

## UNIDAD DE TRABAJO 2: BASES DE DATOS RELACIONALES

### TEMA 4: EL DIAGRAMA ENTIDAD/RELACIÓN

#### 4.1.- INTRODUCCIÓN

El objetivo de este tema es conocer los elementos que podemos utilizar para realizar el modelo conceptual de datos del sistema que estamos estudiando. Se llama modelo conceptual de datos a la representación gráfica del sistema, teniendo en cuenta la información que utiliza dicho sistema y las relaciones que existen entre dicha información. Se puede realizar el modelo de datos de cualquier sistema de información, independientemente de la tecnología que se utilice para realizar los procesos de ese sistema (manual, mecanizado, etc.)

Existen muchos modelos diferentes para realizar la representación de un sistema de información, teniendo en cuenta la información que utiliza, pero todos ellos utilizan mecanismos de abstracción para representar los objetos (cosas) del mundo real y las interrelaciones que existen entre esos objetos.

La construcción del modelo es un paso previo al futuro diseño de las bases de datos en cualquiera de los modelos que seleccionemos para la representación (relacional, jerárquico o en red). Las características principales del modelo son:

- Reflejan tan sólo la existencia de datos, no lo que se va a hacer con ellos, es decir, no los tratamientos.
- No está orientado a aplicaciones particulares.
- Es independiente de las bases de datos y de los sistemas operativos concretos.
- En su confección no se tienen en cuenta restricciones de espacio, almacenamiento ni tiempo de ejecución.
- Está abierto a la evolución del sistema.

#### 4.2.- ABSTRACCIONES QUE SE UTILIZAN EN EL MODELO CONCEPTUAL DE DATOS.

Se llama abstracción a un proceso mental que permite seleccionar algunas características y propiedades de un conjunto de objetos y excluir otras que no son importantes para el tema que se está considerando, a cada abstracción se le asocia un nombre. Por ejemplo, coche es un concepto que todos tenemos y no es necesario definir sus características, además en el concepto coche no consideramos las diferencias de color, modelo etc

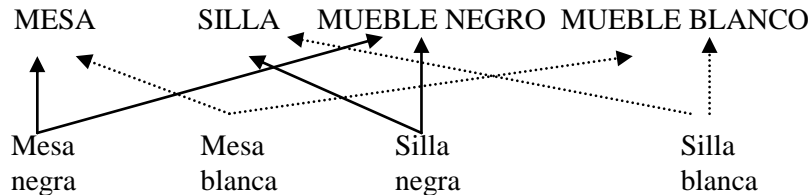
En la realización del modelo conceptual de datos se utilizan tres tipos de abstracciones

- **Abstracción de clasificación:** se utiliza para definir una **clase** como una colección de objetos que se caracterizan por tener propiedades comunes. Se representa mediante un árbol de un nivel, la raíz es la clase y los elementos son las hojas:

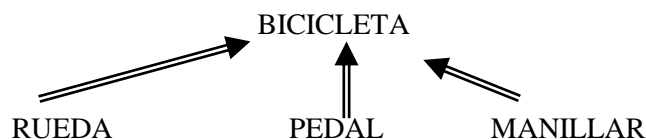


Cada rama del árbol, indica que un nodo hoja **ES-MIEMBRO\_DE** la clase que representa la raíz

Un mismo objeto puede clasificarse de varias maneras o puede pertenecer a varias clases dependiendo de las características que consideremos:

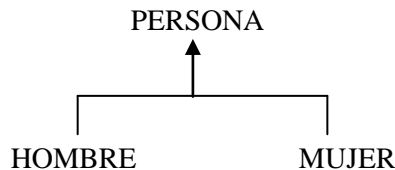


**Abstracción de agregación:** esta abstracción, define una nueva clase a partir de otras clases que son sus componentes:



Se representa mediante un árbol de un nivel en el cual la raíz y todos los nodos hojas son clases, se dice que cada clase hoja **FORMA PARTE DE** la clase raíz.

- **Abstracción de generalización:** define una relación de subconjunto entre elementos de dos o mas clases:



Una generalización se representa mediante un árbol, en el que tanto la raíz como los nodos hojas son clases. Para cada nodo hoja se dice que **ES-UN** subconjunto de la clase raíz.

Es muy importante considerar que en una generalización, todas las propiedades definidas para la clase genérica son heredadas por las clases subconjunto.

#### 4.3.- MODELO CONCEPTUAL DE DATOS: DIAGRAMA ENTIDAD INTERRELACIÓN (E/R)

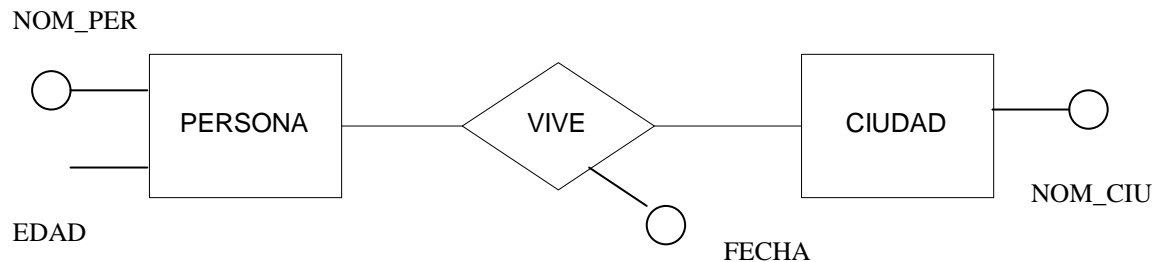
Para la representación gráfica de un sistema de información se utiliza el diagrama Entidad-Interrelación (E/R), en el que aparece la información del sistema y las interrelaciones existentes entre dicha información. El diagrama E/R básico se debe a Peter Chen, que utiliza los siguientes elementos:

