

TEMA 4: EL MODELO ENTIDAD/INTERRELACIÓN

1.- INTRODUCCIÓN

El modelo E/R nace en el año 1976 de la mano de Peter Chen. Dentro del mundo de los modelos de datos es un caso especial debido a que se trata de un modelo *conceptual* en el sentido de que nos permite expresar la semántica de los datos y sus relaciones desde un punto de vista tan ideal que no permite ser implementado, y *estructural* en el sentido de que sólo nos proporciona estructuras y no operaciones sobre ellas. Podemos decir por lo tanto que es un modelo cuyo fin es permitirnos expresar la percepción que tenemos de los datos. ¿Por qué es entonces tan utilizado este modelo? Principalmente por las siguientes razones:

1. Utiliza una notación muy sencilla, fácilmente comprensible por los usuarios y que, por lo tanto, permitirá contrastar si la idea que percibimos de los datos manejados por el sistema que estudiamos es fiel a la idea que los usuarios de éste han intentado transmitirnos.
2. Existe una conversión automática de diseños de bases de datos en el modelo E/R a diseños de bases de datos en el modelo Relacional, que sí que es un modelo implementable.

El producto que obtenemos cuando realizamos un diseño de una base de datos según las directrices del modelo E/R es un **Diagrama E/R**. A continuación estudiaremos cuáles son los componentes básicos de estos diagramas.

2.- ELEMENTOS DEL MODELO ENTIDAD/INTERRELACIÓN

Las estructuras del modelo E/R son conocidas como Diagramas Entidad/Interrelación. Los elementos del modelo E/R son las entidades, atributos e interrelaciones.

2.1.- ENTIDADES

Las **entidades** son objetos del mundo real acerca de los cuales se necesita almacenar información. Podemos representar mediante entidades objetos físicos (alumno, profesor...), objetos organizativos (proyecto, departamento...), sucesos (examen, matrícula, viaje...), etc...

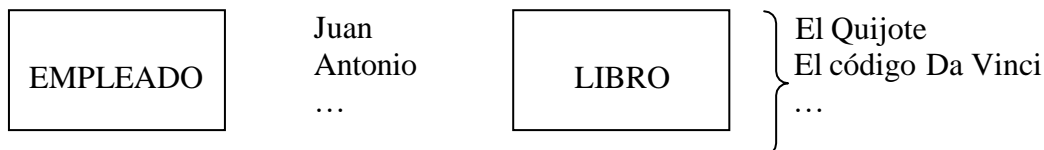
Asociado al concepto de entidad, surge el concepto de ocurrencia de una entidad, que es la realización concreta de una entidad.

Cada entidad debe cumplir las siguientes reglas:

- Tiene que tener existencia propia.
- Cada ocurrencia de un tipo de entidad debe poder distinguirse de las demás.
- Todas las ocurrencias de un tipo de entidad deben tener los mismos tipos de características (atributos).

La representación gráfica de un tipo de entidad es un rectángulo etiquetado con el nombre en mayúsculas del tipo de entidad.

Ej:



Existen dos clases de entidades:

- **Fuertes o regulares:** Las ocurrencias de un tipo de entidad regular tienen existencia propia, es decir, existen por sí mismas.
- **Débiles:** La existencia de cada ocurrencia de un tipo de entidad depende de la existencia de la ocurrencia del tipo de entidad fuerte del cual aquella depende, es decir, si se elimina una ocurrencia del tipo de entidad regular, desaparecerá también con ella todas las ocurrencias de la entidad débil dependientes de la misma. Una entidad débil se representa con un rectángulo con una doble línea y su nombre en el interior.



2.2.- ATRIBUTOS

Un atributo es una propiedad o característica asociada a una determinada entidad y, por tanto, común a todas las ocurrencias de esa entidad. Asociado al concepto de atributo surge el concepto de dominio. Un **dominio** es el conjunto de valores permitidos para un atributo. Por ejemplo, si se tiene el atributo COLOR, el dominio sobre el que se define podría ser: (naranja, blanco, azul y negro).

En función de sus características respecto a la entidad que definen se distinguen dos tipos de atributos:

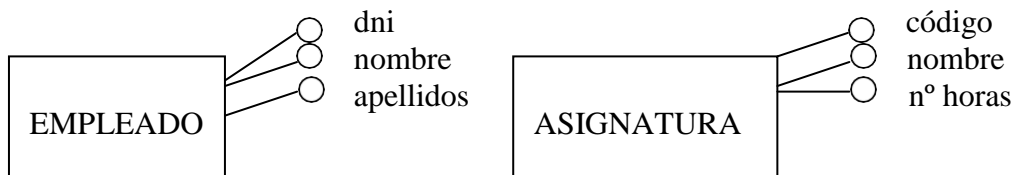
- a) **Atributo Identificador Principal (AIP):** Distingue unívocamente una ocurrencia de entidad del resto de ocurrencias.
- b) **Atributo Descriptor:** Caracteriza una ocurrencia pero no la distingue del resto de ocurrencias de entidad.

De entre todos los atributos de un tipo de entidad, hay que elegir uno o varios que identifiquen unívocamente cada una de las ocurrencias de ese tipo de entidad. Este atributo o conjunto de atributos se denomina **ATRIBUTO IDENTIFICADOR PRINCIPAL (AIP)**, y los atributos que lo componen deben ser mínimos en el sentido de que la eliminación de cualquiera de ellos le hará perder su carácter de identificador.

Puede ocurrir que exista más de un conjunto de atributos que verifiquen la condición de ser identificador unívoco y mínimo de cada ocurrencia del tipo de entidad, por lo que se elegirá a uno de ellos como AIP y el resto serán **ATRIBUTOS IDENTIFICADORES ALTERNATIVOS (AIA)**.

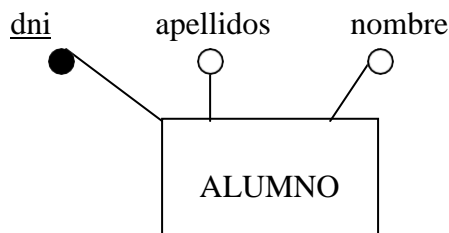
Para representar los atributos de una entidad en un diagrama E/R lo hacemos poniendo sus nombres en minúscula y uniéndolos a la entidad a la que pertenecen mediante una línea acabada en un círculo.

Ej:



El atributo identificador se marca en el diagrama E/R rellenando su círculo correspondiente.

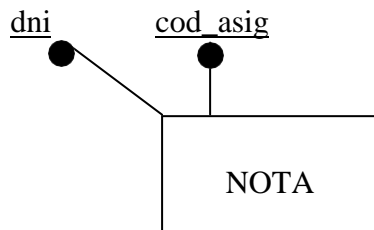
Ej:



Hay dos tipos de atributos identificadores:

- **Simples:** cuando están compuestos por un único atributo (ejemplo anterior).
- **Compuestos:** son aquellos que se componen de varios atributos.

Ej:

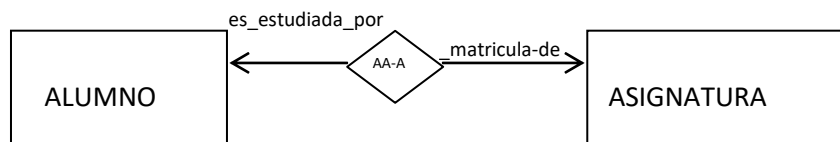


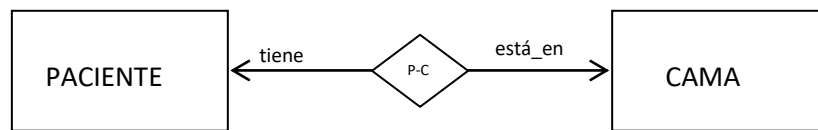
2.3.- INTERRELACIONES

Una interrelación es una asociación o correspondencia entre entidades. Llamamos Tipo de interrelación a la estructura genérica del conjunto de interrelaciones existentes entre dos o más tipos de entidad, mientras que la ocurrencia de una interrelación será la vinculación existente entre las ocurrencias concretas de cada uno de los tipos de entidad que intervienen en la interrelación.

