

Deseño de BBDD: o modelo Entidade-Relación

Sumario

- 1 Introducción
 - ◆ 1.1 **INTERÉSACHE** Actualmente, existen ferramentas que automatizan en gran parte as tarefas de deseño dunha BBDD e que toman como base o modelo E-R. Son o que se chaman **ferramentas CASE** (*Computer Aided Software Engineering* ou Enxeñería do Software Asistida por Computador).
- 2 O modelo Entidade-Relación
 - ◆ 2.1 Entidades
 - ◆ 2.2 Atributos
 - ◆ 2.3 Relacións
- 3 Extensións ao modelo E-R
 - ◆ 3.1 Especialización e xeneralización
 - ◆ 3.2 Entidades débiles

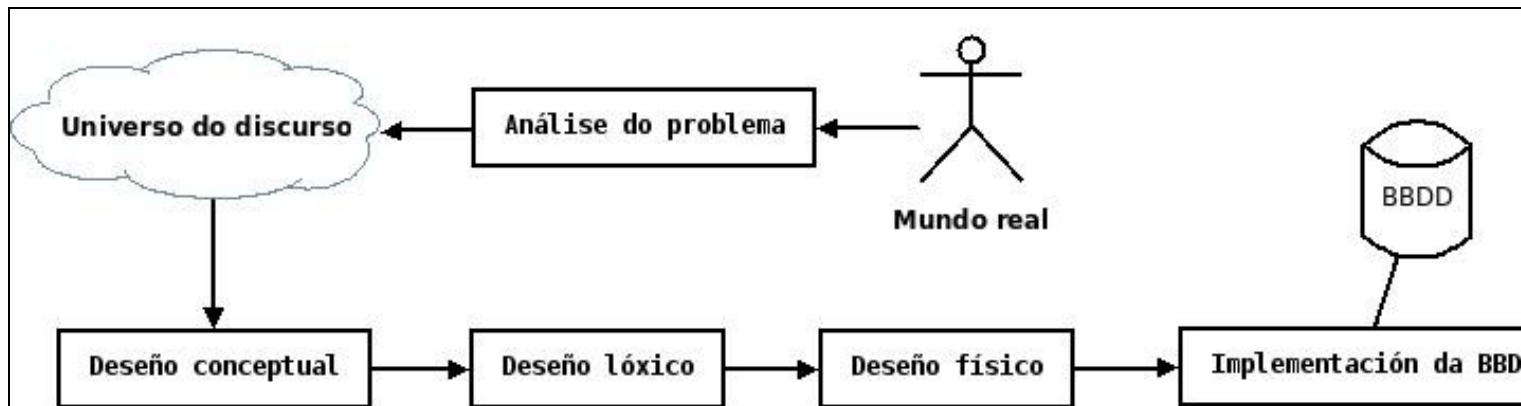
Introdución

Unha das fases fundamentais na construción dunha BBDD, anterior á creación de táboas, campos, etc., é o seu deseño. Este proceso conceptual implica, entre outras cousas, identificar as entidades que nos interesa almacenar e as relacións existentes entre elas, así como definir os atributos que nos resulten útiles.

É importante asegurarnos de que a nosa base de datos está correctamente deseñada para que sexa eficiente e escalable, o cal redundará en que se poida seguir utilizando durante tempo. Por tanto, o que se busca cun bo deseño é conseguir acceso rápido e fácil con redundancia mínima dos datos.

A esta fase da creación da BBDD chámase **esquema ou deseño conceptual**. É fácil intuír que para facer un bo deseño é fundamental o coñecemento do problema a modelar. Isto conséguese mediante reunións cos usuarios finais do sistema que teñen por obxecto conseguir clarificar todas as **especificacións** ou requirimentos do problema. Unha vez clarificados os obxectivos e as necesidades de almacenamento de datos deberase pasar ao deseño propiamente dito.

