

Modelo Entidad-Relación

El modelo de datos de entidad-relación (ER) se basa en una percepción de un mundo real que consiste en un conjunto de objetos básicos llamados entidades y de relaciones entre estos objetos. Se desarrolló para facilitar el diseño de bases de datos permitiendo especificar un esquema empresarial. Este esquema representa la estructura lógica general de la base de datos.

Objetos básicos del modelo ER

Los conceptos básicos previstos por el modelo ER son entidades, relaciones y atributos.

- **Entidades y conjunto de entidades**

Una *entidad* es un objeto que existe y puede distinguirse de otros objetos. La entidad puede ser concreta, por ejemplo: una persona o un libro; o abstracta, por ejemplo un día festivo o un concepto.

Un *conjunto de entidades* es un grupo de entidades del mismo tipo. El conjunto de todas las personas que tienen una cuenta en el banco, por ejemplo, puede definirse como el conjunto de entidades clientes.

Una entidad está representada por un conjunto de *atributos*. Los posibles atributos del conjunto de entidades clientes son nombre, documento, calle y ciudad. Para cada atributo existe un rango de valores permitidos, llamado *dominio* del atributo. El dominio del atributo nombre podría ser el conjunto de todos los nombres de personas de cierta longitud.

- **Relaciones y conjunto de relaciones**

Una *relación* es una asociación entre varias entidades. Por ejemplo es posible definir una relación que asocia al cliente Gutiérrez con la cuenta 401.

Un *conjunto de relaciones* es un grupo de relaciones del mismo tipo. Se definirá el conjunto de relaciones clientecuenta para denotar la asociación entre los clientes y las cuentas bancarias que tienen.

La relación clientecuenta es un ejemplo de una *relación binaria*, es decir, una que implica a dos conjuntos de entidades.

Existen conjuntos de relaciones que incluyen a n-conjuntos de entidades, *relaciones n-arias*, por ejemplo las relaciones tenaria clientecuentasuc que especifica que el cliente Gutiérrez tienen la cuenta 401 en la sucursal Córdoba.

Los *relaciones recursivas* son relaciones binarias que conectan una entidad consigo misma.

Una relación también puede tener *atributos descriptivos o rótulos*. Por ejemplo, fecha podría ser un atributo del conjunto de relaciones clientecuenta. Esto especifica la última fecha en que el cliente tuvo acceso a su cuenta.

- **Cardinalidades de mapeo**

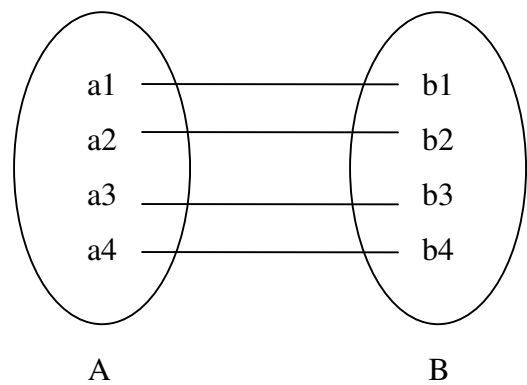
Un esquema ER empresarial puede definir ciertas limitantes con las que deben cumplir los datos contenidos en la base de datos. Una limitante importante es la de las *cardinalidades*

de mapeo que expresan el número de entidades con las que puede asociarse otra entidad mediante una relación.

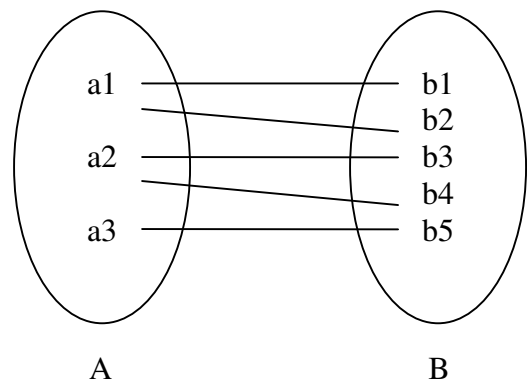
Las cardinalidades de mapeo son más útiles al describir conjuntos binarios de relaciones, aunque también son aplicables a conjuntos n-arios de relaciones.

