

Introducción a las Bases de Datos y Bases de Datos

MODELO DE ENTIDADES Y RELACIONES

2
0
1
7



Tecnicaturas TUPAR y TUDAI

REPRESENTACIÓN DE LA INFORMACIÓN

- La estructura y el contexto le dan **significado** a los datos posibilitando su entendimiento.
- El lenguaje natural es nuestra forma primaria de representación y comunicación de información pero NO es el mejor medio.
- En muchas situaciones resulta más útil establecer formas especializadas de **representación** de la información. Las fórmulas matemáticas, los mapas carreteros y las partituras musicales, constituyen ejemplos de **modelos** especializados de representación simbólica de la información.

Fuente: Apuntes de Cátedra. Bases de Datos. Prof. J. Ale y G. Dejean. Fac. Ingeniería UBA.

REPRESENTACIÓN DE LA INFORMACIÓN

Que es un **dato**? → una pieza atómica que se utilizará para construir el concepto de información.

Una pieza atómica de dato es un valor, que debe ser interpretado como una propiedad perteneciente a un hecho o cosa del mundo real. La interpretación es necesaria para ubicar el **dato** en **contexto**.

Por ej.: 1990 sin un contexto podría significar el nro. de una calle, un año, un precio, etc.

Pieza atómica = significado + valor

A partir de esas piezas atómicas de datos se puede construir una descripción o contexto de un objeto, idea, cosa del mundo real.

REPRESENTACIÓN DE LA INFORMACIÓN

Por ej.: un estudiante de esta facultad podría ser simbolizado de la siguiente forma:

<**LU**:123456; **Nombre y Apellido**: Juan Torres; **Fecha nacimiento**: 23/08/1995; **carrera**: Ing en Sistemas>

Cuando se trata de conjuntos de descripciones como la anterior usualmente se utiliza una notación simplificada que sólo contiene los valores (generalmente diferentes). Esa notación simplificada se denomina **TUPLA o REGISTRO**

Por ej.:

<123456; Juan Torres; 23/08/1995; Ing en Sistemas>

y la interpretación de esos valores es aplicable a todas las tupla por igual. elementos

<**LU**; **Nombre y Apellido**; **Fecha nacimiento**; **carrera**>

MODELO DE DATOS

Es una **herramienta intelectual** que permite plasmar una interpretación de un conjunto de aspectos del mundo real:

- tiene **poder expresivo** para representar cómo están relacionados los datos
- es **abstracta**, para ser mínimamente perturbable ante los cambios concernientes al aspecto evolutivo del mundo real.

Es un **mecanismo de abstracción**, que permite ver el contenido de información de los datos, **en lugar de sus valores individuales**.

Los modelos se usan extensivamente en muchas disciplinas para mejorar la comprensión y abstraer los detalles (modelos físicos, matemáticos, económicos, etc.)

En nuestra disciplina interesa comprender y elaborar **MODELOS DE DATOS**, que puedan ser codificados y manipulados computacionalmente.

DISEÑO DE UN MODELO DE DATOS

Modelado de Datos → organizar los datos para que representen una situación del mundo real con la mayor fidelidad posible, con el objetivo de poder manejarlos computacionalmente ...

No es posible tener un conocimiento completo del mundo real → centrar la atención en la información relevante, ocultando o ignorando otra que **NO** sea relevante o que resulte inadecuada para el propósito requerido.

Comprender las características de la información relevante **para determinar cómo van a ser organizados y procesados los datos** → pueden ser descriptas a través de enunciaciones generales., por ej. lenguaje natural

Un conjunto formal y consistente de tales enunciaciones define un **modelo conceptual de datos**.

DISEÑO DE UN MODELO DE DATOS

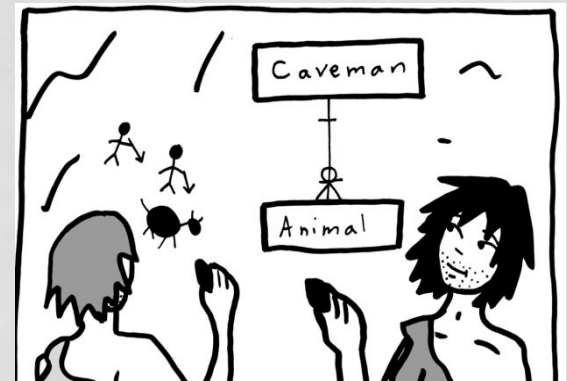
Etapas para encontrar una forma eficaz de representar la información en el mundo computacional:

identificación de los datos de la realidad: actores, recursos, objetos, etc. del mundo real de los cuales interesa guardar información

