



1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación

1.5 El modelo entidad-interrelación

El modelo de datos entidad-interrelación (E-R), también llamado entidad-relación, fue propuesto por Peter Chen en 1976 para la representación conceptual de los problemas del mundo real. En 1988, el ANSI lo seleccionó como modelo estándar para los sistemas de diccionarios de recursos de información. Es un modelo muy extendido y potente para la representación de los datos. Se simboliza haciendo uso de grafos y de tablas. Propone el uso de tablas bidimensionales para la representación de los datos y sus relaciones.

Conceptos básicos

Entidad. Es un objeto del mundo real, que tiene interés para la empresa. Por ejemplo, los ALUMNOS de un centro escolar o los CLIENTES de un banco. Se representa utilizando rectángulos.

Conjunto de entidades. Es un grupo de entidades del mismo tipo, por ejemplo, el conjunto de entidades cliente. Los conjuntos de entidades no necesitan ser disjuntos, se puede definir los conjuntos de entidades de empleados y clientes de un banco, pudiendo existir una persona en ambas o ninguna de las dos cosas.

Entidad fuerte. Es aquella que no depende de otra entidad para su existencia. Por ejemplo, la entidad ALUMNO es fuerte pues no depende de otra para existir, en cambio, la entidad NOTAS es una entidad débil pues necesita a la entidad ALUMNO para existir. Las entidades débiles se relacionan con la entidad fuerte con una relación uno a varios. Se representan con un rectángulo con un borde doble.

Atributos o campos. Son las unidades de información que describen propiedades de las entidades. Por ejemplo, la entidad ALUMNO posee los atributos: número de matrícula, nombre, dirección, población y teléfono. Los atributos toman valores, por ejemplo, el atributo población puede ser ALCALÁ, GUADALAJARA, etcétera. Se representan mediante una elipse con el nombre en su interior.

Dominio. Es el conjunto de valores permitido para cada atributo. Por ejemplo el dominio del atributo nombre puede ser el conjunto de cadenas de texto de una longitud determinada.

Identificador o superclave. Es el conjunto de atributos que identifican de forma única a cada entidad. Por ejemplo, la entidad EMPLEADO, con los atributos Número de la Seguridad Social, DNI, Nombre, Dirección, Fecha nacimiento y Tlf, podrían ser identificadores o superclaves los conjuntos Nombre, Dirección, Fecha nacimiento y Tlf, o también DNI, Nombre y Dirección, o también Num Seg Social, Nombre, Dirección y Tlf, o solos el DNI y el Número de la Seguridad Social.

Clave candidata. Es cada una de las superclaves formadas por el mínimo número de campos posibles. En el ejemplo anterior, son el DNI y el Número de la Seguridad Social.

Clave primaria o principal (primary key): Es la clave candidata seleccionada por el diseñador de la BD. Una clave candidata no puede contener valores nulos, ha de ser sencilla de crear y no ha de variar con el tiempo. El atributo o los atributos que forman esta clave se representan subrayados.



Clave ajena o foránea (*foreign key*): Es el atributo o conjunto de atributos de una entidad que forman la clave primaria en otra entidad. Las claves ajenas van a representar las relaciones entre tablas. Por ejemplo, si tenemos por un lado, las entidades ARTÍCULOS, con los atributos código de artículo (clave primaria), denominación, stock. Y, por otro lado, VENTAS, con los atributos Código de venta (clave primaria), fecha de venta, código de artículo, unidades vendidas, el código de artículo es clave ajena pues está como clave primaria en la entidad ARTÍCULOS.

A. Relaciones y conjuntos de relaciones

Definimos una **relación** como la asociación entre diferentes entidades. Tienen nombre de verbo, que la identifica de las otras relaciones y se representa mediante un rombo. Normalmente las relaciones no tienen atributos. Cuando surge una relación con atributos significa que debajo hay una entidad que aún no se ha definido. A esa entidad se la llama *entidad asociada*. Esta entidad dará origen a una tabla que contendrá esos atributos. Esto se hace en el *modelo relacional a la hora de representar los datos*. Lo veremos más adelante.

Un **conjunto de relaciones** es un conjunto de relaciones del mismo tipo, por ejemplo entre ARTÍCULOS y VENTAS todas las asociaciones existentes entre los artículos y las ventas que tengan estos, forman un conjunto de relaciones.

