

# Modelado entidad relación conceptual

## Objetivo

---

Los modelos de datos conceptuales son medios para representar la información de un problema en un alto nivel de abstracción.

A partir de la definición de modelos de datos conceptuales es posible captar las necesidades del cliente/usuario respecto del SI que necesita. Esta descripción de la realidad permite mejorar la interacción usuario-desarrollador, dado que disminuye la brecha existente entre la realidad del problema a resolver y el sistema a desarrollar.

Los modelos conceptuales ER utilizan las bases planteadas por Chen en 1976, para construir la representación básica de las estructuras de datos que darán, posteriormente, soporte a la BD a construir.

En este capítulo, se desarrollan los conceptos necesarios para generar un modelo conceptual ER, con ejemplos de cada una de las situaciones generadas. Para generar los modelos se utilizará una herramienta denominada CASER, la cual fue desarrollada como trabajo de investigación por docentes y alumnos relacionados con la asignatura Introducción a las Bases de Datos, de la Facultad de Informática de la Universidad Nacional de La Plata.

Esta herramienta, de uso libre, permite generar modelos conceptuales de datos, con la metodología, la sintaxis y la semántica definidas en este capítulo.

La creación de la herramienta se debe a que, en la actualidad, no se dispone de ningún *software* que sea de utilidad para la definición de modelos conceptuales en términos estrictamente académicos.

## Características del modelo conceptual

Un modelo conceptual debe poseer cuatro características o propiedades básicas. Estas son (por orden alfabético, no de importancia):

- **Expresividad**, es decir, capturar y presentar de la mejor forma posible la semántica de los datos del problema a resolver.
- **Formalidad**, que requiere que cada elemento representado en el modelo sea preciso y bien definido, con una sola interpretación posible. Esta formalidad es comparable a la formalidad matemática.
- **Minimalidad**, característica que establece que cada elemento del modelo conceptual tiene una única forma de representación posible y no puede expresarse mediante otros conceptos.
- **Simplicidad**, que establece que el modelo debe ser fácil de entender por el cliente/usuario y el desarrollador.

Entre las propiedades definidas es posible analizar que algunas de ellas son opuestas. Por ejemplo, un modelo expresivo (rico en conceptos) puede no ser simple; es decir, en pos de expresar mejor los datos de un problema se puede perder simpleza.

Por último, sería deseable incorporar un rasgo adicional: que el modelo sea fácil de leer. Este concepto podría considerarse un corolario de todas las propiedades anteriores. Un modelo expresivo, formal, mínimo y simple debería ser fácil de leer.

## Componentes del modelo conceptual

En 1976, Chen planteó tres constructores básicos para generar el modelo de datos ER. Estos tres constructores son: entidades, relaciones y atributos. A lo largo de este apartado, se desarrollarán con detalle estos tres conceptos y sus características asociadas. Además, serán generados una serie de ejemplos que ilustren su uso y comportamiento. Nuevamente, para los ejemplos de este apartado se utilizará CASER, herramienta desarrollada en conjunto por los mismos autores de este libro.

### Entidades

Una entidad representa un elemento u objeto del mundo real con identidad, es decir, se diferencia únicamente de cualquier otro objeto o cosa, incluso siendo del mismo tipo. Ejemplos:

- Un alumno (se diferencia de cualquier otro alumno).
- Un vehículo (aunque sean de la misma marca o del mismo modelo, tendrán atributos diferentes, por ejemplo, el número de motor).
- Una materia de una carrera universitaria.

En los ejemplos anteriores se observa que hay entidades que representan objetos o elementos reales y concretos (autos, alumnos), o de existencia virtual o abstracta (materia).

A continuación, se presentan algunas entidades posibles.

*Juan García, Calle 6 N° 1213, La Plata*

*Andrea Castro, Calle 56 N° 231, Berisso*

*Inés Tartaila, Montevideo N° 123, Ensenada*

*Renault Simbol, azul, BXE 123*

*Peugeot 307, verde, HHI 343*

*Fiat Palio, rojo, IAA 990*

*Análisis Matemático I, primer año*

*Introducción a las Bases de Datos, segundo año*

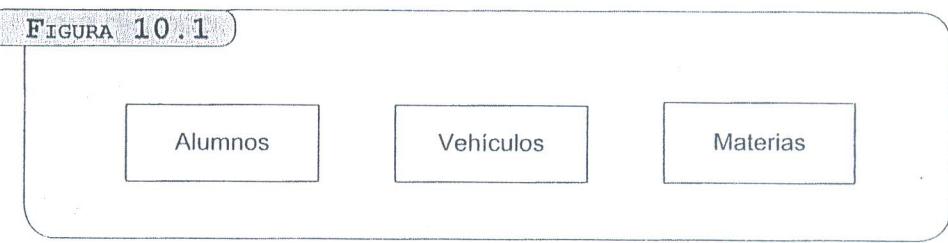
*Ingeniería de Software, tercer año*

Durante el proceso de modelización de datos, es necesario identificar las entidades, pero resulta muy complejo representar cada una de esas entidades que se distinguen en el problema a resolver. Si se desea modelizar todos los alumnos de un curso numeroso, es poco viable referenciar a cada uno de ellos. Por este motivo, una vez reconocidas las entidades, el diseñador del modelo de datos busca las propiedades comunes de estas, y genera una representación que aglutina a esas entidades. Esta representación se denomina conjunto de entidades.

Un **conjunto de entidades** es una representación que, a partir de las características propias de cada entidad con propiedades comunes, se resume en un núcleo.

La forma de representar un conjunto de entidades es mediante un rectángulo. La Figura 10.1 muestra la representación de tres conjuntos de entidades que resumen a los alumnos, vehículos y materias universitarias antes descriptos.

FIGURA 10.1



Para reconocer las entidades del problema que permitan luego modelar los conjuntos de entidades, en general se analizan los sustantivos que conforman la especificación del problema.

## Relaciones

Las relaciones representan agregaciones entre dos (binaria) o más entidades. Describen las dependencias o asociaciones entre dichas entidades. Por ejemplo, la entidad Dolores García cursa (relación) la materia Introducción a las Bases de Datos (otra entidad).

Siguiendo el análisis expresado anteriormente para las entidades, en el proceso de modelado es complejo distinguir y representar cada una de las relaciones. Aquellas que tienen propiedades comunes se agrupan en lo que se denomina conjunto de relaciones.

**Un conjunto de relaciones** es una representación que, a partir de las características propias de cada relación existente entre dos entidades, las resume en un núcleo.

La forma de representar un conjunto de relaciones es mediante un rombo. La Figura 10.2 a) muestra la representación del conjunto de relaciones existente entre el conjunto de entidades alumnos y materias. La Figura 10.2 b) de dicho gráfico muestra el conjunto de relaciones propiedad, existente entre los conjuntos de entidades vehículo y propietario.

FIGURA 10.2



Figura 10.2 a)



Figura 10.2 b)

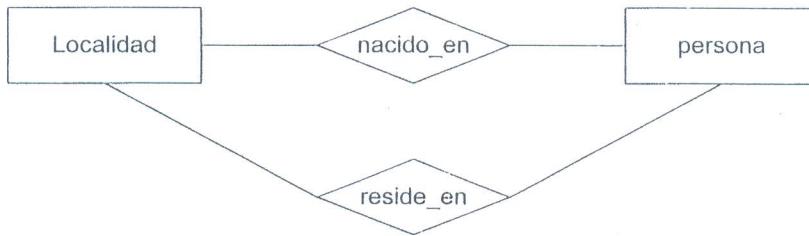


Figura 10.2 c)

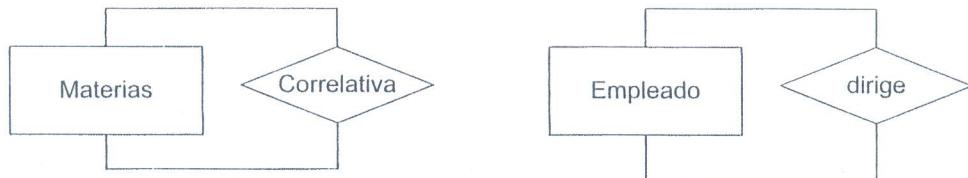


Figura 10.2 d)

Para reconocer las relaciones del problema, que permitan luego modelar los conjuntos de relaciones en general, se analizan los verbos que conforman la especificación del problema.

Pueden existir más de un conjunto de relaciones entre dos conjuntos de entidades. La Figura 10.2 c) presenta un ejemplo, donde los conjuntos de entidades persona y conjuntos de entidades ciudad poseen dos conjuntos de relaciones posibles: nacido\_en y reside\_en. El primero de ellos relaciona cada persona con su lugar de nacimiento, en tanto que el segundo establece el nexo entre cada persona y su lugar actual de residencia.

En el resto de este libro, por razones prácticas, cada vez que se haga referencia a un conjunto de entidades o a un conjunto de relaciones, se lo denominará entidad o relación directamente.