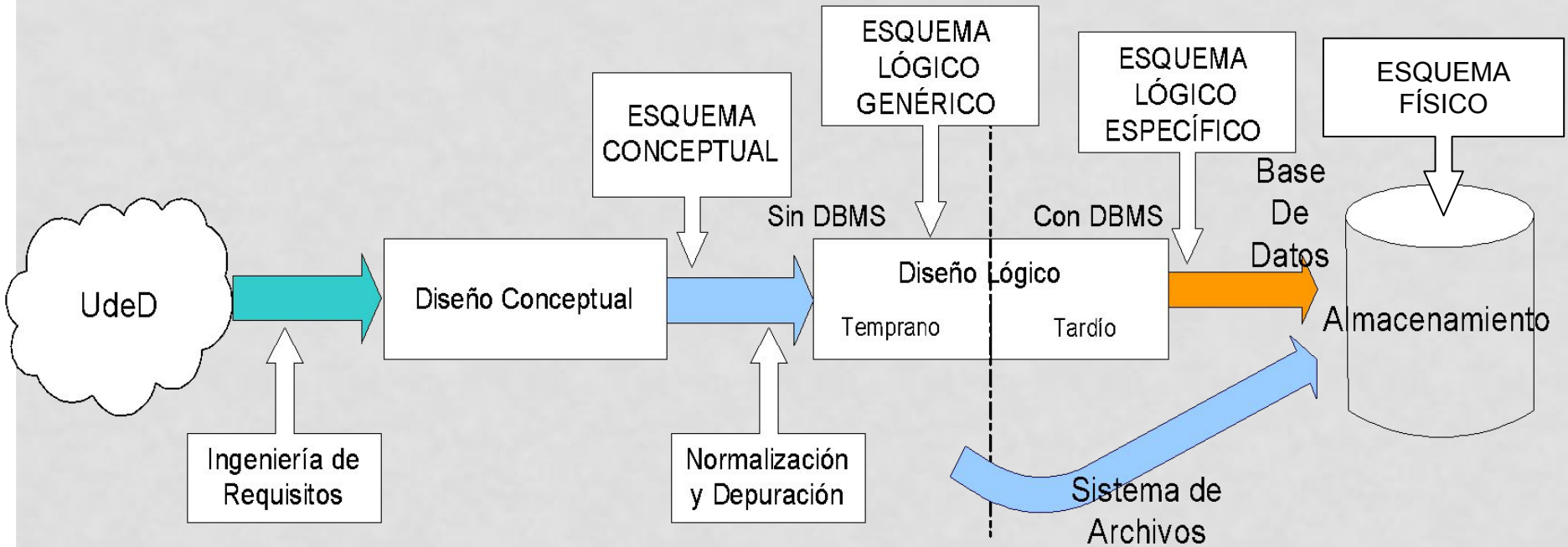
identificación de relaciones entre datos: detección de los vínculos significativos que se dan entre los elementos de las etapas anteriores.

abstracción de datos y relaciones: representación simbólica sólo de los elementos detectados en las etapas anteriores

MODELO CONCEPTUAL DE DATOS (MCD)



ETAPAS EN EL DISEÑO DE DATOS



MODELO CONCEPTUAL DE DATOS

- Se destaca el **Modelo de Entidades y Relaciones** (Modelo E/R - MER), propuesto por Chen en dos artículos ya icónicos, 1976 y 1977.
- Según Chen, “El Modelo E/R puede ser usado como una base para una vista unificada de los datos”, adoptando “el enfoque más natural del mundo real que consiste en entidades e interrelaciones (relaciones)”.
- Posteriormente otros autores lo han extendido con importantes aportes, lo que ha dado lugar a una familia de Modelos de Datos (MER Extendido – MERExt).

MODELO DE ENTIDADES Y RELACIONES EXTENDIDO

- Es un modelo de datos semántico, capaz de describir los requerimientos de datos para un sistema de información de forma directa, con una notación gráfica fácil de entender.
- El MERE ha tenido una gran difusión en la comunidad informática dedicada al desarrollo de aplicaciones generales sobre bases de datos → ha sido el modelo más extendido en las herramientas CASE de ayuda al diseño de BD.
- Es simple y poderoso para modelar abstracciones del mundo real.
- Es una representación gráfica de un modelo de datos, fácilmente traducible a un esquema de BD:

Reglas de transformación → esquema relacional → esquema lógico

- Se ha convertido en un estándar 'de facto', incluso muchas herramientas de diseño de BD utilizan sus conceptos.

CONCEPTOS BÁSICOS

El Modelo E/R (MER), tal como fue propuesto por Chen, distinguía los siguientes elementos estáticos:

- **Entidad:** objeto real o abstracto que existe en la realidad y acerca del cual se desea almacenar información.
- **Relación:** (o Interrelación) asociación o vinculación entre entidades.
- **Atributos:** Los atributos son características de las entidades y de las relaciones, que proveen detalles descriptivos acerca de ellas.

Diferentes tipos de interrelaciones, entidades y características de los atributos conducen al MERE.

REPRESENTACIÓN DE ELEMENTOS BÁSICOS

- DIAGRAMA DE ENTIDADES Y RELACIONES

Entidades



Atributos



Relaciones
entre entidades



ENTIDADES

Otras definiciones de Entidad:

- “Cualquier objeto (real o abstracto) que existe en la realidad y acerca del cual queremos almacenar información en la base de datos”.
- “Algo con realidad objetiva que existe o puede ser pensado”.
- “Una persona, lugar, cosa, concepto o suceso, real o abstracto, de interés para la empresa”.
- “Objetos (hechos, cosas, personas,...) que tienen propiedades en común y una existencia autónoma.”

CONJUNTO ENTIDAD

Abstrayendo las características comunes de un conjunto de entidades se obtiene:

- **Conjunto Entidad:** Representa el tipo de entidades o la estructura genérica que describe un conjunto de entidades. Se lo identifica con un nombre (en singular por lo general) que caracteriza el conjunto entidad (aunque generalmente es mencionado simplemente como entidad).
- Ej. todos los Alumnos → Entidad **ALUMNO**
 - Todas las entidades de un conjunto tienen el mismo conjunto de atributos que interesa modelar. Todos los ejemplares de ALUMNO, tendrán LU, Nombre, Apellido.

Cada conjunto entidad necesita un **identificador**: atributo o conjunto de ellos que permita identificar a cada uno de los ejemplares que componen el conjunto entidad. La Entidad ALUMNO tiene como atributo identificador de los ejemplares a LU.