La mayoría de los conjuntos de relaciones en un sistema de BD son binarias (dos entidades) aunque puede haber conjuntos de relaciones que implican más de dos conjuntos de entidades, por ejemplo, una relación como la relación entre cliente, cuenta y sucursal. Siempre es posible sustituir un conjunto de relaciones no binario por varios conjuntos de relaciones binarias distintos. Así, conceptualmente, podemos restringir el modelo E-R para incluir sólo conjuntos binarios de relaciones, aunque no siempre es posible.

La función que desempeña una entidad en una relación se llama *papel*, y normalmente es implícito y no se suele especificar. Sin embargo, son útiles cuando el significado de una relación necesita ser clarificado.

Una relación también puede tener atributos descriptivos, por ejemplo, la FECHA_OPERACIÓN en el conjunto de relaciones CLIENTE_CUENTA, que especifica la última fecha en la que el cliente tuvo acceso a su cuenta (ver Figura 1.3).



Figura 1.3. Relación con atributos descriptivos.

Diagramas de estructuras de datos en el modelo E-R

Los diagramas Entidad-Relación representan la estructura lógica de una BD de manera gráfica. Los símbolos utilizados son los siguientes:

- Rectángulos para representar a las entidades.



1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación

- Elipses para los atributos. El atributo que forma parte de la clave primaria va subrayado.
- Rombos para representar las relaciones.
- Las líneas, que unen atributos a entidades y a relaciones, y entidades a relaciones. Si la flecha tiene punta, en ese sentido está el uno, y si no la tiene, en ese sitio está el muchos. La orientación señala cardinalidad.
- Si la relación tiene atributos asociados, se le unen a la relación.
- Cada componente se etiqueta con el nombre de lo que representa.

En la Figura 1.4 se muestra un diagrama E-R correspondiente a PROVEEDORES-ARTÍCULOS. Un PROVEEDOR SUMINISTRA muchos ARTÍCULOS.

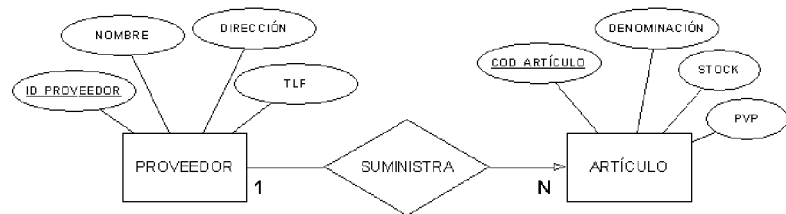


Figura 1.4. Diagrama E-R, un proveedor suministra muchos artículos.

Grado y cardinalidad de las relaciones

Se define **grado de una relación** como el número de conjuntos de entidades que participan en el conjunto de relaciones, o lo que es lo mismo, el número de entidades que participan en una relación. Las relaciones en las que participan dos entidades son *binarias* o de grado dos. Si participan tres serán *ternarias* o de grado 3. Los conjuntos de relaciones pueden tener cualquier grado, lo ideal es tener relaciones binarias.

Las relaciones en las que sólo participa una entidad se llaman *anillo* o de grado uno; relaciona una entidad consigo misma, se las llama *relaciones reflexivas*. Por ejemplo, la entidad EMPLEADO puede tener una relación JEFE DE consigo misma: un empleado es JEFE DE muchos empleados y, a la vez, el jefe es un empleado.

Otro ejemplo puede ser la relación DELEGADO DE los alumnos de un curso: el delegado es alumno también del curso. Ver Figura 1.5.

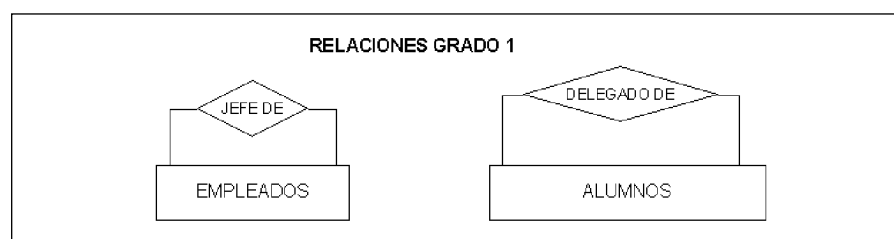


Figura 1.5. Relaciones de grado 1.