**Entidad:** es la representación gráfica de objetos reales o abstractos de los que se guarda información en el sistema, PERSONA, EMPLEADO, CIUDAD. Se representan gráficamente mediante un rectángulo con su nombre dentro.



**Interrelaciones o relaciones:** son asociaciones entre dos o más entidades. Se utiliza un rombo con su nombre dentro, que debe ser significativo de la interrelación que representa. La interrelación se une a las entidades mediante una línea:

**Atributos:** cada una de las propiedades de las entidades y de las interrelaciones se denomina atributo. Los atributos se representan mediante un círculo unido a la entidad o interrelación mediante una línea recta:

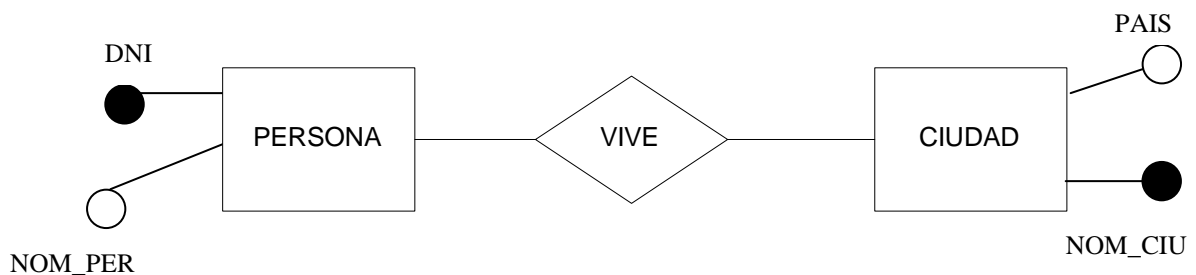


Se llama dominio de un atributo al conjunto de valores que puede tomar dicho atributo, que depende del problema que estemos considerando.

Se llama **dominio de un atributo**: conjunto de valores y reglas de formato, que puede tomar dicho atributo, dependerá del sistema que estemos considerando.

**Designer:** los detalles que se definen para un dominio incluyen el tipo de datos y tamaño del atributo y un tipo de datos y tamaño diferentes para la columna. El mismo dominio puede describir tanto atributos como columnas, pudiendo tener cada una un tipo distinto. Por ejemplo, para un dato monetario, el tipo de datos del atributo puede ser Money, pero como no existe un tipo de datos Money para una columna en una tabla de Oracle, pues el tipo de datos para dicha columna será **number**.

**Atributo superclave o identificador:** es un conjunto de atributos que identifican de forma única a cada ocurrencia de entidad. Por ejemplos, son superclaves de la entidad persona el conjunto de campos formado por: nombre + apellidos + dirección + fecha de nacimiento. Otra superclave estaría formada por: nombre + apellido + fecha de nacimiento + teléfono. Otra sería el DNI, otra en número de la S.S. etc.



**Clave candidata:** es cada una de las superclaves que además son mínimas, es decir, está formada por el mínimo número de campos posibles. Dentro de las superclaves que hemos dicho antes, serán claves candidatas el DNI o el número de la S.S.

**Clave primaria o principal:** es la clave candidata seleccionada por el diseñador de la base de datos. Las características de una clave primaria son:

- No puede contener valores nulos. Un atributo o campo contiene valores nulos cuando no hay información en dicho campo o atributo. No es lo mismo un campo con valor NULL que con un espacio en blanco.
- Ha de ser conocida y sencilla de crear.
- No ha de variar con el tiempo.

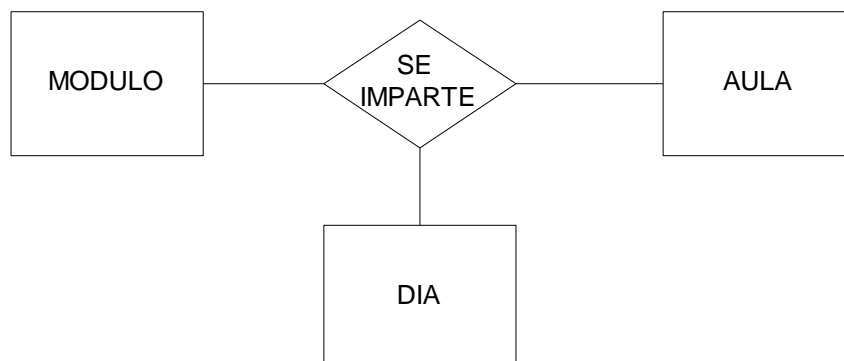
**Clave alternativa:** es cada una de las claves candidatas no seleccionadas. En nuestro ejemplo, serán claves alternativas: nombre + apellidos + dirección + fecha de nacimiento, o la formada por los campos: nombre + apellidos + fecha de nacimiento + teléfono, o si se ha elegido como clave principal el DNI pues se quedaría como clave candidata el número de la S.S.