CONJUNTO ENTIDAD

Está formado por ejemplares o instancias, que son

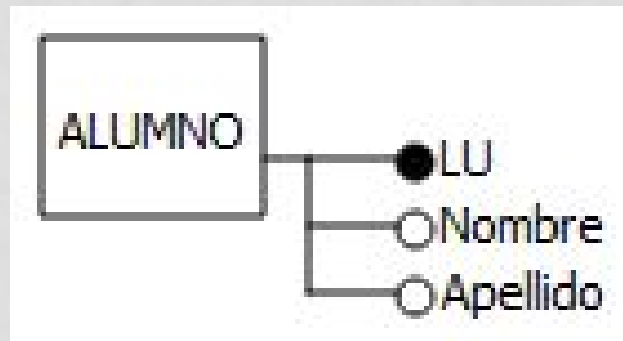
→ entidades, cada una representa simbólicamente a

→ Un objeto del mundo real, distinguible entre otros

Se describe por medio de sus propios atributos.

Los objetos individuales son **instancias** de la entidad

(Conjunto) Entidad



Instancias o ejemplares
(individuales)

123, Carlos, Sánchez
124, Miguel, Rodríguez
125, José, González
126, Agustín, García

.....

ENTIDADES

Existen dos categorías de tipos de entidades:

- **Regulares o fuertes**, que son aquellas cuyos ejemplares tienen existencia por sí mismos (ej. Los ejemplares de ALUMNO)
- **Débiles**, en las cuales la identificación y existencia de un ejemplar dependen de la identificación y existencia de un ejemplar de otro tipo de entidad, por ejemplo:
 - la existencia e identificación de una COPIA_LIBRO depende de la identificación y existencia de un ORIGINAL_LIBRO;
 - el RENGLON_REMITO depende de la identificación y existencia de un ejemplar de REMITO;
 - la identificación y existencia de una PROVINCIA depende de la identificación y existencia del PAÍS al cual pertenece.....

IDENTIFICACIÓN DE ENTIDADES

...problema para identificarlas

- Uno de los problemas que existirán en el diseño del MER es la decisión de si un determinado objeto o concepto se modela como una entidad o no.
- Por ejemplo, el color es habitualmente una propiedad de una entidad (como es el caso del color de un auto), pero en una fábrica de pinturas probablemente sería apropiado modelar el color como una entidad con sus propias propiedades.

Las tres propiedades inherentes de las entidades:

1. Tiene sentido que exista (existencia propia), evaluado esto siempre en el contexto de un sistema destinado a manejar información.
2. Cada ejemplar de un tipo de entidad debe poder distinguirse de los demás,
3. Todos los ejemplares de un tipo de entidad deben tener las mismas propiedades.

IDENTIFICACIÓN DE ENTIDADES

Pero ...

- La primera de estas reglas no sería aplicable a las entidades débiles → que significa 'existencia'?
- La segunda supone la obligación de un identificador que permita distinguir los distintos ejemplares de un tipo de entidad, lo que tampoco es universalmente aceptado (ni por algunos autores, ni por los modelos, ni por los productos) → que hacer con ejemplares 'iguales'?
- La tercera es relativa: ¿exactamente las mismas?, ¿las mismas entre las que nos interesan? → cómo determinar el conjunto de propiedades?

Respuestas ... → experiencia en el diseño, astucia en la definición de identificadores, y otras destrezas...

ATRIBUTOS

- Son los datos relativos a una entidad o relación
- Cada atributo tiene asociado un **Dominio de definición** (entero, cadena de caracteres, fechas, etc.) y puede tomar un cierto valor dentro del dominio
- Un atributo tiene diferentes características: presencia, cardinalidad y rol, también un *origen* y una composición
- *Se colocan junto a la entidad que describen.*



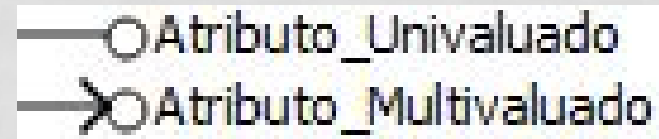
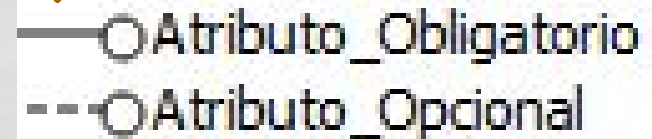
ATRIBUTOS (CONT)

Presencia: indica si

- **siempre** se encontrará un valor para ese atributo (obligatorio)
- habrá casos en los que esté **ausente** (opcional).

Cardinalidad: indica cuántas instancias del atributo pueden encontrarse en la descripción de una entidad o una relación, pueden ser

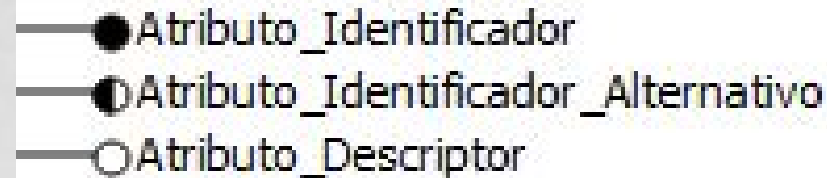
- univaluado un sólo valor para una misma instancia (ej. *Edad*)
- multivaluado si puede haber un conjunto de valores para una misma instancia (ej. *Telefonos, pueden ser varios números*).



ATRIBUTOS (CONT)

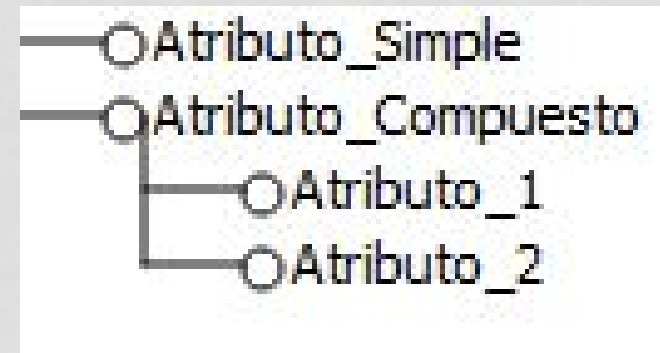
Rol: de acuerdo al rol que tienen dentro de la entidad puede ser:

- Identificador principal (IP) identifica unívocamente cada uno de los ejemplares de la entidad
- Identificador alternativo otro identificador de la entidad que puede cumplir el rol de IP.
- Descriptor atributo que representa una característica de la entidad



Composición: indica si el atributo es

- simple (ej. Nombre, nro.de documento, etc.)
- compuesto por componentes más elementales, interpretándose su valor como la concatenación de los valores correspondientes a los componentes (ej. Dirección: calle, número, piso, dpto.)