1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación



En la Figura 1.6 se muestra una relación de grado dos, que representa un proveedor que suministra artículos, y otra de grado tres, que representa un cliente de un banco que tiene varias cuentas, y cada una en una sucursal:

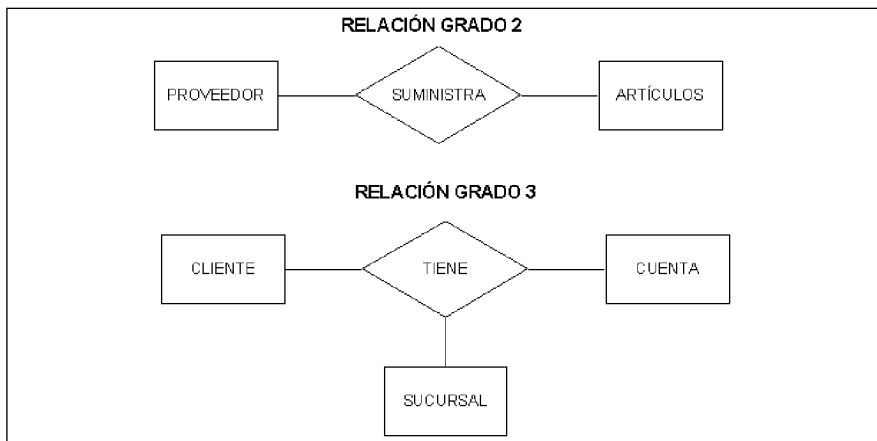


Figura 1.6. Relaciones de grados 2 y 3.

En el modelo E-R se representan ciertas restricciones a las que deben ajustarse los datos contenidos en una BD. Éstas son las restricciones de las cardinalidades de asignación, que expresan el número de entidades a las que puede asociarse otra entidad mediante un conjunto de relación.

Las cardinalidades de asignación se describen para conjuntos binarios de relaciones. Son las siguientes:

- **1:1, uno a uno.** A cada elemento de la primera entidad le corresponde sólo uno de la segunda entidad, y a la inversa. Por ejemplo, un cliente de un hotel ocupa una habitación, o un curso de alumnos pertenece a un aula, y a esa aula sólo asiste ese grupo de alumnos. Ver Figura 1.7:

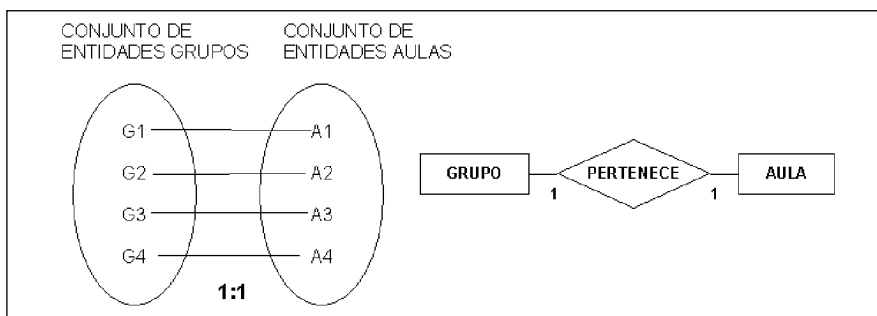


Figura 1.7. Representación de relaciones uno a uno.

- **1:N, uno a muchos.** A cada elemento de la primera entidad le corresponde uno o más elementos de la segunda entidad, y a cada elemento de la segunda entidad le corresponde uno sólo de la primera entidad. Por ejemplo, un proveedor suministra muchos artículos (ver Figura 1.8).



1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación

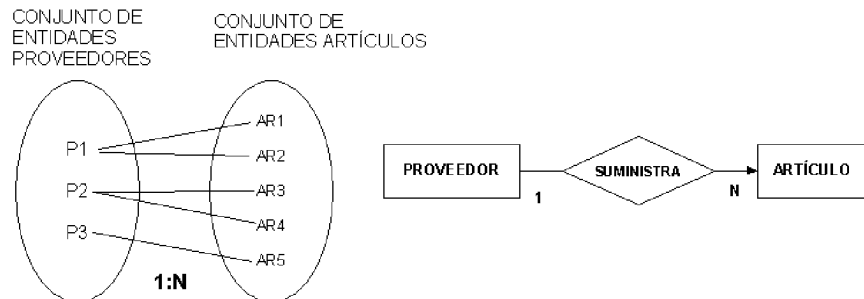


Figura 1.8. Representación de relaciones uno a muchos.

