propietario puede vender varias parcelas).
- Dibujar el diagrama Entidad/Interrelación.



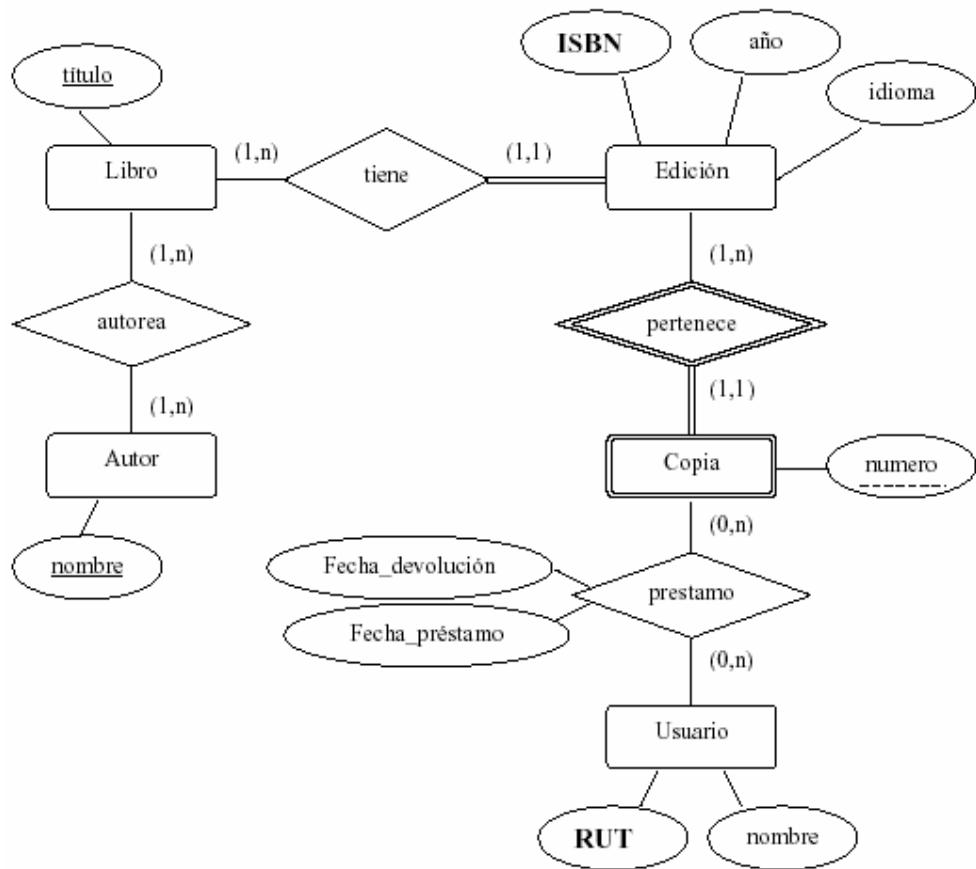
- Determinar los atributos de cada entidad.
- Definir la clave primaria (única) de cada entidad.

### Paso del modelo E-R al diseño de la base de datos

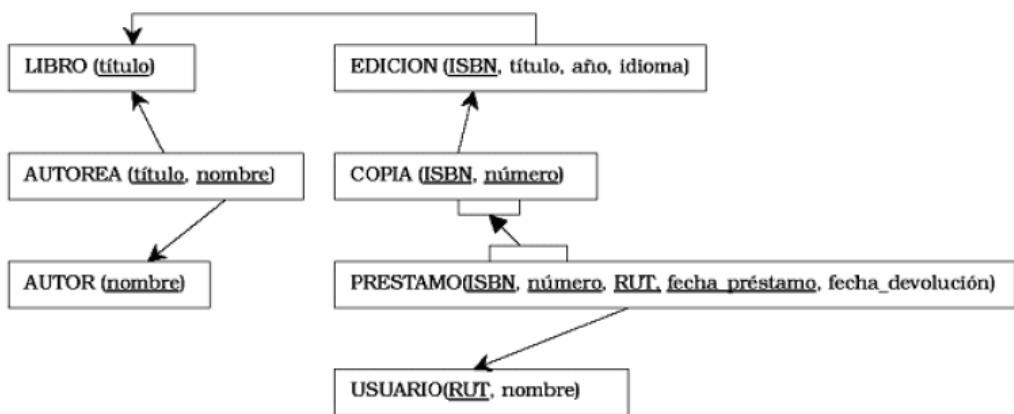
- Las entidades entre las que hay una interrelación **uno a uno** se deben fusionar en una sola entidad.
- Una vez hecho esto, cada una de las entidades que quedan se convierte en una tabla con una clave primaria y una serie de atributos, de los cuales algunos pueden ser claves secundarias.
- Las interrelaciones **uno a muchos** se transforman en atributo y clave secundaria de la tabla que representa a la entidad situada del lado de la interrelación correspondiente a **muchos**
- Las interrelaciones **muchos a muchos** entre dos entidades pasan a ser una tercera tabla con claves secundarias procedentes de ambas entidades. Estas claves secundarias deberán formar parte de la clave primaria de la tabla en la que se convierte la interrelación, cuando corresponda.
- Hay una serie de herramientas disponibles en el mercado que pueden automatizar el proceso de conversión de un modelo E/R en un esquema de base de datos.

Por ejemplo:

## Modelo E / R

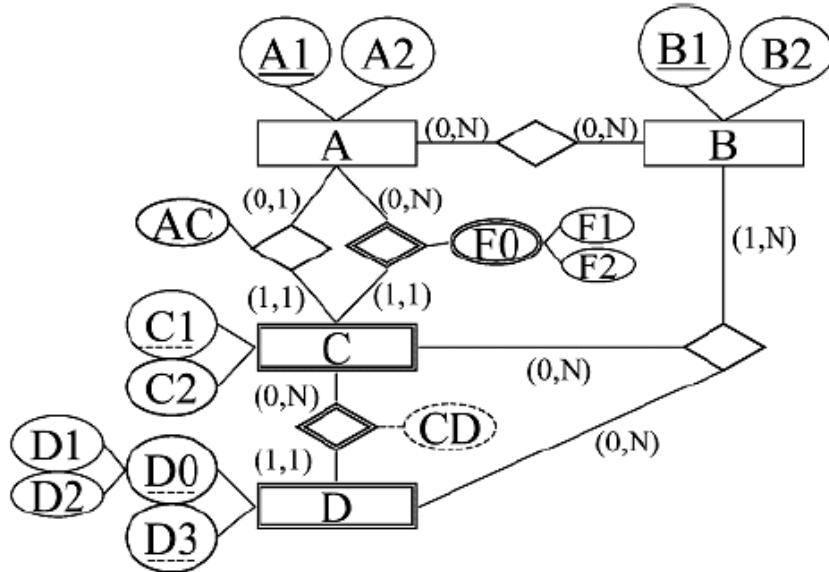


## Modelo Relacional

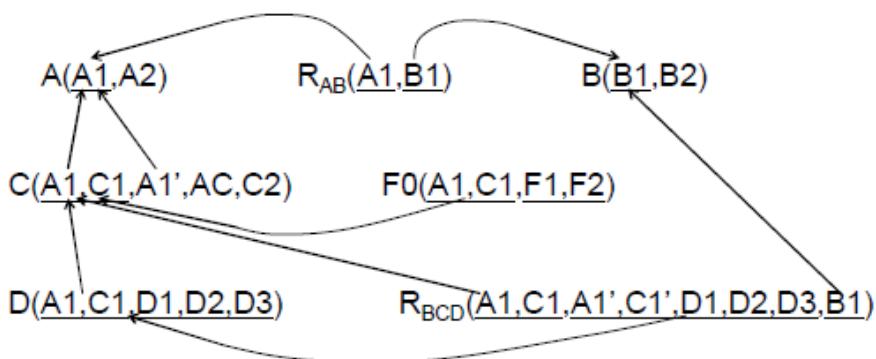


## Ejemplo 2

### Modelo E / R



### Modelo Relacional



Algunas Bases de Datos Relacionales más populares son MS Access, dBase, FoxPro, Paradox, Approach, Oracle y Open Office Base.