Por último, pueden existir relaciones recursivas entre entidades. Se denomina relación recursiva a aquella relación que une dos entidades particulares del mismo conjunto. La Figura 10.2 d) presenta dos ejemplos de dicha situación. En el primero de ellos se define la relación correlatividad. Esta relación establece que para cursar una materia se deben tener aprobadas otras materias. El segundo ejemplo presenta la relación dirige o es\_jefe, entre dos entidades empleados de una empresa.

En el Capítulo 9, se desarrolló el tema de abstracciones. Una de dichas abstracciones es la agregación. Cuando se definió la agregación, se incorporó una propiedad que se denominó cardinalidad. La cardinalidad define el grado de relación existente en una agregación. Este concepto está presente nuevamente, cuando se definen las relaciones existentes entre entidades. Así, cada relación (como es una agregación hasta ahora binaria) debe tener definida la cardinalidad máxima y mínima. La forma de representar esta cardinalidad se exemplifica en la Figura 10.3. Esta figura repite los ejemplos de la figura anterior, mostrando en cada caso la cardinalidad máxima y mínima existente.

FIGURA 10.3

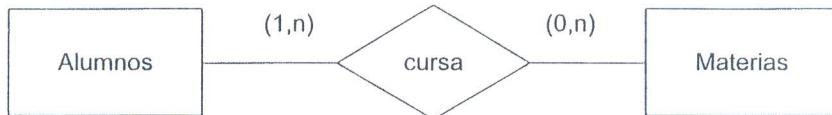


Figura 10.3 a)



Figura 10.3 b)

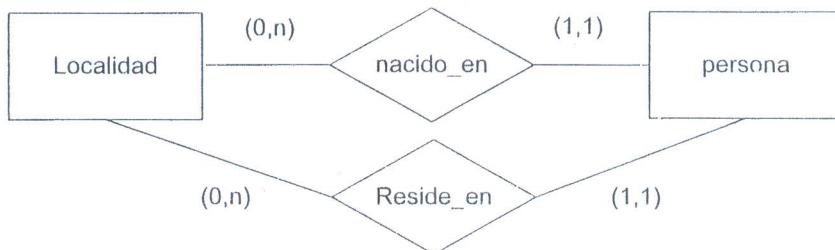


Figura 10.3. c)

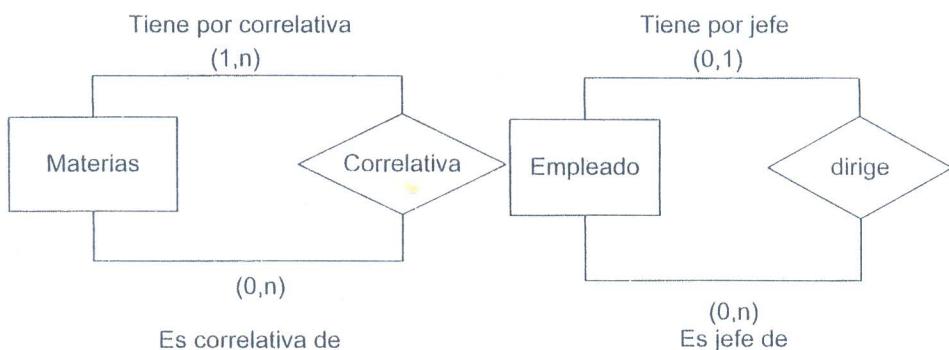


Figura 10.3 d)

La Figura 10.3 a) indica que:

$C_{\text{Min}}(\text{alumnos, cursa}) = 1$

$C_{\text{Max}}(\text{alumnos, cursa}) = n$

$C_{\text{Min}}(\text{materia, cursa}) = 0$

$C_{\text{Max}}(\text{materia, cursa}) = n$

Esto muestra que un alumno debe cursar al menos una materia (obligatoriamente), pero puede cursar varias. Además, una materia puede no ser cursada (opcional) por ningún alumno o ser cursada por varios.

Se debe notar que la cardinalidad define el grado de relación existente. Es fundamental, en la construcción de un modelo, definir con precisión la cardinalidad existente. Se podría dudar sobre algún valor de cardinalidad; en ese caso, el analista está obligado a consultar al cliente/usuario que plantea el problema, ya que la decisión tomada no debe ser unilateral.

La Figura 10.3 b) es similar al caso anterior presentado. En la Figura 10.3 c) existen dos relaciones; por lo tanto, se deben analizar más cardinalidades:

$C_{\text{Min}}(\text{persona, nacido\_en}) = 1$

$C_{\text{Max}}(\text{persona, nacido\_en}) = 1$

$C_{\text{Min}}(\text{localidad, nacido\_en}) = 0$

$C_{\text{Max}}(\text{localidad, nacido\_en}) = n$

$C_{\text{Min}}(\text{persona, reside\_en}) = 1$

$C_{\text{Max}}(\text{persona, reside\_en}) = 1$

$C_{\text{Min}}(\text{localidad, reside\_en}) = 0$

$C_{\text{Max}}(\text{localidad, reside\_en}) = n$

En este ejemplo, una persona debe haber nacido necesariamente en una sola ciudad y, además, debe residir (en términos legales) en una sola ciudad.

Por último, en la Figura 10.3 d) se debe definir la cardinalidad en una relación recursiva. La forma de proceder es similar a lo visto hasta aquí. Sin embargo, se puede observar en la figura respectiva que, asociado a la relación, además de la cardinalidad, aparece un texto o rótulo que ayuda a la interpretación de la relación. El problema que tiene la relación recursiva es cómo debe ser leída o interpretada posteriormente. En estos ejemplos, dependiendo de la forma en que se lea el resultado, se puede llegar a una interpretación errónea.

Una materia tiene por correlativa al menos a otra, pero puede tener varias; en tanto que una asignatura puede no ser correlativa de ninguna otra o puede serlo de varias. En la relación dirige, un empleado puede

no tener jefe o a lo sumo tener uno; en tanto que un empleado puede no ser jefe o tener varias personas a cargo.

En los ejemplos vinculados con la Figura 10.3, se observan relaciones uno a muchos o muchos a muchos. Quedan sin exemplificar las relaciones uno a uno; un caso de ellas es presentado en la Figura 10.4 a).

La Figura 10.4 b) muestra un ejemplo de una relación entre tres entidades, denominada ternaria. La expresividad de la cardinalidad, en este caso, se ve afectada. Por ejemplo, suponga que una materia se cursa siempre en un aula, pero por varios alumnos a la vez; la cardinalidad de (materia, cursa) respecto de alumnos indica que varios alumnos la pueden cursar, pero la cardinalidad de (materia, cursa) respecto de aula es uno, es decir, se cursa solo en uno. En este caso, para que el modelo exprese sin errores el problema, se debe optar por la solución (0,N).

Para evitar este problema, algunos autores plantean solamente representar relaciones binarias entre entidades. Para ello, aportan otras herramientas que pueden utilizarse cuando se requieren soluciones ternarias. No obstante, en este libro no se hará referencia a estas soluciones. Se debe considerar que cualquier problema puede ser resuelto con agregaciones binarias, si se toman las decisiones de diseño correctas.

**FIGURA 10.4**



Figura 10.4 a)

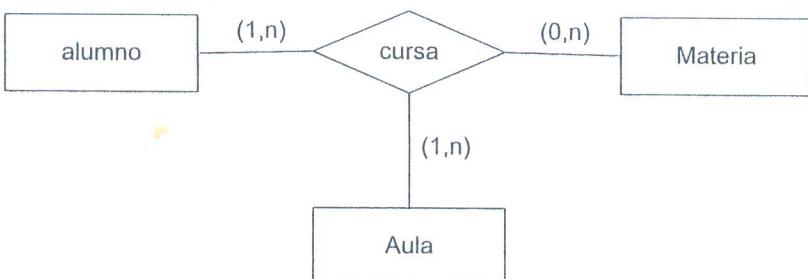


Figura 10.4 b)

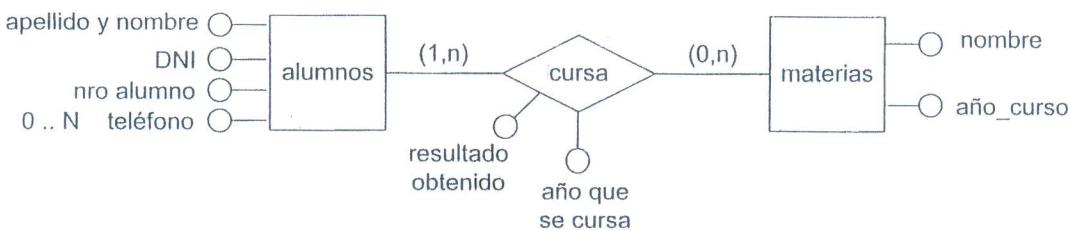
## Atributos

Un atributo representa una propiedad básica de una entidad o relación. Es el equivalente a un campo de un registro. Para la entidad alumno, mediante los atributos se representan, por ejemplo, el nombre del alumno, su DNI o su número de alumno.

