

UT 3 - Modelo Entidad/Relación

[3.1] INTRODUCCIÓN

Ya hemos visto anteriormente que existen varios esquemas a realizar para poder representar en forma de base de datos informática un problema procedente del ordenador.

El primero de esos esquemas es el llamado **esquema conceptual**, que representa la información de forma absolutamente independiente al Sistema Gestor de Base de Datos.

Los esquemas internos de las diferentes bases de datos no captan la semántica del mundo real, ya que se centran más en estructura de almacenamiento más cercanas a la computadora; de ahí que primero haya que pasar por uno o dos esquemas previos, más cercanos al mundo real y que permitirían adaptar el problema a todo tipo de sistemas.

El hecho de saltarse el esquema conceptual conlleva un problema de pérdida de percepción con el problema real, ya que nos aproximamos demasiado al ordenador y nos alejamos de la información como la entiende el ser humano y eso, a la larga, provoca problemas de incoherencia en los datos. El esquema conceptual debe reflejar todos los aspectos relevantes del mundo a real a modelar.

[3.1.1] PETER P. CHEN Y EL MODELO ENTIDAD/RELACIÓN

En 1976 y 1977 dos artículos de **Peter P. Chen** detallaron un modelo para realizar esquemas con la idea de proveer una visión unificada de los datos de un sistema de base de datos. Este modelo es el **modelo entidad/interrelación** (*entity/relationship* en inglés) que actualmente se conoce más con el nombre de **entidad/relación** (**Modelo E/R** o **ME/R**, en inglés *E/RM*).

Posteriormente otros autores han añadido mejoras a este modelo, lo que ha producido toda una familia de modelos basados en el modelo Entidad/Relación original. La más aceptada actualmente es el modelo **entidad/relación extendido (ERE)** que complementa algunas carencias del modelo original. No obstante las diversas variantes del modelo hacen que los esquemas que dibujan los profesionales no sigan un verdadero estándar y sean dispares, aunque hay ideas muy comunes a todos los “*dialectos*” del modelo entidad/relación.

Hay que insistir en que este modelo no tiene nada (o muy poco) que ver con las bases de datos relacionales, los esquemas entidad/relación se pueden utilizar (en principio) con cualquier SGBD ya que son conceptuales. Puede que nos confunda el uso de la palabra *relación*, ya que está también presente en el modelo relacional de las bases de datos de **Edgar F. Codd**; pero el concepto de relación en este esquema no tiene nada que ver con la idea de relación expuesta por **Codd** en su modelo relacional.

[3.2] COMPONENTES DEL MODELO

[3.2.1] ENTIDAD

Se trata de cualquier objeto u elemento (real o abstracto) acerca del cual se pueda almacenar información en la base de datos. Es decir cualquier elemento informativo que tenga importancia para una base de datos.

Ejemplos de entidades son: una persona que se llama Pedro, la factura número 32456, el coche matrícula 3452BCW, etc.

Una entidad no es una propiedad concreta, sino un objeto que puede poseer múltiples propiedades (**atributos**). Es decir “*Sánchez*” es el contenido del atributo *Primer Apellido* de la entidad que representa a la persona *Leire Sánchez Crespo* con DNI 12766374,...

Las entidades son objetos completos, con todos los valores de las propiedades de dicho objeto. Descubrir entidades es la tarea principal del diseño de esquemas Entidad/Relación

[3.2.2] CONJUNTOS DE ENTIDADES

Las entidades que poseen las mismas propiedades, porque se refieren al mismo tipo de ente, forman conjuntos de entidades. Ejemplos de conjuntos de entidades son: *personas*, *facturas*, *coches*, ...

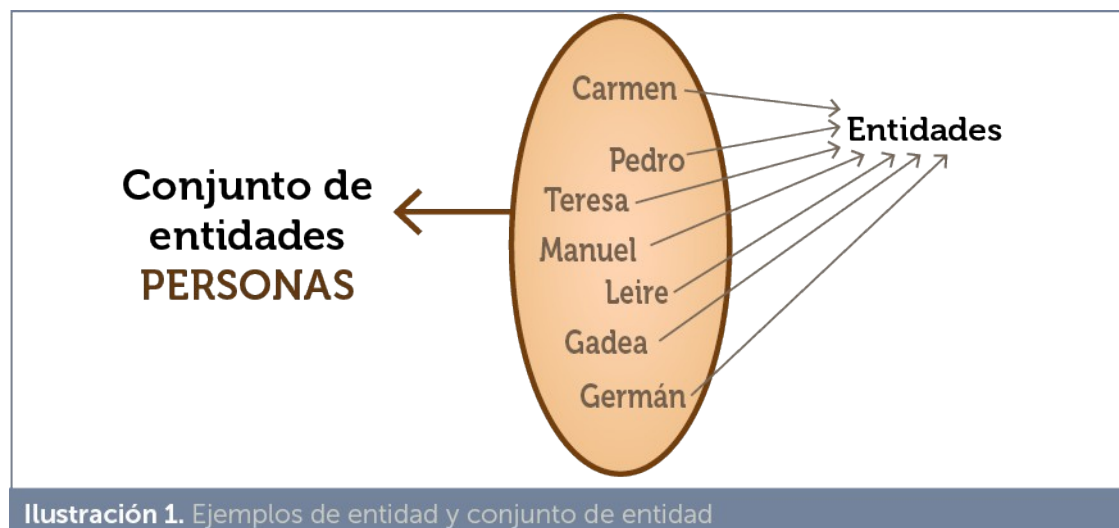


Ilustración 1. Ejemplos de entidad y conjunto de entidad

En la actualidad se suele llamar **entidad** a lo que anteriormente se ha definido como conjunto de entidades, seguramente por una cuestión de ahorro en el lenguaje. De este modo hablaríamos de la entidad *PERSONAS*. Mientras que cada persona en concreto sería una **ocurrencia** o un **ejemplar** de la entidad *persona*.

Esta terminología es la que actualmente vamos a utilizar en este manual: los conjuntos de entidades serán entidades a secas, y lo que Chen llamaba entidades, serán ejemplares de dicha entidad.