Por tanto, existe unha fase de análise, outra de deseño e finalmente a construción da BBDD.



A fase de deseño está dividida en tres fases:

- **Deseño conceptual**, propiamente dito, onde usamos o modelo Entidade-Relación, que veremos a continuación, para modelar o Universo do discurso, é dicir, a visión do mundo real mediante uns determinados obxectivos.
- **Deseño lóxico**, onde "traducimos" o modelo Entidade-Relación ao modelo relacional, xa con táboas e relacións.
- **Deseño físico**, onde se especifica como se almacenarán os datos, é dicir, índices para acceder a rexistros, organizacións secuenciais, etc.



Actualmente, existen ferramentas que automatizan en gran parte as tarefas de deseño dunha BBDD e que toman como base o modelo E-R. Son o que se chaman **ferramentas CASE** (*Computer Aided Software Engineering* ou Enxeñería do Software Asistida por Computador).

O modelo Entidade-Relación

O modelo Entidade-Relación (E-R, en diante) é un modelo conceptual ideado por **Peter Chen** nos anos 70. Ten un nivel de abstracción suficiente como para poder deseñar calquera base de datos con independencia do SXBD e do sistema operativo. **Nós usaremos o modelo E-R conxuntamente co modelo relacional.** De feito, xa vimos na primeira práctica que unha entidade no modelo E-R represéntase mediante unha táboa no modelo relacional.

O modelo E-R consiste nun conxunto de regras e notacións que permiten formalizar a semántica do mundo real que se pretende modelar (tamén denominada Universo do discurso ou minimundo) mediante unha representación gráfica ou diagrama. O modelo xira exclusivamente ao redor de dous conceptos: as entidades e as relacións.

Entidades

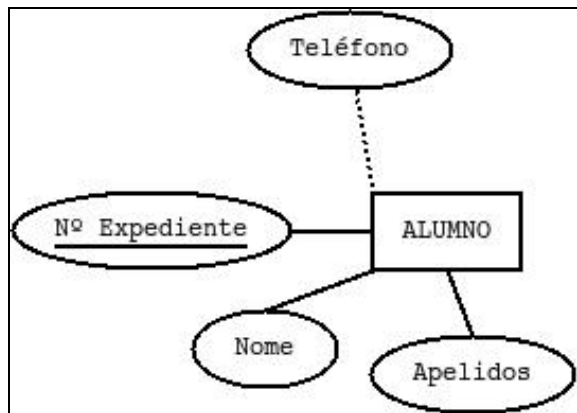


As entidades representan conxuntos de elementos con existencia propia e que se caracterizan polas mesmas propiedades. Xeralmente son persoas, cousas, lugares,...., é dicir, conceptos sobre os que necesitamos gardar información. A súa representación gráfica faise por medio dun rectángulo dentro do cal se escribe o nome da entidade en maiúsculas (xeralmente un **substantivo**).

Atributos

Hai determinadas características dunha entidade que nos pode interesa recoller dentro do noso deseño para que se almacenen na BBDD. Chámanselles atributos. Estas propiedades non teñen o "rango" de entidade, é dicir, só teñen sentido formando parte dunha entidade ou, como veremos máis adiante, dunha **interrelación**.

Un mesmo concepto pode modelarse de distintas formas (por exemplo, como unha entidade ou como un atributo). Así, se estivésemos modelando unha BBDD para unha tenda de roupa, probablemente teríamos unha entidade denominada PEZA e un dos seus atributos podería ser Cor (vermella, negra, etc.). Con todo, se estivésemos a falar dunha BBDD para xestionar a información dun taller de vehículos dedicado a traballos de chapa e pintura, o concepto de cor pode ter tal importancia que pase a ser unha entidade COR, pois ten existencia propia e un conxunto de propiedades (código de cor, textura, tipo de mestura, etc.).