Las diferencias entre las entidades o relaciones se distinguen fácilmente por los atributos de cada una de ellas. En general, de la misma forma que las entidades, los atributos están indicados por sustantivos que definen el problema.

La Figura 10.5 presenta un ejemplo de atributos. Se presenta, además, la sintaxis gráfica que deberá utilizarse para definir los atributos.

**FIGURA 10.5**



Los atributos también tienen asociado el concepto de cardinalidad. Esto es, cuando se define un atributo se debe indicar si es o no obligatorio y si puede tomar más de un valor (polivalente). Las expresiones posibles son:

1..1 monovalente obligatorio; en caso de que un atributo presente esta cardinalidad, no debe ser incluida explícitamente en el modelo. Por ende, un atributo que solamente está definido por su nombre tiene asociada la cardinalidad 1..1. Atención que la cardinalidad existe y está presente; solamente en este caso no se indica en forma explícita. Un caso que ejemplifica esta situación es el DNI de un alumno.

0..1 monovalente no obligatorio.

1..n polivalente obligatorio (podrían existir múltiples valores para este atributo). Por ejemplo, el atributo título de un profesor de una asignatura.

0..n polivalente no obligatorio. Por ejemplo, el e-mail de un alumno.

En la Figura 10.5, el atributo teléfono es un atributo polivalente no obligatorio; el resto de los atributos es monovalente obligatorio.

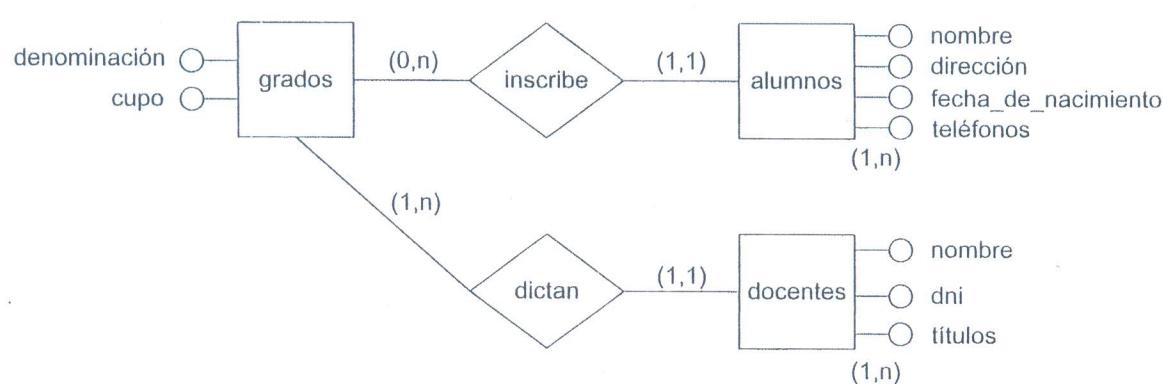
Además, cada atributo debe tener definido un dominio. El dominio de un atributo define el conjunto de valores posibles que el atributo puede tomar. El concepto de dominio está asociado al de tipo de datos. Por cuestiones de practicidad, la definición de dominio no se vuelve al modelo de datos; este concepto debe ser definido como parte de la especificación de requerimientos del problema. Recién en las etapas de modelado físico, se describe explícitamente el dominio de cada atributo.

## Construcción del diagrama conceptual. Ejemplo

En este apartado, se describirán someramente dos situaciones reales a partir de las cuales serán definidos los modelos de datos respectivos.

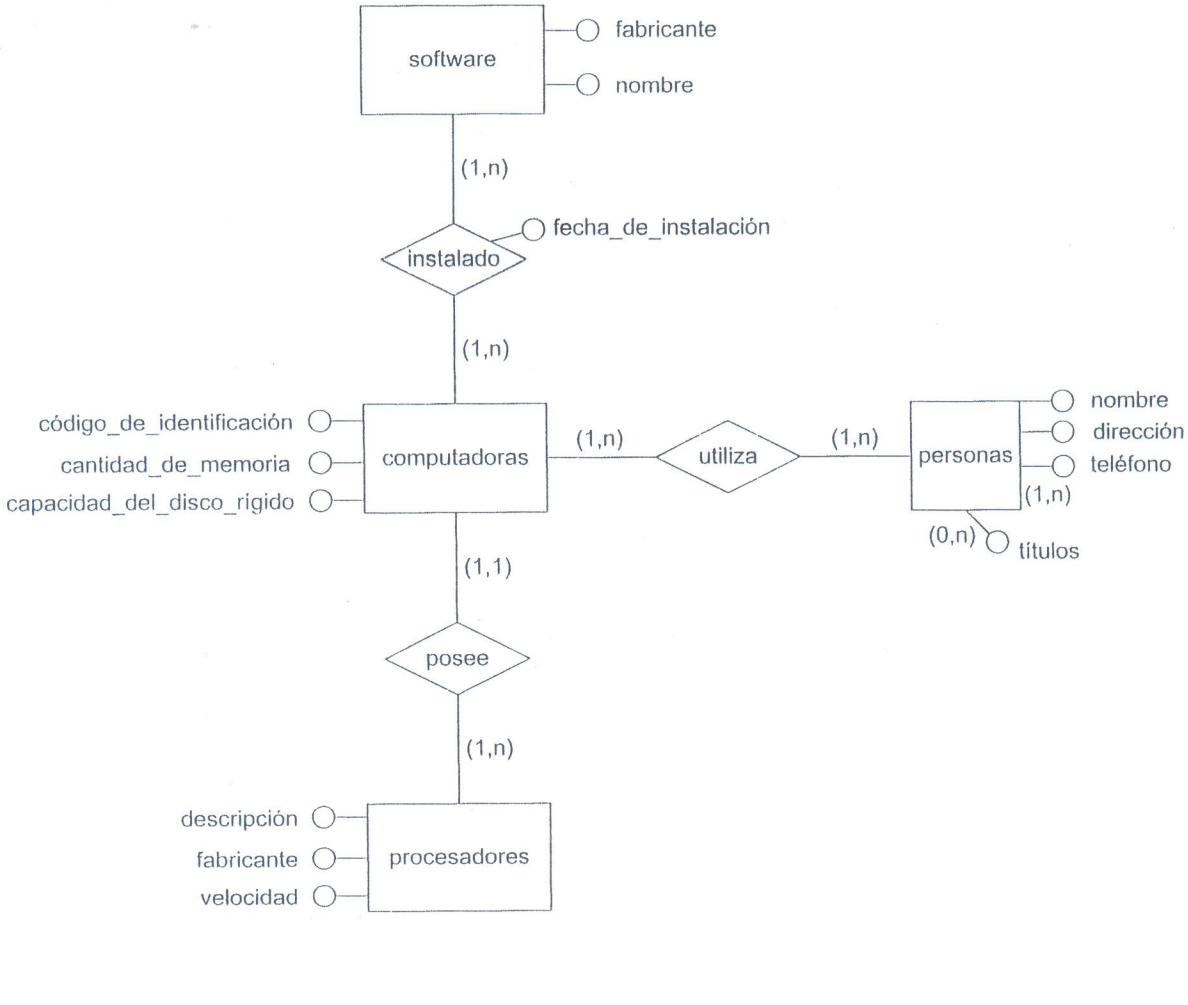
La primera de ellas presenta el caso de un colegio. Es necesario que el modelo represente a los alumnos que concurren al establecimiento. De cada alumno interesan su nombre, fecha de nacimiento, dirección y teléfonos. Los alumnos se inscriben en grados, de acuerdo con su nivel alcanzado. Cada grado tiene una denominación posible y, además, se conoce el cupo máximo para cada grado. A fines prácticos del colegio, asimismo se debe administrar el plantel docente. Los docentes son los responsables de dictar cursos en cada grado. Puede ocurrir que un grado disponga de más de un docente, pero necesariamente debe tener al menos uno asignado. Se desea conocer de cada docente su nombre, su DNI y los títulos que posea. Por último, cada docente es asignado solamente a un grado. La Figura 10.6 presenta el modelo de datos que resuelve el problema y que está representado con la herramienta de modelado CASER.

FIGURA 10.6



En el segundo ejemplo propuesto, el esquema representa el parque de computadoras de un instituto. De cada computadora se debe guardar un código de identificación, la tipo de procesador que tiene, la cantidad de memoria RAM y la capacidad del disco rígido. Además, cada computadora puede ser utilizada por varios miembros del instituto. Para ello, el problema también debe representar a estas personas (se conocen sus datos personales) y el/los títulos que pueden tener. Asimismo, cada computadora puede tener instalado diferente *software*; es importante tener registro de esta información, y desde qué fecha dicho *software* está instalado. La Figura 10.7 resume el modelo en cuestión.