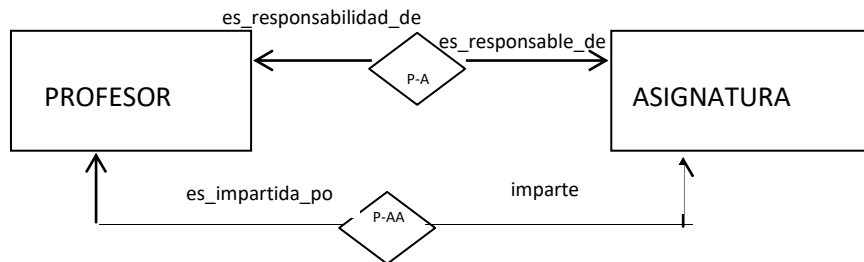
Los tipos de interrelación se representan mediante un rombo etiquetado con una inicial para cada una de las entidades, y dos nombres, para cada uno de los arcos que señalan a las entidades. Estos nombres deben ser verbos o frases que contengan un verbo y que expresen el significado de la interrelación.

Ej:





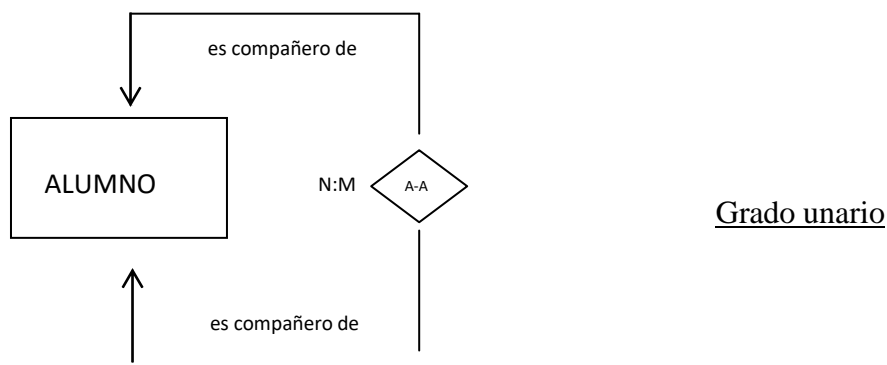
Entre dos entidades puede haber más de una interrelación:



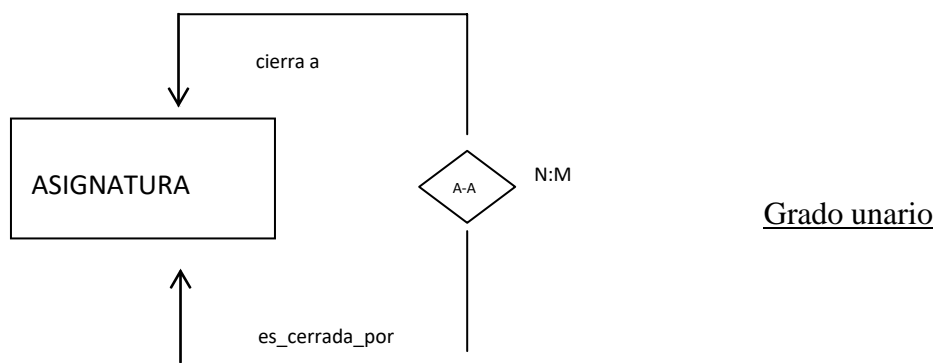
Los tipos de interrelación se caracterizan por tres propiedades:

- Nombre:** Identifica de forma única el tipo de interrelación.
- Grado:** Número de tipos de entidad que participan en un tipo de interrelación. Podemos hablar de grado unario (interrelación reflexiva), binario, ternario...

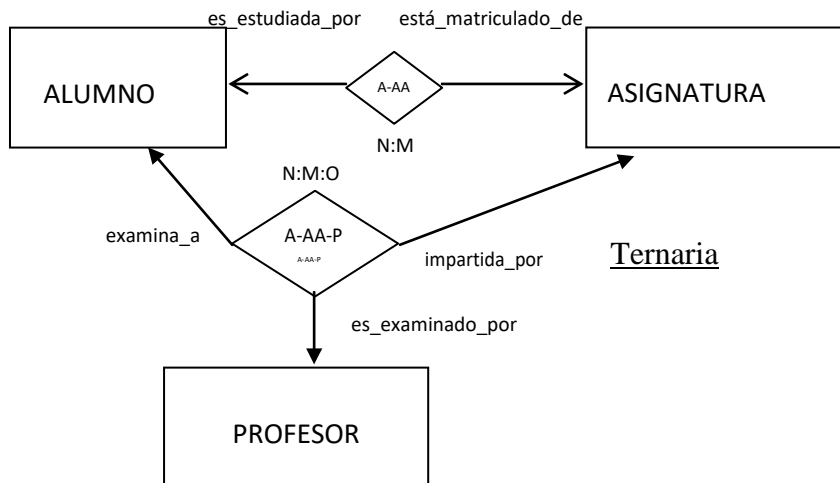
Ej:



Ej:



Ej:

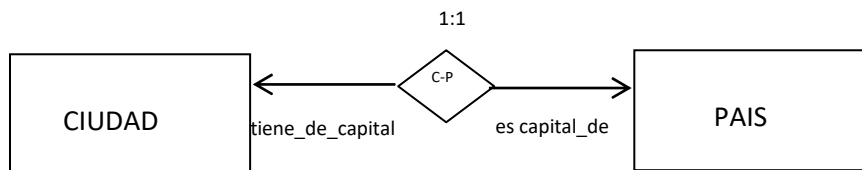


Las interrelaciones que más se suelen utilizar son las de grado binario (intervienen dos entidades).

- c) **Tipo de correspondencia:** número de instancias (ocurrencias) de una entidad que se relacionan con las instancias de otra entidad. Puede ser:

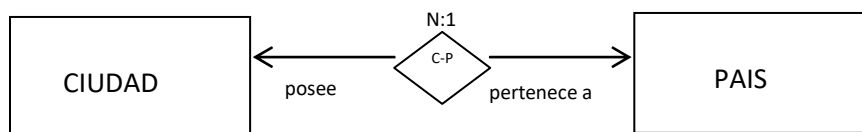
- **Uno a uno (1:1):** Una ocurrencia de la entidad A se asocia como máximo con una única ocurrencia de la entidad B y viceversa.

Ej:



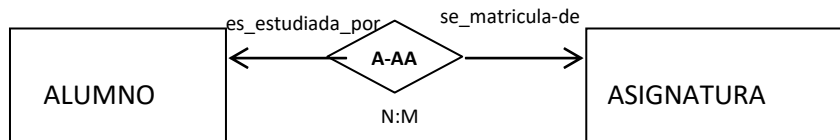
- **Uno a muchos (1:N):** Una ocurrencia de la entidad A se asocia con un número indeterminado de ocurrencias de la entidad B, pero una ocurrencia de la entidad B se asocia como máximo con una ocurrencia de A.

Ej:



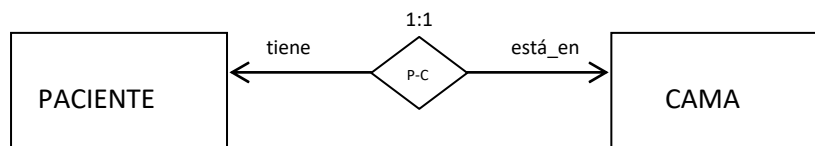
- **Muchos a muchos (M:N):** Una ocurrencia de la entidad A se asocia un número indeterminado de ocurrencias de la entidad B y viceversa.