Para un conjunto binario de relaciones R entre los conjuntos de entidades A y B, la cardinalidad de mapeo puede ser:

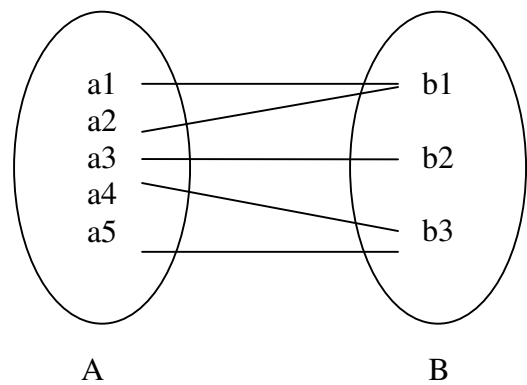
Una a una: una entidad de A está asociada únicamente con una entidad de B y una entidad de B está asociada solo con una entidad de A.



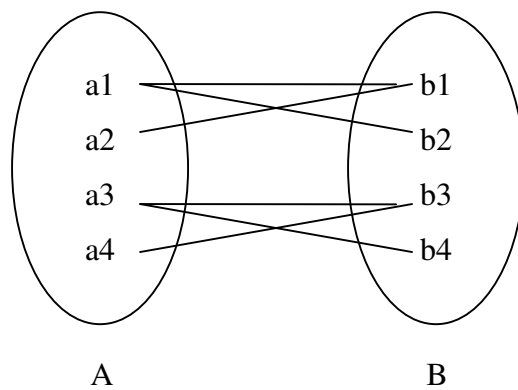
Una a muchas: una entidad en A está asociada con varias entidades de B, pero una entidad de B puede asociarse únicamente con una entidad de A.



Muchas a una: una entidad de A está asociada únicamente con una entidad en B, pero una entidad de B está relacionada con varias entidades de A.



Muchas a muchas: una entidad en A está asociada con varias entidades de B y una entidad en B está vinculada con varias entidades de A.



Para ilustrar lo anterior, considérese el conjunto de relaciones clientecuenta. Si en un banco dado una cuenta puede pertenecer únicamente a un cliente y un cliente puede tener varias cuentas, entonces el conjunto de relaciones clientecuenta es una a muchas, de cliente a cuenta. Si una cuenta puede pertenecer a varios clientes, entonces el conjunto de relaciones clientecuenta es una a muchas, de cuenta a cliente, entonces en definitiva el conjunto de relaciones clientecuenta es muchas a muchas.

Las **dependencias de existencia** constituyen otra clase importante de limitantes. Si la existencia de la entidad x depende de la existencia de la entidad y, entonces se dice que x es dependiente por existencia de y. Funcionalmente esto quiere decir que si se elimina y, también se eliminará x. Se dice que la entidad y es una entidad dominante y que x es una entidad subordinada. Por ejemplo supongamos que tenemos los conjuntos de entidades cuenta y transacción. Se forma la relación cuentatransac entre estos dos conjuntos es decir que para una cuenta determinada pueden existir varias transacciones. Esta relación es una a muchas de cuenta a transacción. Cada entidad transacción debe estar relacionada con una entidad cuenta. Si se elimina una entidad cuenta, entonces deben eliminarse también todas las entidades transacción vinculada con esa cuenta. Por lo contrario pueden eliminarse entidades transacción de la base de datos sin afectar ninguna cuenta. Por lo tanto, el conjunto de entidades cuenta es dominante y transacción es subordinada en la relación cuentatransac.

- **Llaves primarias**

Una tarea muy importante dentro de la modelación de bases de datos consiste en especificar cómo se van a distinguir las entidades y las relaciones. Conceptualmente, las entidades individuales y las relaciones son distintas entre sí, pero desde el punto de vista de una base de datos la diferencia entre ellas debe expresarse en términos de sus atributos. Para hacer estas distinciones, se asigna una **llave primaria** a cada conjunto de entidades, esta, es un conjunto de uno o más atributos que, juntos, permiten identificar en forma única a una entidad dentro del conjunto de entidades. Por ejemplo: el atributo documento del conjunto entidades cliente es suficiente para distinguir a una entidad cliente de otra, por lo tanto puede ser la llave primaria de ese conjunto de entidades.