En el siguiente TEMA veremos más en detalle al Modelo Relacional.

## ANEXOS

### Ejemplo Modelo entidad-relación

En este apartado se estudia un modelo de datos de alto nivel que nos permite diseñar el esquema conceptual de una DB.

El modelo entidad-relación se incluye a veces entre los modelos orientados a objetos por la noción de distinguibilidad de entidades (similar a la identidad de objetos).

Se basa en los conceptos de: entidad, tipo de entidad, atributo y relación.

Toda esta información se representará en los diagramas entidad-relación.

#### Conceptos

- Entidad:

Def.: Menor objeto con significado en una instancia.

Por Ej.: para el análisis de la DB secretaría, el alumno con los siguientes datos:

DNI = 01234567Z,

Nombre y apellidos = Manuel Vázquez Prieto,

Teléfono = 91-12345678

Domicilio = Calle del Jazmín 7, 4 Izq.

COU = SI

<e> = elemento de

- Atributo:

Def.: Componentes que determinan una entidad. Cada atributo tiene asociado un dominio:

Conjunto de valores que puede tomar.

Ej.: La entidad del Ej.: anterior viene determinada por los valores de sus atributos DNI, Nombre y Apellidos, Teléfono, Domicilio y COU.

- Atributos monovalorados y multivalorados:

Def.: Se llaman atributos multivalorados a aquellos que pueden contener más de un valor simultáneamente y monovalorados a los que sólo pueden contener un valor.

Ej.:

Una persona puede tener varios números de teléfono (casa, trabajo, móvil) y puede que no se interese tenerlos todos. En este caso haremos de teléfono un atributo multivalorado.

- Atributos simples y compuestos:

Def.: Se dice que un atributo es compuesto cuando puede descomponerse en otros componentes o atributos más pequeños, y simple en otro caso.

Ej.:

En el caso del domicilio puede que nos interese descomponerlo a su vez en calle, el número y la ciudad por separado.

- Clave:

Def: Atributo o conjunto de atributos cuyos valores identifican únicamente cada entidad.

Ej.: DNI es un atributo clave de la entidad Alumno.

- Tipo de entidad:

Es el conjunto de entidades que comparten los mismos atributos (aunque con diferentes valores para ellos).

Ej.: En nuestro caso Alumnos será un tipo de entidad que representa cualquier multiconjunto de entidades en el que todas tengan como conjunto de atributos {DNI, Nombre y Apellidos, ...} y valores dentro de los dominios correspondientes. Asignaturas será otro tipo de entidad, etc.

- Relaciones

Def: Conjuntos de la forma  $\{(e_1, \dots, e_n) | e_1 \in E_1, e_2 \in E_2, \dots, e_n \in E_n\}$  con  $e_i$  entidades y  $E_i$  conjuntos de entidades del mismo tipo.

Ej.: Sea  $\{a_1, a_2, a_3, a_4\}$  un conjunto de entidades de tipo alumno ( $i \in$  alumnos concretos) y  $\{b_1, b_2, b_3\}$  3 asignaturas concretas. Una posible relación:  $\{(e_1, b_1), (e_2, b_1), (e_1, b_2)\}$

Diciendo que  $e_1$  está matriculado tanto en  $b_1$  como en  $b_2$  y  $e_2$  en  $b_1$ .

- Tipos de relación:

Representan a todas las posibles relaciones entre conjuntos del mismo tipo. Se identifican mediante los tipos de entidades que relacionan (y los atributos si tienen). Es el producto cartesiano  $E_1 \times E_2 \times \dots \times E_n$ , siendo  $E_i$  conjuntos de entidades.

Ej.: El tipo de relación matrícula relaciona el tipo de entidad alumnos con el tipo de entidad asignaturas.

- Jerarquías Isa:

Def: Se dice A isa B si el conjunto de entidades B es una generalización del conjunto de entidades A.  $\text{attrib}(B) \subset \text{attrib}(A)$ . A hereda de B (en el mismo sentido de la programación orientada a objetos).

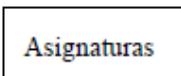
- Atributos clave prestados (borrowed key attributes)

En una jerarquía Isa A isa B, los atributos clave de A pueden serlo también de B.

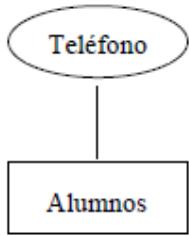
### Diagramas entidad relación

Los componentes básicos de los diagramas ER son los atributos, los tipos de entidades y los tipos de relaciones.

- Tipos de entidades: Rectángulos.



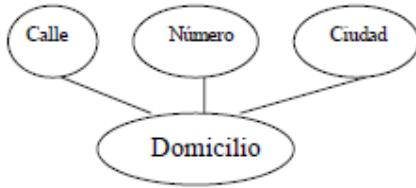
- Atributos: Elipses. Se conectan mediante líneas a los tipos de entidades o tipos de relación.



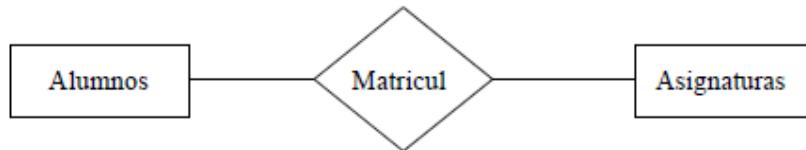
- Atributos multivalorados: Una elipse con doble línea:



- Atributos compuestos. Los componentes de un atributo se representan a su vez como atributos:



- Tipos de Relación: Rombos conectados a los tipos de entidades que relacionan.