- **N:1, muchos a uno.** Es el mismo caso que el anterior pero al revés; a cada elemento de la primera entidad le corresponde un elemento de la segunda, y a cada elemento de la segunda entidad, le corresponden varios de la primera.
- **M:N, muchos a muchos.** A cada elemento de la primera entidad le corresponde uno o más elementos de la segunda entidad, y a cada elemento de la segunda entidad le corresponde uno o más elementos de la primera entidad. Por ejemplo, un vendedor vende muchos artículos, y un artículo es vendido por muchos vendedores (ver Figura 1.9).

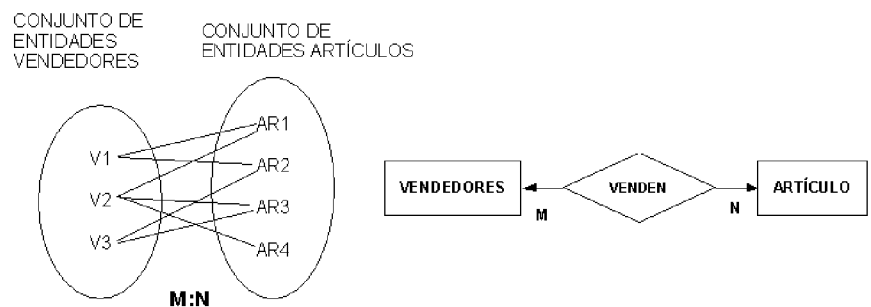


Figura 1.9. Representación de relaciones muchos a muchos.

La **cardinalidad de una entidad** sirve para conocer su grado de participación en la relación, es decir, el número de correspondencias en las que cada elemento de la entidad interviene. Mide la obligatoriedad de correspondencia entre dos entidades.

La **representamos entre paréntesis indicando los valores máximo y mínimo: (máximo, mínimo)**. Los valores para la cardinalidad son: (0,1), (1,1), (0,N), (1,N) y (M,N). El valor 0 se pone cuando la participación de la entidad es opcional.

En la Figura 1.10, que se muestra a continuación, se representa el diagrama E-R en el que contamos con las siguientes entidades:

- EMPLEADO está formada por los atributos N°. Emple, Apellido, Salario y Comisión, siendo el atributo N°. Emple la clave principal (representado por el subrayado).
- DEPARTAMENTO está formada por los atributos N°. Depart, Nombre y Localidad, siendo el atributo N°. Depart la clave principal.

1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación



- Se han definido dos relaciones:
 - La relación «PERTENECE» entre las entidades EMPLEADOS y DEPARTAMENTO, cuyo tipo de correspondencia es 1:N, es decir, a un departamento le pertenecen cero o más empleados (0,N). Un empleado pertenece a un departamento y sólo a uno (1,1).
 - La relación «JEFE», que asocia la entidad EMPLEADO consigo misma. Su tipo de correspondencia es 1:N, es decir, un empleado es jefe de cero o más empleados (0,N). Un empleado tiene un jefe y sólo uno (1,1). Ver Figura 1.10:

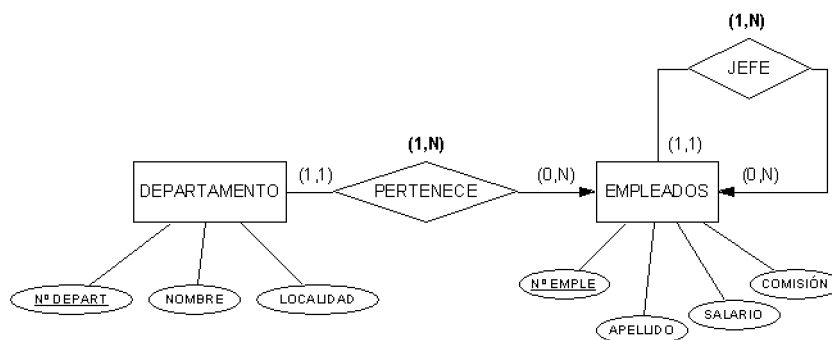


Figura 1.10. Diagrama E-R de las relaciones entre departamentos y empleados.