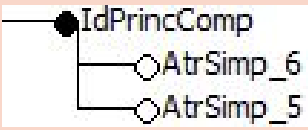


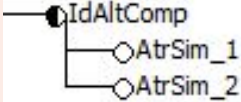
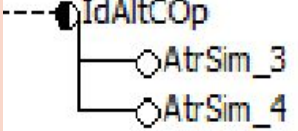


ATRIBUTOS (CONT)

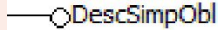
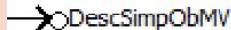
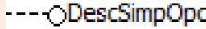
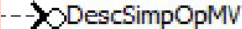
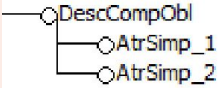
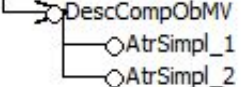

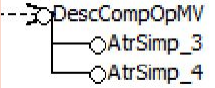
Origen: indica la procedencia del atributo, puede ser

- nativo si pertenece a la descripción de la entidad o relación
- derivado si su valor se obtiene mediante cálculo o algoritmo a partir de otros atributos (x ej. edad obtenido a partir de la fecha de nacimiento y la fecha actual)

ATRIBUTOS IDENTIFICADORES

CASOS				REPRESENTACION
Rol	Composición	Presencia	Cardinalidad	
Identificador Principal	Simple	Obligatorio	Univaluado	 IdPrincSimp
	Compuesto			 IdPrincComp AtrSimp_6 AtrSimp_5
Identificador Alternativo	Simple	Obligatorio		 IdAltSimp
		Opcional		 IdAltSOpc
	Compuesto	Obligatorio		 IdAltComp AtrSim_1 AtrSim_2
		Opcional		 IdAltCOpc AtrSim_3 AtrSim_4

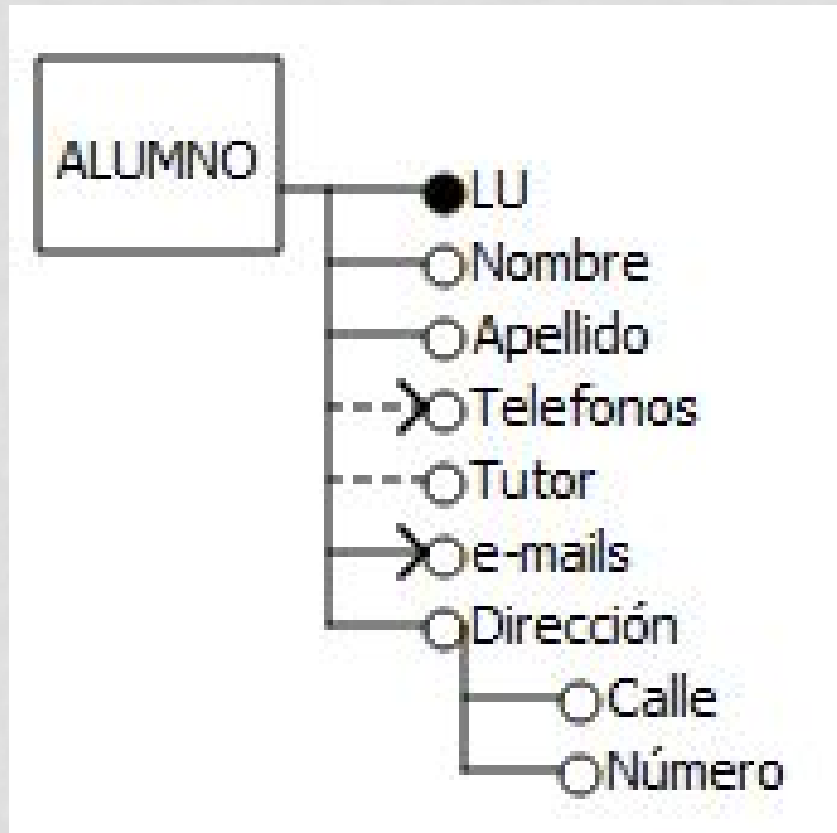
ATRIBUTOS DESCRIPTORES

CASOS				REPRESENTACION
Rol	Composición	Presencia	Cardinalidad	
Descriptor	Simple	Obligatorio	Univaluado	
			Multivaluado	
		Opcional	Univaluado	
			Multivaluado	
	Compuesto	Obligatorio	Univaluado	
			Multivaluado	
		Opcional	Univaluado	
			Multivaluado	

ATRIBUTOS (CONT)

- Todas las instancias de una entidad se describen mediante el **mismo** conjunto de atributos ... (3a. Propiedad de Entidades)
- (Casi) Siempre hay un atributo cuyo valor es distinto para cada instancia de una entidad → **atributo identificador principal** (ej. *Número de libreta de un Alumno, patente de un Automóvil*) ... (2a. Propiedad de Entidades)
- Algunas entidades pueden tener más un atributo identificador → **identificadores alternativos** (ej. *documento del Alumno*) ... (2a. Propiedad de Entidades)

ATRIBUTOS (CONT)



Cada conjunto de entidades debe tener **al menos** un identificador principal.

Ese identificador principal podría ser un atributo compuesto.

¿Podría ser un atributo multivaluado?

¿Podría ser un atributo opcional?

EJEMPLO

Vamos a identificar **de forma precisa** las entidades y sus atributos para el siguiente enunciado:

“Se quiere registrar información para un sistema de compras de productos que realizan clientes de un negocio.”