[3.2.3] REPRESENTACIÓN GRÁFICA DE LAS ENTIDADES

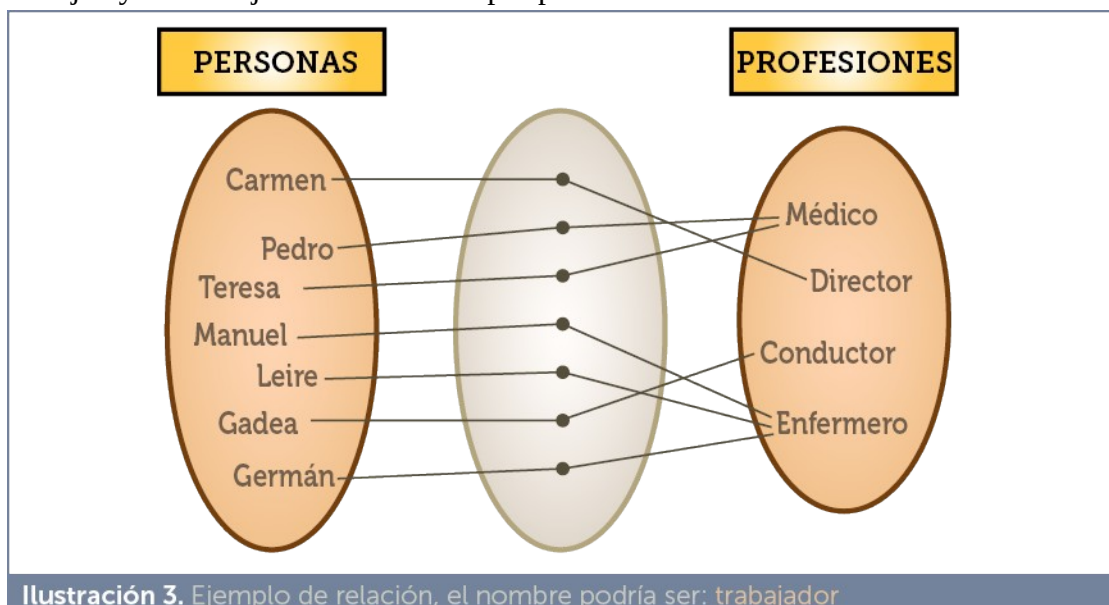
En el modelo entidad relación los conjuntos de entidades se representan con un rectángulo dentro del cual se escribe el nombre de la entidad:



[3.3] RELACIONES

[3.3.1] QUÉ ES UNA RELACIÓN

Representan **asociaciones** entre entidades, es decir entidades de un conjunto que tienen contacto con entidades de otro conjunto. Es el elemento del modelo que permite relacionar en sí los datos del mismo; de otro modo tendríamos información aislada. Por ejemplo, en el caso de que tengamos una entidad *personas* y otra entidad *trabajos*. Ambas se relacionan ya que las personas trabajan y los trabajos son realizados por personas.



En el dibujo anterior, *trabajar* podría ser el nombre del conjunto de relaciones entre las entidades *personas* y *trabajos*.

En una relación (Chen llamaba *conjunto de relaciones* a lo que ahora vamos a llamar *relación* a secas) cada ejemplar (cada relación en la terminología de Chen) asocia un elemento de una entidad con otro de la otra entidad.

Una regla fundamental es que **en una relación no pueden aparecer dos veces relacionados los mismos ejemplares de cada entidad**. Una relación establece una asociación entre un ejemplar de

una entidad con un ejemplar de otra entidad, una vez. Es decir; en el ejemplo anterior, en la relación no puede aparecer dos veces el mismo trabajador asociado al mismo trabajo.

Si detectamos que nuestra base de datos requiere asociar dos veces a los mismos ejemplares y solo hemos hecho una relación, no hay que dudar: nuestro diseño no es correcto.

[3.3.2] REPRESENTACIÓN GRÁFICA

La representación gráfica de las entidades se realiza con un rombo al que se le unen líneas que se dirigen a las entidades, las relaciones tienen nombre (se suele usar un verbo). En el ejemplo anterior podría usarse como nombre de relación, trabajar:



Ilustración 4. Representación gráfica de una relación en el modelo E/R

[3.3.3] EJEMPLOS DE RELACIONES

- **Relaciones Binarias.** Son las relaciones típicas. Se trata de relaciones que asocian dos entidades.
- **Relaciones Ternarias.** Relacionan tres entidades. A veces se pueden simplificar en relaciones binarias, pero no siempre es posible.
- **Relaciones n -arias.** Relacionan n entidades (por ejemplo relaciones cuaternarias, quinquenarias,...). Son muy raras
- **Relaciones dobles.** Se llaman así a dos relaciones distintas que sirven para relacionar a las mismas relaciones. Son las más difíciles de manejar ya que al manipular las entidades hay que elegir muy bien cuál es la relación adecuada para hacerlo.
- **Relación reflexiva.** Es una relación que sirve para relacionar dos ejemplares de la misma entidad (personas con personas, piezas con piezas, etc.)

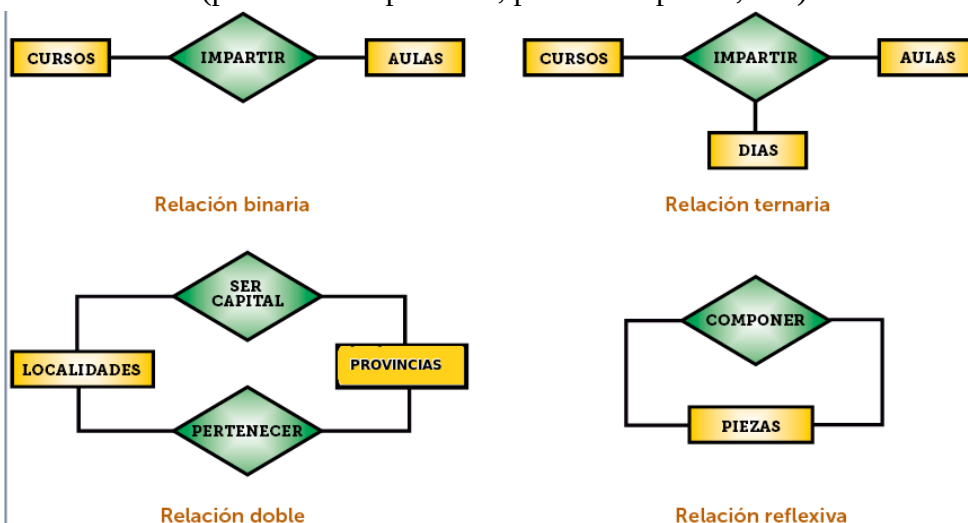


Ilustración 5. Tipos de relaciones

[3.3.4] CARDINALIDAD

Indica el número de relaciones en las que una entidad puede aparecer. Se anota en términos de:

- **Cardinalidad mínima.** Indica el número mínimo de asociaciones en las que aparecerá cada ejemplar de la entidad (el valor que se anota es de **cero o uno**, aunque tenga una cardinalidad mínima de más de uno, se indica sólo un uno)
- **Cardinalidad máxima.** Indica el número máximo de relaciones en las que puede aparecer cada ejemplar de la entidad. Puede ser uno, otro valor concreto mayor que uno (tres por ejemplo) o muchos (se representa con n). Normalmente la cardinalidad máxima es **1 ó n**

En los esquemas entidad / relación la cardinalidad se puede indicar de muchas formas. Quizá la más completa (y la que se utiliza en este documento es ésta) consiste en anotar en los extremos la cardinalidad máxima y mínima de cada entidad en la relación.