Caso práctico



- 1** Vamos a realizar el diagrama de estructuras de datos en el modelo E-R. Supongamos que en un centro escolar se imparten muchos cursos. Cada curso está formado por un grupo de alumnos, de los cuales uno de ellos es el delegado del grupo. Los alumnos cursan asignaturas, y una asignatura puede o no ser cursada por los alumnos.

Para su resolución, primero identificaremos las entidades, luego las relaciones y las cardinalidades y, por último, los atributos de las entidades y de las interrelaciones, si las hubiera.

1. Identificación de entidades: una entidad es un objeto del mundo real, algo que tiene interés para la empresa. Se hace un análisis del enunciado, de donde sacaremos los candidatos a entidades: CENTROS, CURSOS, ALUMNOS, ASIGNATURAS, DELEGADOS. Si analizamos esta última veremos que los delegados son alumnos, por lo tanto, los tenemos recogidos en ALUMNOS. Esta posible entidad la eliminaremos. También eliminaremos la posible entidad CENTROS pues se trata de un único centro, si se tratara de una gestión de centros tendría más sentido incluirla.
2. Identificar las relaciones: construimos una matriz de entidades en la que las filas y las columnas son los nombres de entidades y cada celda puede contener o no la relación, las relaciones aparecen explícitamente en el enunciado. En este ejemplo, las relaciones no tienen atributos. Del enunciado sacamos lo siguiente:
 - Un curso está formado por muchos alumnos. La relación entre estas dos entidades la llamamos PERTENECE, pues a un curso pertenecen muchos alumnos, relación 1:M. Consideramos que es obligatorio que existan alumnos en un curso. Para calcular los máximos y mínimos hacemos la pregunta: a un CURSO, ¿cuántos ALUMNOS pertenecen, como mínimo y como máximo? Y se ponen los valores en la entidad ALUMNOS, en este caso (1,M). Para el sentido contrario, hacemos lo mismo: un ALUMNO, ¿a cuántos CURSOS va a pertenecer? Como mínimo a 1, y como máximo a 1, en este caso pondremos (1,1) en la entidad CURSOS.

(Continúa)



1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación

(Continuación)

- De los alumnos que pertenecen a un grupo, uno de ellos es DELEGADO. Hay una relación de grado 1 entre la entidad ALUMNO que la podemos llamar ES DELEGADO. La relación es 1:M, un alumno es delegado de muchos alumnos. Para calcular los valores máximos y mínimos preguntamos: ¿un ALUMNO de cuántos alumnos ES DELEGADO? Como mínimo es 0, pues puede que no sea delegado, y como máximo es M, pues si es delegado lo será de muchos; pondremos en el extremo (0,M). Y en el otro extremo pondremos (1,1), pues obligatoriamente el delegado es un alumno.
- Entre ALUMNOS y ASIGNATURAS surge una relación N:M, pues un alumno cursa muchas asignaturas y una asignatura es cursada por muchos alumnos. La relación se llamará CURSA. Consideramos que puede haber asignaturas sin alumnos. Las cardinalidades serán (1:M) entre ALUMNO-ASIGNATURA, pues un alumno, como mínimo, cursa una asignatura, y, como máximo, muchas. La cardinalidad entre ASIGNATURA-ALUMNO será (0,N), pues una ASIGNATURA puede ser cursada por 0 alumnos o por muchos.

En la Tabla 1.2 se muestra la matriz de entidades y relaciones entre ellas:

	CURSOS	ALUMNOS	ASIGNATURAS
CURSOS	-----	PERTENECE (1:M)	-----
ALUMNOS	x	ES DELEGADO(1:M)	CURSA(N:M)
ASIGNATURAS	-----	x	-----

Tabla 1.2. Matriz de entidades y relaciones entre ellas.

Las celdas que aparecen con una x indican que las relaciones están ya identificadas. Las que aparecen con guiones indican que no existe relación. En la siguiente figura se muestra el diagrama de las relaciones y las cardinalidades.