1. Se necesita guardar los datos de los clientes, los mismos son, un identificador, nombre y apellido, fecha de nacimiento, dirección y su teléfono.
2. Opcionalmente una persona pueda tener otro número de teléfono y varios números de teléfono celular.
3. Los datos de las facturas que se registran son: su tipo y número (que identifica a cada factura), su fecha, el cliente y el importe total. Además se necesita guardar los datos de los productos que compondrán dicha factura.

RELACIONES (INTERRELACIONES)

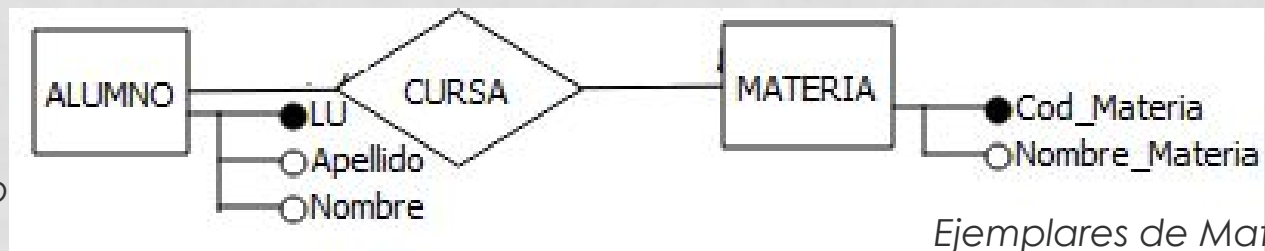
Una relación (interrelación) es una asociación, vinculación o correspondencia entre conjuntos de entidades, y se materializa en un conjunto de asociaciones entre dos o más instancias del mismo o diferente tipo.

Igual que en el caso de las entidades, distinguiremos entre:

- **Conjunto Relación:** el tipo de relación o estructura genérica que describe un conjunto de relaciones, y
- Cada relación, o **instancia de relación** es decir, cada uno de los ejemplares concretos.

RELACIONES (INTERRELACIONES)

CURSA es un tipo de relación que vincula los tipos de entidades ALUMNO y MATERIA.



Ejemplares de Alumno
123, Carlos, Sánchez
124, Miguel, Rodríguez
125, José, González
.....

Ejemplares de Materia
EdD, Estructuras de Datos
BD, Bases de Datos
.....

Un ejemplar del tipo de relación CURSA es la vinculación entre el alumno “123, Carlos, Sánchez” y la materia “EdD, Estructuras de Datos” dado que satisface la frase “123, Carlos Sánchez cursa la materia EdD, Estructuras de Datos”.

TIPO DE RELACIÓN

Características

- **Nombre** al igual que las entidades debe ser único
- **Grado u orden** es el número de tipos de entidades que participan
- **Tipo de correspondencia o cardinalidad o multiplicidad** es el número máximo de ejemplares de un tipo entidad que pueden estar asociados , en una determinada relación, con un ejemplar de otro(s)
- Puede tener atributos propios

Los términos más usuales son los subrayados

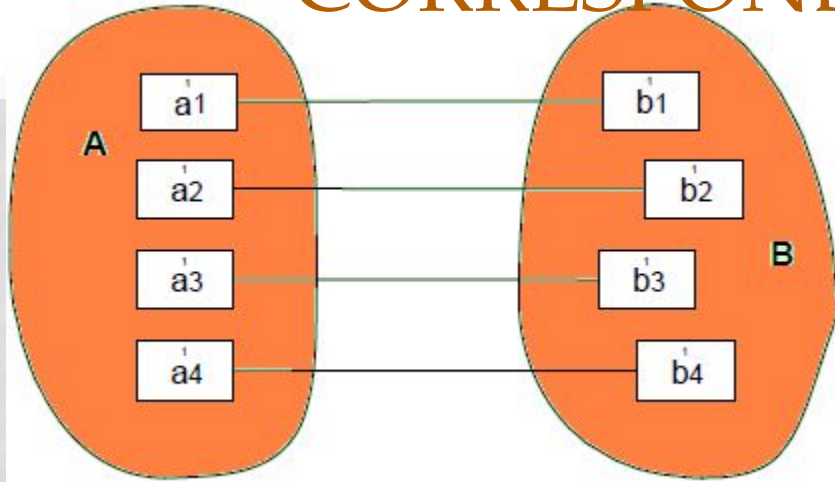
RELACIONES

- Una relación R de orden n (n -aria) relaciona n conjuntos de entidades $E_1 \dots E_n$.
- Cada instancia en R involucra las entidades E_1, \dots, E_n

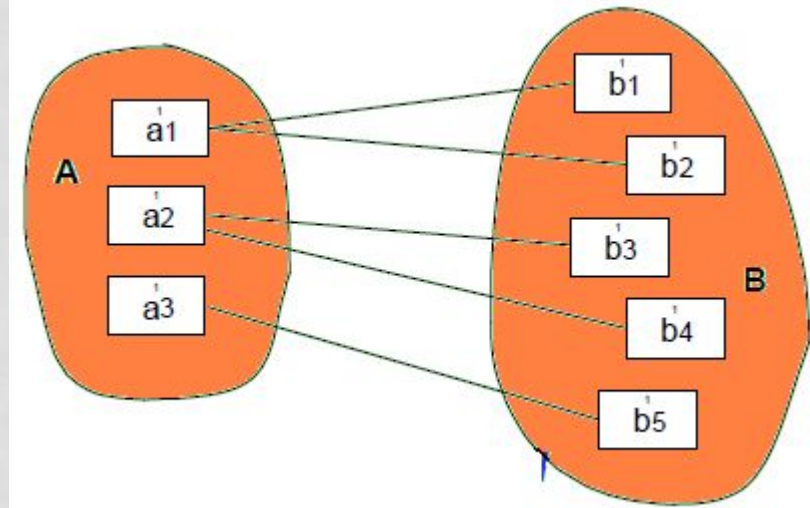
Entonces si:

- $n=1$ la relación se denomina **UNARIA**: una entidad involucrada;
- $n=2$ la relación se denomina **BINARIA**: dos entidades involucradas;
- $n=3$ la relación se denomina **TERNARIA**: tres entidades involucradas;
- ...