Ej:

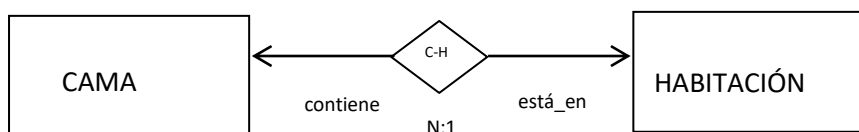


Vamos a ver ahora un ejemplo de los tres tipos de conectividad en el marco de un hospital:

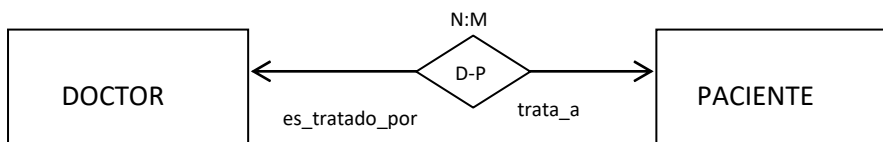
- a. Nos han surgido dos entidades, una que contiene información acerca de los pacientes que se encuentran en cada instante en el hospital, y otra que recoge la información de todas las camas del hospital. Cada paciente ocupa una cama en el hospital con la restricción de que no puede haber más de un paciente en cada cama, por lo que cada instancia paciente se asocia con una y sólo una instancia cama, y cada instancia cama se asocia con una y sólo una instancia paciente. Tenemos una interrelación 1:1 entre las entidades PACIENTE y CAMA.



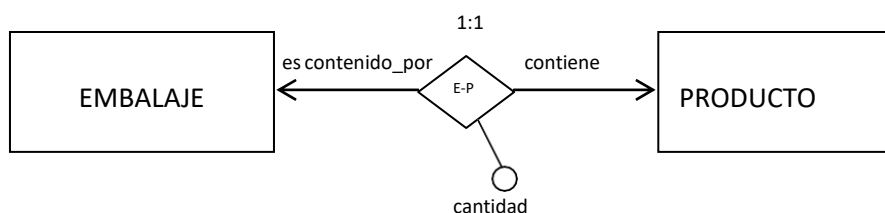
- b. Imaginemos ahora que tenemos una nueva entidad llamada HABITACIÓN, que recoge la información de todas las habitaciones del hospital. Cada cama está situada en una habitación, pero en cada habitación puede haber más de una cama. En este caso, cada instancia cama se asocia con una sola instancia habitación, y cada instancia habitación puede asociarse con varias instancias cama. Decimos entonces que entre HABITACIÓN y CAMAS existe una asociación 1:N.



- c. Por último, supongamos que detectamos una nueva entidad DOCTOR, que recoge la información de todos los médicos del hospital. Cada doctor puede estar tratando a varios enfermos, y cada enfermo puede estar siendo tratado por varios doctores. En este caso, cada instancia doctor puede estar asociada con varias instancias pacientes, y cada instancia paciente puede estar asociada con varias instancias doctor.



Las relaciones pueden tener atributos propios.



3.- DEPENDENCIA EN EXISTENCIA Y DEPENDENCIA EN IDENTIFICACIÓN

El concepto de entidad débil está directamente relacionado con las restricciones de tipo semántico del modelo E/R y, más concretamente, con la denominada restricción de existencia. Esta restricción establece el hecho de que la existencia de una entidad no tiene sentido sin la existencia de otra, es decir, una entidad tiene dependencia de existencia de otra cuando sin la primera la segunda carecería de sentido.

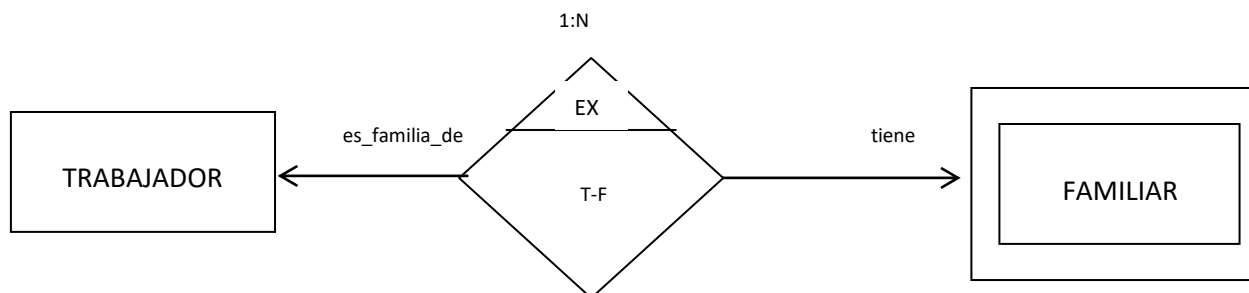
La pregunta correcta para saber si una entidad tiene dependencia de existencia respecto a otra sería la siguiente: *¿Se debe borrar alguna ocurrencia de la entidad A si se borra una ocurrencia de la entidad B?*

Si la respuesta es afirmativa la entidad tiene dependencia de existencia, por lo que se le denomina entidades débiles.

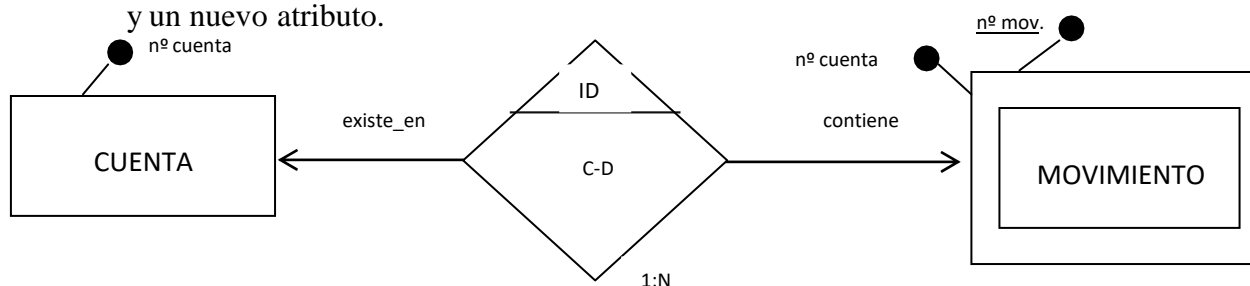
Los tipos de interrelación se clasifican, según el tipo de entidades que vinculan, en **regulares** si asocian tipos de entidad fuertes y **débiles** si asocian un tipo de entidad débil con un tipo de entidad regular. Un tipo de interrelación débil exige siempre que las cardinalidades del tipo de entidad fuerte sean (1,1). Entre una entidad fuerte y su entidad débil **siempre hay una relación 1:N**

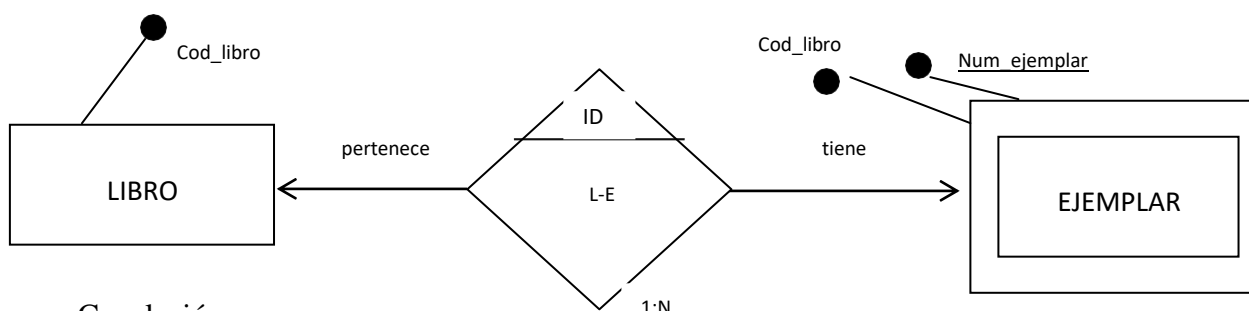
Dentro de los de interrelación débil se pueden distinguir:

1. **Dependencia en existencia:** Se dice que hay una dependencia en existencia cuando una interrelación está vinculando un tipo de entidad fuerte con uno débil, de forma que las ocurrencia del tipo de entidad débil no pueden existir sin la ocurrencia de la entidad fuerte de la que dependen. Si desaparece una ocurrencia de un tipo de entidad fuerte, todas las ocurrencias de la entidad débil que dependen en existencia de la misma desaparecen con ella.