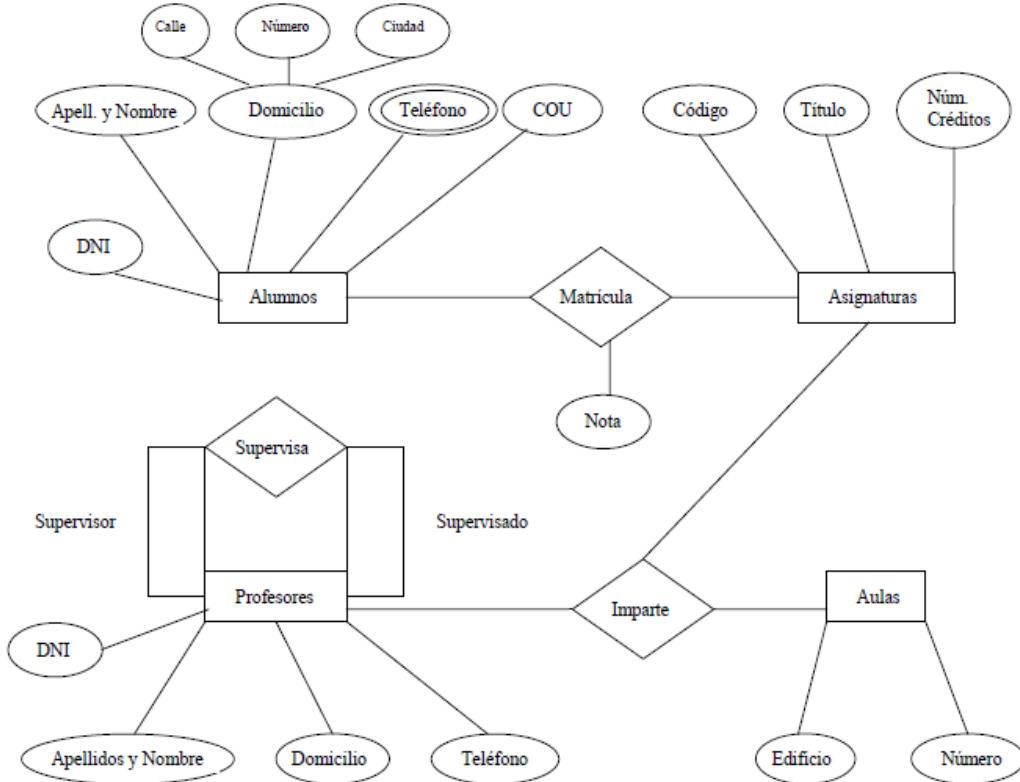


## Cuestiones de diseño

Pasos en el diseño de un diagrama E-R:

- Elección de los tipos de entidad y sus atributos.
- Elección de los tipos de relación.

Ejemplo:



## Restricciones

Con los elementos anteriores tenemos una primera aproximación a los diagramas ER, en la que tenemos definidos los elementos principales de los diagramas. Sin embargo, en el modelo ER también se pueden definir numerosas *restricciones* sobre los tipos de entidades y tipos de relaciones.

Ej.: En la relación *supervisa* un profesor puede tener a lo sumo un supervisor, pero el diagrama anterior permite

### **SUPERVISOR    SUPERVISADO**

({DNI=666666,...}, {DNI=444444,...})  
 ({DNI=000001,...}, {DNI=444444,...})

Que no debería ser una instancia válida de la relación.

Def:

Las *restricciones* son propiedades que se asocian a un tipo de entidad o de relación.

Las *instancias válidas* del tipo de entidad o relación son aquellas en las que se verifique el conjunto de restricciones asociadas.

### **Cardinalidad de un tipo de relación**

**Def:** Cardinalidad de una entidad en una relación (nivel de instancias)

Sea  $r$  una relación entre tipos de entidades  $E_1, E_2, \dots, E_k$ , entonces se dice que:

- El cardinal de  $E_i$  en  $r$  es  $n$  si dados  $e_1 < e > E_1, \dots, e_{i-1} < e > E_{i-1}$ ,  $e_{i+1} < e > E_{i+1}, \dots, e_k < e > E_k$  cualesquiera, se verifica que existen exactamente  $n$   $e_i < e > E_i$  tales que:  
 $< e_1, \dots, e_{i-1}, e_i, e_{i+1}, \dots, e_k > < e > r$ .
- El cardinal de  $E_i$  en  $r$  es menor o igual que  $n$  si dados  $e_1 < e > E_1, \dots, e_{i-1} < e > E_{i-1}$ ,  $e_{i+1} < e > E_{i+1}, \dots, e_n < e > E_k$  cualesquiera, se verifica que existen a lo sumo  $n$   $e_i < e > E_i$  tales que:  
 $< e_1, \dots, e_{i-1}, e_i, e_{i+1}, \dots, e_k > < e > r$

Ej.: Consideremos la siguiente relación (instancia) de tipo AxBxC:

### **A    B    C**

A1 B1 C1 1  
 A1 B1 C2 2  
 A2 B2 C1 3  
 A2 B2 C2 4

A3 B1 C1 5

A3 B1 C2 6

Para A:

(B1,C1): A1 (1), A3 (5)

(B1,C2): A1 (2), A3 (6)

(B2,C1): A2 (3)

(B2,C2): A2 (4)

Para B:

(A1,C1): B1 (1)

(A1,C2): B1 (2)

(A2,C1): B2 (3)

(A2,C2): B2 (4)

(A3,C1): B1 (5)

(A3,C2): B1 (6)

Para C:

(A1,B1): C1 (1), C2(2)

(A2,B2): C1 (3), C2(4)

(A3,B1): C1 (5), C2(6)

Si asumimos que en la instancia de la BD se tiene A = {A1,A2,A3},  
B={B1,B2}, C={C1,C2}

El cardinal de A en esta instancia es  $\leq 2$ , la de B=1 y la de C=2.

Esto vale para instancias pero las restricciones se aplican sobre los tipos de relaciones.

**Def:** Restricciones de cardinalidad. (nivel de esquema)

Se dice que un tipo de entidad participa en un tipo de relación con una restricción de cardinalidad  $=n$  o  $\leq n$  si en todas las instancias válidas del tipo de relación se verifica la restricción. En el caso en que un tipo de entidad participe varias veces en el tipo de relación se podrá establecer una restricción para cada *papel*.

Ejs.:

1.



2. Fijado un alumno puede haberse matriculado en cualquier número de asignaturas no hay restricción sobre *asignatura* en la relación *matrícula*.

Fijada una asignatura, puede haberse matriculado sobre ella un número cualquiera de alumnos no hay restricciones sobre el tipo de entidad *alumnos* en la relación *matrícula*.

3. El supervisor de un profesor, si lo tiene, es único. El tipo de entidades *profesor*, en el papel *supervisor* tiene cardinal  $\leq 1$ . El tipo de entidades *profesor*, en el papel *supervisado* no tiene ninguna restricción de cardinal: un profesor puede supervisar a un número indeterminado de profesores.

4. Tipo de relación *imparte*:

- Dado un profesor y una asignatura, existe a lo sumo (si es profesor de la asignatura) un aula en la que se imparte clase. Restricción de cardinal para *aulas*:  $\leq 1$ .
- Dada una asignatura y un aula, puede haber varios profesores (ej.: uno de mañana y otro de tarde) o ninguno.

## Diagramas ER

Hay dos formas de expresar las restricciones de cardinalidad sobre tipos de relaciones en los diagramas:

O bien poniendo la restricción directamente sobre la línea ( $=1$ ,  $\leq 10\dots$ )

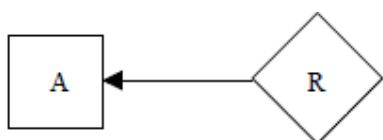
O, más común:

a) Restricciones de una a muchas, de muchas a una, de una a una

Concepto:

Dado un tipo de relación R entre tipos de entidad E1, E2, ..., En se puede especificar una restricción de cardinalidad distinguiendo si el tipo de entidad tiene restricción  $\leq 1$  o  $=1$  diremos que participa con cardinalidad **una** o no tiene ninguna restricción (cardinalidad muchas).

Si A participa con cardinalidad **una** en el tipo de relación R lo representaremos



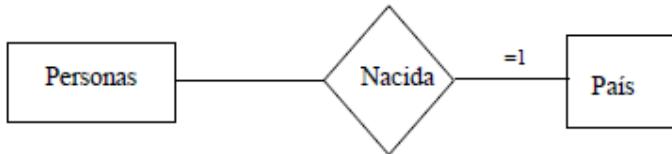
El caso de muchas se representa con una línea sin flecha (como hasta ahora, porque no tiene restricción).

