

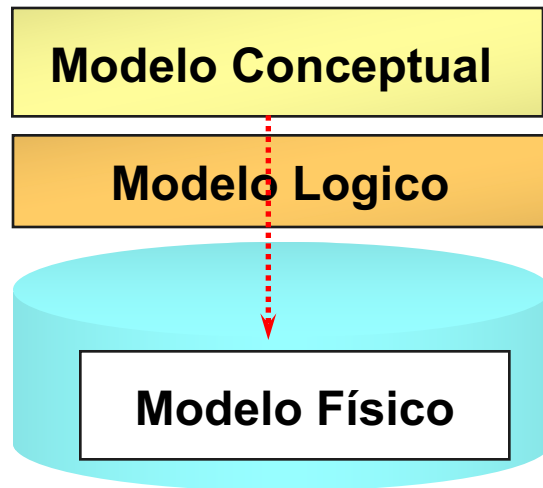


Modelado multidimensional. Parte I : Diseño conceptual

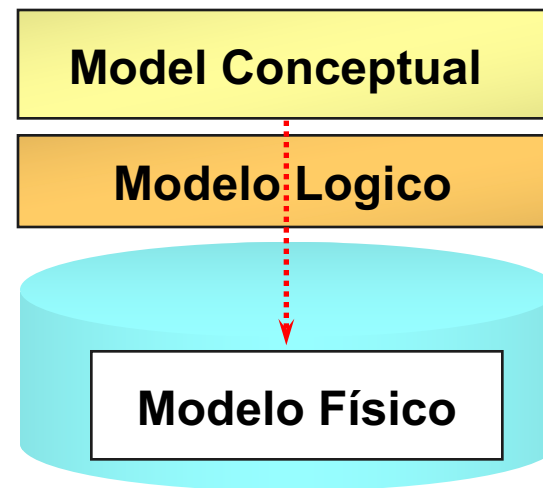
Tema 3

Modelos utilizados en BD

- Modelo Conceptual
- Modelo Logico
- Modelo Físico



**Sistemas
Operacionales**

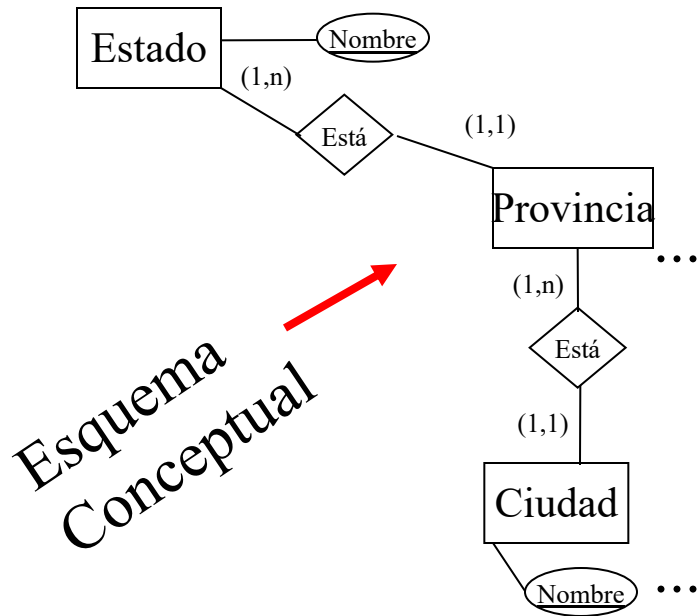


Almacenes de Datos

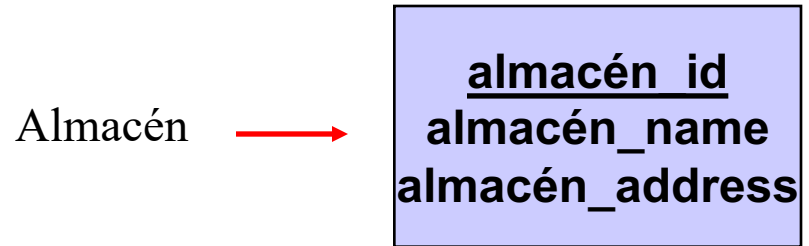
Modelos utilizados en BD

Sistemas operacionales

- Normalización
 - Sist. Transaccionales



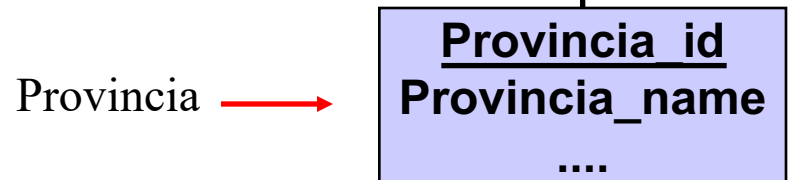
Almacén



Ciudad



Provincia



Estado





Diseño conceptual de DW

- **Los modelos conceptuales**

- Permiten una mejor comunicación entre diseñadores y usuarios para comprender los requisitos de aplicación
- Más estable que un esquema orientado a la implementación (lógica), que cambia con la plataforma
- Proporcionar un mejor soporte para interfaces de usuario visuales
- No hay un modelo conceptual bien establecido para datos multidimensionales
- Varias propuestas basadas en UML, en el modelo ER, o en el uso de notaciones específicas
- Problemas:
 - Enfocados a modelar sistemas transaccionales
 - No reflejan las propiedades multidimensionales (hechos, dimensiones, etc.)
 - No se puede expresar tipos complejos de jerarquías
 - La falta de un mapeo a la plataforma de implementación



Diseño conceptual de DW

- En la actualidad, los almacenes de datos están diseñados utilizando en su mayoría modelos lógicos (esquemas estrella y copo de nieve)
 - Difícil para expresar requisitos (se requieren conocimientos técnicos)
 - Limita a usuarios para definir sólo elementos que los sistemas de implementación subyacentes pueden gestionar