2. **Dependencia por identificación:** Hay una dependencia en identificación cuando, además de la dependencia en existencia, las ocurrencias del tipo de entidad débil no se pueden identificar sólo mediante sus propios atributos, sino que se tiene que añadir la clave de la ocurrencia de la entidad fuerte de la cual dependen. El atributo identificador de una entidad débil por identificación va a estar formado por el atributo identificador de la entidad fuerte de la que depende y un nuevo atributo.





Conclusión:

1. La dependencia en existencia no implica una dependencia en identificación, hecho que si sucede en el caso inverso, pues una entidad que depende de otra por su AIP no tendrá sentido sin la existencia de esta última.
2. En una interrelación con cardinalidad N:M nunca habrá entidades débiles. La razón es que la supuesta ocurrencia de la entidad débil que se tuviera que borrar podría estar asociada a más de una ocurrencia de la supuesta entidad fuerte, lo que implicaría la imposibilidad de su borrado.

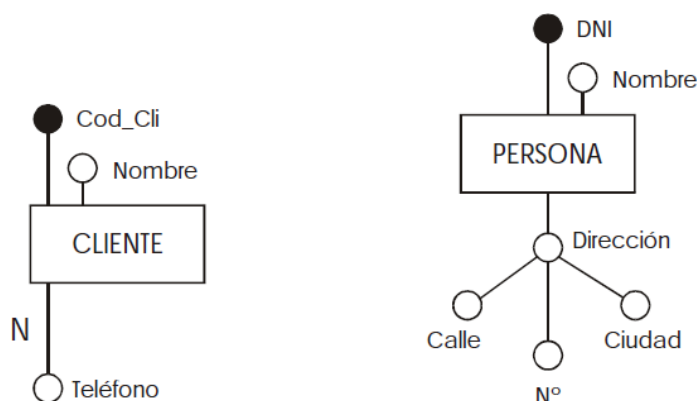
4.- ATRIBUTOS MULTIOCURRENTES Y COMPUESTOS

Existe una restricción que en el modelo E/R que afecta a la tipología de los diferentes atributos. Se pueden distinguir dos tipos diferentes de atributos:

- a) **Atributos multiocurrentes o multivaluados:** Son aquellos atributos que para una misma ocurrencia de la entidad toman más de un valor. Por ejemplo, si cada cliente puede tener más de un teléfono y es de interés guardar todos sus posibles valores, el atributo teléfono sería multiocurrente.
- b) **Atributos compuestos:** Son aquellos que agrupan en sí mismos, por afinidad o por forma de uso, más de un atributo, por ejemplo:
 - Por su forma habitual de utilización, el atributo “dirección” engloba los atributos calle, número, ciudad, provincia y código postal.
 - Por su significado, el atributo “nombre” de una entidad PERSONAS engloba nombre de pila, primer apellido y segundo apellido.

En el Diagrama E/R estos dos conceptos se reflejan como sigue:

- a) Si un atributo es multiocurrente o multivaluado se etiquetará su arco con un valor de cardinalidad N.
- b) Si un atributo es compuesto, se especificarán sus atributos componentes rodeando al mismo y enlazándolos al símbolo del atributo compuesto mediante arcos.



5.- MODELO ENTIDAD/INTERRELACIÓN EXTENDIDO(E/RE)

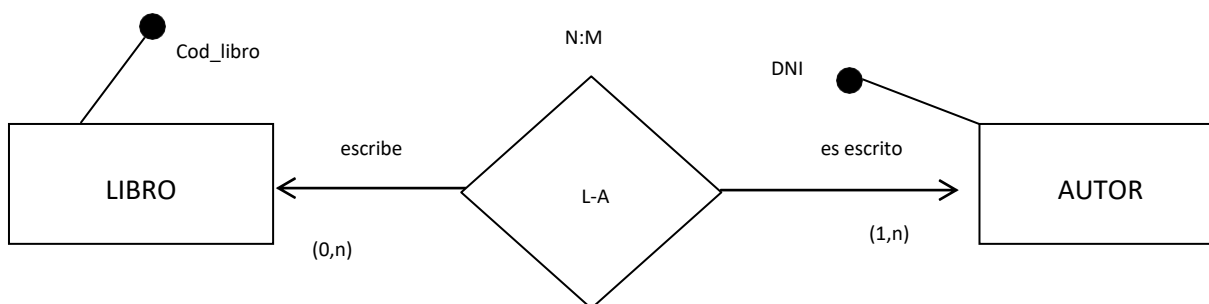
El modelo E/R con el paso del tiempo ha sufrido una serie de modificaciones tanto en su simbolismo gráfico, como en la ampliación de sus elementos.

5.1.- CARDINALIDAD

La cardinalidad permite especificar si todas las ocurrencias de una entidad participan o no en la interrelación establecida con otra u otras entidades:

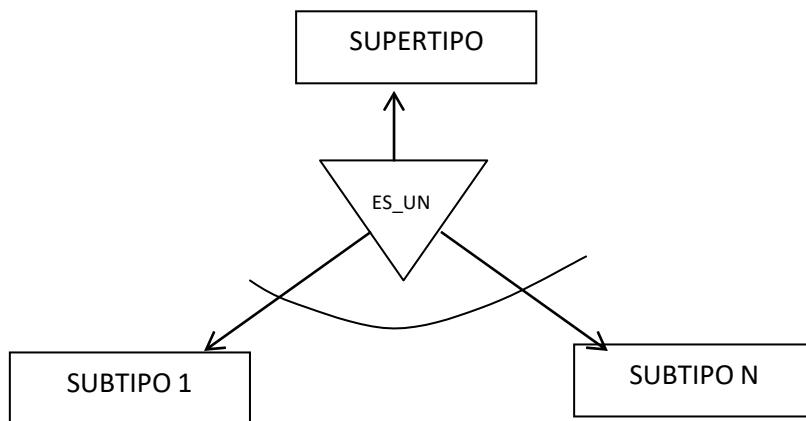
- Si toda ocurrencia de la entidad A debe estar asociada con al menos una ocurrencia de la entidad B a la que está asociada por una determinada interrelación, se dice que la clase de pertenencia es obligatoria, es decir, la cardinalidad mínima es 1.
- Si no toda ocurrencia de la entidad A necesita estar asociada con alguna ocurrencia de la entidad B asociada, se dice que la clase de pertenencia es opcional, es decir, la cardinalidad mínima es 0.

Se puede definir la cardinalidad de un tipo de entidad como el número mínimo y máximo de ocurrencia de un tipo de entidad que pueden estar relacionadas con una ocurrencia del otro, u otros tipos de entidad que participante en el tipo de interrelación. Su representación gráfica es una etiqueta del tipo (0,1), (1,1), (0,n) ó (1,n) según corresponda.