Ejemplo de uso de cardinalidad:

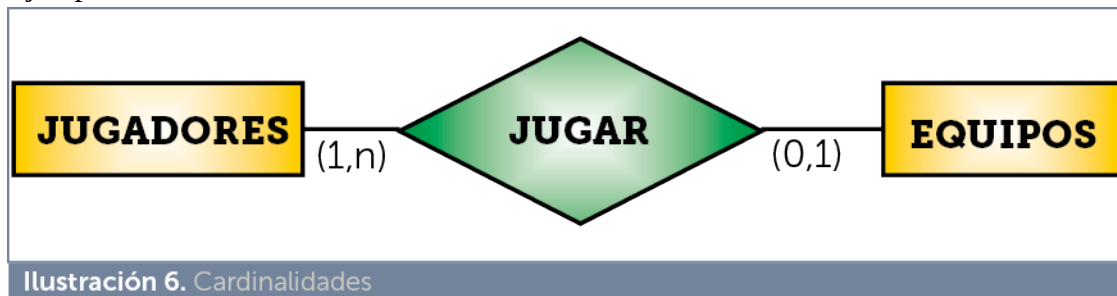


Ilustración 6. Cardinalidades

En el ejemplo un jugador tiene una cardinalidad mínima de 0 (puede no estar en ningún equipo) y una máxima de 1 (como mucho está en un equipo, no puede estar en dos a la vez). Cada equipo tiene una cardinalidad mínima de uno (en realidad sería una cardinalidad mínima más alta, pero se anota un uno) y una máxima de n (en cada equipo hay muchos jugadores)

[3.3.5] ROLES

A veces en las líneas de la relación se indican **roles**. Los roles representan el papel que juega una entidad en una determinada relación. Son imprescindibles cuando las relaciones son complejas, ya que ayudan a entender el sentido de la relación.

Ejemplo:

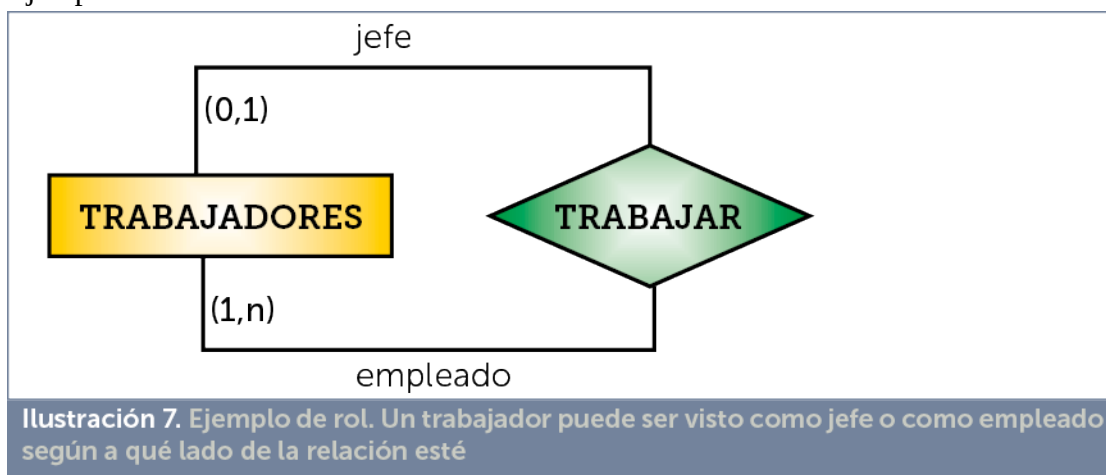


Ilustración 7. Ejemplo de rol. Un trabajador puede ser visto como jefe o como empleado según a qué lado de la relación esté

[3.4] ATRIBUTOS

Describen propiedades de las entidades y las relaciones. Son fundamentales y establecen la información que deseamos almacenar de cada objeto de la base de datos. El modelo Entidad/Relación clásico los representa con elipses, dentro de las cuales se coloca el nombre del atributo. La elipse se une con una línea a las entidades. Ejemplo:

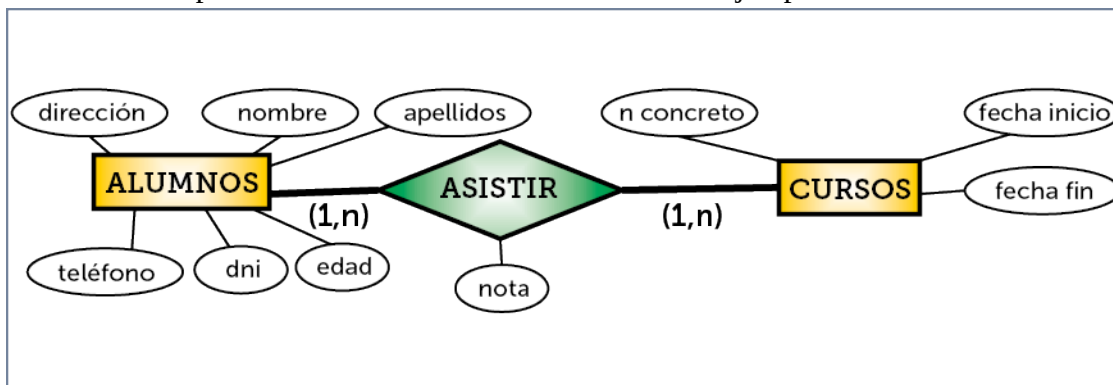
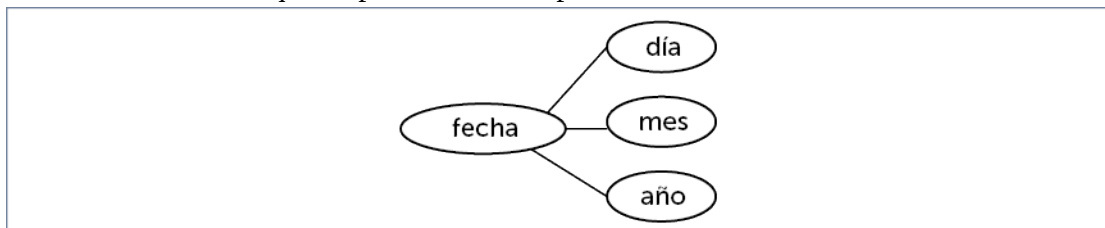


Ilustración 8. Extracto de esquema E/R con atributos

[3.4.1] TIPOS DE ATRIBUTOS

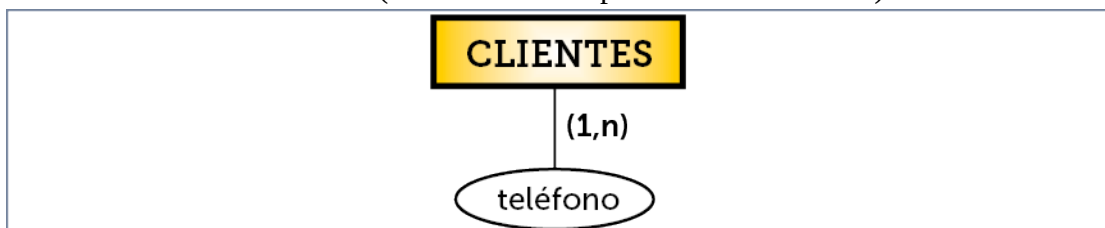
COMPUESTO

Se trata de atributos que se pueden descomponer en otros más sencillos:



MÚLTIPLES

Pueden tomar varios valores (varios teléfonos para el mismo cliente):



Otra forma de representar atributos que pueden tomar múltiples valores es ésta:



OPCIONALES

Lo son si pueden tener valor nulo (es decir, si pueden quedar vacíos, sin valor):



Otra forma de representarlos es:



[3.4.2] IDENTIFICADOR PRINCIPAL O CLAVE

Se trata de uno o más atributos de una entidad cuyos valores son únicos en cada ejemplar de la entidad. Es decir todos los elementos de una entidad tienen en ese (o esos) atributo, un valor diferente (y nunca vacío).