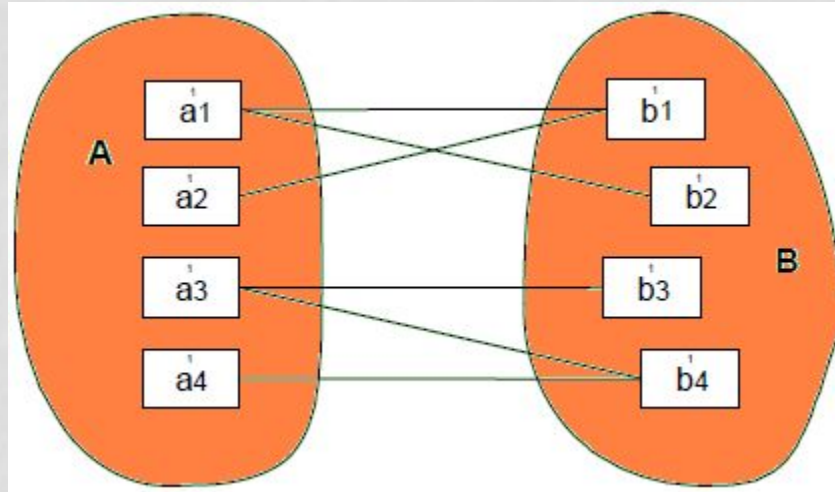
RELACIONES: TIPOS DE CORRESPONDENCIAS



Uno a uno – 1 : 1



Uno a muchos – 1 : N



Muchos a muchos – N : N

RELACIONES: CARDINALIDADES

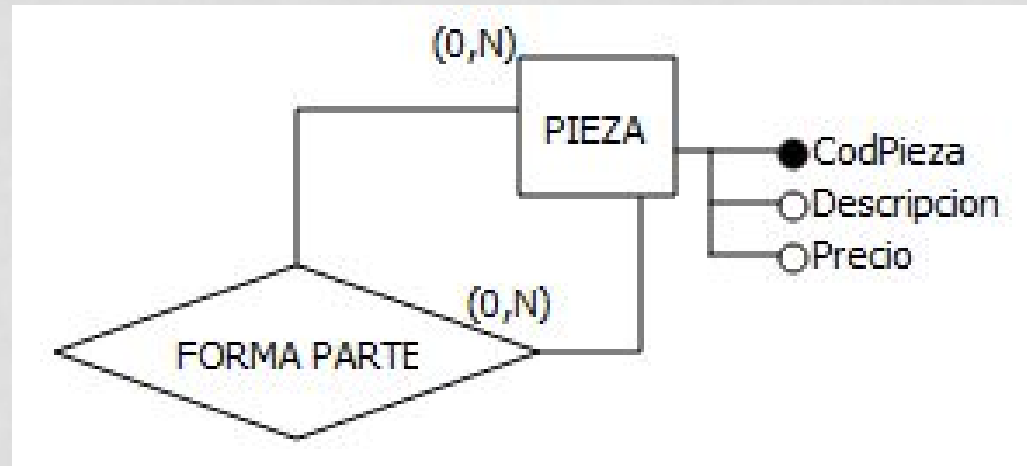
- Esta información se coloca sobre los vínculos (líneas), en el ejemplo encerrado entre paréntesis.
- La lectura que se hace de las cardinalidades se denomina **Look-Across (LA)** o Chen-Style, se lee sobre la línea de la 'entidad destino'
- ¿Cuántos ejemplares de la entidad E1 puede relacionarse con cada ejemplar de la entidad E2, como máximo y como mínimo?



RELACIONES: CARDINALIDADES

- La **cardinalidad máxima** representa el máximo número de ejemplares de una entidad con los que se puede relacionar otra entidad
 - al menos **1** (puede ser 0 o 1)
 - como máximo **N** (muchos o varios, es una cantidad variable)
 - como máximo **a** (es una cantidad fija de ejemplares)
- La **cardinalidad mínima** representa el mínimo número de ejemplares de una entidad con los que se puede relacionar otra entidad
 - cardinalidad mínima **0**, un ejemplar de una entidad **puede** estar relacionado con otro,
 - cardinalidad mínima **1**, un ejemplar de una entidad **debe** estar relacionado al menos con un ejemplar
 - cardinalidad mínima **a**, un ejemplar de una entidad **debe** estar relacionado al menos con **a** ejemplares

RELACIONES UNARIAS (REFLEXIVAS/RECURSIVAS)



Cual es la semántica de esta relación y de sus cardinalidades máximas?