 - Ejemplo: Usuarios obligados a usar sólo jerarquías simples apoyadas en las herramientas actuales



Diseño conceptual de DW

- Modelado multidimensional (MD)
 - Parte estructural
 - Parte dinámica
- Parte estructural
 - Hechos y dimensiones
- Parte dinámica
 - Operaciones de consulta al modelo MD



Modelado multidimensional

- Tal y como el usuario percibe el mundo real objeto de estudio
 - Perspectiva estructural
 - Modelado o Modelo Multidimensional (MD)
 - Hechos y Dimensiones
 - Perspectiva dinámica
 - Definición de requerimientos iniciales sobre el modelo MD
 - Operaciones de consulta avanzada



Modelado multidimensional

Parte estructural

- Perspectiva estructural → modelo MD
 - Hechos
 - Objeto de análisis
 - Ej. Ventas de productos, compras, alquileres, transportes
 - Dimensiones
 - Diferentes perspectivas para analizar los hechos
 - Ej. Productos, almacenes, tiempo, vehículos, etc.



Modelado multidimensional

Parte estructural

- Hechos representan normalmente relaciones *muchos a muchos* con todas las dimensiones y, *muchos a uno* con cada dimensión en particular
 - Ej. Ventas de productos (H) por producto (D), almacenes (D) y tiempo (D)
 - Un producto (D) → varias ventas (H)
 - Una venta (H) → un solo producto (D) y almacén (D)



Modelado multidimensional

Parte estructural

- Sin embargo, a veces hechos son *muchos a muchos* con dimensiones en particular
 - Ej. Tickets emitidos (H) por
 - Un ticket (H) puede contener muchos productos
- Hechos y Dimensiones se caracterizan por atributos
 - Hechos → atributos de hecho o medidas
 - Dimensiones → atributos de dimensión