- Identificar los atributos, como el enunciado no explicita ningún tipo de característica de las entidades nos imaginamos los atributos, que pueden ser los siguientes:

CURSOS - COD_CURSO (clave primaria), DESCRIPCIÓN, NIVEL, TURNO y ETAPA

ALUMNOS - NUM-MATRÍCULA (clave primaria), NOMBRE, DIRECCIÓN, POBLACIÓN, TLF y NUM_HERMANOS

ASIGNATURAS - COD-ASIGNATURA (clave primaria), DENOMINACIÓN y TIPO

En la Figura 1.11 se representa el diagrama de estructuras del ejercicio:

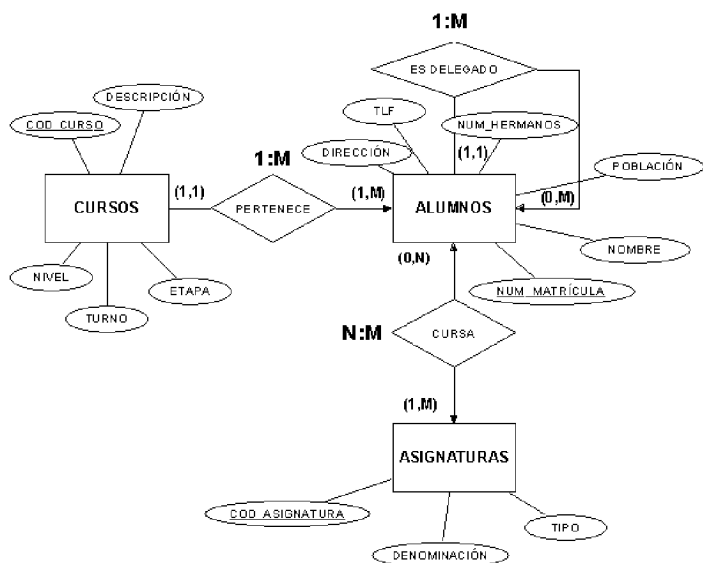


Figura 1.11. Diagrama de estructuras en el modelo E-R.



Actividades propuestas



1 Se desea realizar el diagrama de estructuras de datos en el modelo E-R, correspondiente al siguiente enunciado:

Supongamos el bibliobús que proporciona un servicio de préstamos de libros a los socios de un pueblo. Los libros están clasificados por temas. Un tema puede contener varios libros. Un libro es prestado a muchos socios, y un socio puede coger varios libros. En el préstamo de libros es importante saber la Fecha de préstamo y la Fecha de devolución. De los libros nos interesa saber el título, el autor y el número de ejemplares.

Generalización y jerarquías de generalización

Las generalizaciones proporcionan un mecanismo de abstracción que permite especializar una entidad (que se denominará *supertipo*) en subtipos, o lo que es lo mismo generalizar los subtipos en el supertipo.

Una generalización se identifica si encontramos una serie de atributos comunes a un conjunto de entidades, y unos atributos específicos que identificarán unas características.

Los atributos comunes describirán el supertipo y los particulares los subtipos. Una de las características más importantes de las jerarquías es la herencia, por la que los atributos de un supertipo son heredados por sus subtipos. Si el supertipo participa en una relación los subtipos también participarán.

Por ejemplo, en una empresa de construcción podremos identificar las siguientes entidades:

- EMPLEADO, con los atributos N_EMPLE (clave primaria,) NOMBRE, DIRECCIÓN, FECHA_NAC, SALARIO y PUESTO.
- ARQUITECTO, con los atributos de empleado más los atributos específicos: COMISIONES, y NUM_PROYECTOS.
- ADMINISTRATIVO, con los atributos de empleado más los atributos específicos: PULSACIONES y NIVEL.
- INGENIERO, con los atributos de empleado más los atributos específicos: ESPECIALIDAD y AÑOS_EXPERIENCIA.

En la Figura 1.12 se representa este ejemplo de generalización.

La generalización es total si no hay ocurrencias en el supertipo que no pertenezcan a ninguno de los subtipos, es decir, que los empleados o son arquitectos, o son administrativos, o son aparejadores, no pueden ser varias cosas a la vez. En este caso, la generalización sería también exclusiva. Si un empleado puede ser varias cosas a la vez la generalización es *solapada* o *superpuesta*.

La generalización es *parcial* si existen empleados que no son ni ingenieros, ni administrativos, ni arquitectos. También puede ser *exclusiva* o *solapada*. Las cardinalidades en estas relaciones son siempre (1,1) en el supertipo y (0,1) en los subtipos, para las exclusivas. (0,1) o (1,1) en los subtipos para las solapadas o superpuestas.