Cada **pieza** forma-parte de otra u otras **piezas**

Cada **pieza** esta-formada-por otra u otras **piezas**

Notar que se describe con hechos afirmativos

RELACIONES BINARIAS 1:N



Cual es la semántica de esta relación y de sus cardinalidades máximas? **1: N**

Cada **carrera** pertenece a un único **departamento**

Cada **departamento** posee muchas **carreras**

RELACIONES BINARIAS N:N

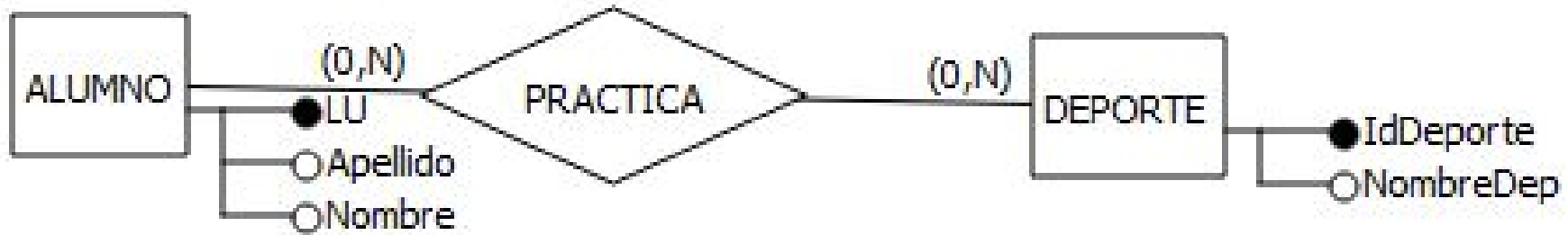


Cual es la semántica de esta relación y de sus cardinalidades máximas? **N:N**

Cada **alumno** practica varios **deportes**

Cada **deporte** es practicado por varios **alumnos**

RELACIONES OPCIONALES VS. OBLIGATORIAS



Cual debería ser la semántica 'COMPLETA' de esta relación, interpretada en el mundo real?

Un alumno puede practicar un deporte? O podría no practicar ninguno?

Un deporte puede ser practicado por varios alumnos? O podría ocurrir que nadie lo practicara?

Hasta ahora sólo se ha indicado que un deporte podría ser practicado por varios alumnos y que un alumno podría practicar varios deportes... los casos más restrictivos no están representados.



RELACIONES OPCIONALES VS. OBLIGATORIAS

CARDINALIDADES MÍNIMAS



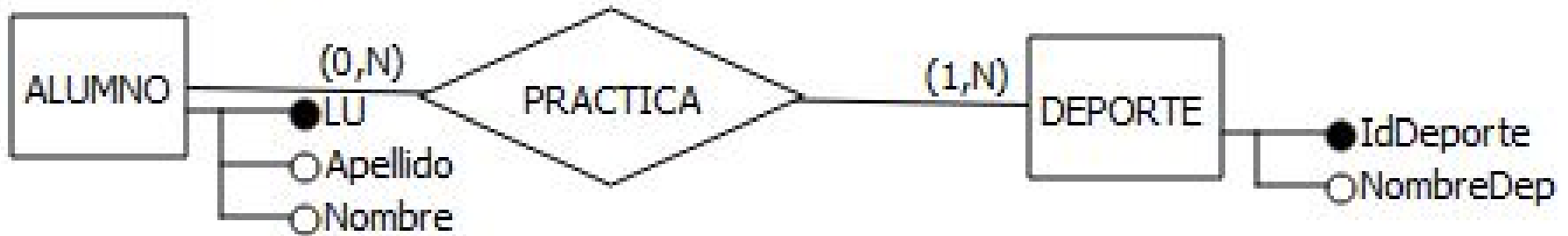
Si un alumno practica al menos un deporte?
 $* , N \longrightarrow 1 , N$

O podría no practicar ninguno?
 $* , N \longrightarrow 0 , N$

Un deporte es practicado por al menos un alumno?
 $1 , N \longrightarrow * , N$

O podría ocurrir que nadie lo practicara?
 $0 , N \longrightarrow * , N$

RELACIONES OPCIONALES VS. OBLIGATORIAS CARDINALIDADES MÍNIMAS



Este diagrama expresa que:

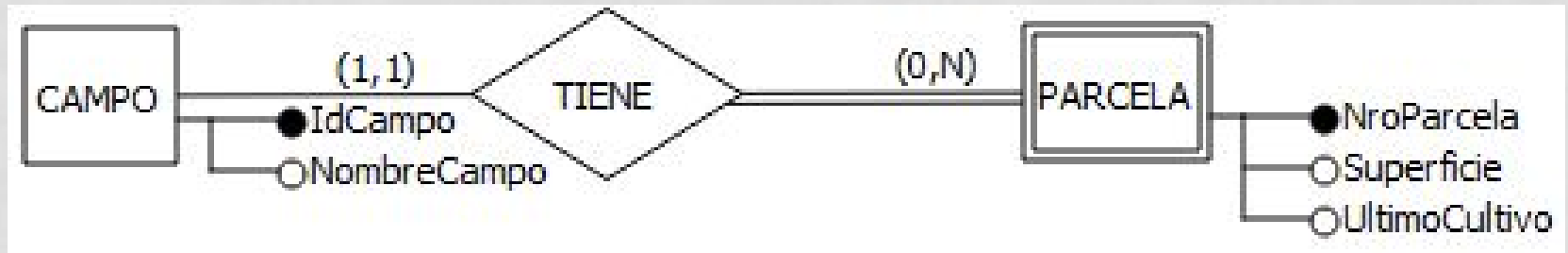
- Un alumno practica **al menos** un deporte y podría practicar **varios**.
- Un deporte **podría no ser practicado por ningún** alumno, pero puede ser practicado por **uno o más alumnos**.

En la cardinalidad mínima

- 0 indica **OPCIONALIDAD**
- 1 indica **OBLIGATORIEDAD** (relación mandatoria)

RELACIÓN ENTIDAD DÉBIL - ENTIDAD FUERTE

- Una *entidad débil* puede ser unívocamente identificada sólo en el contexto de otra *entidad fuerte* o propietaria.
- Entidades fuerte y débil están vinculadas por una relación binaria $(1,1):(*,N)$. Siempre la cardinalidad del lado 1 es 1. (Porqué?)
- La entidad débil tiene una dependencia de existencia y de identificación respecto de la entidad fuerte.