Ej.:s

1.

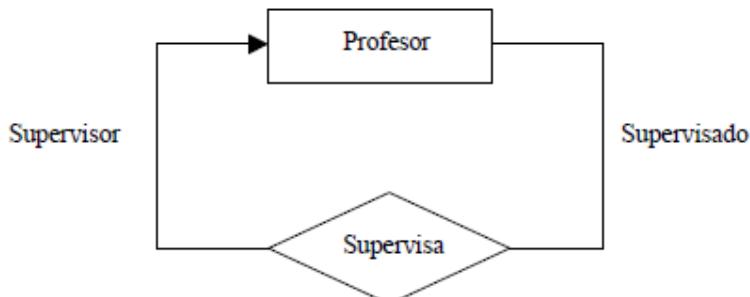


O bien

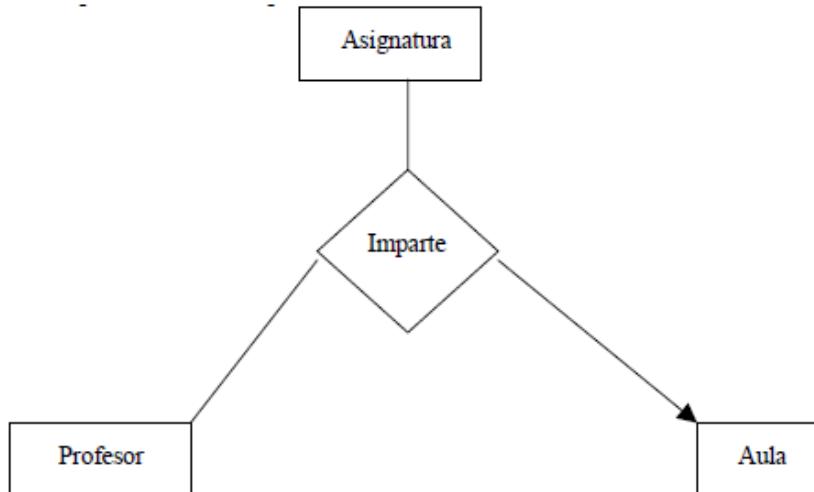


2. Matrícula: Se queda como está.

3. Profesores y supervisores



#### 4. Tipo de relación *imparte*:



#### **Participación de una entidad en una relación**

Sea  $r$  una relación definida sobre los tipos de entidades  $E_1, \dots, E_m$  y sea  $E_j \in \{E_1, \dots, E_m\}$ :

- Participación:

Def.: Se dice que la participación de la entidad  $e_j \in E_j$  en  $r$  es  $n$  ( $n \in \mathbb{N}$ ) si  $e_j \in E_j$  aparece en  $n$  tuplas de la relación.

- Participación total:

Def.: Se dice que  $E_j$  tiene participación total en  $r$  si cada entidad  $e_j \in E_j$  se encuentra en alguna tupla de  $r$ . En otro caso se dice que la participación es *parcial*.

Ej.: Consideremos la siguiente relación  $r$  de tipo  $A \times B \times C$ :

**A B C**

A1 B1 C1

A1 B1 C2

A2 B2 C1

A2 B2 C2

A3 B1 C1

## A3 B1 C2

Con los multiconjuntos de entidades  $A = \{A1, A2, A3\}$ ,  $B = \{B1, B2\}$ ,  $C = \{C1, C2\}$

La participación de  $A1, A2, A3$  en esta instancia es =2, la de  $B1=4$ , la de  $B2=2$  y la de  $C1=3$  y la de  $C2=3$ .

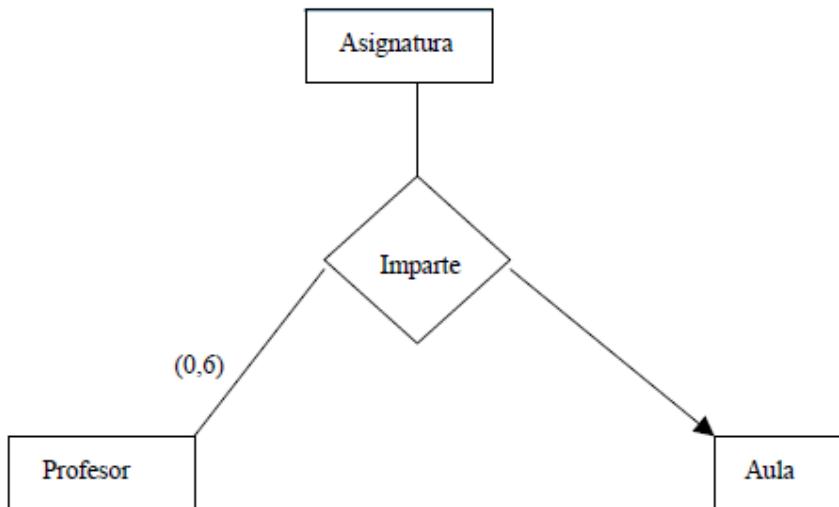
- Restricciones de participación en los esquemas
- Def.: Una restricción de participación (min,max) ( $\text{min} <e> N, \text{max}<e> N$ ) de un tipo de entidades Ej en un tipo de relación R indica que en todas las instancias válidas de la BD se verifica:

$e <e> \text{ Ej } " \text{ participación de } e \text{ en } R \text{ está entre } \text{min} \text{ y } \text{max}.$

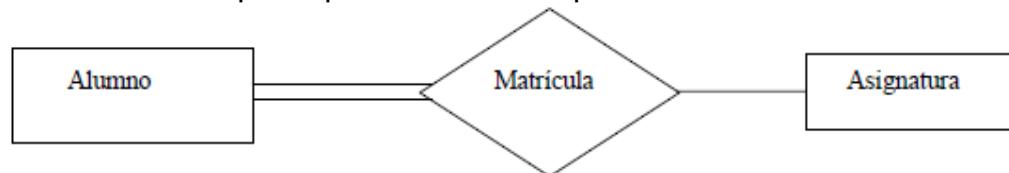
- Def.: Una restricción de participación *total* de un tipo de entidades Ej en un tipo de relación R indica que en todas las instancias válidas de la BD, se verifica que Ej tiene una participación total.

**Diagramas ER**

La restricción de participación (min,max) se representa



La restricción de participación total se representa como:



## Unicidad de entidades

### Claves

- Superclave.

Def.: Dado un tipo de entidades E en una BD, se llama superclave a cualquier conjunto de atributos que permita distinguir a todas las entidades de cualquier instancia válida de E en la BD.

Si alguno de los atributos de la superclave corresponde a otro tipo de entidad F se debe verificar:

- E y F deben participar en un tipo de relación binaria R en la que F debe tener una restricción de cardinalidad  $\leq 1$ .
- Los atributos que F aporta para la clave candidata de E deben ser atributos de una clave candidata de F.
- La participación de E en R debe ser total.

### Propiedad:

Si S es una superclave y  $S \subsetneq S'$ , entonces  $S'$  superclave

Ej.:

En el caso de asignaturas tenemos en realidad 6 superclaves  
{título}, {código}, {título, num.creditos}, {código, num.creditos}, {título, código}, {título, código, num.creditos}.

- Clave candidata.