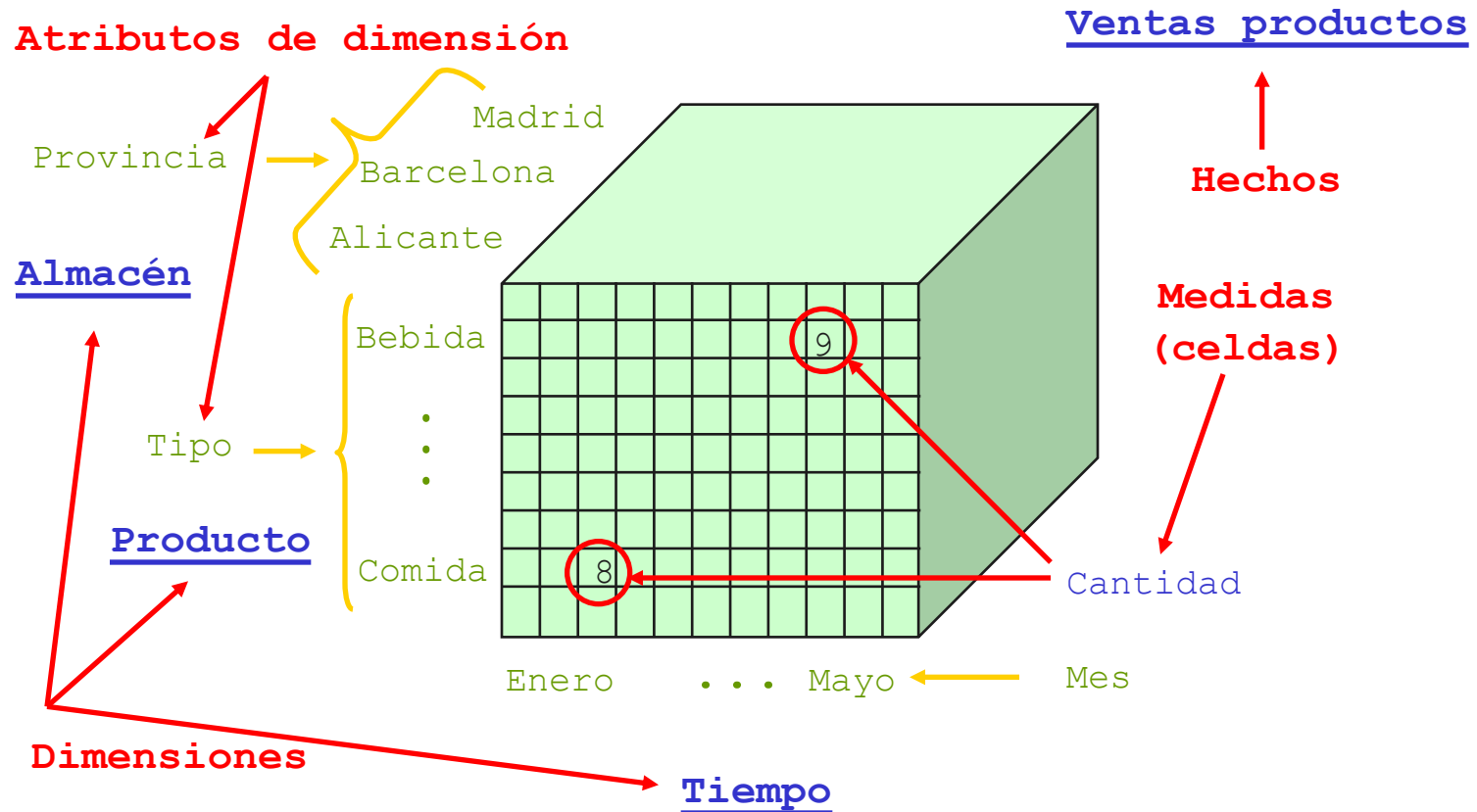
Modelado multidimensional

Parte estructural

- ¿Cómo se representa el modelo MD intuitivamente?
 - Cubos
 - Hipercubos (Cubos sobre cubos)
 - Tablas multidimensionales tipo hoja de cálculo, etc.

Modelado multidimensional

Parte estructural. Cubo



Modelado multidimensional

Parte estructural. Tablas MD.

■ Tablas multidimensionales

Ventas			Producto.Grupo = “Supermercado”			
			Comida		Bebida	
			Cong	Fresco	Refresco	Alcohol
Almacén. comunidad = “Comunidad Valenciana”	Alicante	Albatera	100	200	300	400
		Elche	500	600	700	800
	Valencia	Burjasot	900	1000	1100	1200
		Cullera	1300	1400	1500	1600



Modelado multidimensional

Parte estructural. Dimensiones

- Puede haber alto grado de categorización
 - Atributos en función de instancias
 - Ej. Volumen y porcentaje de alcohol sólo para bebidas
 - Ej. Tiempo y modo preparación sólo para comidas
- Atributos dimensión → jerarquías clasificación
 - Los niveles de jerarquía serán usados para la agregación de las medidas
 - Ej. Ciudad, comunidad, tipos de productos, etc.



Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación

- Las instancias de niveles \equiv miembros

- Clasificación

- Cardinalidad

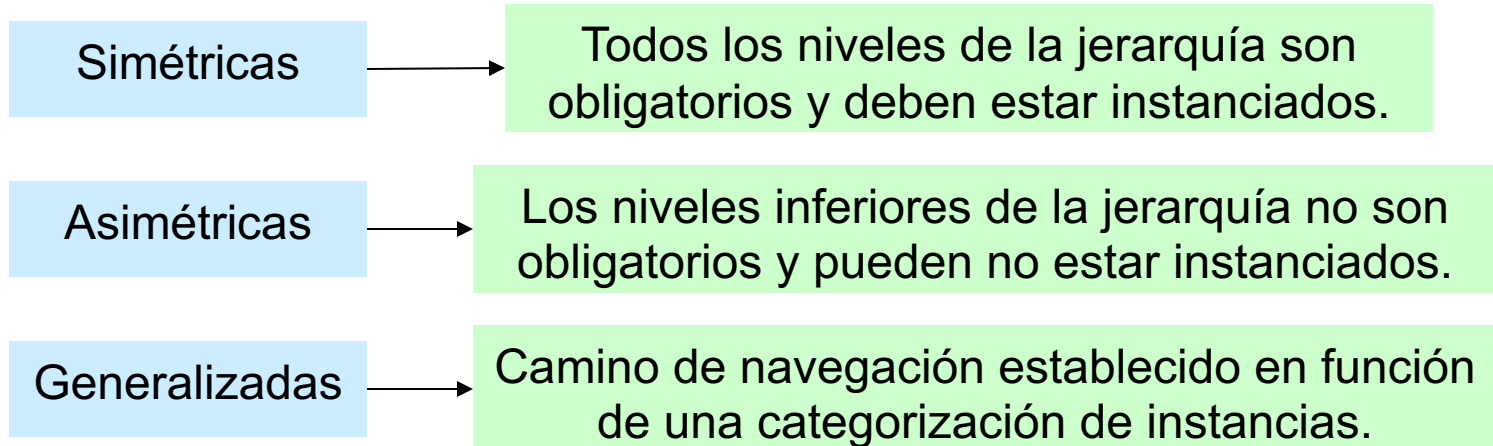
- Por defecto \rightarrow estrictas ($1-m$)
 - Una instancia sólo se relaciona con una instancia del nivel superior de jerarquía
 - Ej. Un almacén está ubicado en una sola ciudad
- Sin embargo algunas pueden ser **no estrictas (n-m)**
 - Ej. Un almacén pertenece a más de una zona de ventas

Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación

■ Caminos

■ Simples → Representación mediante árbol



■ Múltiples → Representación mediante grafo

Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación simples.