Un atributo que identifica univocamente a cada rexistro ou **exemplar** dunha entidade chámase **chave primaria**. Poden existir varios atributos que cumpran esta condición (por exemplo, o DNI e o número da Seguridade Social). A estes atributos chámase **chaves candidatas**.

Existen algún atributos dos que podemos non coñecer o seu valor para un rexistro concreto (por exemplo, un número de teléfono dun alumno cando se matricula). A estes chámase **atributos opcionais** e represéntanse mediante unha liña de puntos.

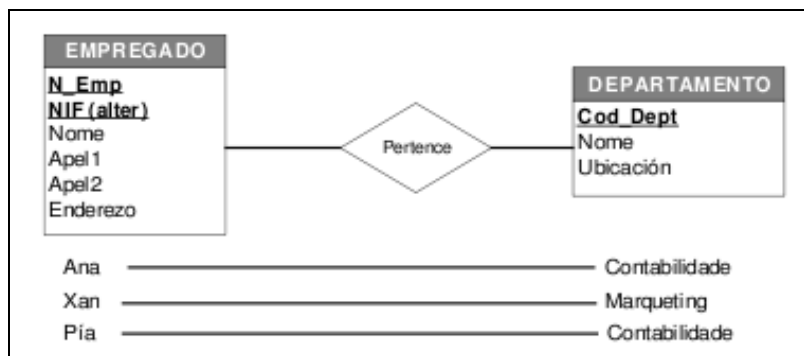
Veremos máis tipos de atributos no modelo E-R estendido.

Para cada atributo hai un conxunto de valores permitidos chamado **dominio**. Por exemplo, o dominio dun atributo idade son enteiros positivos menores que 120. Os dominios non se representan no modelo E-R.

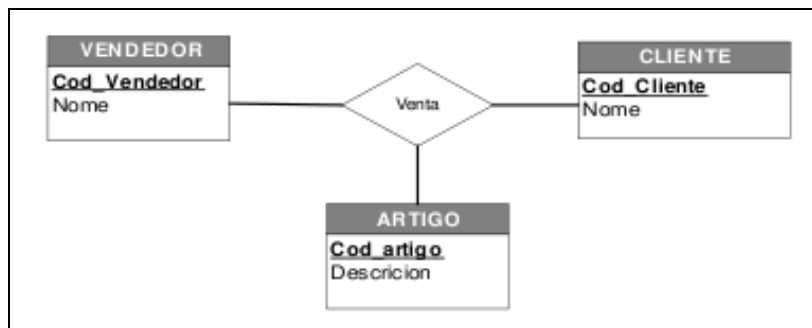
Relacións

As entidades non están illadas no mundo real. Por exemplo, un empregado traballa nun departamento. As relacións (ou interrelacións) establecen

vínculos entre entidades. No modelo E-R represéntase cun rombo unido ás entidades mediante liñas. Dentro do rombo escríbese o nome da interrelación en minúsculas, que en xeral, adoita coincidir cun **verbo en infinitivo**.

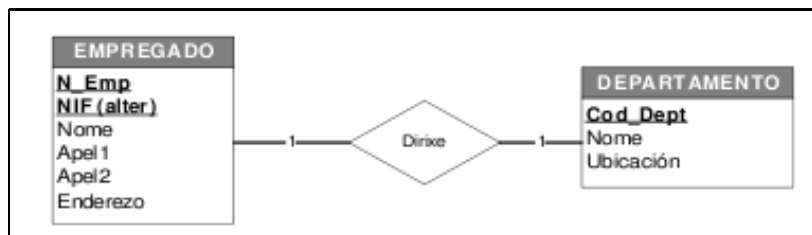


O **grao** dunha relación indica o número de entidades que relaciona. A anterior é unha relación binaria (grao 2). Se relacionamos tres entidades falamos dunha relación ternaria ou de grao 3. A relación que aparece de forma habitual no modelado dunha BBDD é a binaria. No seguinte gráfico vemos unha relación de grao 3.