Este tipo de atributos son fundamentales y se marcan en el esquema subrayando el nombre del identificador.

Para que un atributo sea considerado un buen identificador tiene que cumplir con los siguientes requisitos:

- [1] Deben distinguir a cada ejemplar de la entidad o relación. Es decir no puede haber dos ejemplares con el mismo valor en el identificador.
- [2] Todos los ejemplares de una entidad deben tener el mismo identificador.
- [3] Un identificador puede estar formado por más de un atributo.
- [4] Puede haber varios identificadores **candidatos**, en ese caso hay que elegir el que tenga más importancia en nuestro sistema (el resto pasan a ser **alternativos**).

Todas las entidades deben de tener un identificador, en el caso de que una entidad no disponga de un identificador (puede ocurrir, pero hay que ser cauteloso, a veces se trata de entidades que están mal modeladas) entonces hay que añadir un atributo que haga de identificador.

Los futuros valores de este atributo no nos interesan, lo único que necesitamos de él es que disponga de un valor diferente para cada ejemplar de la entidad. El nombre de este atributo artificial es la palabra **id** seguida del nombre de la entidad. Por ejemplo *id_personas*.

[3.4.3] IDENTIFICADOR ALTERNATIVO.

Se trata de uno o más atributos en la entidad cuyos valores son únicos para cada ejemplar de una entidad, pero que no son identificadores ya que hay atributos que resultan ser mejores identificadores. Los identificadores alternativos se marcan con un subrayado discontinuo (ejemplo de subrayado discontinuo)

[3.4.4] ELECCIÓN DE BUENOS IDENTIFICADORES PRINCIPALES

Como hemos visto en los dos puntos anteriores, es habitual disponer de varios identificadores candidatos para la misma entidad. Por ello ¿cuál elegir como identificador principal?

[1]Siempre debemos elegir el candidato que tenga más que ver con el problema que estamos resolviendo. Es decir entre un documento nacional de identidad y un código que sólo se usa en la empresa (código cliente, código de socio, nº de personal,...), hay que elegir la segunda opción. La razón es que seguro que en la empresa se tienen más en consideración este segundo número. Como razón técnica : los códigos internos siempre son más cortos que los generales.

[2]Realmente en el modelo conceptual la única razón a tener en cuenta es la expuesta en el punto anterior. Sin embargo, podemos ahorrar trabajo cuando hagamos el esquema lógico si ya cumplimos estas reglas (que en realidad son técnicas):

- Sólo debemos elegir como identificadores principales, a atributos (sea uno o varios cuyos valores son únicos) a fechas, números enteros y textos cortos y de tamaño fijo.
- De los posibles candidatos observados (que cumplan el punto anterior), elegir el que tenga más relación con el problema que estamos resolviendo. Si no encontramos una diferencia conceptual, entonces elegir el que tenga un tamaño más corto.
- Si ningún candidato cumple estas reglas, es mejor incluso inventarse un identificador (al final contendrá un número entero diferente para cada elemento de la entidad).

[3]En todo caso, cuando pasemos el modelo E/R a su forma lógica (por ejemplo utilizando el modelo relacional) podemos cambiar de idea respecto al identificador; por lo que lo dicho en el apartado anterior puede no ser tenido en cuenta.

[3.5] MODELO ENTIDAD RELACIÓN EXTENDIDO

En el modelo entidad relación extendido aparecen nuevos tipos de relaciones. Son las **relaciones ISA** (*es un*) y las **entidades débiles**

[3.5.1] ENTIDADES DÉBILES

Se dice que una entidad débil es aquella cuya existencia depende de otra (considerada su **entidad fuerte**). Se trata de entidades totalmente supeditadas a otras, de modo que si un ejemplar de la entidad fuerte desaparece, todos los ejemplares de la entidad débil relacionados, desaparecerán también del sistema.

Las entidades débiles ocurren cuando hay una entidad más fuerte de la que dependen, en el sentido de que la propia existencia de la entidad débil está supeditada a la existencia de su entidad fuerte. Lógicamente tienen relación con esa entidad, y es esa relación es la que marca el hecho de que una entidad es débil y la otra fuerte.

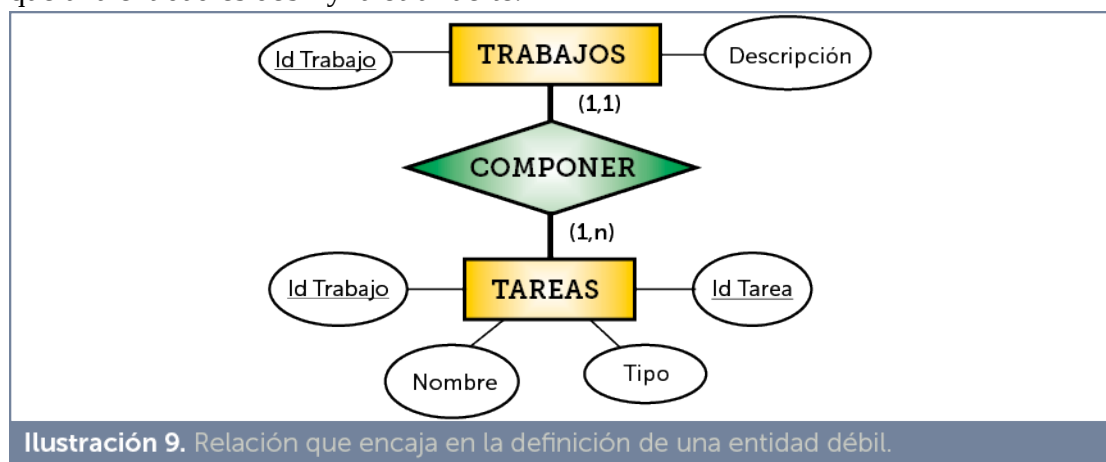


Ilustración 9. Relación que encaja en la definición de una entidad débil.