Def.: Se llama clave candidata de un tipo de entidad a una superclave que no contiene ningún subconjunto que también sea superclave.  
(Conjunto mínimo de atributos que forma una superclave).

- Clave primaria.

Se llama *clave primaria* a la clave candidata seleccionada por el diseñador para distinguir entre las entidades de cada instancia.

## Diagramas ER

En los diagramas ER los atributos de la clave primaria se representan con sus nombres subrayados.

## Tipos de entidad débiles

Un tipo de entidades que no tiene suficientes atributos para formar una clave primaria se denomina tipo de entidades débil.

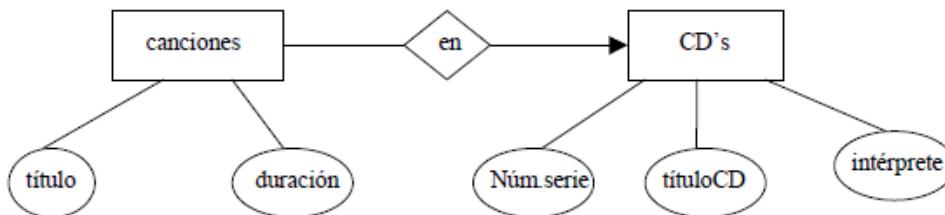
Ej.:

Supongamos que estamos diseñando una BD para CDs de música.

Vamos a utilizar la siguiente información:

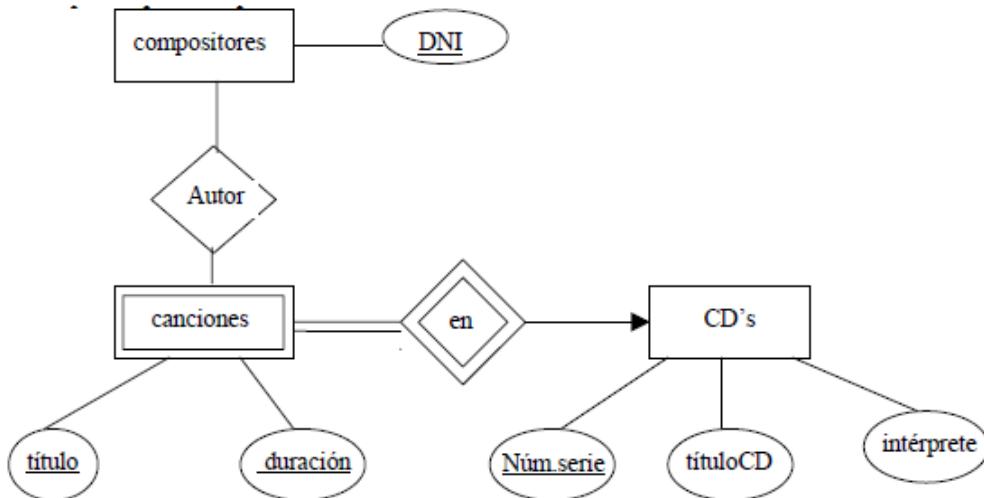
CD : Título del CD, intérprete, númer. Serie Canción: Título, duración

También deseamos relacionar las canciones con el CD al que pertenecen. Esta relación será de muchas a una entre canciones y CD's (a cada canción le corresponde un CD).



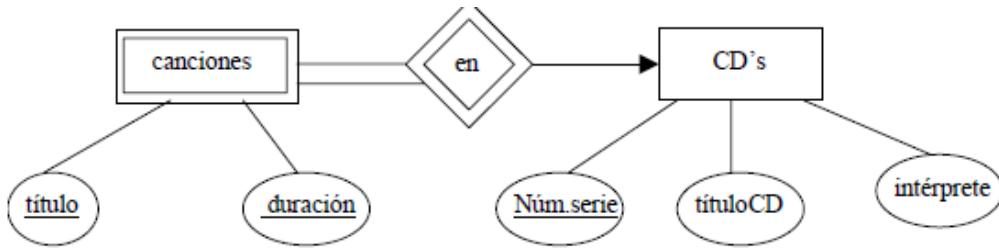
### Utilidad de las entidades débiles

También queremos relacionar cada canción con su autor o autores. Un autor viene dado por su DNI que no puede repetirse.



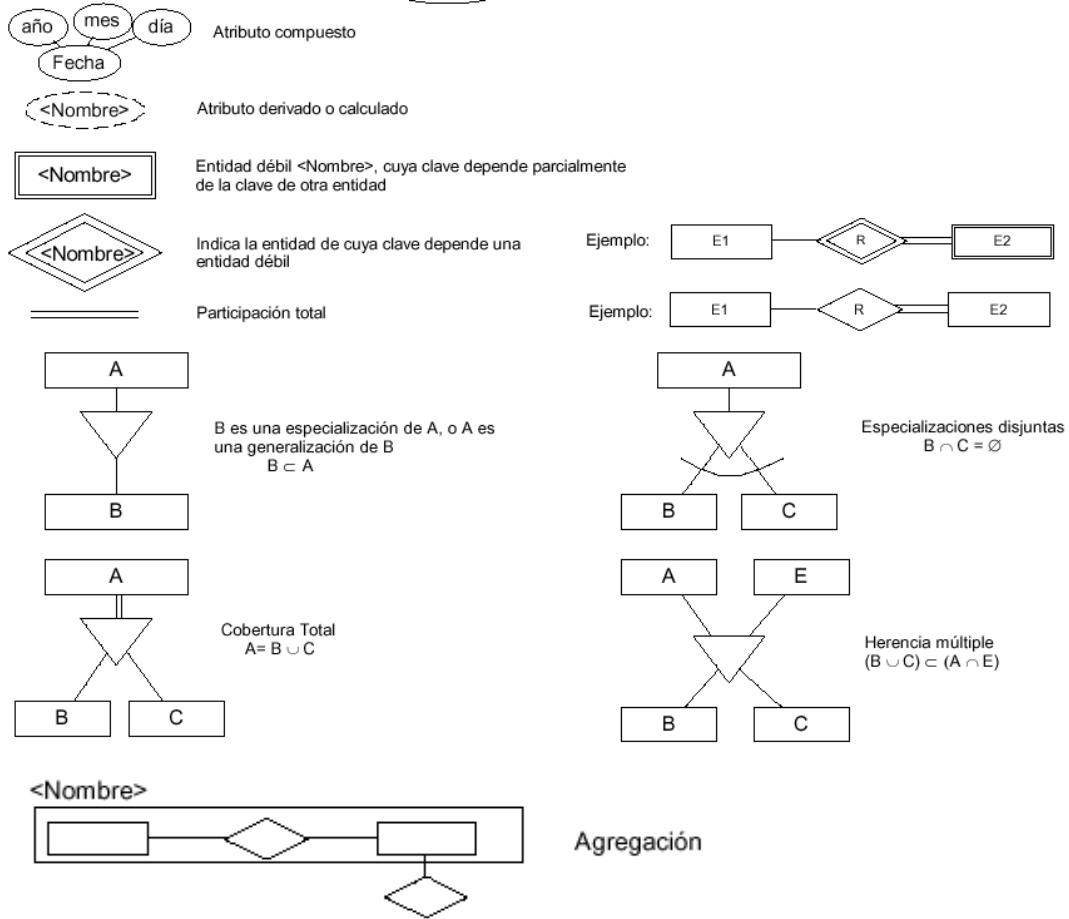
### Diagrama ER

Los tipos de entidad débiles se representan con rectángulos dobles, y el tipo de relación (o los tipos) que permiten formar la clave se indican con un doble rombo.



