



Universidad
Francisco de Vitoria
UFV Madrid

Fundamentos de la ingeniería informática

Ingeniería de sistemas industriales

Curso 2019-2020

Bases de datos

1. Base de datos

Una base de datos es un conjunto de datos relativos al mismo tema, organizada sistemáticamente de modo que sea posible recuperar información bien de los datos simples, bien de la conjunción de esos datos, en razón de las relaciones que se establezcan entre ellos.

Una base de datos no necesariamente se refiere a datos en soporte digital, bibliotecas, hemerotecas, registros (de la propiedad, de la propiedad intelectual, de desarrollo, de patentes...) son bases de datos también.

Evidentemente, dada la capacidad de proceso (recuperación de información) que tienen las bases de datos digitales, todas las demás han migrado a este formato o a un modelo mixto en el que conviven elementos digitales y elementos físicos (papel por lo general).

Las unidades básicas de información de una base de datos se denominan registros. Un registro representa digitalmente un “objeto” del mundo real con expresión de los parámetros que los caracterizan (atributos)

De este modo una persona podría venir caracterizada por su NIF, nombre, apellidos, sexo, nacionalidad, fecha de nacimiento, calle domicilio, ciudad domicilio, provincia domicilio, ciudad de nacimiento, provincia de nacimiento, nombre del padre y nombre de la madre.

Así Juan Pérez Gómez podría tener un registro como el siguiente

Juan Pérez Gómez = {
NIF = 01234567A
nombre = Juan
apellidos = Pérez Gómez
sexo = V
nacionalidad = ESP
fecha de nacimiento = 1/10/2000
calle domicilio = Paseo de la Castellana 12 3º A
ciudad domicilio = Madrid
provincia domicilio = Madrid
ciudad de nacimiento = Getafe
provincia de nacimiento = Madrid
nombre del padre = Juan
nombre de la madre = Maria

Juan Pérez Gómez es un registro que tiene un formato (estructura de los atributos) que denominamos **tipo de registro**, que en este caso llamaremos **tipo de registro: Ciudadano**

Son tres los modelos básicos de organización de los registros una base de datos: jerárquicas, en red y relacionales.

1.1. Bases de datos jerárquicas

En una base de datos jerárquica los registros se organizan jerárquicamente en una estructura en forma de árbol: Existe un registro raíz del cuál cuelgan registros hijos, de cada uno de los cuales pueden colgar otros registros, cadena que se puede prolongar todo lo que sea necesario. Lo normal es que cada generación se corresponda con un tipo de registros.



1.2. Bases de datos en red

En las bases de datos jerárquicas cada registro, a excepción del registro raíz, tiene un único antecesor y 0, 1 o más registros descendientes. Ningún registro puede ser antecesor de ninguno de sus antecesores. Cuando no se cumplen estas reglas nos encontramos en una base de datos en RED en la que cualquier registro puede estar vinculado a cualquier otro registro de la red siempre cada vínculo tenga sentido en el contexto de la base de datos.

1.3. Bases de datos relacionales

El modelo de bases de datos relacionales maneja conceptos completamente distintos. Una base de datos modela los distintos grupos de entidades del mundo real y modela también las relaciones entre ellos (de ahí su nombre). Estas son las más populares en la actualidad y por ello ocuparán el resto presente tema.

Cada base de datos implementa un modelo de diseño de un tipo denominado Modelo Entidad-Relación que representa gráficamente, mediante los diagramas entidad-relación, los grupos entidades y las relaciones entre ellas, permitiendo encontrar incoherencias y redundancias en la información contenida lo que permite optimizar su funcionamiento.

2. Modelo Entidad-Relación

El propósito del modelo entidad-relación es permitir la descripción conceptual de una organización que puede ser diseñada sin tener que atender a temas de eficiencia de la base física, cosa que tiene principal importancia en los otros modelos.

Una **entidad** modela un “objeto” existente en la realidad y es distinguible de otras. Por “objeto” queremos decir cualquier objeto físico, pero también cualquier concepto abstracto que sea distinguible de otros de su tipo. Por ejemplo: cada ciudadano es una entidad, cada coche también lo es, pero de igual forma es una entidad cada el estado del tiempo atmosférico cada día a cada hora.

El conjunto de todas las **entidades similares** constituye un **conjunto de entidades**. Son conjuntos de entidades: todos los ciudadanos, todos los ciudadanos pelirrojos, todos los Ferrari Testarossa, todos los registros pluviométricos de Boadilla, todas las obras literarias publicadas en España en 2015, todas las sinfonías de Beethoven.

2.1. Atributos

Una de las etapas clave para definir un modelo para representar la realidad (un aspecto de la realidad) es definir los conjuntos de entidades vinculados a esa realidad. Será necesario caracterizar a todos los miembros de un conjunto de entidades mediante una colección de características que llamamos **atributos**. Bajo esta perspectiva, “similitud de entidades” implica que es posible determinar todos y cada uno de esos atributos para cada una de las entidades que componen el grupo; y al mismo tiempo “distinguibilidad entre entidades” implica que no es posible encontrar dos entidades que tengan exactamente iguales todos sus atributos.