1. Sistemas gestores de bases de datos

1.5 El modelo entidad-interrelación

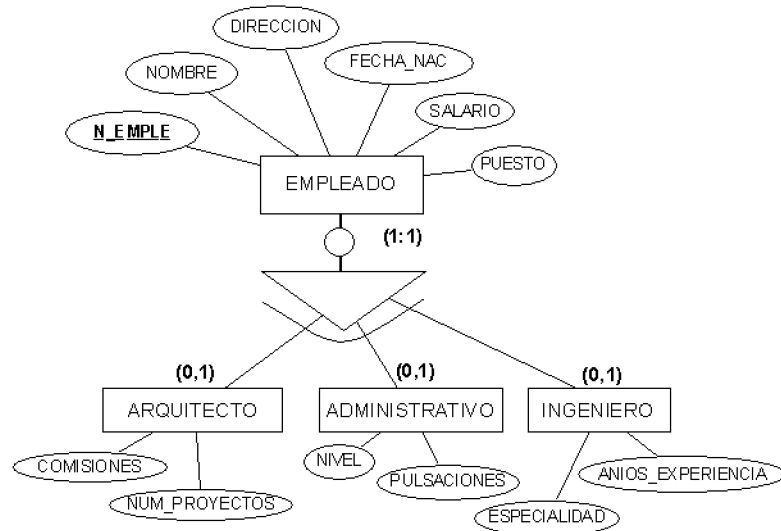


Figura 1.12. Representación de una generalización.

Así pues, habrá jerarquía solapada y parcial (que es la que no tiene ninguna restricción) solapada y total, exclusiva y parcial, y exclusiva y total. En la Figura 1.13 se muestran cómo se representan.

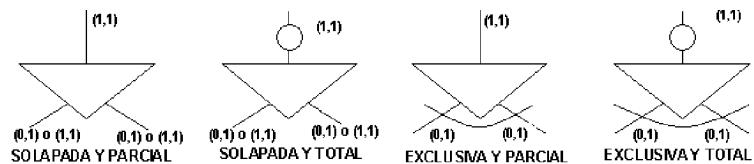


Figura 1.13. Tipos y representación de jerarquías.



Actividades propuestas

2 Representa las siguientes jerarquías e indica el tipo de generalización:

Un concesionario de coches vende coches nuevos y usados. Los atributos específicos de los nuevos son las unidades y el descuento; de los usados son los kilómetros y el año de fabricación.

Consideramos el conjunto de personas de una ciudad, distinguimos a los trabajadores, estudiantes y parados. De los trabajadores nos interesa el número de la Seguridad Social, la empresa de trabajo y el salario. De los estudiantes, el número de matrícula y el centro educativo, y de los parados la fecha del paro.

En un campo de fútbol los puestos de los futbolistas pueden ser: portero, defensa, medio y delantero.



Agregación

Una limitación del modelo E-R es que no es posible expresar relaciones entre relaciones. En estos casos se realiza una agregación, que es una abstracción a través de la cual las relaciones se tratan como entidades de nivel más alto. Por ejemplo, consideramos una relación entre EMPLEADOS y PROYECTOS, un empleado trabaja en varios proyectos durante unas horas determinadas y en ese trabajo utiliza unas herramientas determinadas. La representación del diagrama de estructuras se muestra en la Figura 1.14 siguiente:

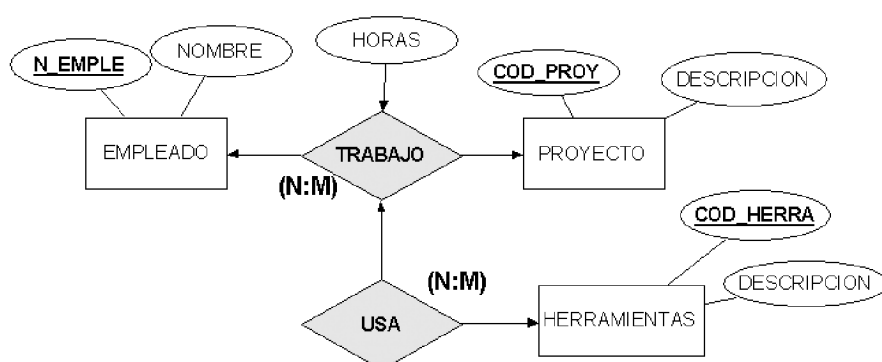


Figura 1.14. Diagrama E-R de una relación entre otra relación.