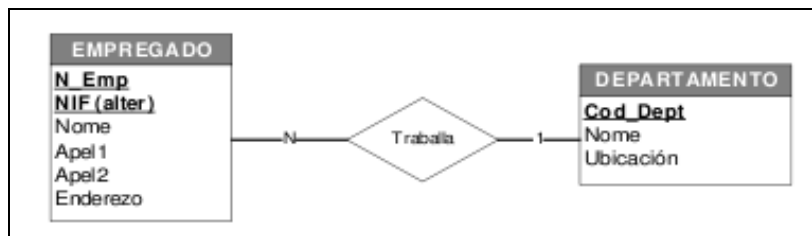


Existen diferentes relacións en función do número de exemplares ou instancias das entidades que se relacionan:

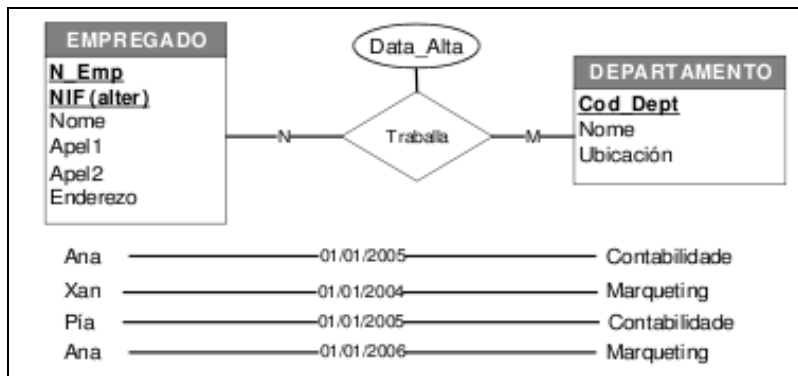
- **Relación 1 a 1.** Por exemplo, no caso de que un DEPARTAMENTO teña un único EMPREGADO e que ese EMPREGADO só traballe nun DEPARTAMENTO.



- **Relación 1 a N.** Por exemplo, un caso parecido ao anterior, pero noutro universo de discurso, no que un DEPARTAMENTO teña varios EMPREGADOS e que eses EMPREGADOS só traballe nun DEPARTAMENTO.

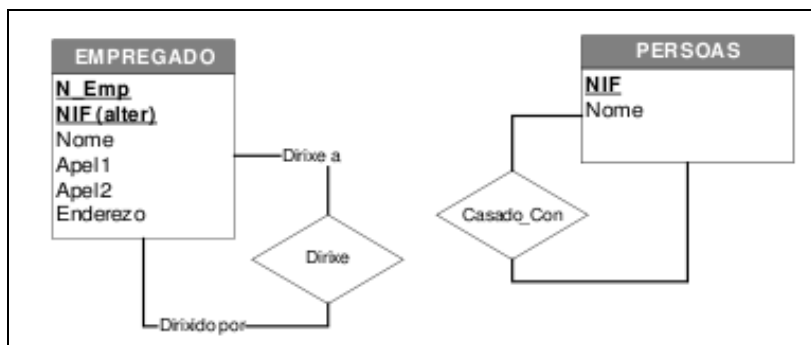


- **Relación N a M.** Por exemplo, un EMPREGADO puido traballar en varios DEPARTAMENTOS ao longo da súa vida e nun DEPARTAMENTO traballan moitos EMPREGADOS. Pódense especificar límites concretos, tanto para N como para M. Por exemplo, un curso pode non ter máis de 30 alumnos matriculados como veremos en breve.