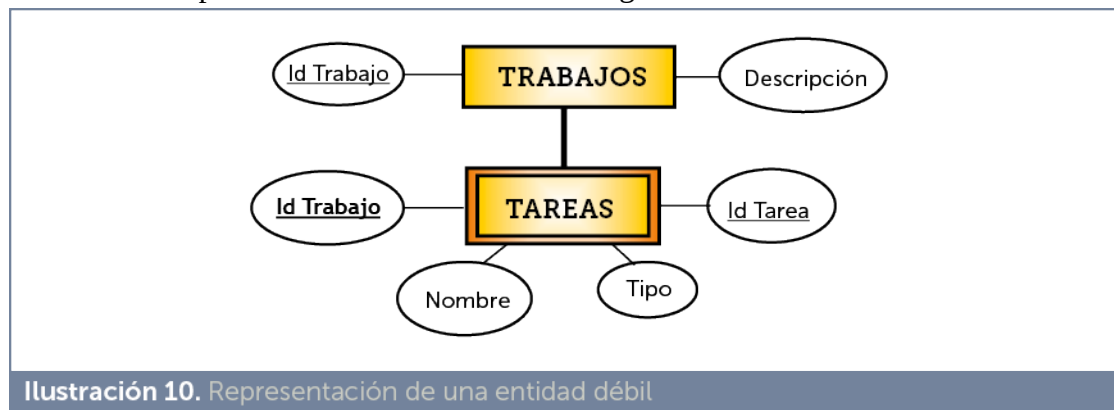
En la [Ilustración 9](#), la relación entre las tareas y los trabajos es 1 a n (cada trabajo se compone de n tareas). Una tarea obligatoriamente está asignada a un trabajo, es más no tiene sentido hablar de tareas sin hablar del trabajo del que forma parte. La relación que tienen las tareas con los trabajos es de debilidad, puesto que no podemos referirnos a una tarea de forma independiente.

En el ejemplo anterior hay incluso (aunque no siempre) una **dependencia de identificación** ya que las tareas se identifican por un número de tarea y el número de trabajo al que se asignan. Cuando eso ocurre, es un síntoma definitivo de que se trata de una entidad débil. Por ello es tan importante determinar correctamente los identificadores de las entidades.

La cardinalidad de una entidad débil siempre es la siguiente:

- **(1,1)**. En el lado de la entidad fuerte, haciendo referencia a que cada elemento de la entidad débil se relaciona con uno y solamente uno de los ejemplares de la entidad fuerte.
- **(0,n)** o **(1,n)**. En el lado de la entidad fuerte. Lo habitual es $(0,n)$ indicando que los ejemplares de la entidad fuerte no tienen por qué relacionarse con ninguno de los ejemplares de la entidad débil. De todos modos las cardinalidades en este tipo de relaciones no se indican, por lo que no es tan importante acertar con ellas.

La forma de representar estas relaciones es la siguiente:



No hace falta dibujar el rombo de la relación ni la cardinalidad, se sobreentiende el tipo y cardinalidad (1 a n) que posee. En este caso además hay dependencia de identificación (no siempre la hay)

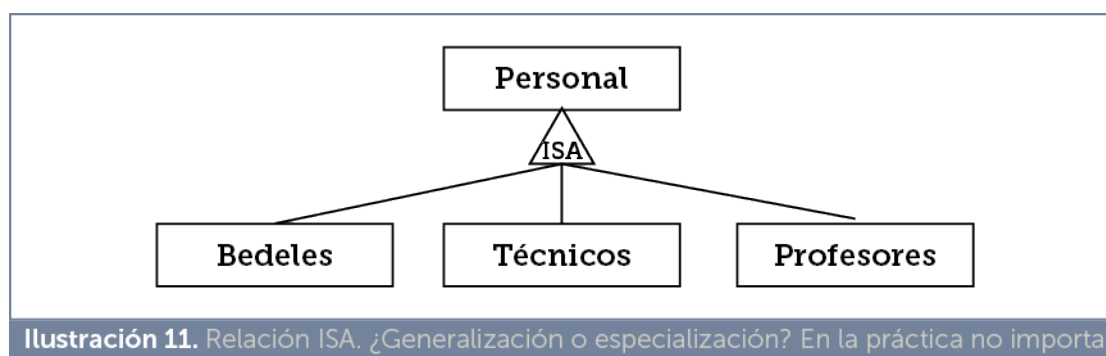
[3.5.2] RELACIONES ISA O RELACIONES DE HERENCIA

Son relaciones que indican relaciones que permiten distinguir **tipos** de entidades, es decir tendremos entidades que son un (**is a**, en inglés) tipo de entidad respecto a otra entidad más general. Se utilizan para unificar entidades agrupándolas en una entidad más general (**generalización**) o bien para dividir una entidad en entidades más específicas (**especificación**): aunque hoy en día a todas ellas se las suele llamar generalización e incluso (quizá incluso más adecuadamente) **relaciones de herencia**.

Se habla de **superentidad** refiriéndonos a la entidad general sobre las que derivan las otras (que se llaman **subentidades**). En la superentidad se indican los atributos comunes a todas las subentidades, se sobreentiende que las subentidades también tienen esos atributos, pero no se indican de nuevo esos atributos en el diagrama.

Normalmente cuando tenemos una especialización, las subentidades comparten clave con la **superentidad** (además de los atributos comunes); esto es muy matizable y de hecho hoy en día ningún diseñador intenta distinguir entre si tenemos una especialización o una generalización, porque al final ambas implican el mismo esquema interno en la base de datos.

La forma clásica de representar una ISA en el modelo Entidad/Relación es:



En general se suelen indicar las cardinalidades en las relaciones ISA, pero se suele sobreentender (cuando no se indican explícitamente) que hay un (0,1) encima de cada subentidad (que significa que cada ejemplar de la subentidad solo puede relacionarse como mucho con uno de la subentidad e incluso con ninguno; un empleado de *personal* podría ser o no ser un *profesor*). Pero se puede perfectamente indicar la cardinalidad (se usa ya la notación de ISA con triángulo hacia abajo que es la más popular en España actualmente):

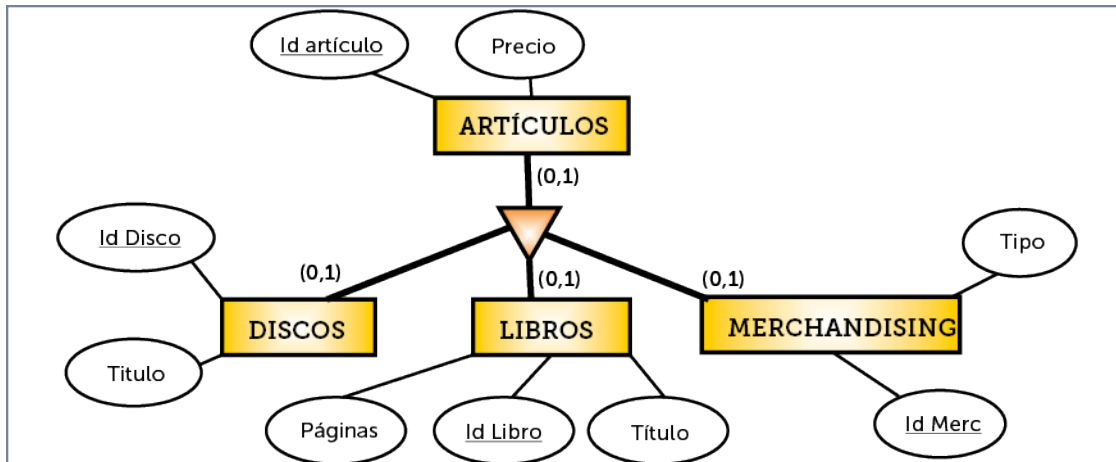


Ilustración 12. Ejemplo de relación ISA

Las cardinalidades que aparecen en el esquema anterior son las habituales (de hecho no se suele indicar ya que se sobreentiende), pero cualquier cardinalidad sería válida.