FIGURA 10.7



## Componentes adicionales del modelo conceptual

A fin de poder dotar al modelado conceptual con las características descriptas al principio de este capítulo, se han incorporado otros elementos que permiten aumentar la expresividad del modelo construido. Estos elementos son cuatro: jerarquías, subconjuntos, atributos compuestos e identificadores, y se presentan en este apartado.

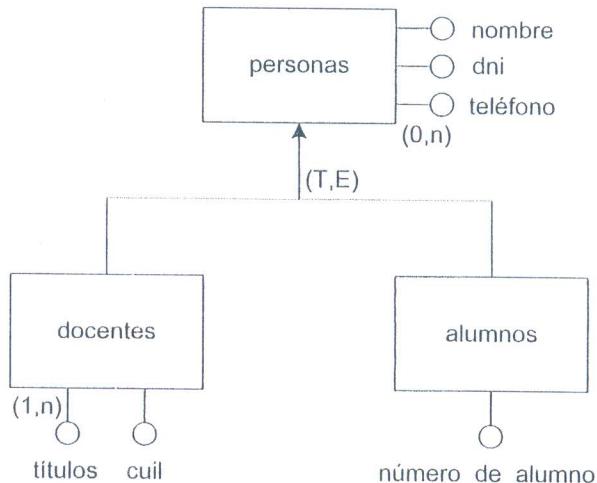
### Jerarquías de generalización

Siguiendo el concepto de abstracción de generalización presentado en el Capítulo 9, es posible que un modelo conceptual ER incorpore el concepto de jerarquías de generalización de entidades o de relaciones.

La generalización permite extraer propiedades comunes de varias entidades o relaciones, y generar con ellas una superentidad que las aglutine. Así, las características compartidas son expresadas una única vez en el modelo, y los rasgos específicos de cada entidad quedan definidos en su sub entidad.

La Figura 10.8 presenta un ejemplo que ilustra esta situación. Al representar docentes y alumnos de una facultad, es lógico suponer que ambas entidades comparten atributos tales como: nombre, dirección y teléfono. Estos atributos son definidos para la superentidad persona. Además, el alumno tiene rasgos propios, como por ejemplo el número de alumno, y el docente posee CUIL y títulos. Nótese, asimismo, que en la Figura 10.8 se expresa la cobertura ( $T, E$ ) que posee la jerarquía definida, de acuerdo con las pautas ya establecidas.

FIGURA 10.8





La cobertura, al igual que la cardinalidad, depende directamente del problema. Nuevamente el analista no debe tomar decisiones unilaterales al respecto. Si no tiene suficiente información, debe recurrir a la fuente, el cliente o el usuario.

Se debe recordar la propiedad básica de la abstracción de generalización: el concepto de herencia. Las subentidades o especializaciones (alumnos y docentes en el ejemplo) heredan los atributos de la superentidad o generalización (persona).

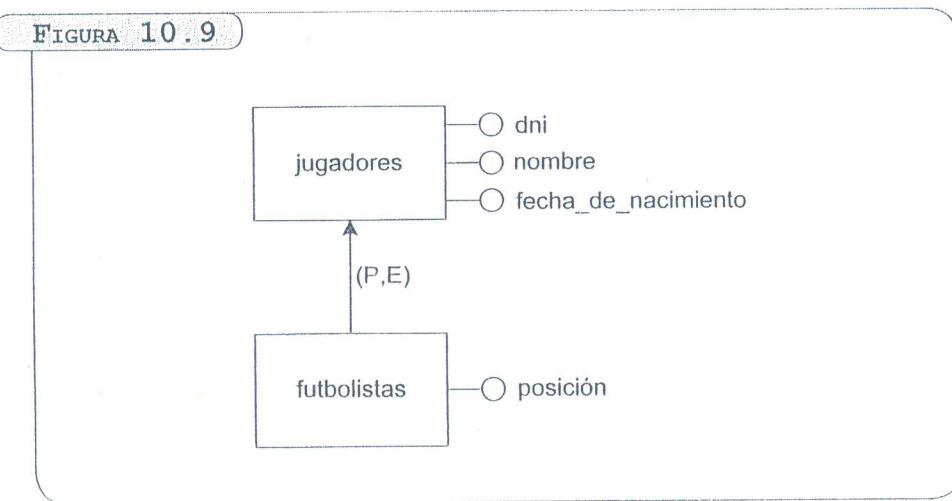
## Subconjuntos

Los subconjuntos representan un caso especial de las jerarquías de generalización. Hay problemas donde se tiene una generalización de la que se desprende solamente una especialización. Este es el caso de los subconjuntos. La representación es similar al caso anterior.

Solamente hay que considerar que no es necesario indicar la cobertura para los subconjuntos. Esto se debe a que no puede tratarse de una cobertura total; si no, la especialización y la generalización serían lo mismo, representarían los mismos atributos. Además, no puede ser superpuesta, dado que no hay una segunda especialización con la cual superponerse. Por lo tanto, un subconjunto siempre indica cobertura parcial y exclusiva. Sin embargo, en la herramienta CASER se indica la cobertura aunque no sea necesario.

El ejemplo de la Figura 10.9 representa a los jugadores de todos los deportes de un club cualquiera. Aquí, cada jugador es reconocido por su nombre, DNI y fecha de nacimiento. En particular, para cada uno de los deportistas que juegan al fútbol, se debe definir la posición que le corresponde en la cancha.

FIGURA 10.9

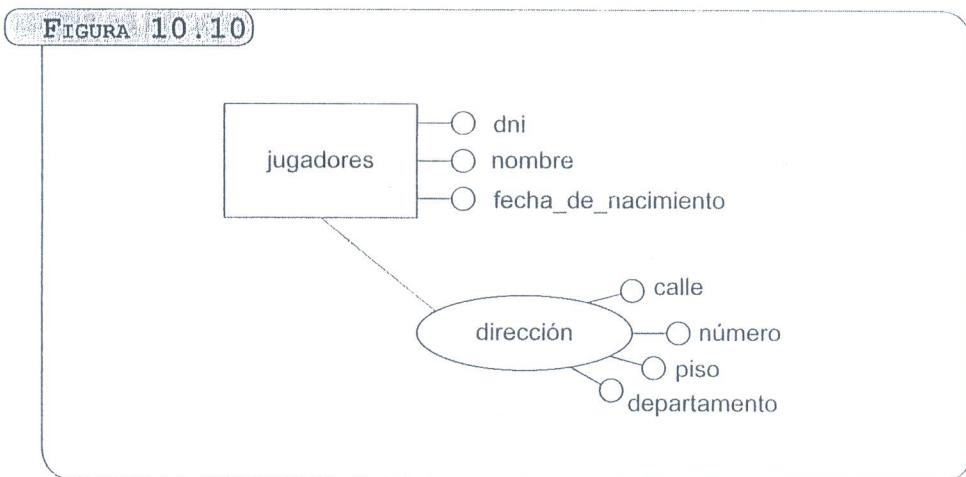


## Atributos compuestos

Los atributos compuestos representan a un atributo generado a partir de la combinación de varios atributos simples. Un ejemplo para ilustrar esta situación puede ser la dirección de una persona. Se podría optar por modelar la dirección como un solo atributo simple donde se indican la calle, el número y, eventualmente, piso y departamento. Así, el dominio podría ser una cadena de 50 caracteres. Sin embargo, es interesante poder mantener la calle, el número, piso y departamento por separado. De esta forma, como queda indicado en la Figura 10.10, se establece la dirección como atributo compuesto por los cuatro ítems mencionados.

Un atributo compuesto podría ser polivalente o no obligatorio. Lo mismo podría ocurrir con los atributos simples que lo componen.

**FIGURA 10.10**



## Identificadores

Un identificador es un atributo o un conjunto de atributos que permite reconocer o distinguir a una entidad de manera única dentro del conjunto de entidades. El concepto de identificador está ligado a los conceptos de claves primarias y candidatas, vertidos en capítulos previos de este libro.

Para el modelado conceptual, se utiliza el concepto de identificador. Los identificadores pueden ser de dos tipos:

- **Simples o compuestos:** de acuerdo con la cantidad de atributos que lo conforman, el identificador es simple si está conformado por solo un atributo y es compuesto en el resto de los casos.

- **Internos o externos:** si todos los atributos que conforman un identificador pertenecen a la entidad que identifica, es interno; en su defecto, es externo.

La Figura 10.11 presenta diversas situaciones que ejemplifican atributos simples y compuestos, que en todos los casos son atributos internos. Se debe notar la representación gráfica que se utiliza para definirlos.