**Clave ajena o foránea:** es un conjunto de atributos de una entidad que forman la clave primaria en otra entidad. Por ejemplo, si se ha definido en la entidad FACTURAS el campo NIF, como este campo es clave primaria en la entidad cliente, da origen a una clave ajena en la entidad FACTURAS.

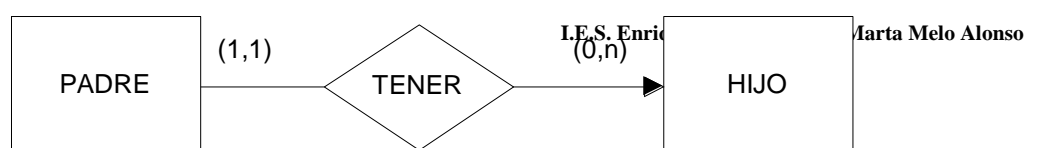
**Clave artificial:** es la clave primaria que ha sido creada independientemente del contenido de los atributos que pudieran formar la clave primaria. Sirve para crear claves primarias más sencillas que una formada por la unión de varios atributos. En su creación se ha de tener en cuenta que ha de ser única e invariable. Habitualmente, las claves artificiales poseen nombres como: id\_cli, codigo\_cli, cod\_cli, etc.

**Dominio de un atributo:** conjunto de valores permitidos y reglas de formato de dicho atributo. En Designer: los detalles que se definen para un dominio incluyen el tipo de datos y tamaño del atributo y un tipo de datos y tamaño diferentes para la columna. El mismo dominio puede describir tanto atributos como columnas, pudiendo tener cada una un tipo distinto. Por ejemplo, para un dato monetario, el tipo de datos del atributo puede ser Money (dinero), pero como no existe un tipo de datos Money para una columna de una tabla en Oracle, pues el tipo de datos para dicha columna será number.

**Grado o dimensión de una interrelación:** es el número de entidades que participan en ella: cuando asocia dos entidades se dice que la interrelación es de grado 2 (binaria), si asocia más de dos entidades se llama n-aria:



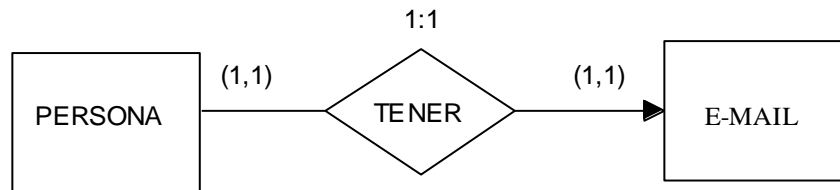
**Cardinalidad de las entidades en una relación:** en las relaciones, también se representa su cardinalidad representada por el par (cardinalidad-mínima, cardinalidad-máxima). Cuando la cardinalidad máxima es



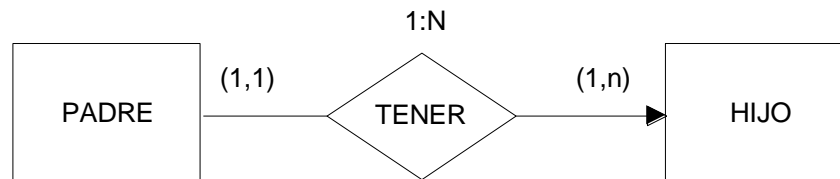
n, la línea que une la interrelación con la entidad termina en punta de flecha. La cardinalidad de una entidad, se escribe sobre la línea que une la interrelación con la otra entidad:

**Tipo de la relación:** es el número máximo de correspondencias en las que puede intervenir un elemento de cada entidad, pueden ser 1:1, 1:N o N:N:

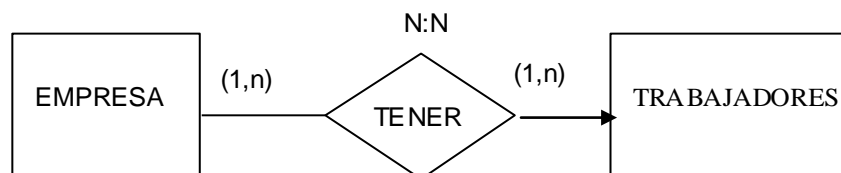
1:1 → A cada elemento de la primera entidad le corresponde no más de un elemento de la segunda entidad y viceversa:



1:N → A cada elemento de la primera entidad le pueden corresponder uno o más elementos de la segunda entidad, y a cada componente de la segunda entidad le corresponde un solo elemento de la primera, por ejemplo:



N:N → A cada elemento de la primera entidad le pueden corresponder varios elementos de la segunda y viceversa. Por ejemplo: una empresa puede tener varios trabajadores y un trabajador puede pertenecer a varias empresas:

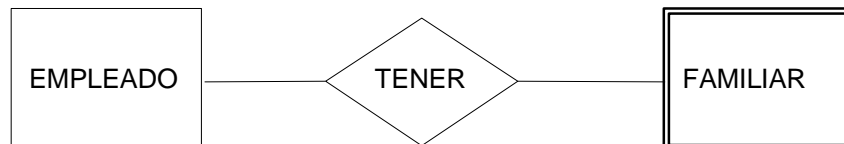


#### **4.4.- AMPLIACIÓN DEL MODELO E/R DE PETER CHEN**

Diferentes autores amplían el modelo de Peter Chen con el objetivo de poder representar en el modelo conceptual de datos otros conceptos que aparecen en el estudio de los sistemas (orientado a datos) y que no era posible representarlos mediante el modelo establecido por Peter Chen. Las nuevas aportaciones son:

##### **4.4.1.- CLASES DE ENTIDADES: REGULARES Y DÉBILES:**

Una entidad es regular cuando tiene existencia por si misma y es débil si su existencia depende de otra entidad. Por ejemplo si consideramos en una empresa la entidad EMPLEADO, estaríamos ante una entidad regular, ya que existe por si misma. Pero si también consideramos la entidad FAMILIAR de los empleados, tendríamos que la existencia de los elementos de esta entidad están condicionados a la existencia de EMPLEADO de la empresa. De igual forma si desaparece un determinado empleado, desaparecerán todos los elementos de FAMILIAR ligados a ese empleado. Por tanto la entidad FAMILIAR, sería una entidad débil y se representan con un rectángulo doble:



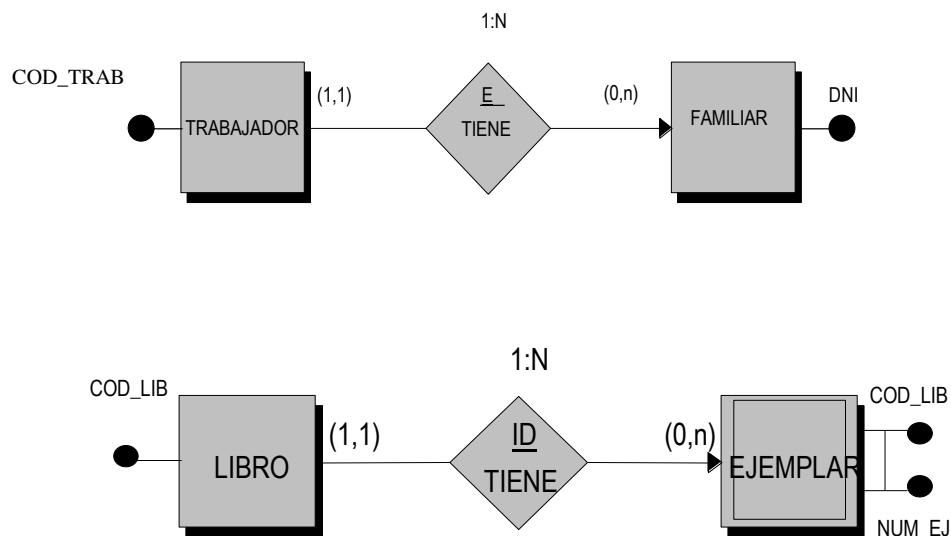
#### 4.4.2.- INTERRELACIONES REGULARES Y DÉBILES

Las interrelaciones se clasifican en **regulares** si asocian entidades regulares, y **débiles** si asocian una entidad regular con una débil.

En las interrelaciones débiles podemos distinguir la **dependencia en existencia** y la **dependencia en identificación**.

La dependencia en existencia ocurre cuando los elementos de la entidad débil no pueden existir si desaparece la entidad regular de la que dependen. Y la dependencia en identificación ocurre, si además de lo anterior la identificación de los elementos de la entidad débil no se puede realizar con sus propios atributos, sino que hay que incluir el atributo identificador principal de la entidad regular. Toda dependencia en identificación es una dependencia en existencia pero el caso contrario no tiene porque ser cierto. Por ejemplo, si para identificar los ejemplares de un libro, utilizamos el código del libro y un contador: AAAA/01, AAAA/02, AAAA/03, ...

Para señalar que existe una dependencia en existencia en el diagrama E/R, se escribe una “E” dentro del rombo de la interrelación, si la dependencia es en identificación, se escribe “ID”

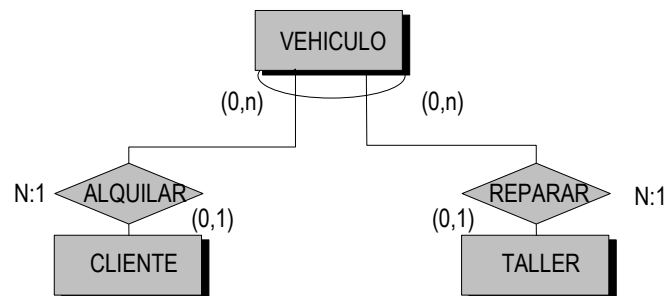


#### 4.4.3.- INTERRELACIONES EXCLUSIVAS

Una misma entidad puede participar en dos o más interrelaciones en un sistema de información. Se dice que dos o más interrelaciones son exclusivas, cuando las ocurrencias de una entidad que participa en todas esas interrelaciones, solamente puede tener correspondientes a través de una de ellas.