Es posible que un conjunto de entidades no tenga suficientes atributos para formar una llave primaria. Por ejemplo: el conjunto entidades transacción tiene tres atributos: numtransac, fecha e importe. Aunque cada entidad transacción es distinta, dos transacciones hechas en cuentas diferentes pueden tener el mismo número de transacción, entonces el conjunto entidades transacción no tienen una llave primaria. Una entidad de un conjunto de este tipo se denomina **entidad débil** y una entidad que puede tener una llave primaria recibe el nombre de

entidad fuerte. El concepto de entidades fuertes y débiles está relacionado con el de “dependencia por existencia”.

Un conjunto de entidades débiles no tienen una llave primaria sin embargo es preciso tener una forma de distinguir entre todas las entidades del conjunto, aquella que depende de una entidad fuerte de otro conjunto relacionado. El **discriminador** de un conjunto de entidades débiles es un conjunto de atributos que permite hacer esta distinción. Por lo tanto para nuestro ejemplo el discriminador es numtransac ya que para cada cuenta estos números identifican en forma única cada una de las transacciones.

La llave primaria de un conjunto de entidades débiles está formada por la llave primaria de la entidad fuerte de la que dependen su existencia y su discriminador. En el caso del conjunto de entidades transacción, su llave primaria es (numcuenta, numtransac), donde numcuenta identifica a la entidad dominante de una transacción y numtransac distingue a las entidades transacción dentro de la misma cuenta.

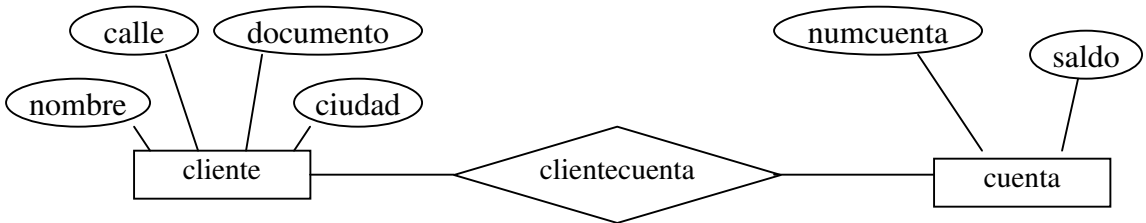
Los conjuntos de relaciones también tienen llaves primarias. Sus llaves primarias se forman tomando todos los atributos que constituyen las llaves primarias de los conjuntos de entidades que definen el conjunto de relaciones. Por ejemplo: documento es la llave primaria de cliente y numcuenta es la llave primaria de cuenta. Por lo tanto, la llave primaria del conjunto de relaciones clientecuenta es (documento, numcuenta).

Diagrama entidad-relación

La estructura lógica general de una base de datos puede expresarse en forma gráfica por medio de un **diagrama ER** que se integra con los siguientes componentes:

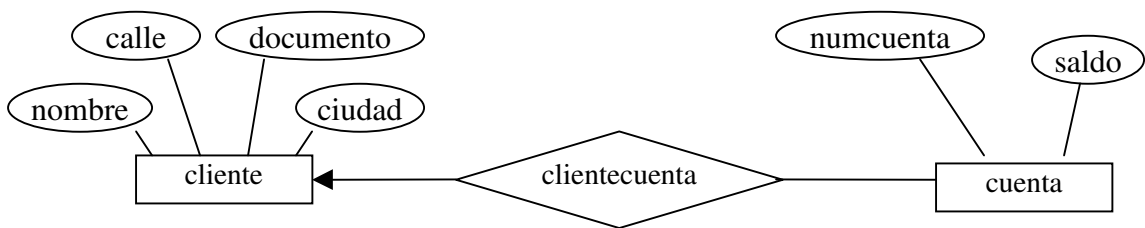
- Rectángulos:** representan conjuntos de entidades.
- Elipses:** representan atributos.
- Rombos:** representa conjuntos de relaciones.
- Líneas:** conectan los atributos a los conjuntos de entidades, y los conjuntos de entidades a los conjuntos de relaciones.