Cada atributo de una entidad toma valores dentro de un determinado dominio. Mayoritariamente el dominio para un atributo será un conjunto de enteros, números reales o cadenas de caracteres, pero son posibles otros tipos de valores.

2.2. Claves

La selección de los atributos relevantes para los conjuntos de entidades es otra etapa relevante del diseño de una base de datos. Existen atributos individuales o subconjuntos de atributos que por sí solos pueden caracterizar (distinguir de las otras) cada entidad. Por ejemplo: nacido en Getafe por sí sólo no distingue a un ciudadano de todos los demás, en particular no lo distingue de los demás ciudadanos nacidos en Getafe; sin embargo, su NIF sí pues es único para cada ciudadano español. El nombre por sí solo no lo es, pero en combinación puede serlo (dependiendo del alcance de la base de datos). El atributo o conjunto de atributos que distinguen sin ambigüedad las distintas entidades de un grupo se denomina **clave** del grupo.

El criterio de distinguibilidad exige que todo grupo de entidades tenga al menos una clave. Cuando de los atributos escogidos no es posible encontrar una clave la solución más común es asignar arbitrariamente un código numérico secuencial que funcione como clave.

2.3. Esun

Cuando decimos que A **esun**¹ B, siendo A y B grupos de entidades, queremos decir que B es una generalización de A, lo que es lo mismo que decir que A es una especialización de B, lo que es lo mismo que las entidades de A también son entidades de B, pero que por ser de A tienen algunos atributos específicos de ellas. Por ejemplo: Camión esun Automóvil, pero para los camiones queremos especificar además de los argumentos genéricos de los automóviles (combustible, nº de marchas, potencia, par motor), por ejemplo, la carga máxima autorizada, tipo de carné necesario, etc.

Esun es entonces una relación entre grupos de entidades. La razón principal para declarar la relación esun entre dos grupos de entidades A y B es declarar que A hereda los atributos de B, es decir todos los miembros de A tienen los atributos de B y alguno más que no tiene sentido para todos los miembros de B. Técnicamente cada miembro a de A está relacionado *exactamente* con un miembro b de B, ya que en la realidad a y b son la misma entidad. Ningún b puede estar relacionado con varios a , pero un b puede no estar relacionado con ningún a .

Los atributos clave de A son en realidad los atributos clave de B, y estos atributos no serán parte de la definición de A.

2.4. Relaciones

Una relación entre grupos de entidades es formalmente una lista ordenada de grupos. Un grupo concreto puede aparecer más de una vez en esa lista.

Si existe una relación \mathcal{R} entre una lista de grupos (E_1, E_2, \dots, E_k) una **instancia** (ejemplo) de \mathcal{R} es una **tupla** (conjunto de K valores) donde cada miembro es una entidad de cada grupo (e_1, e_2, \dots, e_k) y por ello diremos que e_1, e_2, \dots, e_k están relacionados por \mathcal{R} .

2.5. Atributos clave prestados

En ocasiones queremos que la clave de un grupo A sean atributos de otro grupo B con el que A esta conectado con una relación distinta de esun. Sólo es necesario que \mathcal{R} asigne a cada a de A un único b de B con el cual esté relacionado.