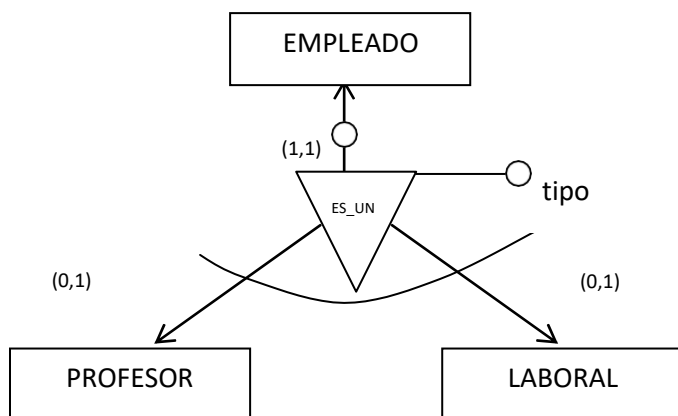


5.2.- GENERALIZACIÓN Y HERENCIA

La **generalización** es un mecanismo de abstracción mediante el cual varias entidades que tienen atributos en común se unen para formar otra entidad, llamada **entidad general, genérica o supertipo**. En concreto expresan una relación de inclusión entre entidades.



Ej:



Todas las instancias de las entidades subtipo se encuentran también en la entidad supertipo. Se supone que partimos de los subtipos y que hacemos una generalización de éstos en el supertipo.

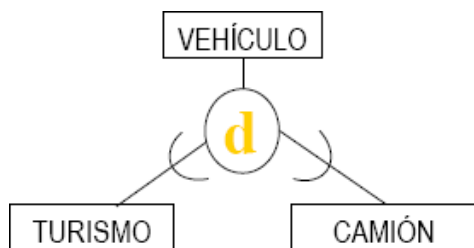
Características:

- ✓ Toda ocurrencia de un subtipo es una ocurrencia del supertipo, las cardinalidades serán siempre (1,1) en el supertipo y (0,1) ó (1,1) en los subtipos.
- ✓ Todo atributo del supertipo pasa a ser un atributo de los subtipos.
- ✓ En este tipo de abstracción, los atributos comunes a todos los subtipos se asignan al supertipo, mientras que los atributos específicos se asocian al subtipo correspondiente.

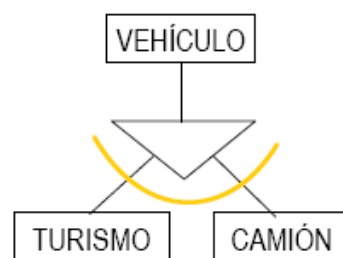
La **especialización** consiste en a partir de una determinada entidad añadir una serie de atributos que se necesitan y entonces obtener varias entidades nuevas llamadas **entidades específicas** o **subtipos**.

Las generalizaciones pueden ser de dos tipos. Se distingue el tipo respondiendo a la pregunta *¿A cuántas subclases pueden pertenecer (a la vez) una instancia de la superclase?*

- a) **Disjuntas o simples.** Ej: las instancias de EMPLEADO estarán en PROFESOR o en LABORAL, pero no en las dos.

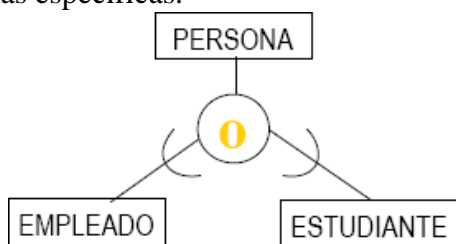


[EN2002]

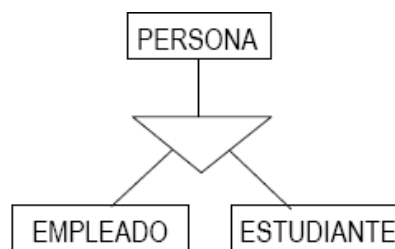


[MPM1999]

- b) **Con solapamiento:** una instancia de la entidad genérica puede estar en varias de las específicas.



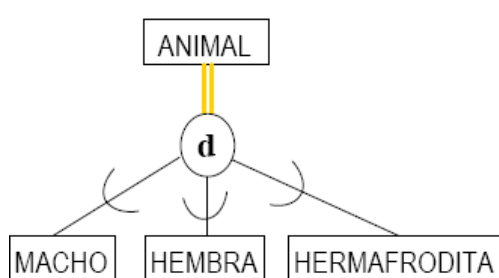
[EN2002]



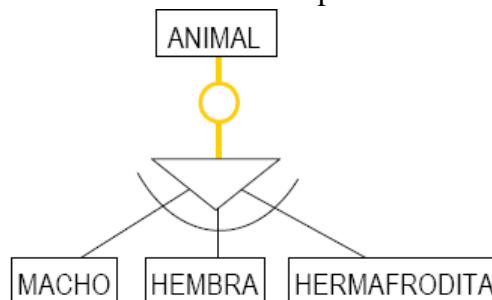
[MPM1999]

La especialización de un tipo de entidad en un conjunto de subtipos puede ser: Se distingue el tipo respondiendo a la pregunta *¿debe TODA instancia de la superclase pertenecer a alguna subclase?*

- a) **Total:** todas las instancias de la entidad pertenecen a alguno de los subtipos, no existiendo ninguna que no pertenezca a al menos uno de los subtipos.

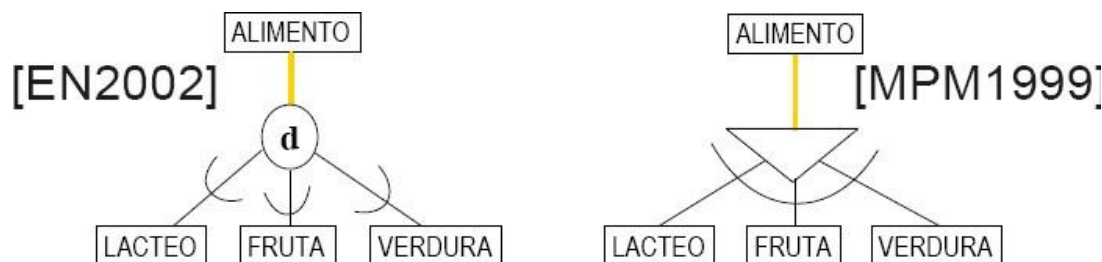


[EN2002]



[MPM1999]

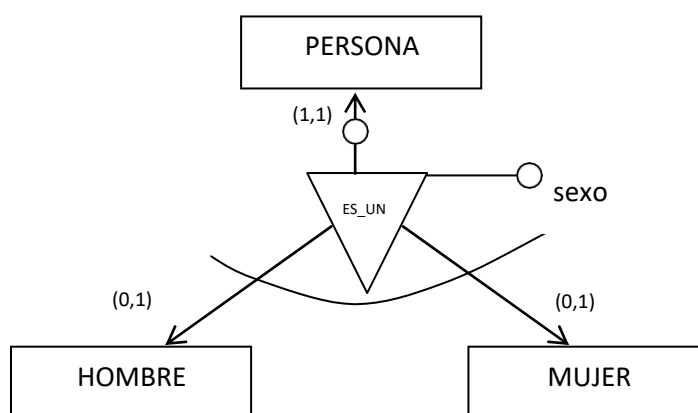
- b) Parcial:** puede haber instancias de la entidad que no pertenezcan a ninguno de los subtipos.