Cada componente se etiqueta con su nombre correspondiente.
Para ilustrar lo anterior veremos los siguientes ejemplos.
En este ejemplo se ve conjuntos de entidades, cliente y cuenta, vinculados entre sí mediante un conjunto binario de relaciones clientecuenta.

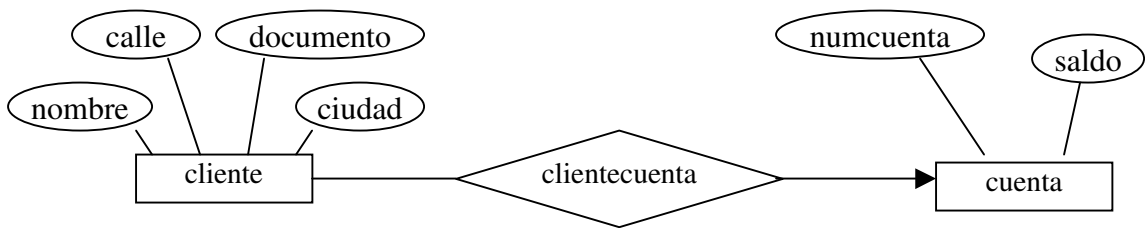


Para distinguir las cardinalidades de las relaciones se dibuja líneas con y sin dirección. En el ejemplo anterior puede verse que el conjunto de relaciones clientecuenta es muchas a muchas (al carecer de dirección, las líneas).

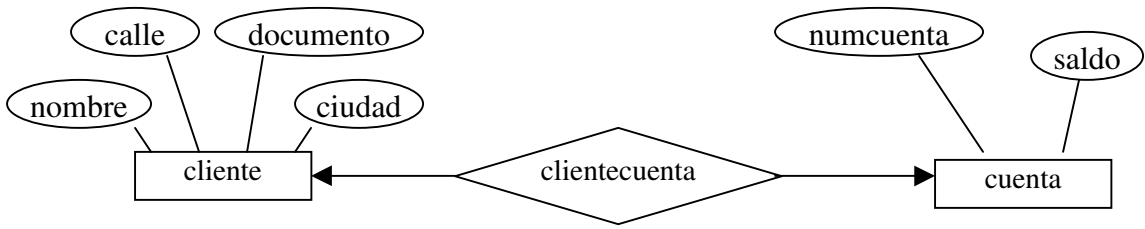
Si el conjunto de relaciones clientecuenta fuera una a muchas, de cliente a cuenta, entonces la conexión cliente tendría una flecha que apuntaría al conjunto de entidades clientes como lo muestra el siguiente gráfico.



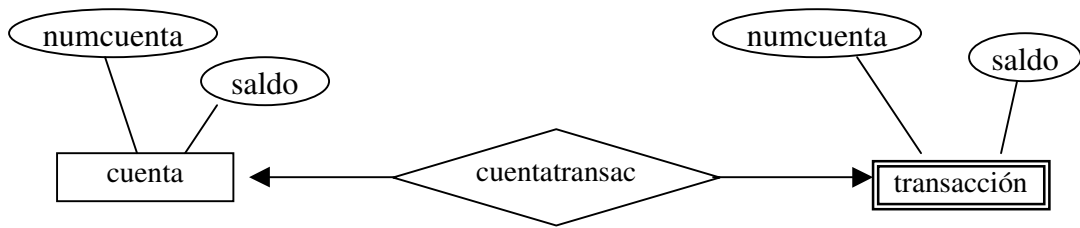
De manera similar, si el conjunto de relaciones clientecuenta fuera muchas a una de cliente a cuenta, entonces la conexión clientecuenta tendría una flecha que apuntaría al conjunto de entidades cuenta.



Por último, si el conjunto de relaciones clientecuenta fuera una a una, entonces la conexión clientecuenta tendría dos flechas, una puntando al conjunto de entidades cuenta y otra al conjunto de entidades clientes.