## NOTACIÓN PARA LOS DIAGRAMAS ENTIDAD/RELACIÓN

<div style="border: 1px solid black; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div>	Entidad <Nombre>
<div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div>	Relación <Nombre> entre entidades
<div style="border: 1px solid black; padding: 2px; display: inline-block;">E1</div> --> <div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div> 1:1 --> <div style="border: 1px solid black; padding: 2px; display: inline-block;">E2</div>	Un elemento de E1 puede estar relacionado como máximo con un elemento de E2, y viceversa
<div style="border: 1px solid black; padding: 2px; display: inline-block;">E1</div> --> <div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div> 1:N --> <div style="border: 1px solid black; padding: 2px; display: inline-block;">E2</div>	Un elemento de E1 puede estar relacionado con varios elementos de E2, pero un elemento de E2 solamente puede estar relacionado con un único elemento de E1
<div style="border: 1px solid black; padding: 2px; display: inline-block;">E1</div> --> <div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div> M:N --> <div style="border: 1px solid black; padding: 2px; display: inline-block;">E2</div>	Un elemento de E1 puede estar relacionado con varios elementos de E2, y cada elemento de E2 puede estar relacionado con varios elementos de E1
<div style="border: 1px solid black; padding: 2px; display: inline-block;">E1</div> --> <div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div> (a,b) --> <div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div> (c,d) --> <div style="border: 1px solid black; padding: 2px; display: inline-block;">E2</div>	Cada elemento de E1 está relacionado con un mínimo de c y un máximo de d elementos de E2, y un elemento de E2 está relacionado con un mínimo de a y un máximo de b elementos de E1
<div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;">&lt;Nombre&gt;</div>	Atributo <Nombre> de la entidad con la que este conectado
<div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;"><u>&lt;Nombre&gt;</u></div>	Atributo clave
<div style="border: 1px solid black; border-radius: 50%; padding: 2px; display: inline-block;"><u><u>&lt;Nombre&gt;</u></u></div>	Atributo multivalorado



## GUÍA DE EJERCICIOS:

### Modelo Entidad/Relación y conversión a Modelo Relacional

#### Modelo Entidad/Relación

##### **Problema 1 (\*): Artículos y encargos**

Una base de datos para una pequeña empresa debe contener información acerca de clientes, artículos y pedidos. Hasta el momento se registran los siguientes datos en documentos varios:

- Para cada cliente: Número de cliente (único), Direcciones de envío (varias por cliente), Saldo, Límite de crédito (depende del cliente, pero en ningún caso debe superar los 3.000.000 pts), Descuento.
- Para cada artículo: Número de artículo (único), Fábricas que lo distribuyen, Existencias de ese artículo en cada fábrica, Descripción del artículo.
- Para cada pedido: Cada pedido tiene una cabecera y el cuerpo del pedido. La cabecera está formada por el número de cliente, dirección de envío y fecha del pedido. El cuerpo del pedido son varias líneas, en cada línea se especifican el número del artículo pedido y la cantidad. Además, se ha determinado que se debe almacenar la información de las fábricas. Sin embargo, dado el uso de distribuidores, se usará: Número de la fábrica (único) y Teléfono de contacto. Y se desean ver cuántos artículos (en total) provee la fábrica. También por información estratégica, se podría incluir información de fábricas alternativas respecto de las que ya fabrican artículos para esta empresa.

*Nota: Una dirección se entenderá como Nº, Calle, Comuna y Ciudad.*

*Una fecha incluye hora.*

Se pide hacer el diagrama ER para la base de datos que represente esta información.

##### **Problema 2 (\*): Sistema de ventas**

Le contratan para hacer una BD que permita apoyar la gestión de un sistema de ventas. La empresa necesita llevar un control de proveedores, clientes, productos y ventas.

Un proveedor tiene un RUT, nombre, dirección, teléfono y página web. Un cliente también tiene RUT, nombre, dirección, pero puede tener varios teléfonos de contacto. La dirección se entiende por calle, número, comuna y ciudad.

Un producto tiene un id único, nombre, precio actual, stock y nombre del proveedor. Además se organizan en categorías, y cada producto va sólo en una categoría. Una categoría tiene id, nombre y descripción.

Por razones de contabilidad, se debe registrar la información de cada venta con un id, fecha, cliente, descuento y monto final. Además se debe guardar el precio al momento de la venta, la cantidad vendida y el monto total por el producto.

**Problema 3: Carreteras**

Diseñar un esquema E/R que recoja la organización de una base de datos para contener la información sobre todas las carreteras del país, sabiendo que se deben cumplir las siguientes especificaciones:

- Las carreteras están divididas en varias categorías (locales, comerciales, regionales, nacionales, autovías, etc).
- Las carreteras se dividen en tramos. Un tramo siempre pertenece a una única carretera y no puede cambiar de carretera.
- Un tramo puede pasar por varias comunas, interesando conocer el Km. de la carretera y la comuna donde empieza el tramo y en donde termina.
- Para los tramos que suponen principio o final de carretera, interesa saber si es que la carretera concluye físicamente o es que confluye en otra carretera. En este caso, interesa conocer con qué carretera confluye y en qué kilómetro, tramo y comuna.

**Problema 4: Sistema de vuelos**

Obtener el diagrama E/R para un sistema de control de vuelos adaptado a las siguientes reglas de gestión (indicar las entidades, interrelaciones, etc., que se deducen de cada una de las reglas):

- a) De cada aeropuerto se conoce su código, nombre, ciudad y país.
  - b) En cada aeropuerto pueden tomar tierra diversos modelos de aviones (el modelo de un avión determina su capacidad, es decir, el número de plazas).
  - c) En cada aeropuerto existe una colección de programas de vuelo. En cada programa de vuelo se indica el número de vuelo, línea aérea y días de la semana en que existe dicho vuelo.
  - d) Cada programa de vuelo despegue de un aeropuerto y aterriza en otro.
  - e) Los números de vuelo son únicos para todo el mundo.
  - f) En cada aeropuerto hay múltiples aterrizajes y despegues. Todos los aeropuertos contemplados están en activo, es decir, tienen algún aterrizaje y algún despegue.
  - g) Cada vuelo realizado pertenece a un cierto programa de vuelo. Para cada vuelo se quiere conocer su fecha, plazas vacías y el modelo de avión utilizado.
  - h) Algunos programas de vuelo incorporan escalas técnicas intermedias entre los aeropuertos de salida y de llegada. Se entiende por escala técnica a un aterrizaje y despegue consecutivos sin altas ó bajas de pasajeros.
  - i) De cada vuelo se quieren conocer las escalas técnicas ordenadas asignándole a cada una un número de orden.
- Por ejemplo, el programa de vuelo 555 de Iberia con vuelos los lunes y jueves despegue de Barajas-Madrid-España y aterriza en Caudell-Sydney-Australia teniendo las siguientes escalas técnicas: 1- Los Pradiños-Sao Paulo-Brasil, 2-El Emperador-Santiago-Chile y 3-Saint Kitts-Auckland-

Nueva Zelanda □ *¿Que cambios se producirán en el caso anterior si en las escalas pudiesen bajar o subir pasajeros?*

Explicar cómo se podría representar esta nueva situación.