Por lo tanto, se pueden representar cuatro tipos de interrelaciones jerárquicas en el modelo EE-R:

a) Total sin solapamiento:

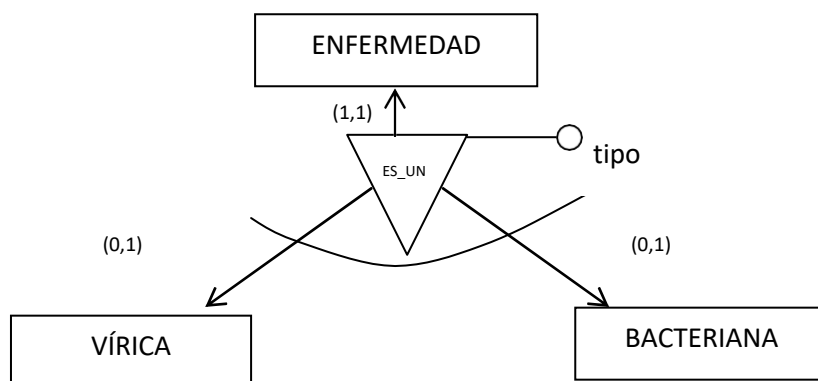
Ej:



- Una persona tiene que pertenecer necesariamente al subtipo HOMBRE o al subtipo MUJER
- No puede haber una persona que pertenezca a la vez al subtipo HOMBRE y al subtipo MUJER
- El atributo *sexo* clasifica a una persona en alguno de los subtipos HOMBRE o MUJER

b) Parcial sin solapamiento:

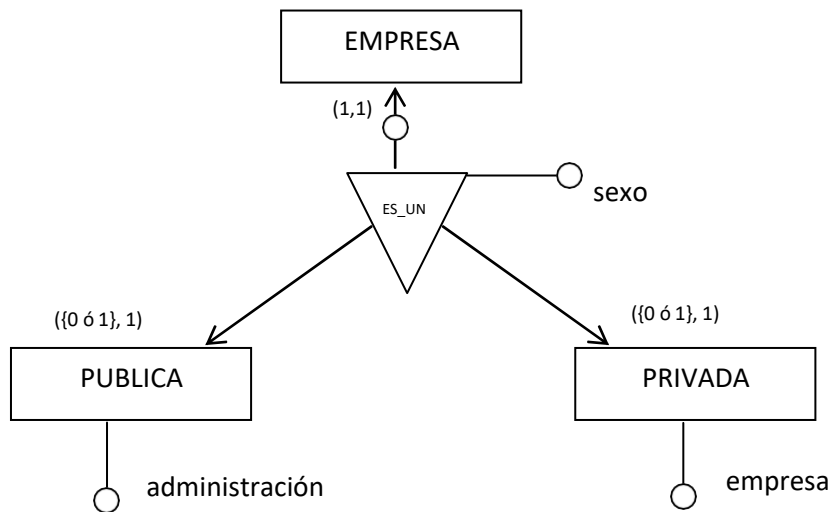
Ej:



- Una enfermedad puede ser vírica o bacteriana, pero también puede haber enfermedades que no pertenezcan a ninguno de los subtipos
- Una enfermedad no puede ser vírica y bacteriana al mismo tiempo.

c) Total con solapamiento:

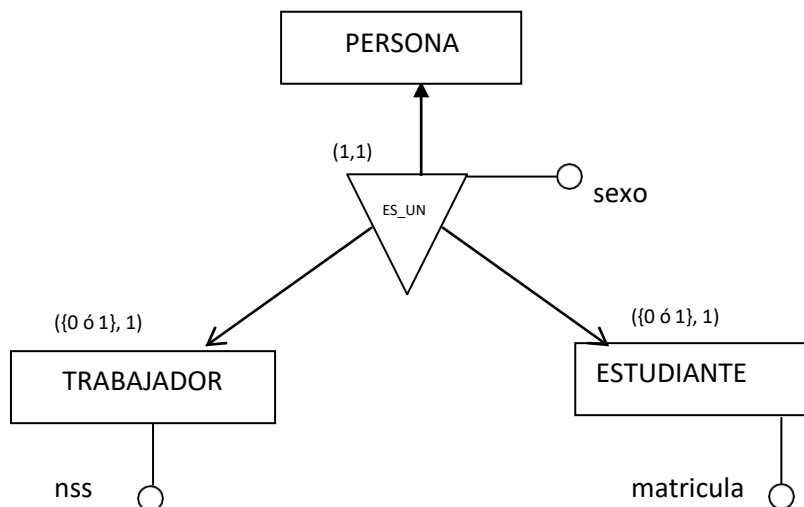
Ej:



- Una empresa ha de especializarse en pública o en privada
- Puede haber instancias de **EMPRESA** que pertenezcan a los dos subtipos (que sean al mismo tiempo públicas y privadas)

d) Parcial con solapamiento:

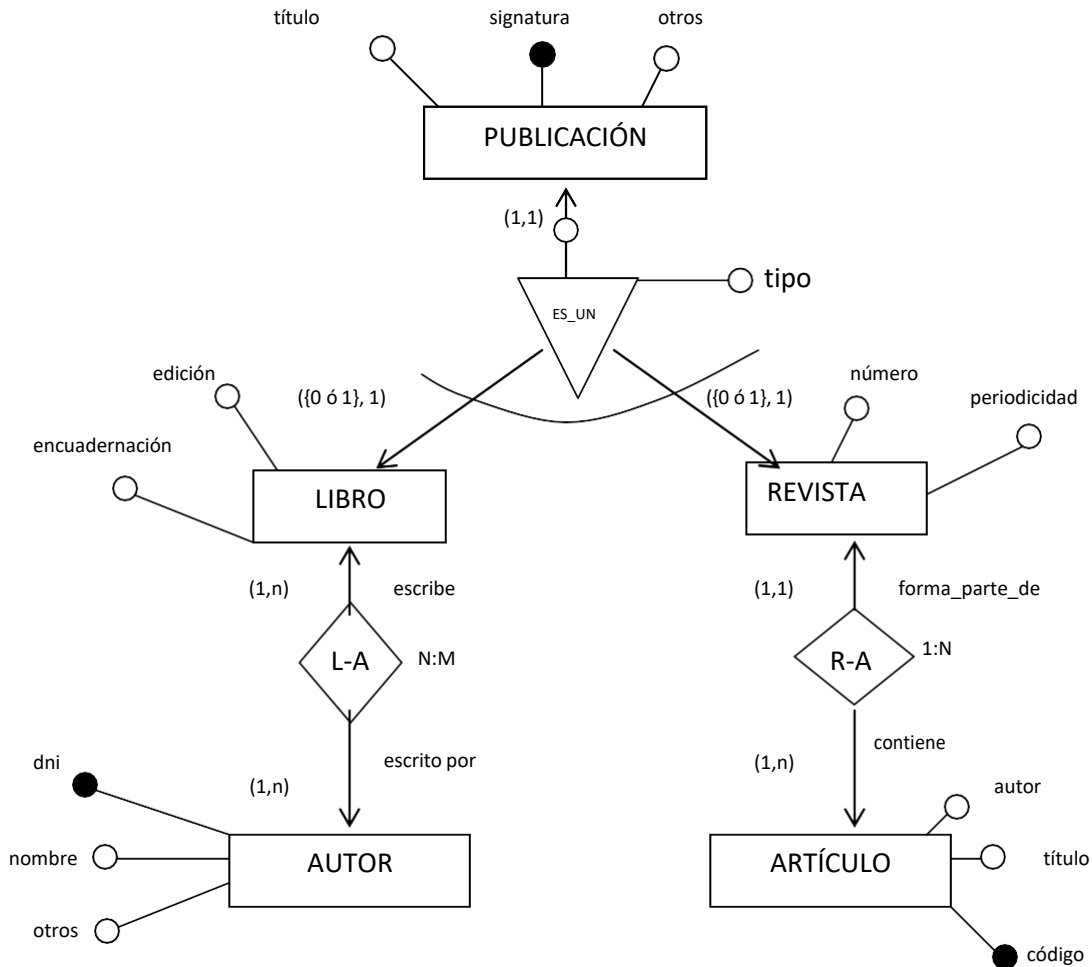
Ej:



- Una persona puede ser del tipo trabajador o del tipo estudiante, o también de ambos a la vez.
- Puede haber alguna instancia de **PERSONA** que no pertenezca a ninguno de los subtipos (que no sean trabajadores ni estudiantes)

Los subtipos tendrán los mismos atributos que el supertipo (aunque no se vuelven a indicar en el diagrama) y los suyos propios (que sí se indican en el diagrama).