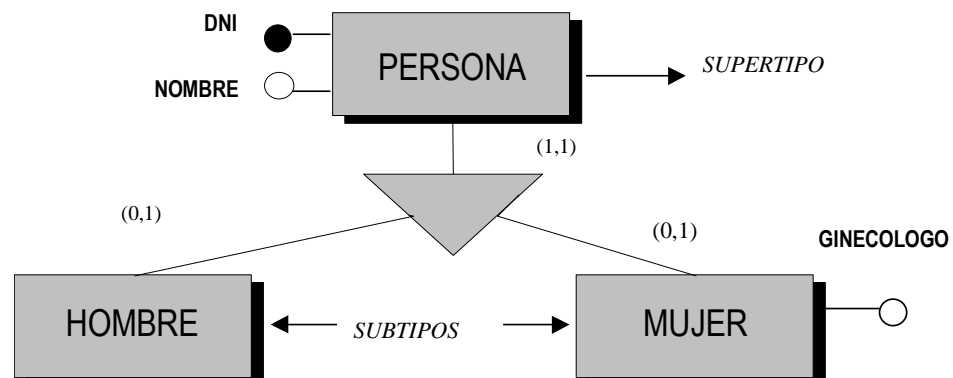
Consideramos los vehículos de una empresa de alquiler de coches y las interrelaciones ALQUILAR y REPARAR, si un vehículo se encuentra en reparación queda excluido de estar alquilado.

La entidad VEHICULO participa en dos interrelaciones, ALQUILAR y REPARAR, sin embargo un determinado coche no puede tener correspondientes a través de las dos interrelaciones. El hecho de tener asociado un cliente le excluye de estar en un taller de reparación. Se representa mediante un arco.



#### 4.4.4.- GENERALIZACIÓN Y HERENCIA

La generalización es otro concepto importante que no se puede representar mediante la simbología de Peter Chen. Entre varias entidades puede existir una interrelación de conjunto a subconjuntos (ES-UN), en los diagramas E/R la generalización se representa mediante un triángulo invertido:



La entidad raíz, recibe el nombre de **supertipo** y las entidades hojas se llaman **subtipos** de la generalización.

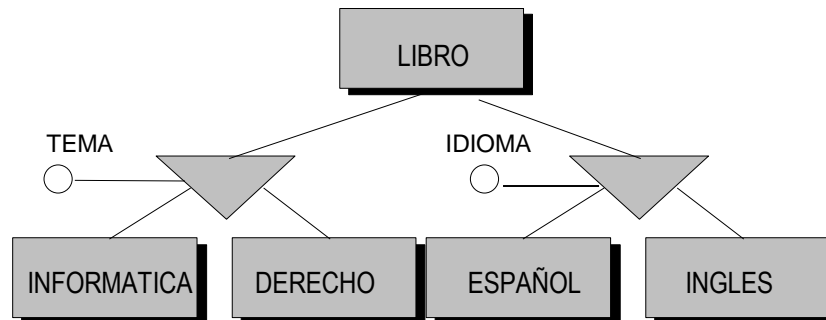
La cardinalidad en el supertipo es siempre (1,1) y las cardinalidades en los subtipos serán (0,1).

La herencia se representa mediante sus atributos (propiedades), todo atributo del supertipo, será un atributo de los subtipos y se escriben en el supertipo. Cada subtipo puede tener sus propios atributos.



De igual forma las interrelaciones que afectan a todos los subtipos, se asocian al supertipo y cada subtipo puede tener sus propias interrelaciones.

La división de una entidad en subtipos (especialización), puede venir determinada por una condición predeterminada, por ejemplo los valores que toma un determinado atributo, llamado discriminante, en este caso el atributo se asocia al triángulo que representa la generalización:



#### 4.4.5.- COBERTURA EN LAS GENERALIZACIONES

**Generalización total:** toda ocurrencia del supertipo es una ocurrencia de uno o varios subtipos. Se representa con un círculo en la línea que une el supertipo con el símbolo de la generalización.

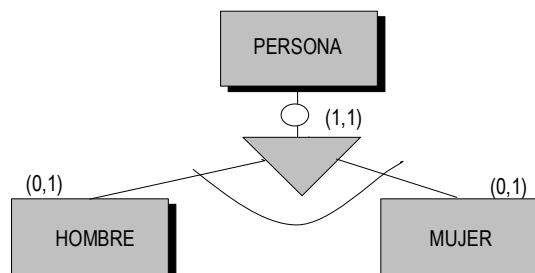
**Generalización parcial:** existen ocurrencias del supertipo que no son ocurrencias de ningún subtipo.

**Generalización exclusiva:** los subtipos no se solapan, una ocurrencia pertenece solamente a uno de ellos. Se representa mediante un arco.

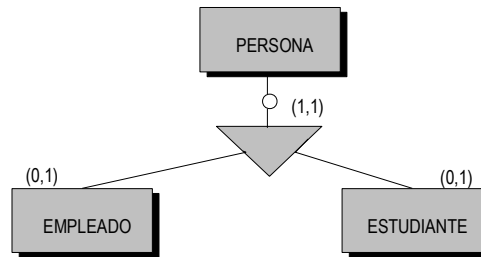
**Generalización superpuesta:** los subtipos se solapan, una ocurrencia puede pertenecer a más de un subtipo.