### **Problema 5 (\*): Olimpiadas**

Las sedes olímpicas se dividen en complejos deportivos. Los complejos deportivos se subdividen en aquellos en los que se desarrolla un único deporte y en los polideportivos. Los complejos polideportivos tienen áreas designadas para cada deporte con un indicador de localización (ejemplo: centro, esquina-NE, etc.). Un complejo tiene una localización, un jefe de organización individual y un área total ocupada.

Los dos tipos de complejos (deporte único y polideportivo) tendrán diferentes tipos de información. Para cada tipo de sede, se conservará el número de complejos junto con su presupuesto aproximado.

Cada complejo celebra una serie de eventos (ejemplo: la pista del estadio puede celebrar muchas carreras distintas.). Para cada evento está prevista una fecha, duración, número de participantes, número comisarios. Una lista de todos los comisarios se conservará junto con la lista de los eventos en los que esté involucrado cada comisario ya sea cumpliendo la tarea de juez u observador. Tanto para cada evento como para el mantenimiento se necesitará cierto equipamiento (ejemplo: arcos, pértigas, barras paralelas, etc).

### **Problema 6: Educando S.A.**

En la Empresa "Educando S.A." se lleva control de sus Bienes y Servicios. El interés primario es poder hacer que los Bienes se manejen de forma rápida y con el menor grado de error. Para esto quien maneja la sección de "Bienes y Suministros" plantea las siguientes condiciones del negocio para la construcción de una base de datos:

- La Sección está dividida en tres (3) áreas: COMPRAS, ALMACEN, INVENTARIO.
- El área de Compras funciona de la siguiente forma:
  - o Recibe las solicitudes de compras de las diferentes áreas de la empresa.
  - o Cada solicitud tiene un responsable.
  - o Cada solicitud es autorizada por el jefe del área y posteriormente por el Director Financiero.
  - o Quien realiza una solicitud puede ser responsable de uno o varios centros de costos, con la salvedad de que él como empleado solo está adscrito a uno.
  - o De la solicitud se debe diligenciar la siguiente información: Número de la solicitud (consecutivo), Fecha, Responsable (nombre y cédula), Centro de Costos, Rubro presupuestal del cual se descargará la compra. En cada solicitud se pueden discriminar uno o muchos ítems con la siguiente información: ítem, nombre del bien, cantidad solicitada, unidad de medida del bien, valor unitario y valor total. Cada solicitud debe ser totalizada.

- o Cada bien es identificado por un código universal que es único y es de carácter devolutivo (suministro) o un bien inmueble.
- o Una vez diligenciada la solicitud es remitida al área de compras para realizar su correspondiente cotización.
- o Las cotizaciones son realizadas con uno o varios proveedores de los bienes solicitados.
- o Una vez la cotización definitiva está lista, se crea una orden contractual que maneja la siguiente información: Número de la orden contractual, *nit* y nombre del proveedor al cual se le va a realizar la compra, fecha de la orden, monto total de la orden, fecha de entrega. Cada orden puede tener asociado uno o varios ítems de la solicitud o solicitudes que van a ser despachadas. Cada ítem tiene la siguiente información: nombre del bien, cantidad solicitada, cantidad despachada, unidad de medida del bien, valor unitario y valor total.
- o La orden de compra es aprobada por el Director Financiero para que sea enviada al proveedor elegido · El área de Almacén funciona de la siguiente forma:
  - o Su función principal es recibir los bienes que llegan de los proveedores y distribuirlos a las correspondientes áreas que realizaron las solicitudes de compras.
  - o Cuando llega un proveedor mercancía, este hace una entrega física de los bienes, los cuales son comparados con la factura que este entrega y con la orden de compra correspondiente. Si esta acción es correcta se registra una entrada de almacén por cada factura relacionada, con la siguiente información: Número de Entrada, Fecha, Número de factura, Proveedor, Total Bienes, Valor Total (los totales deben coincidir con los de la factura). Adjunto a esta se discriminan los ítems recibidos con la siguiente información: nombre del bien, cantidad entregada.
  - o Cuando el almacén decide despachar los bienes a las diferentes áreas solicitantes, registra cada una de las entregas en Salidas de Almacén con la siguiente información:Número de Salida, Empleado responsable del bien a entregar, fecha de salida, fecha de entrega. Por cada entrega se detalla cada uno de los ítems con la siguiente información: nombre del bien, cantidad entregada.
  - o Una entrada de almacén puede generar muchas salidas de almacén, por ejemplo: Pueden ingresar 500 pacas de papel higiénico, pero como se debe repartir entre varias áreas, cada una requiere de una salida de almacén. · El área de inventarios funciona de la siguiente forma:
    - o Es la encargada de administrar y controlar la ubicación de los bienes dentro de la empresa, por esto antes de que el bien salga del almacén debe ser codificado a través de un código único que lo haga identificable dentro de la empresa.
    - o La ubicación del bien se identifica por la siguiente información: responsable del bien, fecha de entrega, dirección del bien (ubicación).

Diseñar modelo ER para la base de datos.

**Problema 7: Torneo de Tenis Grand Slam**

El sistema debe memorizar todos los encuentros que se han desarrollado desde que existe el torneo, así como las siguientes características de estos.

*Descripción:*

El Grand Slam se compone de cuatro torneos anuales que se celebran en Gran Bretaña, Estados Unidos, Francia y Australia. En cada país se pueden desarrollar en distintos lugares (p. ej., en EE. UU. puede desarrollarse en Forest Hill o en Flushing Meadows).

Cada partido tiene asociado un premio de consolación para el perdedor que dependerá de la fase en que se encuentre el torneo (p. ej., el perdedor de octavos de final puede ganar 5.000 dólares). El ganador de la final recibirá el premio correspondiente al torneo.

Cada torneo tiene cinco modalidades: Individual masculino, individual femenino, dobles masculino, dobles femenino y dobles mixtos.

También hay que tener en cuenta la nacionalidad de un jugador, de forma que este puede ser apátrida o tener varias nacionalidades.

*Resultados a considerar:*

El sistema debe dar respuesta a las siguientes preguntas:

1. Dado un año y un torneo, composición y resultado de los partidos.
2. Lista de árbitros que participaron en el torneo.
3. Ganancias percibidas en premios por un jugador a lo largo del torneo.
4. Lista de entrenadores que han entrenado a un jugador a lo largo del torneo y fechas en las que lo hizo.

Ejemplos de acceso a la base de datos.

1. Connors gano Gerulaitis en Roland Garros en 1979 en cuartos de final en individuales masculinos por 6-3 4-6/7-5 6-0.
2. El señor Wilkinson arbitro ese partido.
3. Alemania ha ganado dos veces las individuales masculinas de Wimbledon. Borg ha ganado 2.000.000 de dólares a lo largo de su participación en el Grand Slam.
4. El ganado de Roland Garros de 1987 ganó 20.000 dólares.
5. Noah ha jugado cuatro veces en dobles mixtos con Mandlikova.

**Problema 8:**

Se desea crear un sitio web con información referente a las películas en cartel en las salas de un dudoso cine cercano a la plaza de armas.

De cada película, se almacena una ficha con su título de distribución, su título original, su género, el idioma original, si tiene subtítulos en español o no, los países de origen, el año de la producción, la url de sitio web de la película, la duración (en horas y minutos), la calificación (Apta todo público,+9 años, +15 años,+18 años), fecha de estreno en Santiago, un resumen y un identificador de la película. De cada película interesa conocer la lista de directores y el reparto, es decir para cada actor que trabaja, el nombre de todos los personajes que interpreta.