3. Diagramas entidad-relación

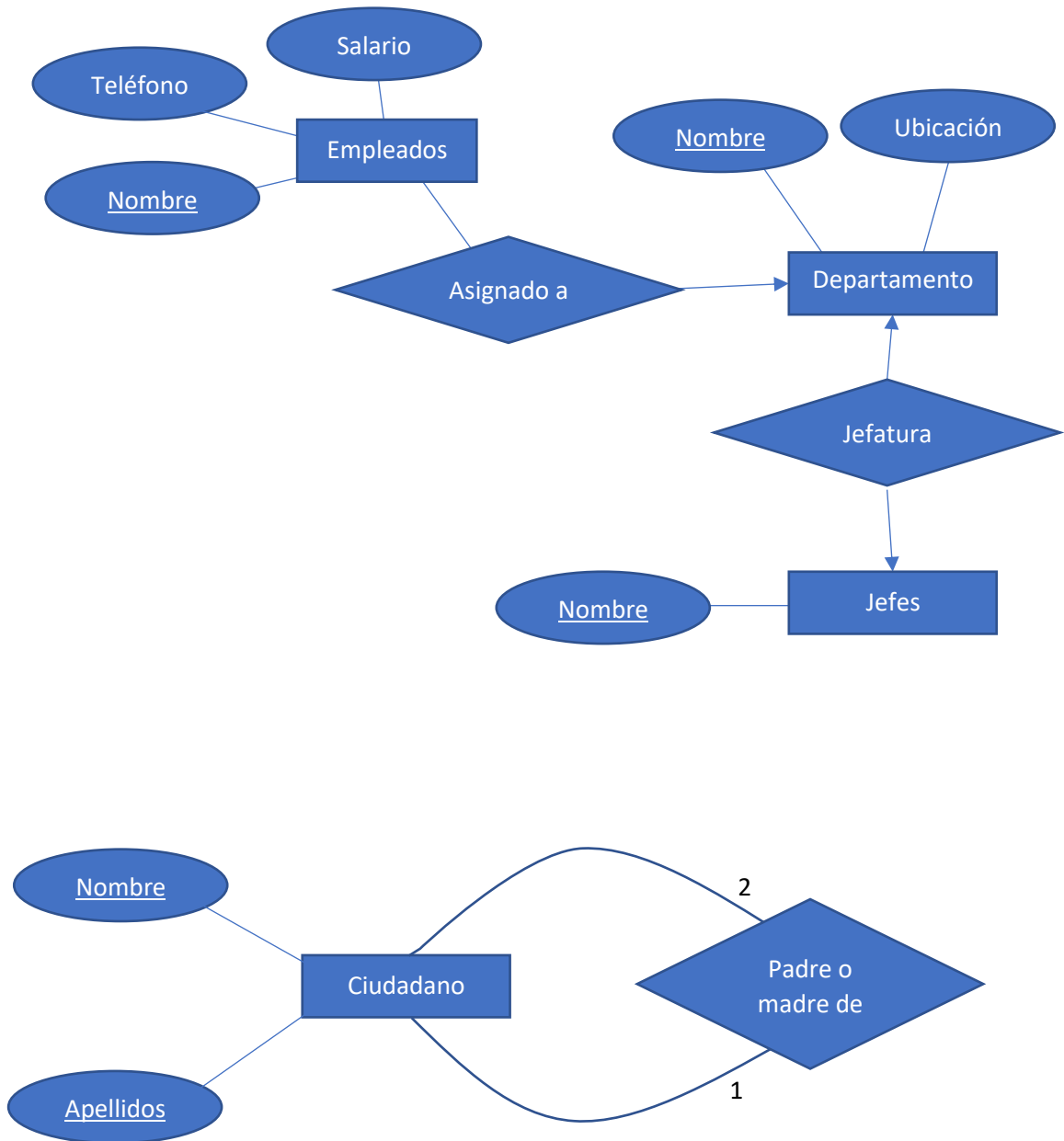
Es muy útil resumir la información del modelo entidad-relación en un diagrama representativo. El convenio de representación es:

- Los rectángulos representan los grupos de entidades
- Los círculos representan atributos enlazados con sus grupos por enlaces no direccionales. Los atributos que forman parte de la clave del grupo aparecerán subrayados.

¹ esun es la traducción literal del término inglés isa (is a)

- Los rombos representan las relaciones enlazados con las entidades que componen la relación mediante enlaces direccionales o no direccionales. El orden en que aparecen las entidades en la relación puede ser indicada numerando los enlaces, aunque ese orden es irrelevante salvo en los casos en los que una entidad aparece múltiples veces en la relación.

Ejemplos:



3.1. Cardinalidad de las relaciones

Para modelar el mundo real adecuadamente, es habitual necesitar clasificar las relaciones en función de cuántas entidades de un grupo están relacionadas con cuántas entidades de otro.

La menos habitual de las formas de relación entre dos grupos es **uno-a-uno**, significando que, si $A \mathcal{R} B$, para cada una de las entidades de A existe como mucho una entidad B relacionada con ella y viceversa, para cada entidad de B como mucho existe una entidad de A relacionada con ella. El ejemplo en los diagramas anteriores es la relación Jefatura, en la cual cada departamento tiene un único jefe y cada jefe jefatura único departamento, pero sin embargo es posible un departamento sin jefe o un jefe que no tenga asignado departamento.

En una relación **uno-a-n** $A \mathcal{R} B$, una entidad de a puede estar relacionada con una, varias o ninguna entidad de B; o a la inversa, en una relación **n-a-uno** una entidad de b estará relacionada con ninguna, una o varias entidades de A. Un ejemplo es la relación "Asignado a" cada empleado está asignado a un departamento, pero un departamento puede tener asignados múltiples empleados.

Finalmente, en una relación **n-a-n** $A \mathcal{R} B$ cada entidad de A puede estar relacionada con ninguna, una o varias entidades de B y viceversa, cada entidad de B puede estar relacionada con ninguna una o varias entidades de A. El ejemplo es la relación "Padre o madre de" en la que un ciudadano puede ser padre o madre de varios y a su vez tener como padre o madre a varios ciudadanos.

Ya que las relaciones muchos-a-muchos aparecen con frecuencia en la práctica, hemos de ser especialmente cuidadosos a la hora de expresarlas en el esquema conceptual. Las bases de datos comunes no permiten implementar directamente este tipo de relaciones por lo que deben ser descompuestas en varias relaciones uno-a-n.

3.2. Expresión de la cardinalidad

Como se puede inducir de los ejemplos, la cardinalidad de la relación se expresa utilizando enlaces direccionales (flechas) o no direccionales (arcos). El concepto n se expresa con un arco entre la relación y la entidad mientras que el concepto de 1 se expresa con una flecha desde la relación hacia la entidad.

uno-a-uno



uno-a-n



n-a-n