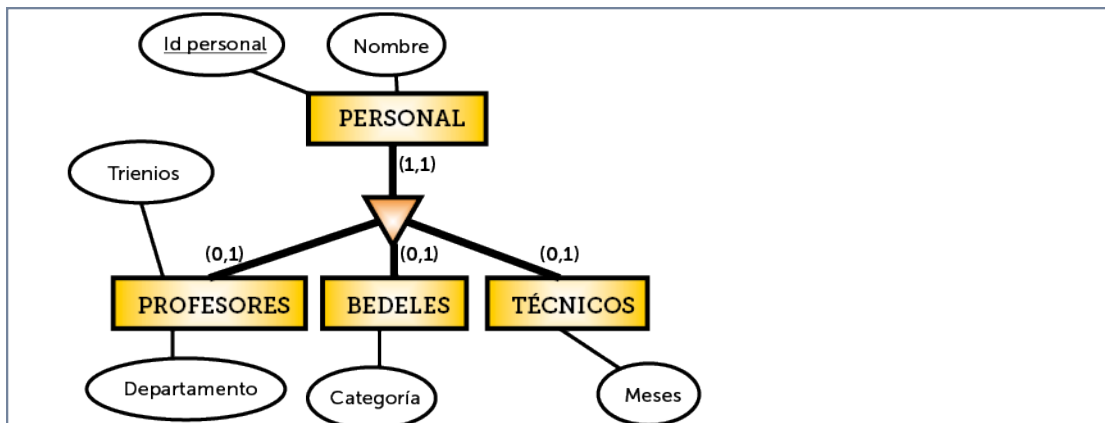
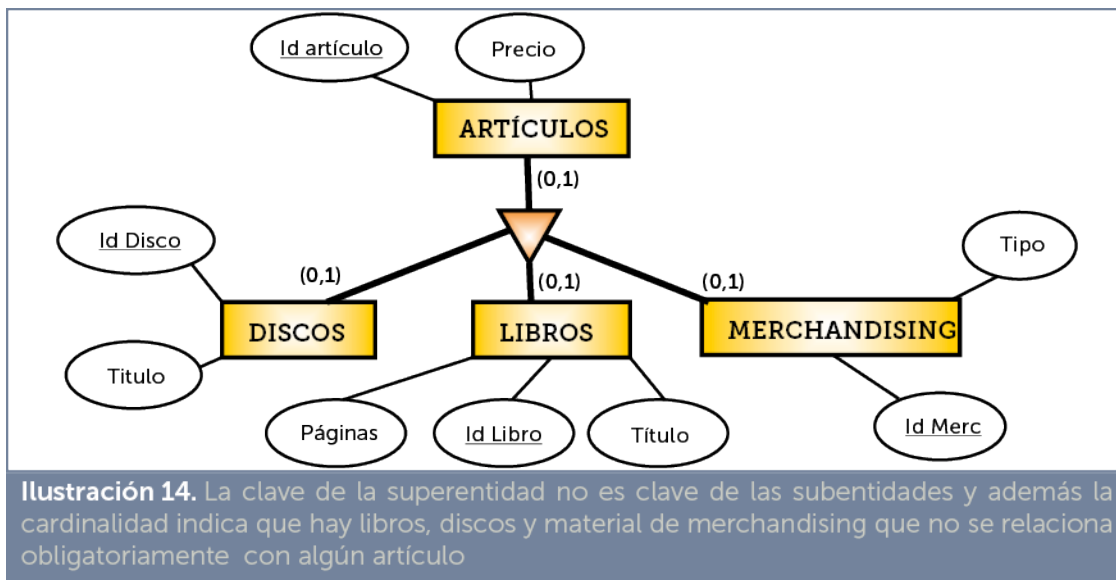


Ilustración 13. En esta relación de herencia, la clave de la superentidad es clave de las subentidades

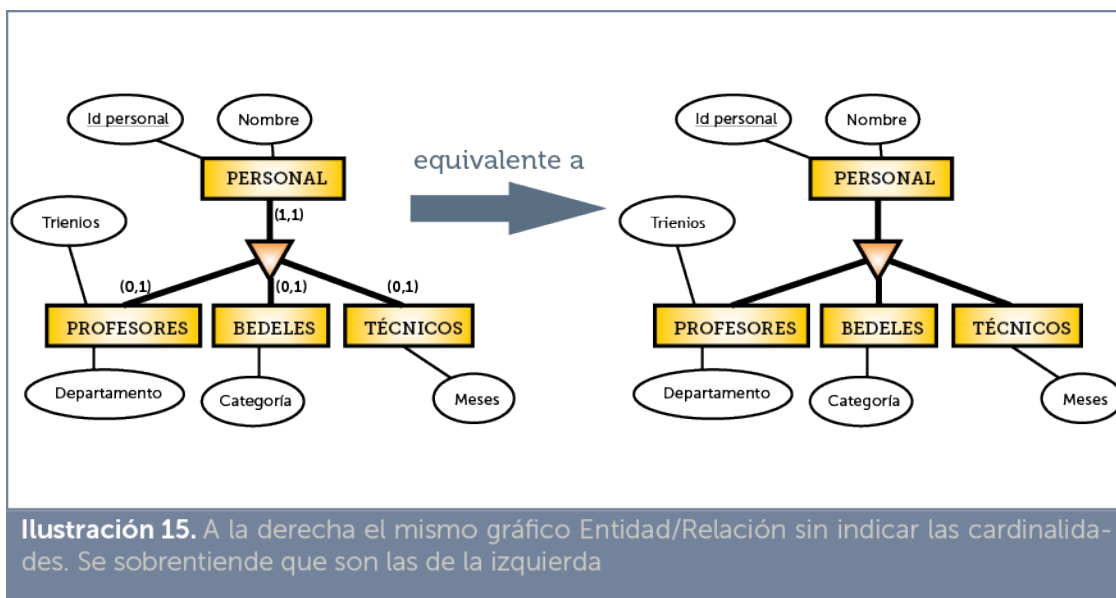
En la relación ISA anterior, los profesores, bedeles y técnicos heredan el atributo *id personal* y el *nombre*, el resto son atributos propios sólo de cada entidad (*trienios* pertenece sólo a los profesores, en este ejemplo)



En la ilustración anterior se utiliza una clave distinta para cada subentidad (es decir, *discos*, *libros* y *merchandising* tienen clave propia), no la heredan.