**FIGURA 10.11**

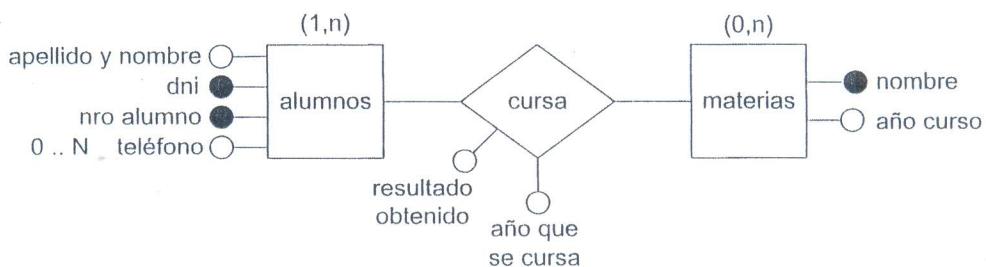


Figura 10.11 a) Identificadores simples

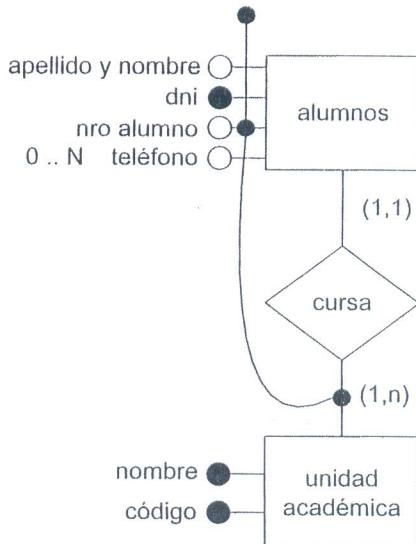


Figura 10.11 b) Identificador Compuesto

Suponga un esquema de una universidad donde los números de alumno no se pueden repetir entre diferentes unidades académicas. Así, el número de alumno 1 está presente en Derecho, Informática, Ciencias Exactas, etcétera. Un ejemplo factible de identificador externo es combinar el número de alumno con la unidad académica de la que proviene. Se puede observar en la Figura 10.12 la notación para esta situación.

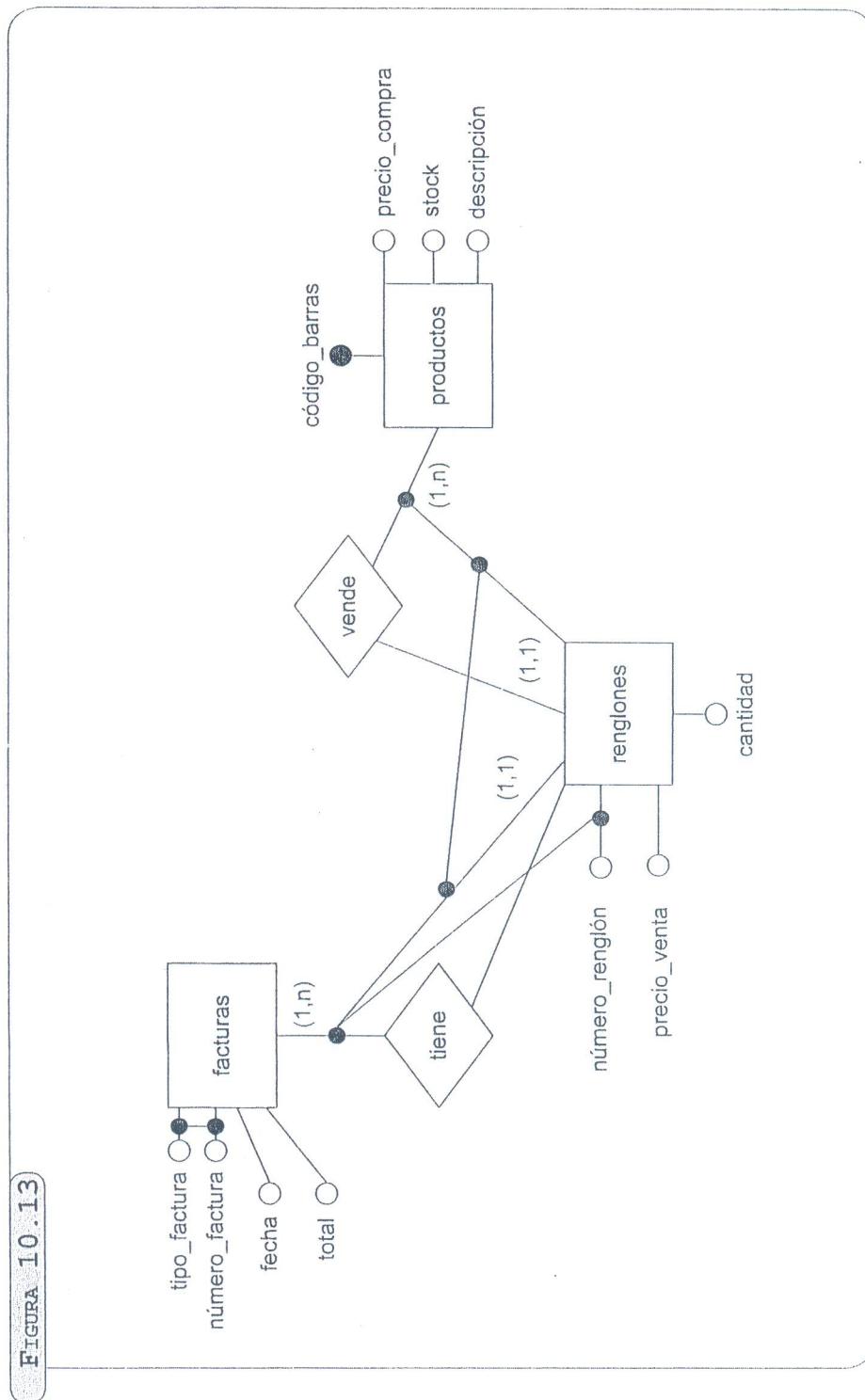
La entidad Alumnos tiene entonces un identificador simple interno (DNI), y otro compuesto y externo, el que se conforma con un identificador de Unidad Académica más el número de alumno.

FIGURA 10.12

Figura 10.12  
Identificador Externo

La Figura 10.13 presenta otro caso para definir identificadores. Se plantea la entidad facturas con los atributos tipo de factura, número, fecha y monto total. El detalle de cada factura está compuesto por renglones; en cada renglón se tiene el precio unitario de venta, el número de renglón y la cantidad vendida. Además, cada renglón representa la venta de un producto. De los productos se conocen el código de barras, la descripción, el stock y el precio de compra.

El identificador de facturas es compuesto, formado por tipo y número de factura; en productos, el identificador es el código de barras. La entidad renglones carece de identificador interno, dado que el número de renglón se repite en cada factura. Por lo tanto, combinando el identificador de facturas y el número de renglón se genera un identificador externo para renglones. Además, en una factura no deberían existir dos renglones con el mismo producto; entonces, los identificadores de facturas y productos combinados forman otro identificador externo para la entidad renglones.



Los identificadores en una jerarquía también son heredados. Esto es, si una generalización posee identificadores, estos son heredados por las entidades que son sus especializaciones. Además, las especializaciones pueden tener otros identificadores definidos.