Los identificadores de las entidades genéricas suelen ser iguales a los de las entidades específicas. A veces puede ser necesario ponerle una categoría al identificador para saber a qué entidades específicas debemos acudir.



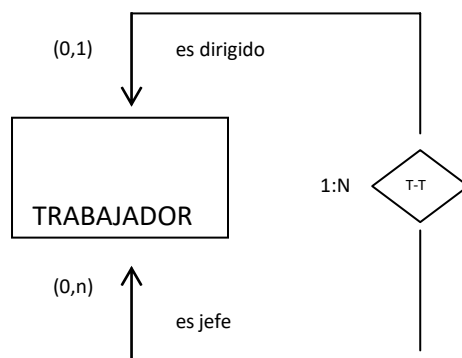
En una generalización se expresa la relación **es_un**.

5.3.- TIPOS DE RELACIONES

5.3.1.- Relaciones Reflexivas

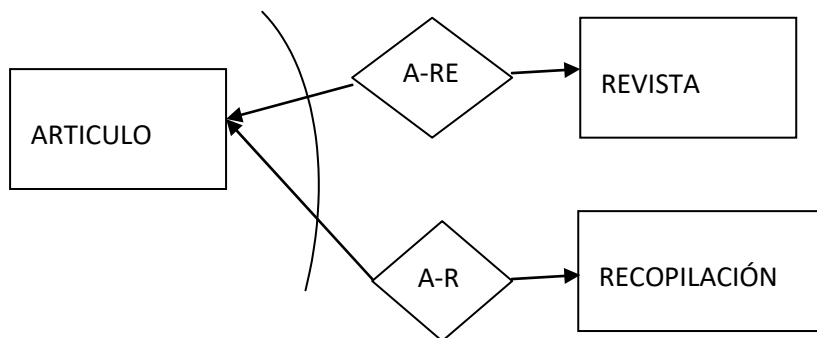
Son relaciones unarias, y, por tanto, en el tipo de interrelación sólo participa un único tipo de entidad.

Ejemplo: Un trabajador puede ser jefe de ningún trabajador o puede serlo de varios trabajadores, mientras que un trabajador sólo es dirigido por ninguno o un trabajador.



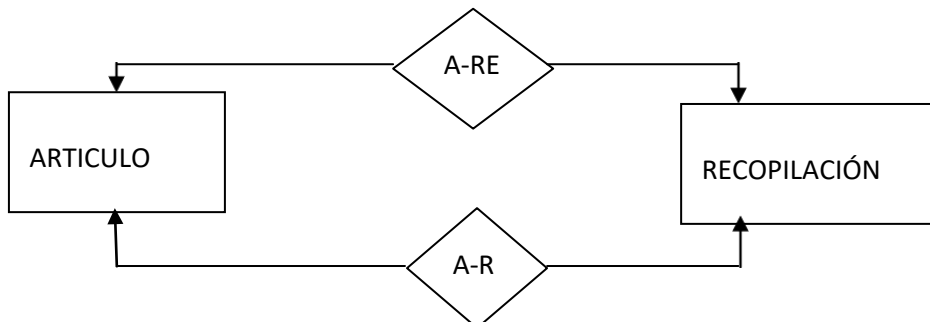
5.3.2.- Relaciones Exclusivas

Se dice que dos o más tipos de interrelación son exclusivas cuando cada ocurrencia de un tipo de entidad sólo puede pertenecer a un tipo de interrelación.



5.3.3.- Interrelaciones entre dos tipos de entidad

Entre dos tipos de entidad puede existir más de un tipo de interrelación:



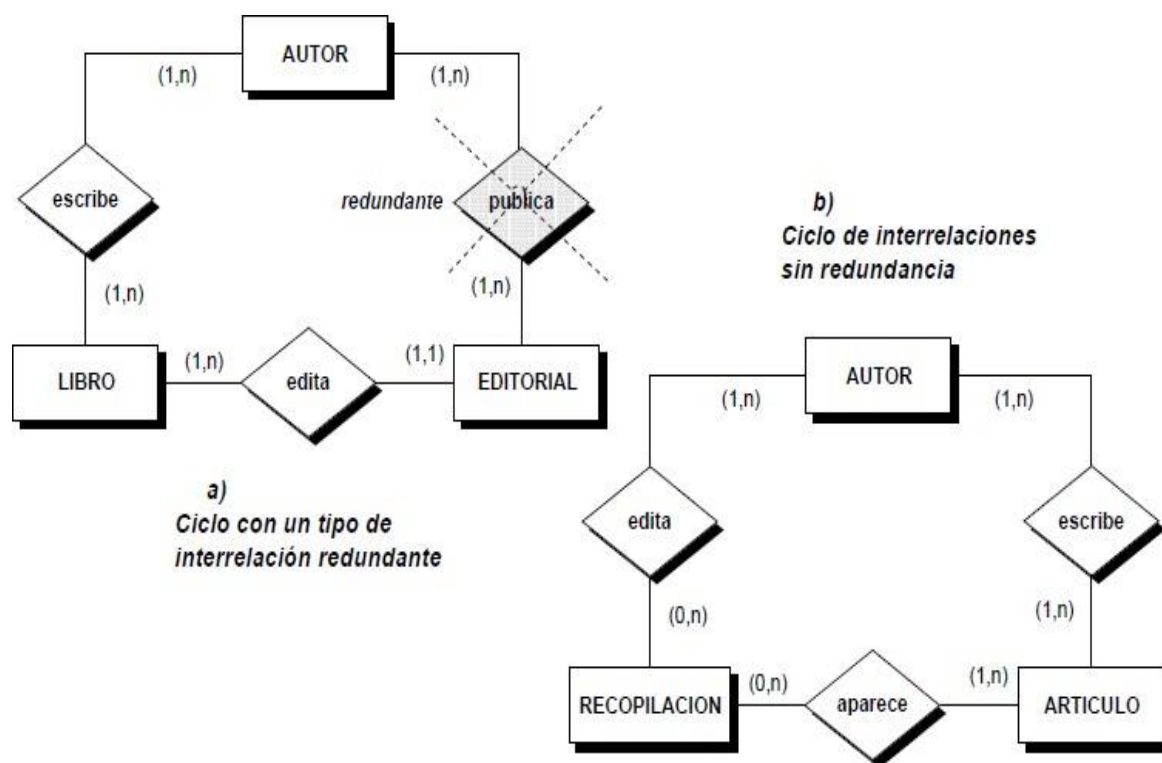
5.3.4.- Control de redundancia

En los esquemas E/R, y en general en los de cualquier MD, es necesario evitar las redundancias para no tener problemas de inconsistencias de la representación.

Un **elemento** de un esquema es **redundante** si puede ser eliminado sin pérdida de semántica.

- Existen dos formas principales de redundancia:

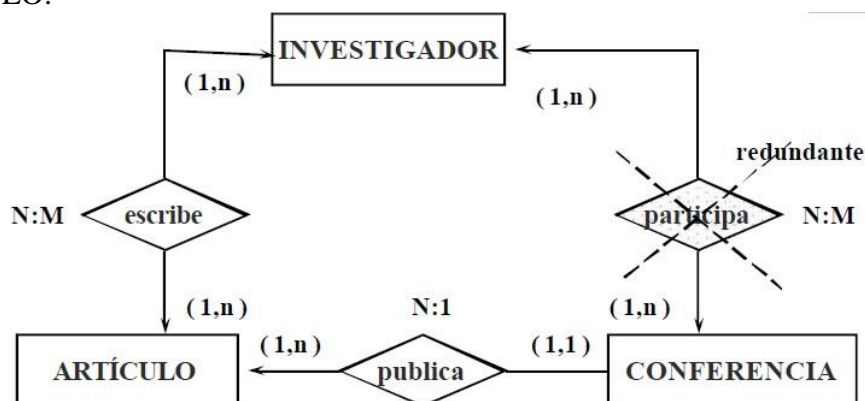
- ✓ En los **atributos** (atributos derivados o calculados): Aunque son redundantes, no dan lugar a inconsistencias siempre que en el esquema se indique su condición de derivados y la fórmula mediante la que han de ser calculados. Por ejemplo, si se almacena la fecha de nacimiento de una persona, no hace falta otro atributo con su edad, ya que esta se puede calcular.
- ✓ En las **interrelaciones** (también llamadas interrelaciones derivadas): Una interrelación es redundante si su eliminación no implica pérdida de semántica porque existe la posibilidad de realizar la misma asociación de ejemplares por medio de otras interrelaciones. Para ello es condición necesaria pero no suficiente, que forme parte de un ciclo => Hay que estudiar detenidamente los ciclos en el diagrama E/R.