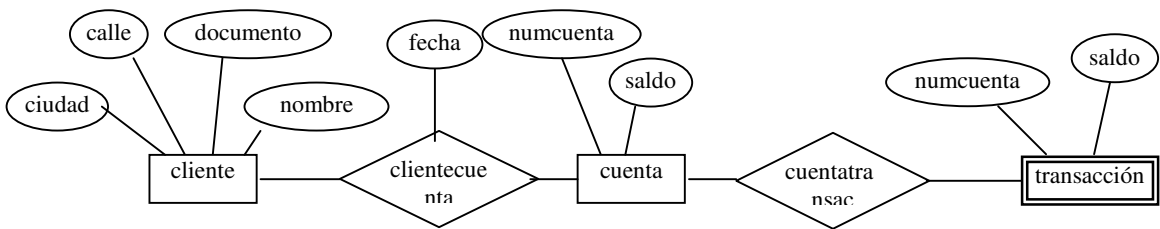
Un conjunto de entidades débiles se indica en los diagramas ER por medio de un rectángulo con doble línea la relación que la conecta al conjunto de entidades fuerte en el que se forma su llave primaria se señala mediante líneas gruesas. En el siguiente ejemplo el conjunto de entidades débiles transacción depende del conjunto de entidades fuertes cuenta, a través del conjunto de relaciones cuentatransac.



Reducción de los diagramas ER a tablas

Una base de datos que se ajuste a un diagrama ER puede representarse por medio de un conjunto de tablas. Para cada conjunto de entidades y de relaciones en la base de datos, existe una tabla única que recibe el nombre del conjunto de entidades o de relaciones correspondiente.

Se utilizará una representación tabular del siguiente diagrama ER .



Representación de conjuntos de entidades fuertes

Explicaremos este tema a través del ejemplo. Consideremos el conjunto de entidades cuenta de diagrama ER anterior, este tiene dos atributos: numcuenta y saldo; Los conjunto de entidades fuerte se representa por medio de una tabla que generalmente se la denomina con el mismo nombre del conjunto de entidades correspondiente y sus columnas corresponden a cada uno de los atributos del conjunto. Cada renglón de esta tabla hace referencia a una entidad del conjunto de entidades cuenta. Entonces la tabla cuenta tiene dos columnas numcuenta y saldo.

El renglón (259,1000) de la tabla cuenta indica que la cuenta número 259 tiene un saldo de 1000. puede agregarse una entidad nueva a la base de datos insertando un renglón en una tabla. También pueden eliminarse o modificarse renglones.

Numcuenta	Saldo
259	1000
630	2000
401	1500
700	1500
199	500
467	900
115	1200

Representación de conjunto de entidades débiles

El conjunto de entidades transacción, del ejemplo, tiene tres atributos: numtransac, fecha e importe. La llave primaria del conjunto de entidades cuenta, del cual depende transacción, es numcuenta. Así, transacción se representa por medio de una tabla con cuatro columnas llamadas numcuenta, numtransac, fecha e importe.

Numcuenta	Numtransac	Fecha	Importe
259	5	11 de mayo 2001	+50
630	11	17 de mayo 2000	+70
401	22	28 de mayo 2000	-300

700	69	3 de junio 2000	-900
-----	----	-----------------	------

Representación de conjunto de relaciones

El conjunto de relaciones clientecuenta del ejemplo implica a los dos siguiente conjuntos de entidades:

- Cliente, cuya llave primaria es documento.
- Cuenta, cuya llave primaria es numcuenta.

Puesto que el conjunto de relaciones tiene un atributo descriptivo fecha, la tabla clientecuenta tiene tres columnas, documento, numcuenta y fecha.

Documento	Numcuenta	Fecha
6.215.487	259	17 de junio 2000
17.302.548	630	12 de mayo 2001
21.215.458	401	23 de mayo 2001
18.124.256	700	3 de enero 2000

Generalización y Especialización

Son relaciones de contención que existen entre un conjunto de entidades de alto nivel y uno o más conjunto de entidades de bajo nivel.