Estos cuatro tipos de generalizaciones se combinan: total exclusiva, total superpuesta, parcial exclusiva, parcial superpuesta.

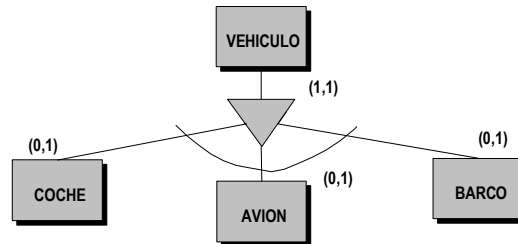
**Generalización total y exclusiva:** toda ocurrencia del supertipo es una ocurrencia de un subtipo y solo de uno:



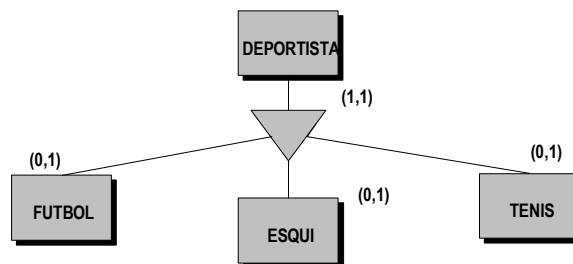
**Generalización total superpuesta:** toda ocurrencia del supertipo es una ocurrencia de uno o más subtipos.



**Generalización parcial exclusiva:** puede haber ocurrencias de la entidad supertipo que no pertenezcan a ningún subtipo, pero si también es una ocurrencia de un subtipo lo es, solo de uno.



**Generalización parcial superpuesta:** no todas las ocurrencias de la entidad supertipo pertenecen a un subtipo, pero algunas pueden pertenecer a más de uno.



#### 4.5.- CONSTRUCCION DE UN ESQUEMA E/R

El modelado de un sistema de información o de una realidad está abierto a diferentes soluciones, no existen unas reglas que nos determinen qué es una entidad, qué es un atributo, etc. Pero podemos enunciar unos principios que nos pueden guiar a elegir los elementos de un diagrama Entidad/Interrelación:

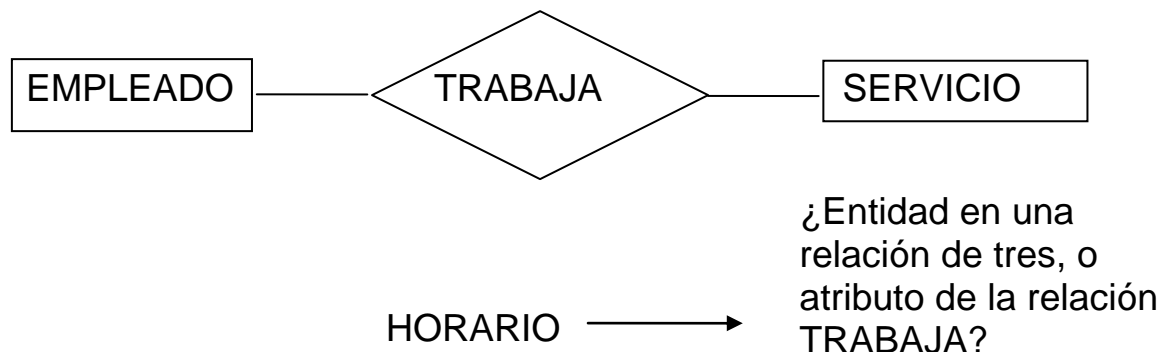
- . Un sustantivo que actúa como sujeto o complemento directo de una frase, en general es una entidad, aunque también podría ser un atributo.

- . Los nombres propios suelen indicar ocurrencias de una entidad.

- . Un verbo es una interrelación.

- . Una preposición suele ser una interrelación, o puede ser la asociación entre una entidad y sus atributos.

Cuando se tiene dudas entre definir un sustantivo como un atributo de una relación o una entidad que participa en una relación de tres, hay que estudiar si al tomarlo como atributo de la relación, éste puede tomar varios valores para una ocurrencia de la relación, en este caso tenemos que establecer una relación de tres.



Si me da más de un horario para un empleado+servicio entonces será una relación.

En cuanto a la estrategia que se puede seguir para realizar un diagrama Entidad/Interrelación, existen varias:

Estrategia descendente: se trabaja como en la descomposición en niveles en los DFD, se parte de una entidad única que representa todo el problema y se van introduciendo los distintos elementos que forman nuestro sistema, en todo momento consideramos el problema de forma global.

Estrategia ascendente: se parte del nivel más bajo, los atributos que se van uniendo para formar las entidades y luego se añaden las interrelaciones entre esas entidades hasta completar el diagrama.

Estrategia de la mancha de aceite: es un método muy utilizado, se comienza por una parte del problema que nos parece más fácil de abordar y lo vamos extendiendo hasta completar el esquema.