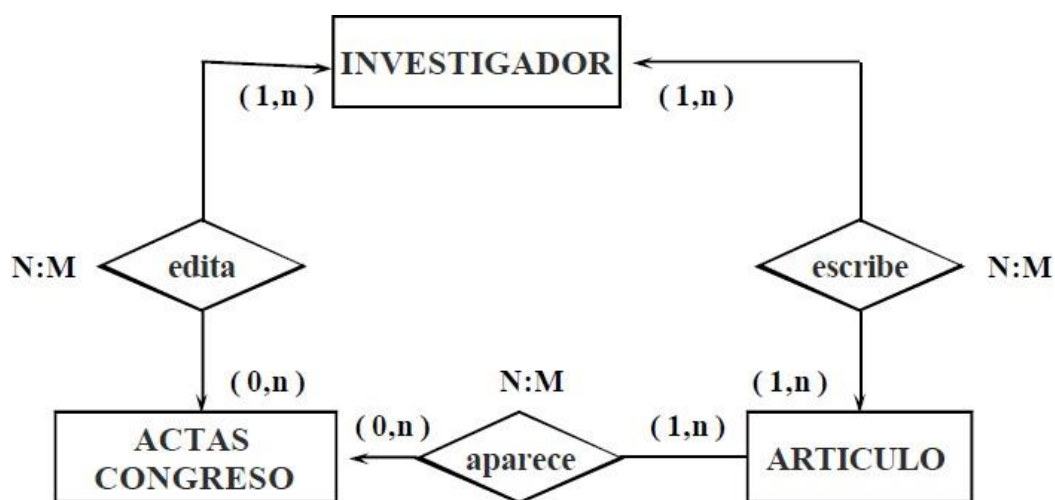
En el ejemplo de la Figura a) se da un ciclo entre AUTOR, LIBRO y EDITORIAL, por lo que en principio es posible que aparezca alguna interrelación redundante. En efecto, si se conocen los libros de un autor y las editoriales que los han editado, se puede deducir fácilmente en qué editoriales ha publicado dicho autor; de forma análoga, dada una editorial si sabemos que libros ha publicado podemos deducir qué autores han escrito para ella, por lo que la interrelación **publica** entre las entidades AUTOR y EDITORIAL es redundante.

Sin embargo, en la Figura b), a pesar de existir un ciclo, no hay ninguna interrelación redundante. En lo que respecta a la interrelación **escribe**, no puede deducirse de las otras dos, ya que aunque sepamos las recopilaciones que ha editado un autor y los artículos que han aparecido en ellas, no podemos saber los artículos que ha escrito dicho autor. La interrelación **edita** tampoco es redundante, ya que un autor escribe varios artículos que pueden aparecer en recopilaciones sin que eso implique que dicho autor edite esas recopilaciones. Por último, la interrelación **aparece** tampoco es redundante, ya que un artículo escrito por un autor no tiene por qué aparecer necesariamente en las recopilaciones que éste edite: puede aparecer en otras recopilaciones o revistas, obsérvese la cardinalidad mínima cero en la entidad RECOPIACION con respecto a la entidad ARTICULO en la interrelación **aparece**.

EJEMPLO:



Ciclo donde aparece una interrelación redundante



Ciclo donde no aparece una interrelación redundante

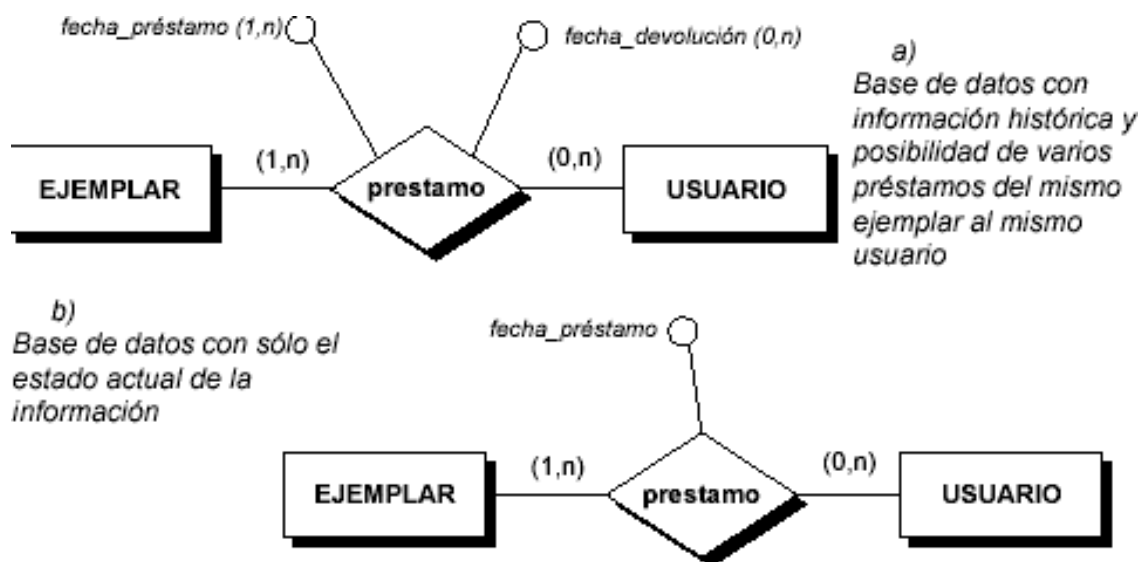
La existencia de un ciclo no implica la existencia de interrelaciones redundantes. Para que una interrelación pueda ser eliminada por redundante se tiene que cumplir:

- que exista un ciclo.
- que las interrelaciones que componen el ciclo sean equivalentes semánticamente.
- que se puedan asociar los ejemplares de las dos entidades que estaban interrelacionadas, aún habiéndose eliminado la interrelación.
- que la interrelación no tenga atributos o que éstos puedan ser transferidos a otro elemento del esquema a fin de no perder su semántica.

6.- DIMENSIÓN TEMPORAL

Es necesario establecer un método semántico y gráfico que recoja en el esquema conceptual el transcurso del tiempo y su influencia en la forma en que cambian los datos:

- Atributos de tipo fecha asociados a entidades o interrelaciones:
 - ✓ Para sucesos instantáneos (sin duración), basta con un sólo atributo de este tipo.
 - ✓ Para poder almacenar hechos que transcurren en un intervalo de tiempo determinado necesitaremos una fecha_inicio y una fecha_fin.
 - ✓ En las bases de datos históricas, en las que una interrelación entre dos ejemplares concretos se pueda repetir en el tiempo, el atributo fecha será multivaluado.
- Cuando es necesario representar la evolución de un tipo de entidad a lo largo del tiempo se utiliza un atributo de estado, que indicará en qué estado concreto se encuentra la entidad. En muchos casos lleva asociado otro atributo, que es la fecha en la que se ha producido el cambio de estado o el intervalo de tiempo en que ha permanecido en dicho estado



Representación del tiempo en esquemas E/R