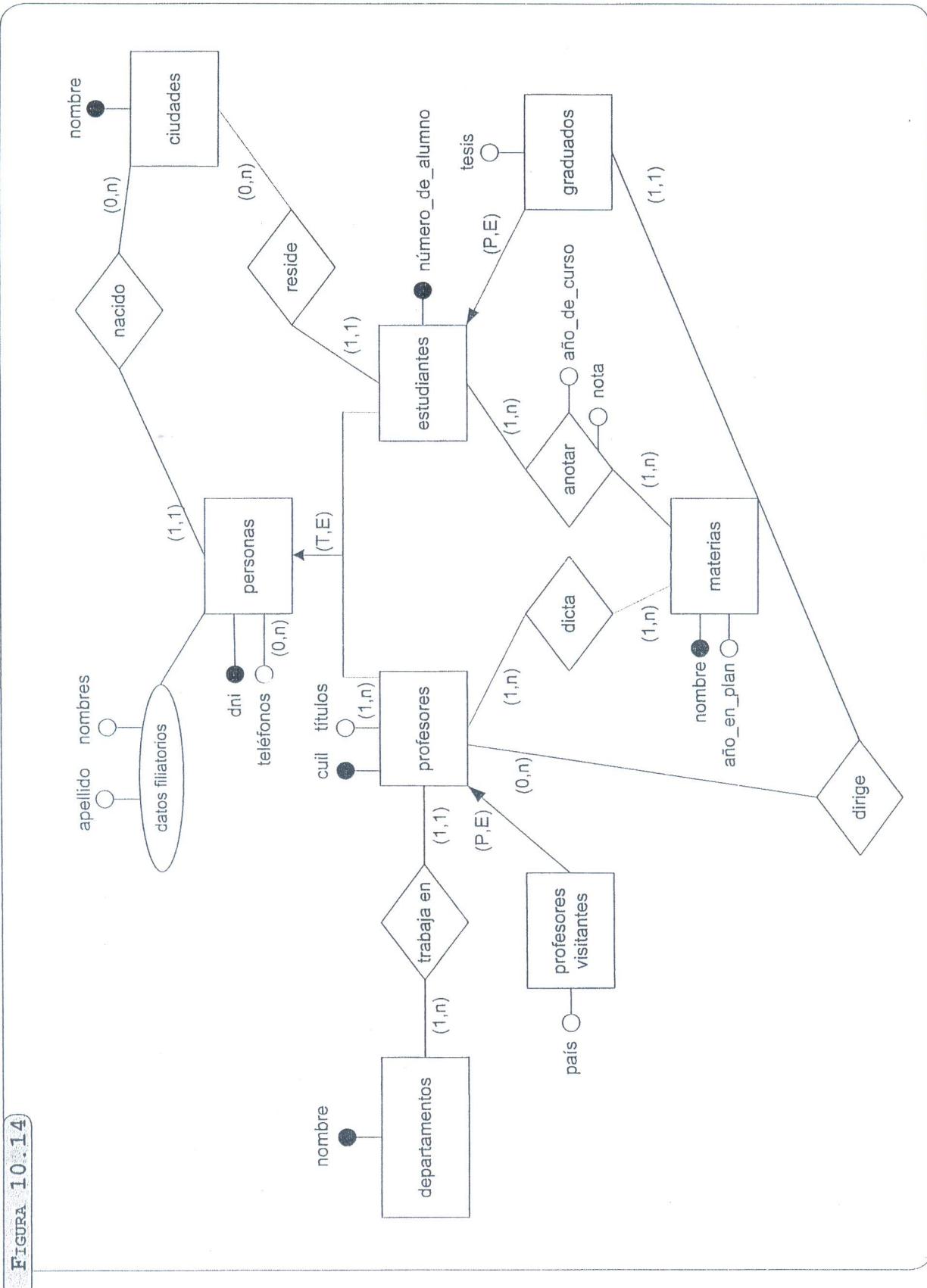
## Ejemplo integrador

A partir del uso de los elementos definidos hasta el momento para la construcción del modelo conceptual, es posible generar cualquier modelo que exprese la realidad de un problema. A continuación, se enuncia un problema y se plantea una solución posible.

Se debe modelizar el esquema de una facultad. En este esquema, los profesores (interesan los datos personales, más el número de CUIL y los títulos que posee cada docente) se relacionan con los departamentos (cada uno tiene un nombre único) donde cumplen sus actividades. Un docente solamente trabaja en un departamento, pero un departamento puede tener varios docentes. Además, cada docente debe indicar en qué ciudad ha nacido. El esquema debe representar asimismo a los estudiantes (descriptos por sus datos personales más el número de alumno, que es único en la facultad). Los estudiantes deben indicar el lugar de nacimiento y de residencia.

Además, se deben señalar las materias que se dictan en la facultad (de cada materia se sabe el nombre, que no puede repetirse, y el año dentro del plan de estudios). Los alumnos se inscriben en materias, las cuales son dictadas por profesores bajo las siguientes suposiciones: un docente puede dictar más de una materia y una materia puede estar dictada por más de un docente. Los alumnos pueden anotarse en más de una materia y se debe indicar el resultado obtenido. Se debe tener en cuenta que si un alumno no aprueba una materia, puede anotarse nuevamente, por lo que es importante conocer el año en que se anota. Los estudiantes pueden ser graduados. En este caso, tendrán definidos un trabajo de tesis y un asesor, que es un profesor de la casa. Pueden existir, además, profesores visitantes, para los cuales debe indicarse el país de procedencia.

La Figura 10.14 muestra el esquema realizado con CASER que resuelve dicho modelo.



## Consideraciones del modelo conceptual

### Mecanismos de abstracción

Los tres mecanismos de abstracción presentados en el Capítulo 9 están presentes en el modelo conceptual ER.

Los tres conceptos básicos (entidades, relaciones y atributos) se basan en la abstracción de clasificación:

- **Entidades:** son una clase de objetos del mundo real con propiedades comunes.
- **Relaciones:** son un conjunto de objetos que relacionan dos o más entidades.
- **Atributos:** son una clase de valores, determinadas por el dominio de cada uno de ellos, que representan propiedades de entidades y/o relaciones.

La abstracción de agregación está presente en:

- **Entidades:** son agregaciones de atributos.
- **Relaciones:** son agregaciones entre entidades y, además, pueden contener agregaciones de atributos.
- **Atributos compuestos:** son agregaciones de atributos simples.

Por último, la abstracción de generalización solo se ha presentado en entidades, pero es posible, además, utilizar generalización de relaciones (aunque muy poco frecuente).

### Ventajas y desventajas del modelo conceptual

El modelo ER ha sido el más utilizado para representar los modelos de datos para los SI. Presenta un conjunto de ventajas que lo hacen aplicable, pero, asimismo, se deben analizar algunas críticas respecto de su construcción.

Las cuatro propiedades indicadas al principio del capítulo –expresividad, formalidad, minimalidad y simplicidad– son ampliamente alcanzadas a partir de los elementos disponibles para construir el modelo.

Lograr que un modelo sea expresivo ayuda a comprender el problema, pero puede atentar contra la simplicidad.

La forma de expresar y determinar la cardinalidad y los identificadores no es sencilla. Se puede comparar el presente libro con otros relacionados con modelado de datos, y comprobar que la forma gráfica y semántica de expresar la cardinalidad varía.

Las relaciones *n*-arias también presentan inconvenientes. Como se discutió anteriormente, la cardinalidad es más compleja de expresar a fin de capturar todas las características del problema.

La minimalidad y la formalidad pueden ser alcanzadas. Obviamente, cada analista puede desarrollar un modelo de manera diferente a partir del mismo problema. Esto se debe a que la forma de representación que cada uno utilice capturará a su mejor entender la naturaleza del problema.

## Revisión del modelo conceptual

### Decisiones respecto de entidades, relaciones o atributos

Para construir un modelo conceptual, hay una serie de decisiones que el analista debe tomar. La decisiones dependen de tres factores fundamentales: el problema, lo que el cliente/usuario espera obtener desde la BD y, por último, la experiencia práctica del analista.

Para exemplificar la situación, suponga que se deben modelizar los datos personales de los empleados de cierta compañía. Este problema ya fue analizado anteriormente, pero esta vez su análisis se detiene en dos conceptos: el tipo de documento de la persona y la localidad donde vive. Una opción podría ser que ambos conceptos sean atributos de persona, como muestra la Figura 10.15 a). Sin embargo, hay otra opción, presentada en la Figura 10.15 b); en este caso, tanto la localidad como el tipo de documento se definen en una nueva entidad que se relaciona con la entidad persona. ¿Qué ventajas presenta esta última solución? Una ventaja directa es que si se quiere saber cuántas personas son originarias de la ciudad de La Plata, bastará con contar las relaciones existentes entre empleados y la ciudad de La Plata. Si se hubiese escrito como un atributo, el valor obtenido dependería de la forma de escritura. Así, escribir todo en mayúscula o todo en minúscula podría afectar el resultado. Algo similar podría ocurrir con el tipo de documento.

Por lo tanto, la decisión respecto de cuál es la mejor forma para modelar depende de los tres factores antes descriptos.

FIGURA 10.15



Figura 10.15 a)

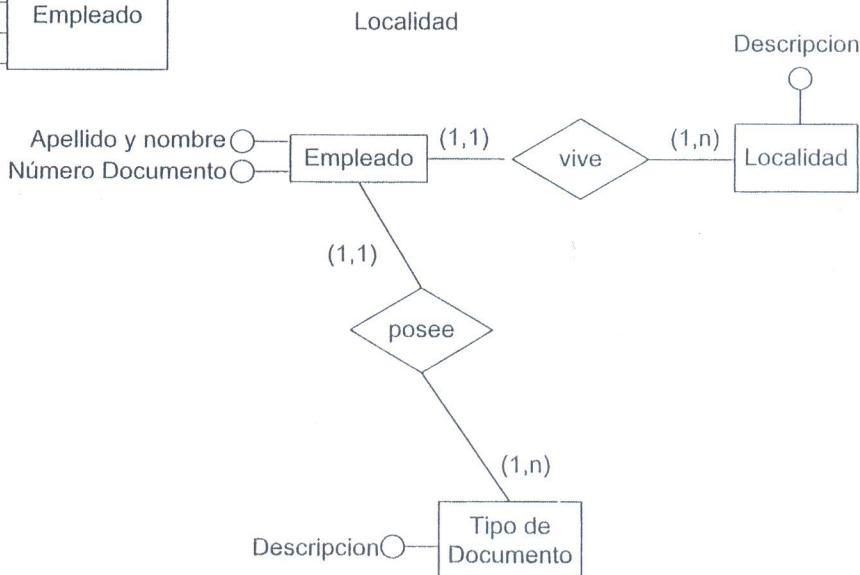


Figura 10.15 b)

La siguiente pregunta a realizar cuando se genera un modelo conceptual tiene que ver con la ventaja o no de definir generalizaciones. Aquí, la decisión debería tomarse considerando si la representación establece una jerarquía o una clasificación. Por ejemplo, si se quiere definir en el modelo de datos solamente el sexo de una persona, la solución se obtiene agregando un atributo a la entidad persona. Sin embargo, si hubiera características propias para los hombres y para las mujeres, sería conveniente generar una jerarquía. Así, la regla a aplicar debería controlar que no queden definidas entidades sin atributos o sin relaciones con otras entidades. Estas entidades (sin atributos y sin relaciones) se denominan entidades colgadas.