#### **4.6.- CUALIDADES DE UN DIAGRAMA ENTIDAD/INTERRELACIÓN**

Una vez que se tiene el diagrama E/R, se deben comprobar las siguientes cualidades:

1. - Comprobar que es completo: quiere decir que está representado todo el sistema. Para ello se miran si todos los requerimientos están representados y se comprueba que todo lo representado forma parte de los requerimientos.

2.- Comprobar que es correcto: es decir que se han utilizado con propiedad los conceptos del modelo E/R. Los errores más comunes son:

- . Utilizar un atributo en vez de una entidad.
- . Olvidar alguna generalización.
- . No representar la herencia en las generalizaciones.
- . Utilizar una interrelación con un número erróneo de entidades (una binaria en lugar de un ternaria).
- . Utilizar una entidad en vez de una interrelación.
- . Olvidar el atributo identificador principal en alguna entidad.
- . Omitir alguna especificación de cardinalidad.

3.- Comprobar que el diagrama obtenido es mínimo: y debemos detectar las posibles redundancias. Esto quiere decir que hemos representado cada requerimiento una sola vez y por tanto no se puede borrar ningún elemento sin perder información. Las redundancias pueden aparecer en los siguientes elementos:

- . Atributos derivados: son los que se pueden obtener a partir de otros atributos ya representados, no es necesario quitarlos pero sí controlarlos.
- . Ciclos de interrelaciones: la redundancia existe cuando una interrelación entre dos entidades posee el mismo significado que una ruta de interrelaciones. No todos los ciclos implican redundancia, pero debe estudiarse con detenimiento.

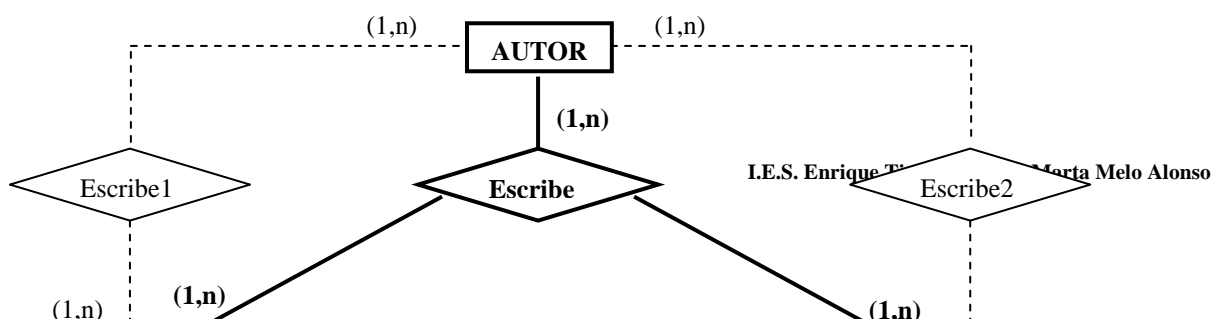
4. - Comprobar su expresividad: los requerimientos deben aparecer expresados de forma que las explicaciones adicionales sean mínimas.

5.- El diagrama debe ser legible: no pueden aparecer líneas cruzadas y a ser posible debe presentar simetría.

#### **4.7 RELACIONES DE GRADO n**

Existen relaciones que asocian más de dos entidades (grado n). Cuando se presenta un tipo de relación de grado n, hay que tener en cuenta que a veces no es propiamente de tal grado, ya que puede descomponerse en varios tipos de relación que asocien entidades dos a dos, es decir, en varias relaciones de grado 2. Sin embargo, otras veces no es posible tal descomposición, ya que la semántica recogida en una y otra solución no es la misma.

Por ejemplo:



(1,n)

Vemos que Escribe asocia tres entidades a la vez: un autor con el tema acerca del cual escribió en un determinado libro (un libro puede estar escrito por varios autores y cada uno puede tratar en él temas distintos).

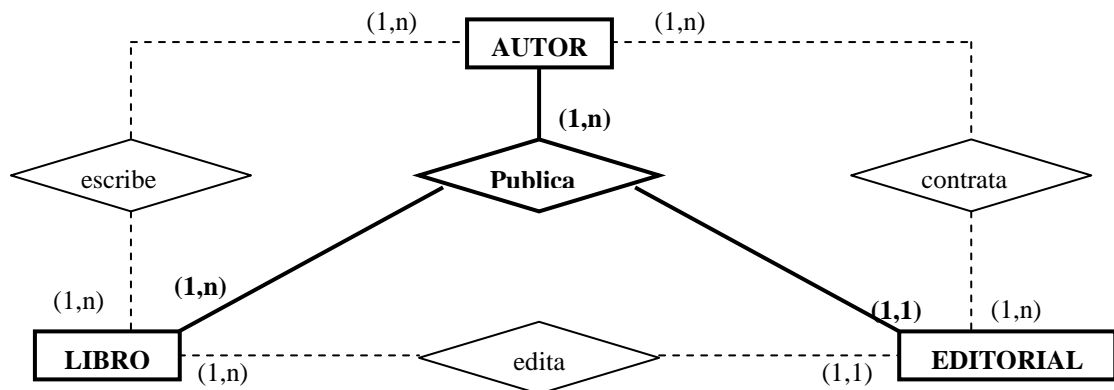
Ej:      Cod\_autor                      Tema                      Libro  
             01                                08                      10

Si sustituimos la relación escribe por las tres: escribe1, escribe2 y trata, no se puede deducir qué tres ocurrencias determinadas de las distintas entidades estaban asociadas en una ocurrencia de relación concreta, por lo que no es posible la descomposición de la relación grado 3 en 3 de grado 2 sin pérdida de semántica.