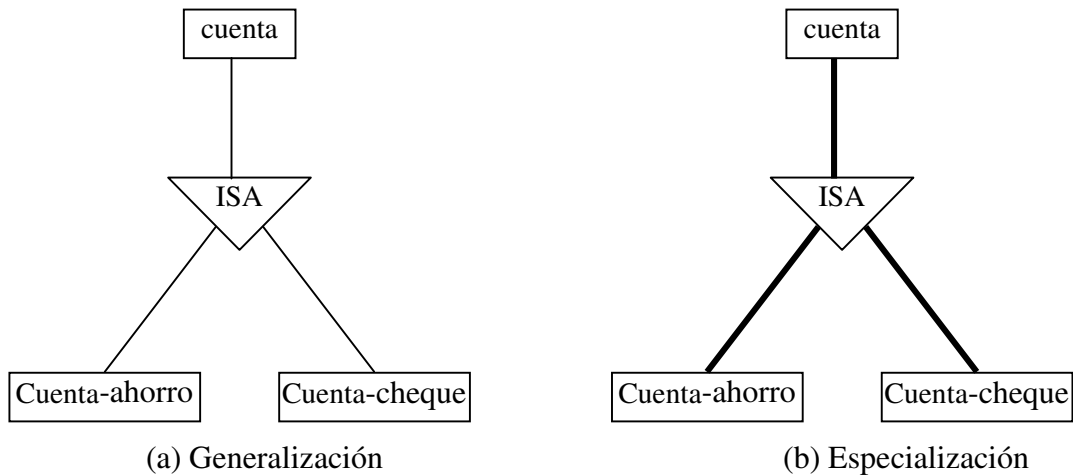
La **generalización** es el resultado de la unión de dos o más conjuntos de entidades, de bajo nivel, para producir un conjunto de entidades de más alto nivel.

La **especialización** es el resultado de tomar un subconjunto de un conjunto de entidades de alto nivel, para formar un conjunto de entidades de más bajo nivel.

Del conjunto de entidades cuenta (alto nivel), clasificamos cada una de estas en cuenta-ahorros (con atributo tasa-interés) y cuentas-cheques (con atributo importe-sobregiro), ambas de bajo nivel. Hay poca diferencia entre la generalización y la especialización. En la primera, cada entidad de alto nivel debe ser una entidad de bajo nivel. La especialización no tiene esta limitante. Por ejemplo, la generalización exigiría que cada cuenta sea una entidad cuenta-cheque o una entidad cuenta-ahorros. La especialización permite la posibilidad de que una entidad cuenta no sea ni una entidad cuenta-cheque ni una entidad cuenta-ahorros (podría ser una cuenta de mercado de valores).

La generalización se utiliza para hacer resaltar las semejanzas entre los tipos de entidades de bajo nivel y para ocultar sus diferencias. La especialización es lo inverso, hace resaltar las diferencias entre los conjuntos de entidades de alto nivel y de bajo nivel. Los atributos son lo que los distinguen. Esto se realiza mediante la herencia de atributos. Los conjuntos de bajo nivel heredan los atributos de los conjuntos de entidades de alto nivel.

En el diagrama ER tanto la generalización como la especialización se representan por medio de un triángulo “ISA” (is a, en inglés, significa “es un”). La generalización se distingue de la especialización en un diagrama ER por el mayor grosor de las líneas que conectan el triángulo ISA y cada una de las entidades.



Por ejemplo, el diagrama (a) representa que una cuenta de ahorros “es una” cuenta.

Existen dos métodos diferentes para transformar a forma tabular un diagrama ER que incluya generalización y especialización.

Primero, crear una tabla para la entidad de más alto nivel y para cada una de las entidades de bajo nivel, crear una tabla que incluya una columna para cada uno de los atributos descriptivos de esa entidad y una columna para cada atributo de la llave primaria de la entidad de alto nivel. Entonces para (a) tendremos:

Cuenta con los atributos numcuenta y saldo.

Cuenta-ahorros con los atributos numcuenta y tasa interés.

Cuenta-cheque con los atributos numcuenta e importe-sobregiro.