La última consideración está ligada, como se explicó anteriormente, a la utilización de atributos compuestos. ¿Conviene un atributo compuesto o directamente se colocan sobre la entidad los atributos simples? Los atributos compuestos se deberían utilizar y luego, cuando se evolucione en el modelo lógico o físico, decidir cómo actuar.

## Transformaciones para mejorar el modelo conceptual

Una vez finalizada la construcción del modelo conceptual, este debe ser puesto en consideración y análisis por parte del cliente/usuario. Si bien este no es un proceso sencillo, con la ayuda del analista es posible que alguien no ligado a la actividad informática pueda analizar el modelo resultante para tratar de determinar faltantes o errores.

Es necesario, entonces, poder lograr un modelo que represente la realidad del problema lo más estrictamente posible, disminuyendo ambigüedades y tomas de decisiones de bajo nivel, y aumentando la expresividad y legibilidad del modelo construido.

Existen una serie de conceptos a revisar:

- **Autoexplicativo:** un modelo se expresa a sí mismo si puede representarse utilizando los elementos definidos, sin necesidad de utilizar aclaraciones en lenguaje natural para expresar características.
- **Compleitud:** un modelo está completo cuando todas las características del problema están contempladas en él. La forma de validar la completitud es revisar la especificación de requerimientos asociada al problema.
- **Corrección:** un modelo es correcto si cada elemento en su construcción fue utilizado con propiedad. Algunos aspectos de fácil resolución en cuanto a correctitud son observar que todas las cardinalidades y coberturas se hayan expresado, haber tenido en cuenta el concepto de herencia en jerarquías, haber expresado todos los identificadores, etcétera.
- **Expresividad:** el modelo conceptual resulta expresivo si a partir de su observación es posible darse cuenta de todos los detalles que lo involucran. Existen varias formas de mejorar la expresividad. Si el lector observa la Figura 10.14, se han utilizado jerarquías para personas, profesores, alumnos y graduados, las que permiten realizar relaciones con otras entidades. Si no se hubiesen utilizado esas jerarquías, la expresividad del modelo se vería seriamente afectada.
- **Extensible:** el modelo conceptual resulta extensible si es fácilmente modificable para incorporar nuevos conceptos en él, resultantes de cambios en los requerimientos del problema.
- **Legibilidad:** un esquema es legible si la representación gráfica es adecuada. Este fue uno de los motivos que llevó a la generación de CASER: disponer de una herramienta para la asistencia en la generación del modelo conceptual, basado en los elementos propios del modelo.

- **Minimalidad:** un esquema es mínimo cuando cada concepto se representa una sola vez en el modelo. Aquí hay dos factores posibles que pueden afectar la minimalidad. Estos factores son: atributos derivados y ciclos de relaciones. Más adelante en este apartado, se ejemplifican estos casos y se analiza la importancia de reconocerlos.

El diseñador de la BD, una vez terminada la construcción del modelo conceptual inicial, debe repasar los puntos anteriores y determinar si fueron cubiertos por el modelo de datos. Las transformaciones a realizar se aplican sobre el esquema obtenido, y generan un nuevo esquema con las características anteriormente analizadas sin afectar el contenido original. Ambos esquemas deben ser en todo momento equivalentes con respecto al contenido de la información.

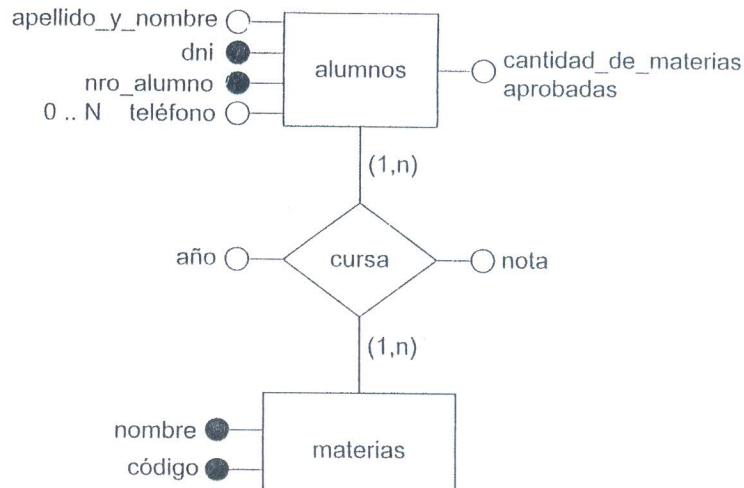
### Transformaciones para minimalidad

Un esquema es mínimo cuando una idea se expresa en él una sola vez. La redundancia de información puede ser un efecto positivo si es conocida y controlada, o un efecto negativo si aparece en el esquema por deficiencia en la modelización. En estos últimos casos, la redundancia puede llevar a pérdida de consistencia en la información.

Hay dos causas que llevan a un modelo a perder la minimalidad. En este apartado, se analizan las situaciones. Se debe notar que el resultado esperado en este punto es detectar las condiciones que afectan la minimalidad, pero no necesariamente solucionarlas. Estas decisiones se pueden tomar cuando se genera el modelo lógico del problema.

Los atributos derivados representan la primera causa que afecta la minimalidad. Un atributo derivado es un atributo que aparece en el modelo, pero cuya información, si no existiera almacenada en él, podría igualmente ser obtenida. La Figura 10.16 presenta un atributo derivado, la cantidad de materias que el alumno tiene aprobadas. Si este atributo no existiera, bastaría con contar las relaciones existentes en cursa con nota superior a 4 para obtener dicho resultado. Si el atributo derivado existe, el esquema no es mínimo; no obstante, las decisiones finales sobre los atributos derivados, como se indicó anteriormente, se tomarán sobre el esquema lógico en el Capítulo 11.

FIGURA 10.16



Los ciclos de relaciones presentan la segunda causa que afecta la minimalidad. Se dice que existe un ciclo cuando una entidad A está relacionada con una entidad B, la cual está relacionada con una entidad C, la que a su vez se relaciona con la entidad A.

No todos los ciclos afectan la minimalidad. Un ciclo afecta la minimalidad cuando una de las relaciones existentes puede ser quitada del esquema, y aun en esas condiciones, el modelo sigue representando la misma información.

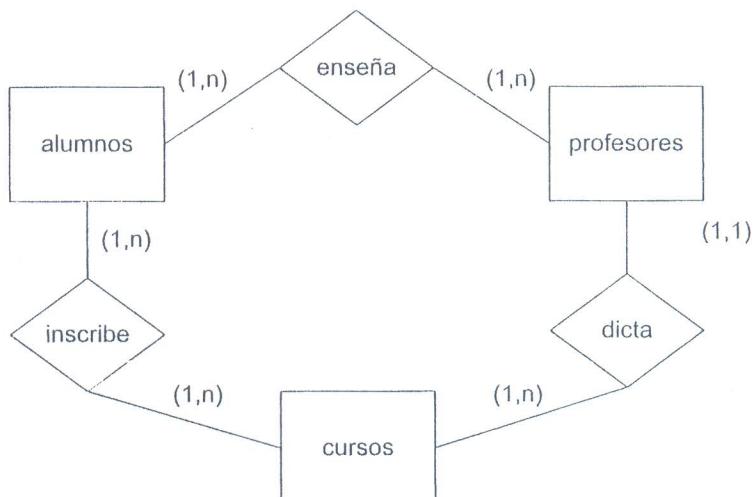
La Figura 10.17 presenta un ejemplo que define las entidades profesores, alumnos y cursos. Los profesores dictan cursos, los cursos son tomados por alumnos y se sabe, además, el dato de los alumnos que tiene cada docente. En este caso, un profesor puede dictar solamente un curso, pero un curso puede ser dictado por varios profesores. Un curso puede ser tomado por varios alumnos, los que a su vez pueden tomar varios cursos. Por último, un alumno puede tener varios docentes y estos, varios alumnos.