Ej:      Cod\_autor                      Tema                      Libro                      → muchos  
             01                                08                      ?                      → muchos

             Cod\_autor                      Libro                      Tema                      → muchos  
             01                                10                      ?                      → muchos

Sin embargo, en el siguiente ejemplo:



Ej:      Cod\_autor                      Cod\_libro                      Cod\_edit  
             01                                10                                13

¿Podríamos descomponerla en las relaciones edita, escribe y contrata?, sí, ya que aportarían la misma semántica que la interrelación de grado 3.

Ej:      Cod\_autor                      Cod\_libro                      Cod\_edit

01

10

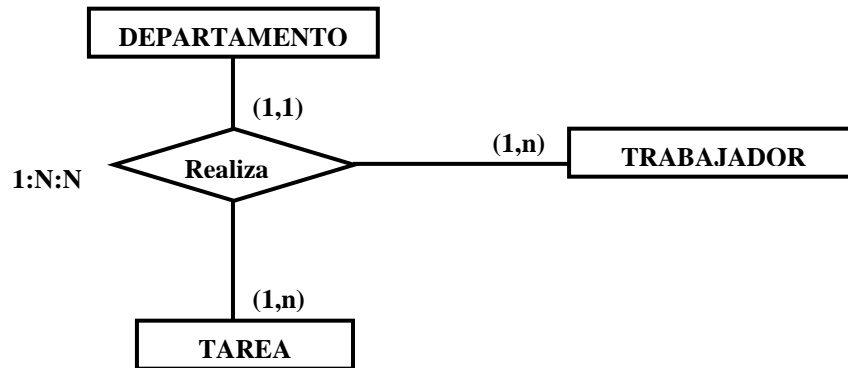
13

→ sí sabemos la editorial.

### **10.7.1 Cardinalidades en una relación n-aria**

Para obtener la cardinalidad de una relación n-aria se cogen las participaciones máximas de cada entidad. Para obtener la participación de una entidad, se fija una ocurrencia del resto de entidades y se observan cuántas ocurrencias, en la entidad en cuestión, le corresponden.

Ej: Se desea guardar las diferentes tareas que realiza un trabajador en cada departamento:



Para la entidad Tarea: dado un departamento y un trabajador ¿cuántas tareas realiza?. Pues n.

Para la entidad Trabajador: dado un departamento y una tarea ¿cuántos trabajadores la realizan? También n.

Para la entidad Departamento: dado un trabajador y una tarea ¿en cuántos departamentos la realiza?, en 1.

#### 4.8.- CONTROL DE REDUNDANCIA

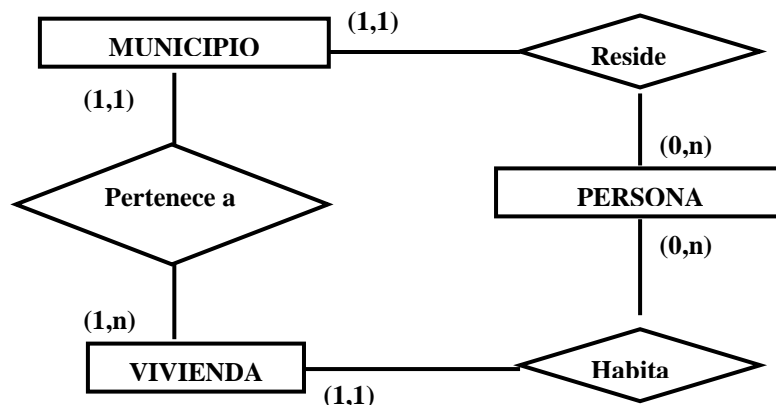
Un esquema es redundante cuando al eliminar un elemento del mismo no se pierde información, se pueden deducir los datos de dicho elemento a partir de los demás.

Condiciones que se deben cumplir para que haya redundancia:

- Debe haber un ciclo en el esquema.
- Las relaciones implicadas en el ciclo deben ser semánticamente equivalentes.
- Las cardinalidades deben ser tales que se pueda eliminar una de las relaciones sin perder información (cardinalidad mínima: 1 en las que dejemos)

Dicho ciclo debe desaparecer, para ello debemos eliminar una de las relaciones que está generando el ciclo.

Ejemplo: Una persona reside un uno y sólo un municipio, en un municipio puede no residir nadie o varias personas. Una persona habita en una y sólo una vivienda, pero en una vivienda puede no habitar nadie o varias personas. Una vivienda pertenece a un único municipio, y en un municipio puede haber una o más viviendas.



Se trata de un ciclo redundante porque cumple las tres condiciones:

- Es un ciclo
- Las relaciones reside y habita son semánticamente equivalentes
- Una persona habita en un única vivienda y esa vivienda se encuentra en un único municipio (cardinalidad mínima: 1)

Si suprimimos la relación reside, podemos saber en qué municipio reside una persona a través de la vivienda.