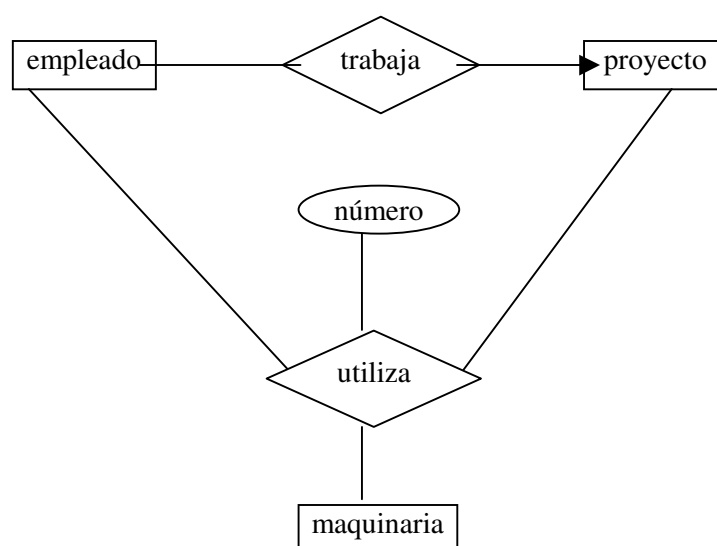
El segundo método sería, no hacer una tabla para la entidad de alto nivel, sino elaborar, para cada entidad de bajo nivel, una tabla que incluya una columna para atributo de esa entidad y una columna por cada uno de los atributos de la entidad de alto nivel, entonces para el mismo ejemplo sería:

Cuenta-ahorros con los atributos numcuenta, saldo y tasa-interés.

Cuenta-cheques con los atributos numcuenta, saldo e importe-sobregiro.

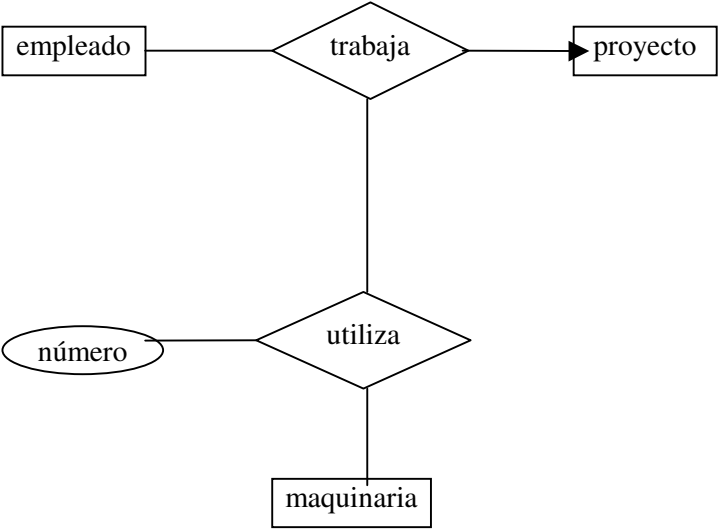
Agregación

Una limitación del modelo ER es que no es posible expresar relaciones entre relaciones. Tomemos el ejemplo de unos empleados que trabajan en un proyecto con varias máquinas diferentes. El diagrama ER sería:



Los conjuntos de relaciones trabaja y utiliza pueden combinarse para formar un solo conjunto de relaciones. Sin embargo, no se podría ver claramente la estructura lógica de este esquema.

La solución es utilizar la *agregación*, una abstracción por medio de la cual las relaciones se tratan como entidades de alto nivel. Entonces en el ejemplo, el conjunto de relaciones trabaja y los conjuntos de entidades empleado y proyecto pueden tratarse como un conjunto de entidades de más alto nivel llamado trabajo. En diagrama ER sería:



La transformación de un diagrama ER que incluya agregación a forma tabular es directa, se crean las siguientes tablas:

- Empleado
- Proyecto
- Trabaja
- Maquinaria
- Utiliza

La tabla del conjunto de relaciones utiliza incluye una columna por cada atributo de la llave primaria del conjunto de entidades maquinaria y la relación trabaja. También incluye una columna para el atributo descriptivo de la relación utiliza.