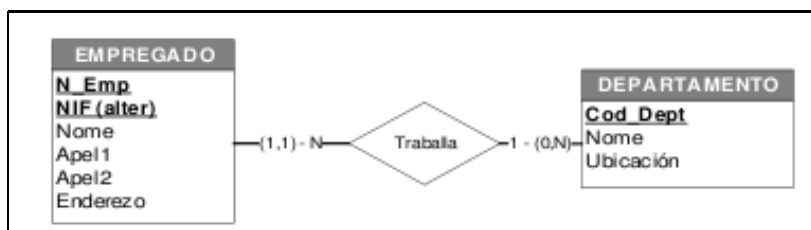


Neste último exemplo, vemos que unha relación tamén pode ter atributos que terán sentido unicamente no caso de que se estableza a relación entre as entidades que as une, como pode ser a datas en que se traballou nun determinado departamento. Se nas especificacións di que interesa rexistrar cando se dá de alta un empregado nun departamento, ese atributo non é do empregado, nin do departamento, senón dos dous en conxunto. A data non tería sentido se o empregado non traballou nunca nun departamento.

Con todo, existe un tipo especial de relación na que só participa un tipo de entidade. A estas relacións chámaselles **recursivas** e asocian entidades do mesmo tipo. Por exemplo, un empregado dunha empresa pode dirixir a outro empregado.



A **cardinalidade** dun tipo de entidade que intervéñen nunha relación defínese como o número mínimo e o número máximo de exemplares dun tipo que poden relacionarse cun elemento do outro tipo de entidade. Para representar as cardinalidades utilizamos un par (x, e) situado sobre a liña que une o tipo de entidade coa interrelación, onde x indica o número mínimo e e o número máximo. Ademais, cando a cardinalidade máxima é N, pódese debuxar unha punta de frecha cara á entidade correspondente.

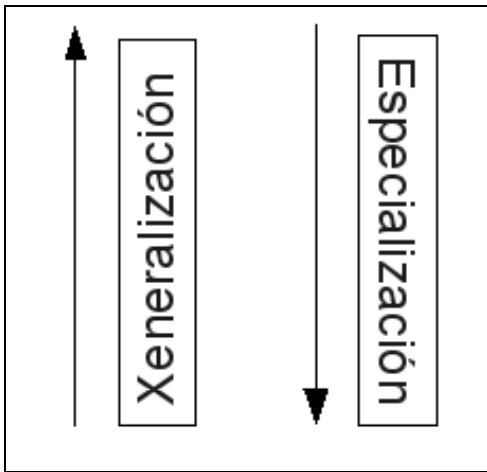


No anterior exemplo un empregado como mínimo traballa en 1 departamento e como máximo tamén en 1. Nos departamentos considerouse a posibilidade de que nalgún momento estean sen empregados (0) e como máximo pode ter N traballadores.

Extensións ao modelo E-R

Ao modelo E/R proposto por Chen engadíronselle algunhas extensións para darlle máis riqueza semántica, é dicir, para que se adaptase mellor á realidade que queremos modelar. A este modelo ampliado chámaselle modelo **E-ER**.

Especialización e xeneralización



A **especialización** proporciónanos un mecanismo de abstracción que permite descompor unha entidade (que se denominará **supertipo**) en **subtipos** ou **subclases**. Así, por exemplo, un "Coche" é un "Vehículo" e un "Camión" é un "Vehículo". Neste caso, "Vehículo" pode considerarse o supertipo, mentres que "Coche" e "Camión" son subtipos de "Vehículo". Os exemplares ou exemplares de "Coche" sono tamén de "Vehículo e igual sucede cos de "Camión".

A **xeneralización** é o proceso inverso á especialización, consiste en identificar as características comúns (atributos e relacións) de varias entidades tipo e xeneralizalas nunha superclase simple. As entidades tipo orixinais dan lugar a subclases.

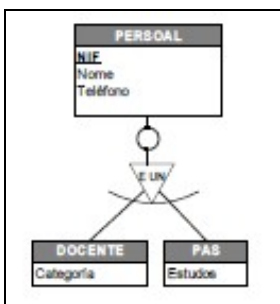
Como vemos, o resultado dunha xeneralización é dunha especialización é o mesmo, unha **xerarquía** na que unha das características máis importantes é a **herdanza**:

- Os **atributos** da superclase son herdados polas subclases.
- As entidades das subclases herdan todas **as relacións** da superclase.
- A **chave** de calquera instancia dunha subclase é a mesma que a da superclase.