Además, es un caso en el que no hay relación obligatoria con la superentidad; es decir, un *disco* podría no ser un *artículo* (porque a lo mejor quiero meter discos en mi base de datos que no vendo en la tienda). No es muy habitual utilizar de esta forma relaciones ISA, pero desde luego es posible.

Normalmente en los esquemas no se indican las cardinalidades de las relaciones ISA y se sobreentiende que la superentidad tiene un 1,1 y las subentidades 0,1; es la cardinalidad habitual.



[3.5.3] ¿CUÁNDO HAY QUE UTILIZAR RELACIONES ISA?

Cuando los diseñadores principiantes conocen las relaciones ISA, suele ocurrir un exceso de uso de ellas en los diseños. No es conveniente abusar de este tipo de relaciones, ya que aumentan en exceso el número de entidades.

No todos los tipos de entidades suponen la aparición de una generalización o una especialización. A veces basta diferenciar los tipos de entidad con el uso de un simple atributo tipo o, como mucho, una entidad que contenga los tipos de entidades.

Lo recomendable es utilizar relaciones ISA cuando ocurre cualquiera de estas situaciones:

- Las subentidades tienen atributos distintos.
- Las subentidades tienen relaciones distintas.

En otros casos, se puede concluir que no resulta necesaria la creación de relaciones ISA.

[3.5.4] EXCLUSIVIDAD

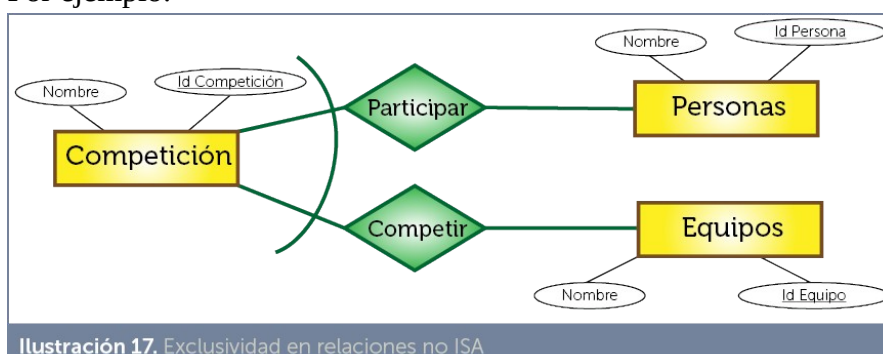
En las relaciones ISA (y también en otros tipos de relaciones) se puede indicar el hecho de que cada ejemplar sólo puede participar en una de entre varias ramas de una relación. Este hecho se marca con un **arco** entre las distintas relaciones. En las relaciones ISA se usa mucho para indicar que cada ejemplar de la superentidad sólo se relaciona con una subentidad, por ejemplo:



En el ejemplo, el *personal* sólo puede ser o *bedel* o *profesor* o *técnico*; una y sólo una de las tres cosas (es por cierto la forma más habitual de relación ISA).

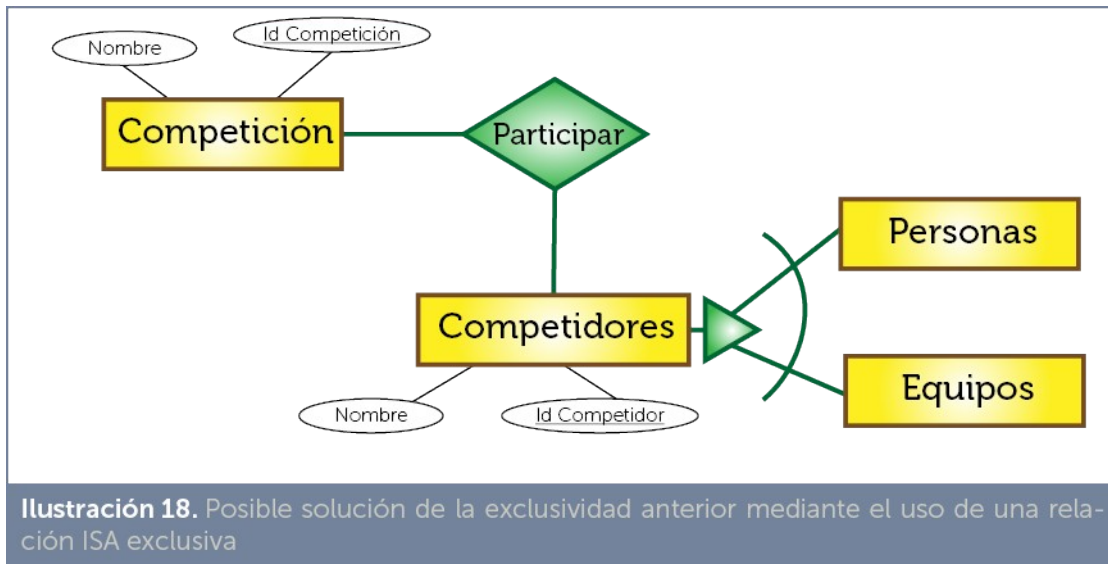
Se pueden marcar relaciones de exclusividad en otras relaciones. Significará lo mismo, es decir que solo uno de las situaciones enmarcadas por un arco se pueden dar a la vez.

Por ejemplo:

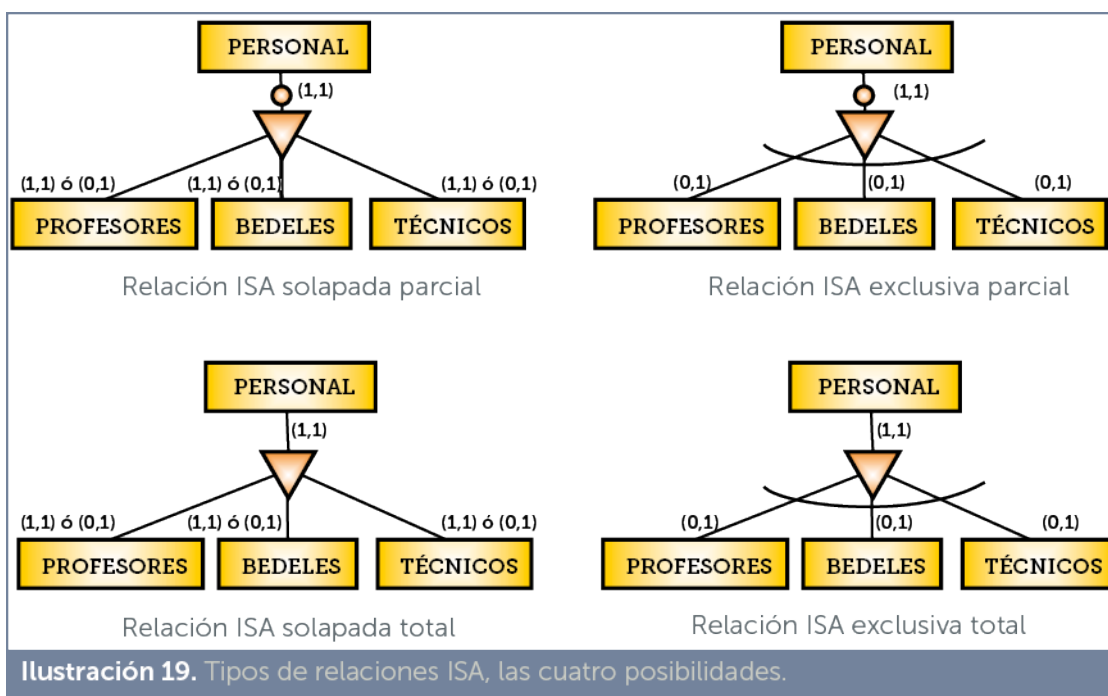


En la Ilustración anterior, solo puede ocurrir a la vez una de las dos relaciones: *Participar* o *Competir*. Significa que en las competencias o participan personas o compiten equipos. Es verdad, que en este caso a lo mejor se podría generalizar indicando una superentidad llamada *Competidores* para las personas y los equipos, que sería mucho más clarificadora.

No obstante no siempre se puede reflejar de otra forma una relación de exclusividad sobre relaciones normales.



[3.5.5] TIPOS DE RELACIONES ISA



Realmente lo que hay que matizar bien en las relaciones ISA es la forma de relacionarse la superentidad con la subentidad. Eso se matiza en base a dos conceptos:

- **Obligatoriedad.** Indica si los ejemplares obligatoriamente se relacionan con ejemplares de las subentidades. Es decir si hay personal que no es profesor ni bedel ni técnico o si fijo es alguna de esas tres profesiones. Hay dos posibilidades:
 - **Relaciones de jerarquía parcial.** Indican que hay ejemplares de la superentidad que no se relacionan con ningún ejemplar de las subentidades (por ejemplo, hay *personal* que no es ni *profesor*, ni *bedel* ni *técnico*). Se indican con un círculo encima del triángulo de la relación ISA.
 - **Relaciones de jerarquía total.** Indican que todos los ejemplares de la superentidad se relacionan con alguna subentidad (no hay personal que no sea ni profesor, ni bedel ni técnico).
- **Número de relaciones.** En este caso se mide con cuántas subentidades se relaciona la subentidad; es decir, si hay personal que pueda ser profesor y bedel a la vez o si sólo puede ser una cosa. Posibilidades:
 - **Relaciones de jerarquía solapada.** Indican que un ejemplar de la superentidad puede relacionarse con más de una subentidad (el personal puede ser profesor y bedel). Ocurren cuando no hay dibujado un arco de exclusividad.
 - **Relaciones de jerarquía exclusiva.** Indican que un ejemplar de la superentidad sólo puede relacionarse con una subentidad (el personal no puede ser profesor y bedel). Ocurren cuando hay dibujado un arco de exclusividad.

[3.6] OTRAS FORMAS DE REPRESENTACIÓN DEL MODELO ENTIDAD/RELACIÓN

Las herramientas CASE actuales no suelen utilizar la forma de dibujar del modelo clásico. La razón fundamental está en la dificultad de dibujar los atributos ya que ocupan un excesivo espacio en el diagrama. Como, además, los analistas están acostumbrados a representar diagramas lógicos basados en el modelo relacional de bases de datos, se usan diagramas entidad relación en los que la representación de entidades y relaciones se hace de esta forma:

- Las entidades son rectángulos en los que la primera fila se dedica al nombre de la entidad y las siguientes filas a cada uno de sus atributos.
- Las relaciones son rombos y sus atributos cuelgan en una caja que contiene el nombre de dichos atributos. En algunos casos ni siquiera se representa la relación con un rombo, sino que simplemente se indica su nombre.

Ejemplo:

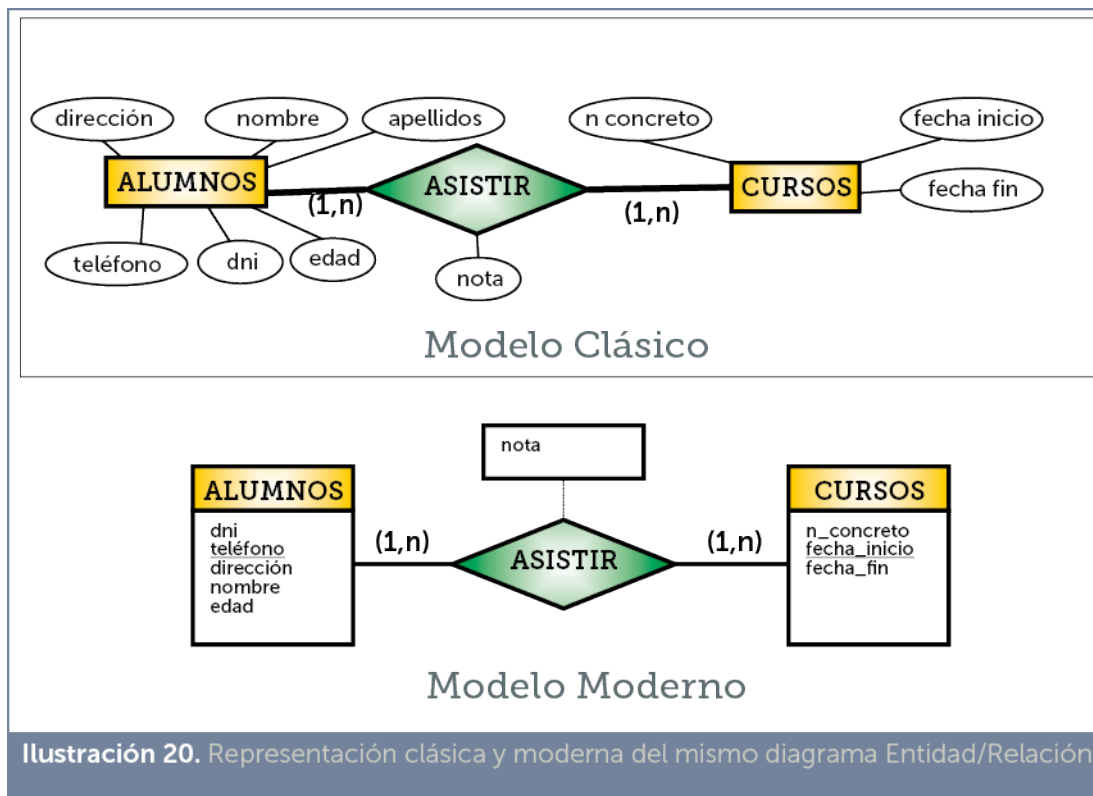


Ilustración 20. Representación clásica y moderna del mismo diagrama Entidad/Relación

- En las entidades débiles es el rombo de la relación el que se dibuja con borde doble y la línea que va hacia la entidad débil, también se dibuja doble.
- Finalmente las relaciones de tipo ISA (relaciones de herencia o generalizaciones) se representan de la misma forma, con un triángulo (aunque en estos diagramas el vértice está al revés)