Además interesa disponer de información sobre los directores y actores que trabajan en cada película.

De ambos, se conoce su nombre (que lo identifica) y su nacionalidad.

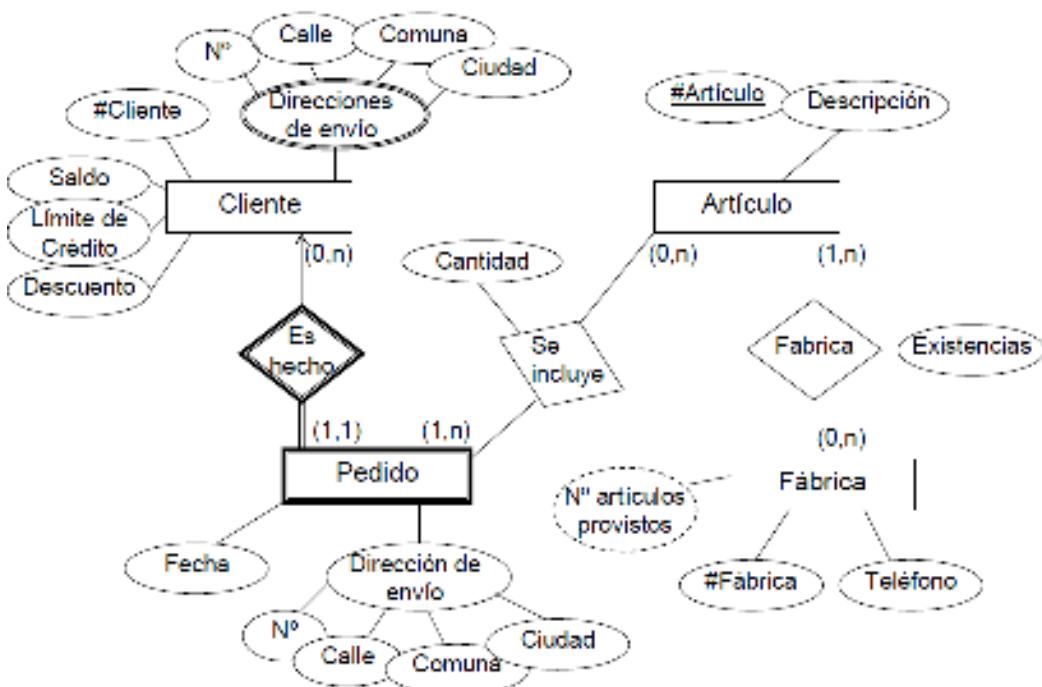
Además se desea conocer la cantidad de películas en las que dirigieron o actuaron. Tenga en cuenta que hay personas que cumplen los dos roles. Los cines pueden tener más de una sala y cada semana cada uno de los cines envía la cartelera para dicha semana, indicando de detalle de las funciones. Para cada función se conoce el día de la semana y la hora de comienzo, y obviamente la sala y la película que exhibe. De cada sala se sabe el nombre, un número que la identifica dentro del cine y la cantidad de butacas que posee. De cada cine se conoce el nombre que lo identifica, su dirección y teléfono para consultas.

Algunos cines cuentan con promociones. Estas promociones dependen de la función. (Ej. De lunes a jueves antes de las 18 50% de descuento en la sala tal del cine tal para la película cual...La función del lunes a las 14 para la película tal en la sala cual, no se cobra a los escolares con túnica...) De cada promoción se conoce una descripción y el descuento que aplica.

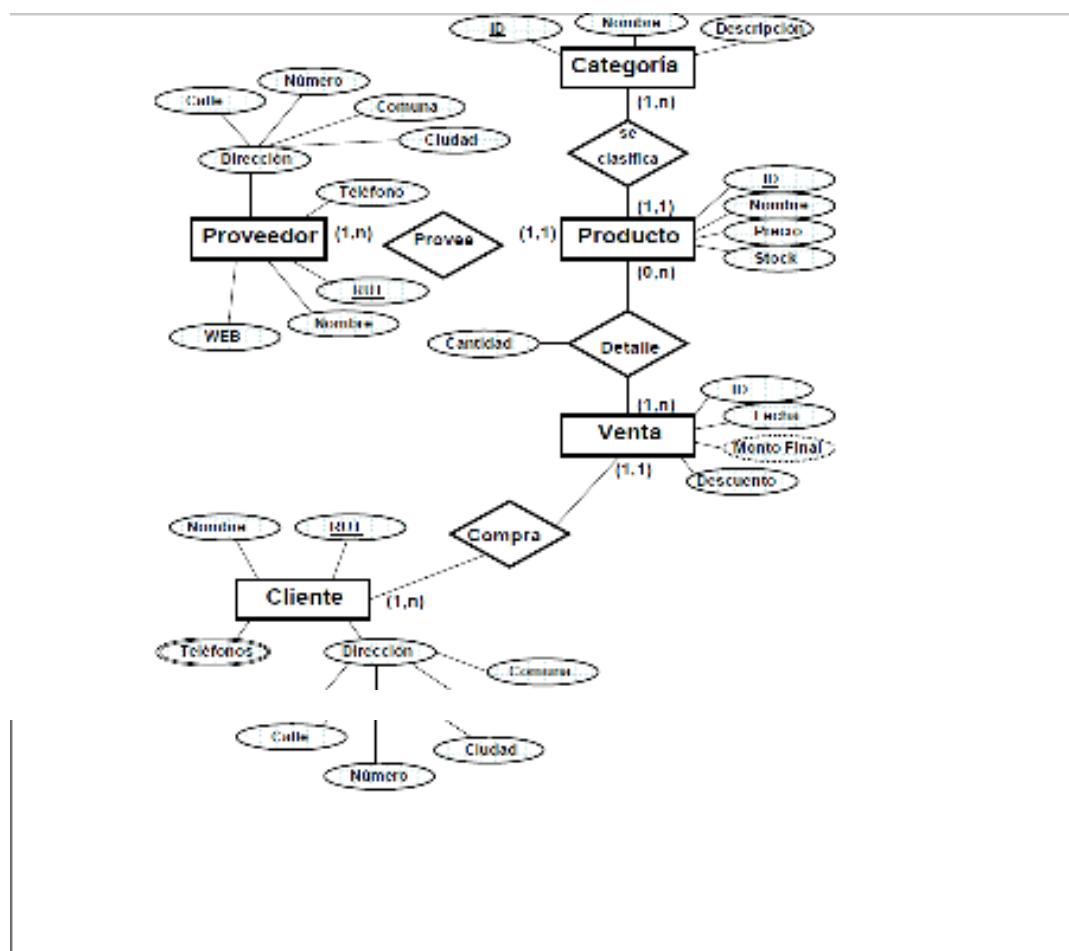
Además del resumen de la película que se incluye en la ficha interesa mostrar la opinión de las personas que vieron la película. De cada opinión se conoce el nombre de la persona que la realiza, su edad, la fecha en que registró su opinión, la calificación que le dio a la película (*Obra Maestra, Muy Buena, Buena, Regular, Mala*) y el comentario propiamente dicho. A cada opinión se le asigna un número que la identifica respecto de la película sobre la cual opina.

## SOLUCIONES A LOS EJERCICIOS (\*):

### Problema 1:



Notas: - El N° de artículos provistos es la suma de las existencias de cada artículo  
 - Se podría almacenar una fábrica de la cual no se tengan

**Problema 2.**

## Problema 5.