Si consideramos la agregación, tenemos que la relación TRABAJO con las entidades EMPLEADO y PROYECTO se pueden representar como un conjunto de entidades llamadas TRABAJO, que se relacionan con la entidad HERRAMIENTAS mediante la relación USA. Ver Figura 1.15:

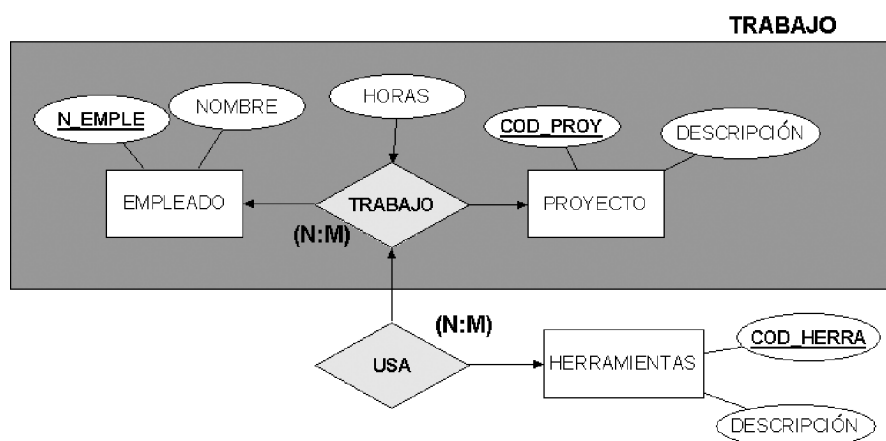


Figura 1.15. Conjunto de entidades y relaciones para representar una relación entre una relación.



1. Sistemas gestores de bases de datos

1.6 Modelo de red



Actividades propuestas

3 Se desea realizar el diagrama de estructuras de datos en el modelo E-R correspondiente al siguiente enunciado:

Una compañía de distribución de productos para el hogar dispone de proveedores que le suministran artículos. Un artículo sólo puede proveerlo un proveedor.

La empresa tiene tres tipos de empleados: oficinistas, transportistas y vendedores. Estos últimos venden los artículos. Un artículo es vendido por varios vendedores, y un vendedor puede vender varios artículos en distintas zonas de venta. De las ventas nos interesa saber la fecha de venta y las unidades vendidas.

1.6 Modelo de red

Este modelo utiliza estructuras de datos en red, también conocidas como *estructuras plex*. Las entidades se representan como registros o nodos, y las relaciones como enlaces o punteros. En una estructura red cualquier componente puede vincularse con cualquier otro. Es posible describirla en términos de padres e hijos, pero, a diferencia del modelo jerárquico, un hijo puede tener varios padres.

Las representaciones lógicas basadas en árboles o en estructuras plex, a menudo, limitan el cambio que el crecimiento de la BD exige, hasta tal punto que las representaciones lógicas de los datos pueden variar afectando a los programas de aplicación que usan esos datos. Los conceptos básicos de este modelo son los siguientes:

- **Elemento:** es un campo de datos. Ejemplo: DNI.
- **Agregados de datos:** conjunto de datos con nombre. Ejemplo: Fecha (día, mes, año).
- **Tipos de registro:** representa un nodo, un conjunto de campos. Cada campo contiene elementos. Es la unidad básica de acceso y manipulación. Se asemeja a los registros en archivos o las entidades en el modelo E-R.
- **Conjunto:** colección de dos o más tipos de registro que establece una vinculación entre ellos. Uno de ellos se llama propietario y el otro, miembro. Tienen una relación muchos a muchos (M:M), que para representarla se necesita un registro conector. Los conjuntos están formados por un solo registro propietario y uno o más registros miembros. Un registro propietario no puede ser a la vez miembro de sí mismo. Una ocurrencia del conjunto está formada por un registro propietario y el resto son registros miembros. Una ocurrencia de registro no puede pertenecer a varias ocurrencias del mismo conjunto.
- **Ciclo:** se forma cuando un registro miembro tiene como descendientes a uno de sus antepasados.
- **Bucle, lazo o loop:** es un ciclo en el que los registros propietarios y miembros son el mismo.