Para analizar si se tienen ciclos no deseados, deberían quitarse del modelo una a una las relaciones para determinar si es posible administrar la información deseada. Suponga que se quita la relación dicta; a partir de la relación inscribe es posible determinar los alumnos de cada curso, y de la relación enseña, los profesores que dictan clases al alumno, pero no es posible afirmar si el profesor dicta o no un curso específico. Ejemplo: el curso es Análisis Matemático I; se

obtiene el listado de alumnos inscriptos: suponga que uno de ellos es José García. De este alumno se obtiene el listado de profesores que dictan varias materias, suponga Gomez, Perez y Díaz; ¿cuál de ellos dicta Análisis Matemático I?, es imposible de determinar; por lo tanto, la relación cursa debe ser parte del modelo.

Si se intenta quitar la relación enseña, se debería poder determinar la lista de alumnos de un profesor. Dado un profesor, Perez, se sabe qué materia dicta, la cual puede ser dictada por varios profesores. Por lo tanto, la lista de alumnos de la materia no determina exactamente con qué profesor cursan. Por consiguiente, la relación enseña no puede ser quitada.

**FIGURA 10.17**



Queda por analizar la relación inscribe; dado un alumno, se puede determinar qué profesores le imparten clase. Como cada profesor puede dictar solamente una materia, es posible saber en qué materia se inscribió el alumno. A partir de la materia se puede conocer la lista de profesores que la dictan, y con esos profesores, la lista de alumnos a los que se imparte clase. Por lo tanto, nuevamente es posible conocer la nómina de alumnos inscriptos. Esto demuestra que si se quita del modelo la relación inscribe, es posible conocer la misma información respecto del problema. Esto significa que el esquema no es mínimo. Nuevamente, las redundancias que se producen con el ciclo se pueden resolver sobre el modelo lógico.

Queda como tarea tomar el ejemplo de la Figura 10.17, y analizar qué sucedería si un docente pudiera dictar varios cursos.

## Transformaciones para expresividad y autoexplicación

Las transformaciones para lograr expresividad y autoexplicación tienen varias causas; se analizan algunas de ellas.

Un modelo es más expresivo si las jerarquías que se representan tienen sentido. Una jerarquía se establece para obtener atributos comunes en una generalización, dejando los atributos específicos para la especialización. El problema surge cuando las especializaciones carecen de dichos atributos. Este ejemplo es bastante común y se observa más cuando el problema determina una clasificación de elementos.

Suponga que los productos que vende cierta compañía se dividen en productos para el hogar, para el auto y para la oficina. Cuando se plantea el esquema conceptual, se establece una generalización de productos (con todos los atributos pertinentes) y especializaciones para productos para el hogar, el auto y la oficina, respectivamente. Pero estas especializaciones no tienen ningún atributo que las conforme. Claramente, se está definiendo en el esquema una clasificación como si fuese una jerarquía, y esto es una solución no deseada.

La Figura 10.18 presenta el mismo problema con dos alternativas; en la segunda se puede observar claramente el significado del atributo tipo que aparece en la relación dirige.

**FIGURA 10.18**

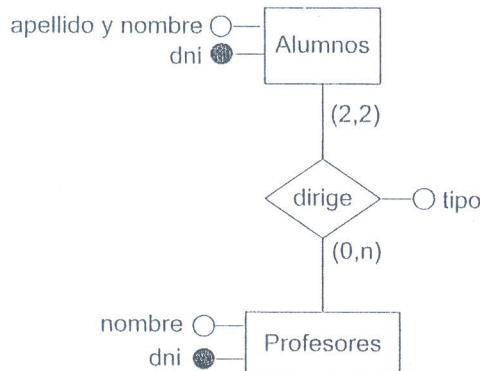


Figura 10.18 a)

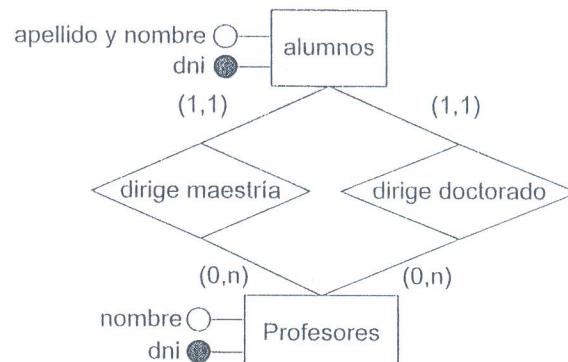


Figura 10.18 b)

## Cuestionario del capítulo

1. ¿Cuál es la finalidad de generar un modelo conceptual de datos?
2. ¿Por qué son necesarias, para un modelo conceptual, las cuatro características definidas?
3. ¿Cuáles son los constructores básicos definidos por Chen para el modelo conceptual?
4. ¿Por qué se agregan nuevos constructores?
5. ¿Qué son las transformaciones de esquemas y qué finalidad persiguen?

## Ejercitación

1. Resuelva los siguientes problemas simples, indicando las entidades, relaciones y atributos involucrados. Además, indique la cardinalidad de los atributos y de las relaciones.
  - a. Se tiene un problema de un banco donde se maneja información de los clientes y las cuentas con las que operan. Esas cuentas pueden ser corrientes, de ahorro o accionarias. De los clientes se conocen su DNI, nombre, dirección y teléfonos; asimismo, el banco les define un código único. Los clientes pueden tener varias cuentas. Las cuentas corrientes tienen un código único, un saldo y un máximo valor en descubier-  
to. Las cajas de ahorro poseen un código único, un saldo, una fecha de vencimiento de la tarjeta de débito asociada, y el monto máximo para operar en cada transacción. Las cuentas accionarias poseen un código y una fecha de apertura, y están relacionadas con todas las acciones que maneja el cliente.
  - b. Historias clínicas. Cada paciente tiene una historia clínica y esta tiene un número identificatorio. De cada paciente se conocen sus datos personales. Cada paciente se atiende con médicos, cuyos datos personales también se conocen. Cada médico tiene una especialidad. Cada paciente puede atenderse con varios médicos.

A partir del modelo anterior, se debería poder responder:

- i. ¿Cuántos médicos hay por cada especialidad?
- ii. ¿Con cuántos médicos se atiende un paciente?

- c. Una biblioteca tiene libros. Los libros tienen un ISBN que los identifica, autores, editorial y año de edición. Por cada libro puede haber varias copias. Cada copia tiene un código único dentro de la biblioteca. Los autores tienen nombre, DNI, código de identificación y nacionalidad.

A partir del modelo anterior, se debería poder responder:

- i. ¿Cuántos autores de nacionalidad argentina aparecen en libros de la biblioteca?
  - ii. ¿Cuántos libros de cada editorial hay?
  - iii. ¿Cuántas copias de cada libro hay?
- d. Se desea administrar información de una facultad. Los datos a gestionar son: alumnos, carreras que se cursan, materias que se dictan y plantel docente. El plantel docente está integrado por profesores y auxiliares docentes. De los alumnos se conocen su DNI, código de alumno, nombre, dirección, e-mail y teléfonos. De cada docente interesan su DNI, nombre, dirección, e-mail, teléfonos, CUIL y títulos obtenidos. Cada docente dicta una asignatura en un año particular. Los alumnos pueden anotarse a cursar (o recursar) varias asignaturas.

A partir del modelo anterior, se debería poder responder:

- i. ¿Cuántos docentes tienen título de Doctor?
  - ii. ¿Cuántos alumnos recursan materias?
  - iii. ¿Qué nacionalidad tiene cada docente y/o alumno?
- e. El esquema representa el menú de platos de una cadena de restaurantes. Para cada restaurante se conocen el nombre, la ubicación, la capacidad y los datos de su responsable. La cadena de restaurantes tiene platos a disposición. Cada plato tiene un nombre, un costo y una serie de ingredientes. Dentro de la cadena, cada restaurante puede o no tener el plato, y puede, además, cobrarlo diferente.
2. Indique en cada ejercicio anterior los posibles identificadores a partir de la forma en que lo haya resuelto y de las siguientes consideraciones: las personas pueden reconocerse por DNI-CUIT-CUIL; los alumnos, por su número.
3. Indique, ante cada problema, si es pertinente o no el uso de una jerarquía. En cada caso, indique la cobertura:

- a. El esquema representa a clientes y proveedores de la empresa. En el caso de los clientes, se debe indicar si la facturación es de tipo A o B. En el caso de los proveedores, es necesario indicar el tipo de servicio que prestan, y si la empresa tiene o no una cuenta corriente.
- b. Un supermercado administra la información de sus productos. De cada producto se conocen el nombre, la presentación del envase, el precio de costo y el precio de venta. Asimismo, es importante conocer quién es el proveedor. Los productos son perecederos o no; además, pueden ser de ferretería, almacén, carnicería, etcétera.
- c. Cierta facultad ordena su plan curricular de la siguiente manera: cada materia tiene un nombre, un año de cursada, uno o más profesores que la dictan y, además, pertenece a un área. Por el momento, las áreas disponibles en la facultad son: Programación Básica; Ingeniería de Software y Bases de Datos; Redes y Sistemas Operativos; y Programación Avanzada.