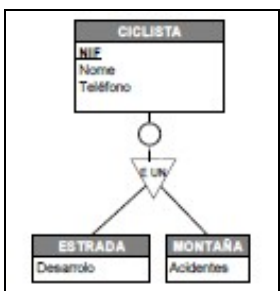
Á asociación entre a superclase e calquera das subclases chámase **É-UN / UNHA**, xa que se di que ?un coche É UN vehículo?.

Hai catro tipos de relacións de xeneralización:

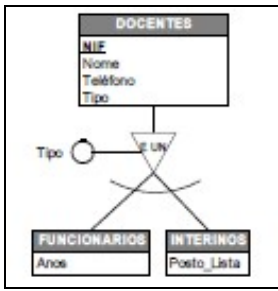
- **Total sen solapamento.** Todo o persoal dun instituto ou é docente ou é PAS (Persoal de Administración e Servizos).



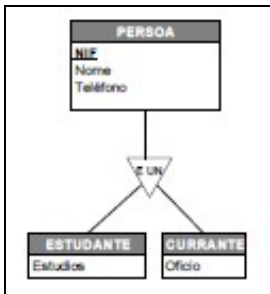
- **Total con solapamento.** Un ciclista, polo mero feito de selo, debe practicar como mínimo unha das modalidades que existen, (aínda que existen moitas, neste exemplo imos supoñer que só existen dúas). Co cal ou fai estrada ou montaña ou fai as dúas.



- **Parcial sen solapamento.** Non todo docente é funcionario ou interino, (pode estar en prácticas, substituto, etc), pero se é funcionario non pode ser interino e viceversa.

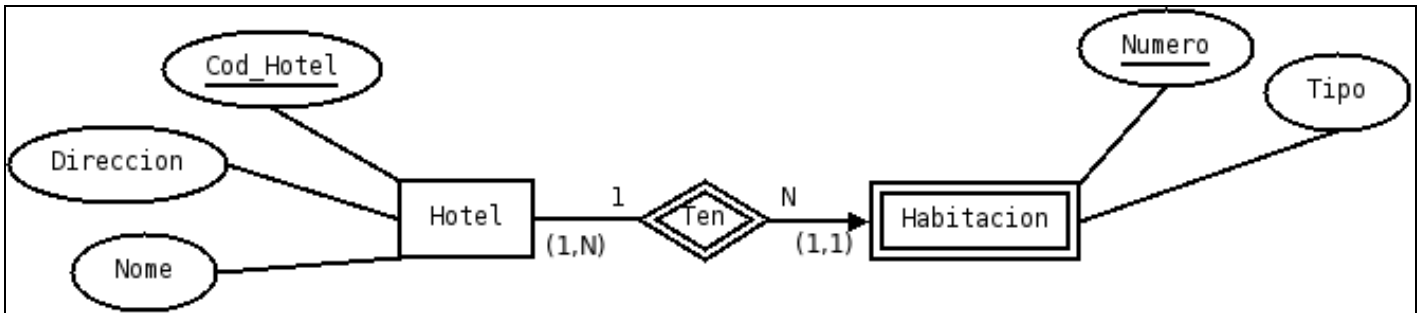


- **Parcial con solapamento.** Toda persoa pode ser estudante e currante ao mesmo tempo, pero tamén pode que estea no paro e non estude. Unha persoa pode non aparecer en ningunha das subclases, nunha delas ou nas dúas simultaneamente.



Entidades débiles

A identificación dunha entidade débil depende doutra entidade, chamada forte. Por exemplo, nunha BBDD de hoteis, as habitación non se poden identificar se non sabemos a que hotel pertencen.



--Arribi 12:09 1 oct 2009 (BST)