Todas

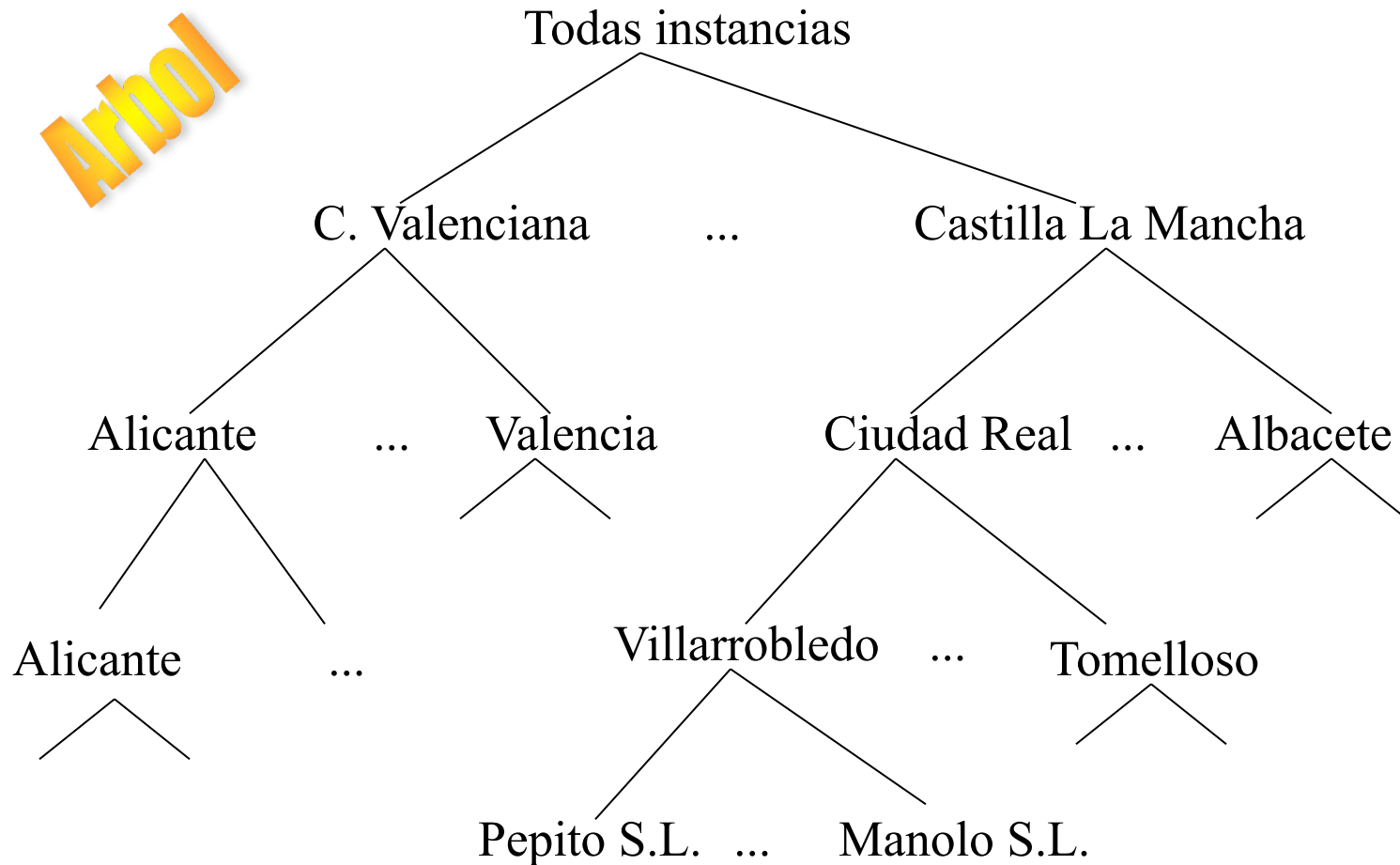
Comunidad

Provincia

Ciudad

Almacén

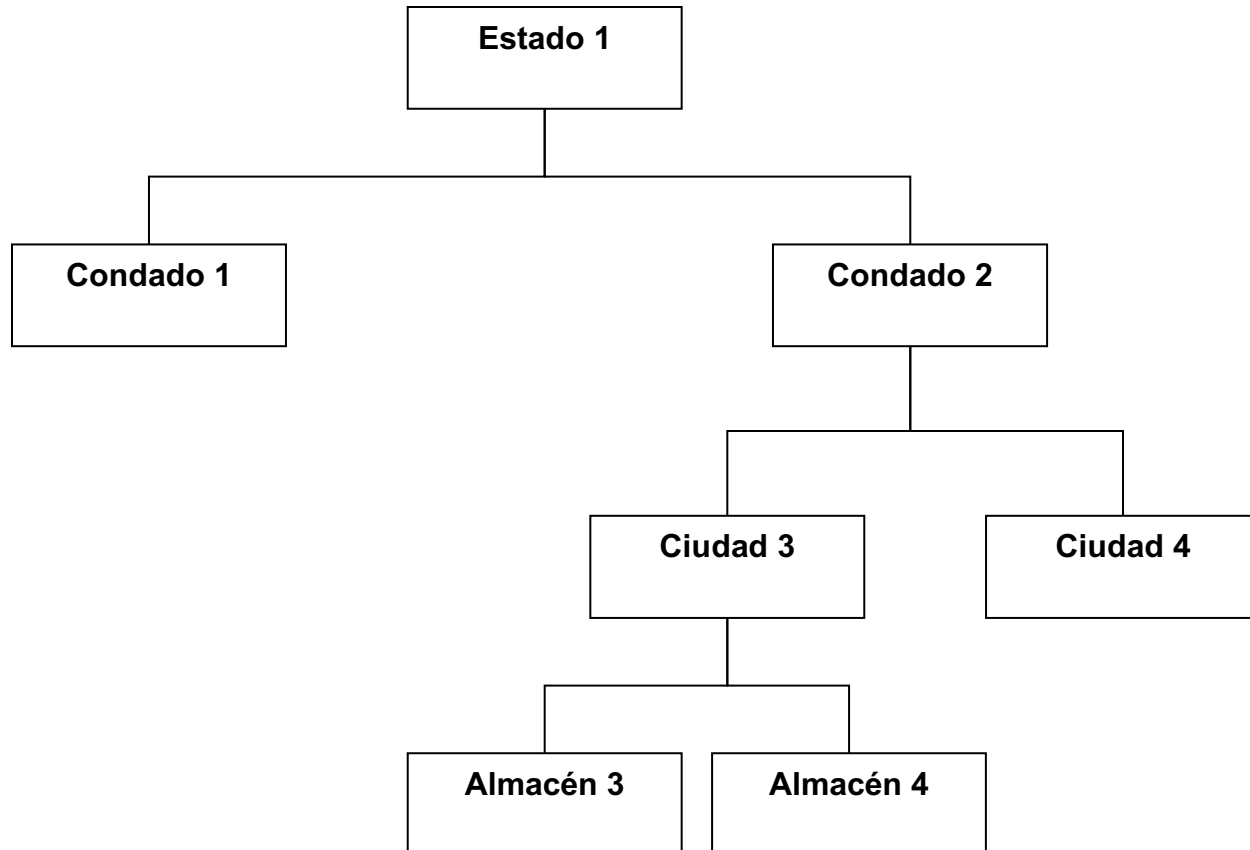
Arbol



Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación simples.

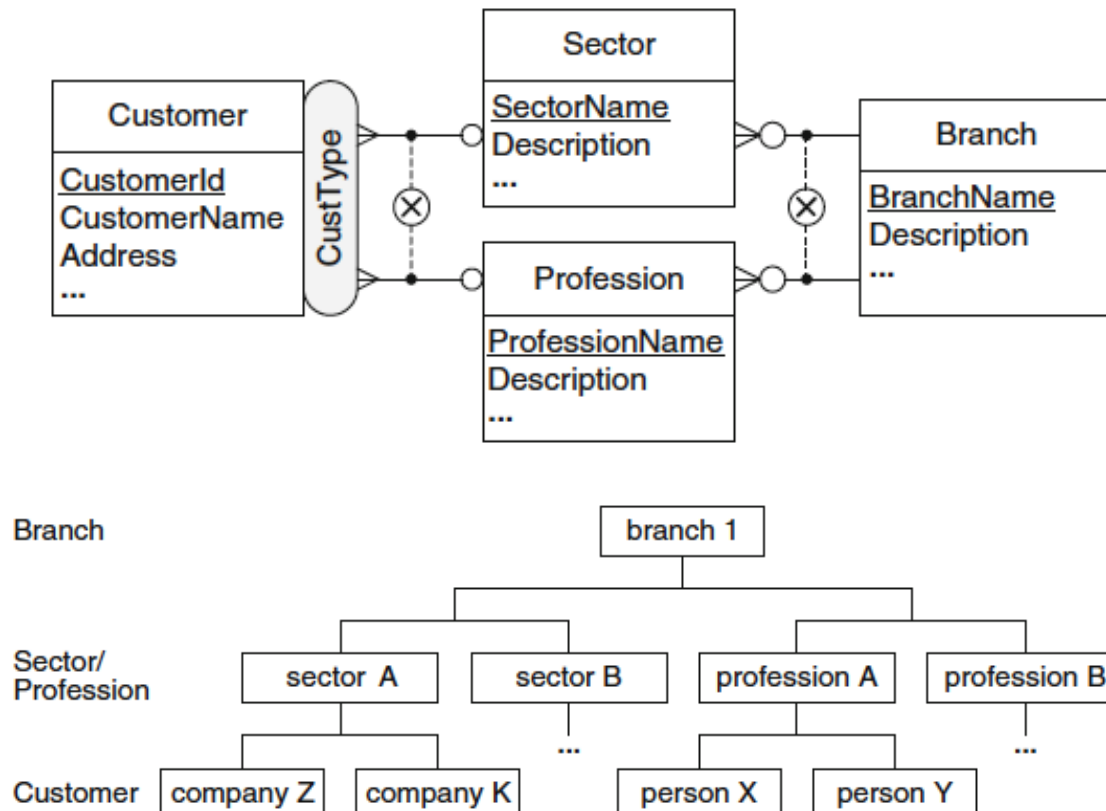
■ Ejemplo de Asimétrica



Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación simples.