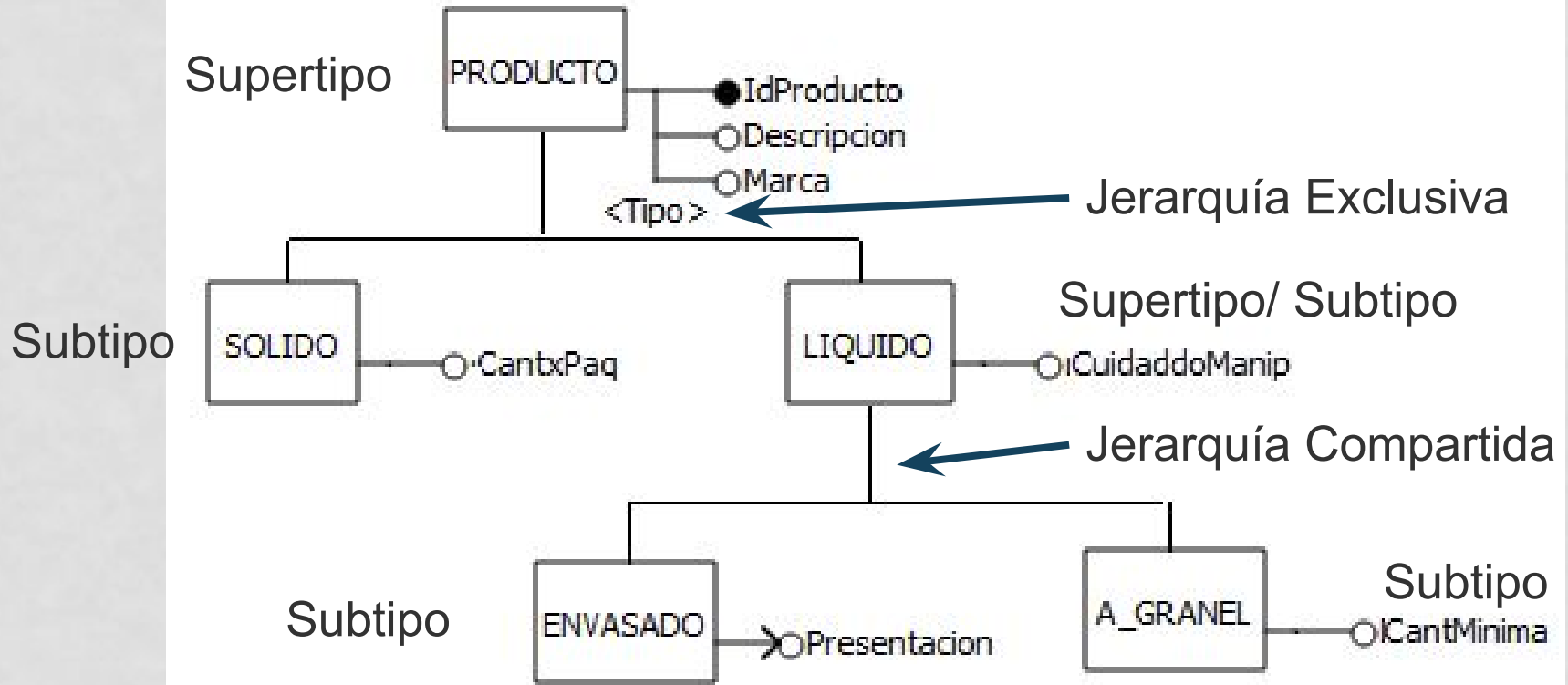


ENTIDAD FUERTE

ENTIDAD DÉBIL

¿Cómo indicar la dependencia de existencia y de identificación? → esquema lógico

JERARQUÍAS (RELACIONES ES-UN O ISA)



Participación total o parcial

¿Cómo representar los distintos casos? → esquema lógico

CONSTRUCCIÓN DEL MER

No existen reglas que indiquen cómo construir un modelo de datos, sólo principios generales a aplicar junto al criterio del diseñador experimentado

- Interpretar las frases expresadas en lenguaje natural (en el relevamiento), identificando cuáles son las *entidades* (datos) y cuales las *relaciones* (entre los datos) en la organización.
- Chen propuso las siguientes heurísticas:
 - En general un **sustantivo** es una entidad, aunque también podría ser un atributo (Ej: “los ALUMNOS cursan MATERIAS”)
 - Un **verbo o frase verbal** puede indicar una relación entre entidades (Ej: “los alumnos CURSAN materias”). Asociaciones entre los datos.
- Que información acerca de las entidades y relaciones deberían registrarse? → determinar los atributos (de entidades y relaciones).
- Respecto de relaciones más complejas (ternarias, agregaciones, etc.) la experiencia del diseñador, las herramientas computacionales que maneje y el conocimiento de las transformaciones en esquemas equivalentes constituyen un recurso fundamental.

RESUMEN: DISEÑO CONCEPTUAL DE DATOS

- El diseño conceptual se guía por el **análisis de los requisitos**
 - Permite obtener una descripción de alto nivel de los datos que finalmente serán almacenados
- **MER/MERExt** es uno de los modelos más populares para el diseño conceptual.
 - Construcciones expresivas, cercanas a la forma en que los usuarios piensan acerca de sus aplicaciones.
 - Numerosas herramientas CASE (que generalmente siguen el modelo binario → no hay relaciones ternarias, que deben ser representadas por 2 o más binarias → no siempre equivalentes!!).
- Algoritmos de derivación 'casi automática' desde DERExt a tablas de la base de datos.

... continuará ...

REPRESENTACIÓN DE LA INFORMACIÓN

- La información valiosa debe tener una representación simbólica adecuada e inteligible para que pueda ser *comunicada*.
- La pieza más elemental de información es el *dato* → aunque no es de utilidad por sí mismo sin un *contexto*, sino vinculado con otros datos → el almacenamiento y la comunicación de información requieren la *estructurar los datos*.