■ Ejemplo de Generalizada





Modelado multidimensional

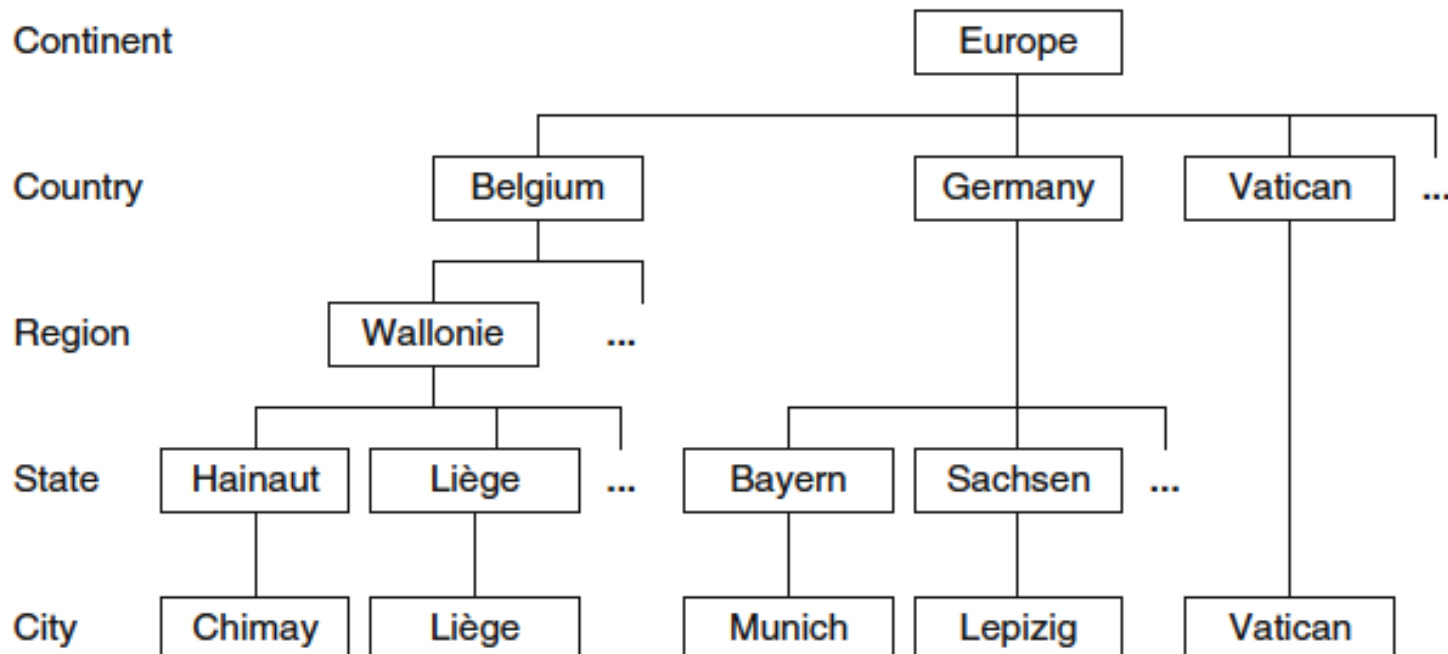
Parte estructural. Dimensiones. Jerarquías de clasificación simples.

- Un caso especial dentro de las jerarquías generalizadas es la jerarquía **desigual**
- Las rutas alternativas se obtienen **saltándose** uno o varios niveles intermedios
- En el nivel de instancia, cada miembro hijo tiene solo un miembro padre, aunque la longitud del camino desde las hojas hasta ese nivel padre puede ser **diferente** para diferentes miembros

Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación simples.

■ Ejemplo de jerarquía desigual





Modelado multidimensional

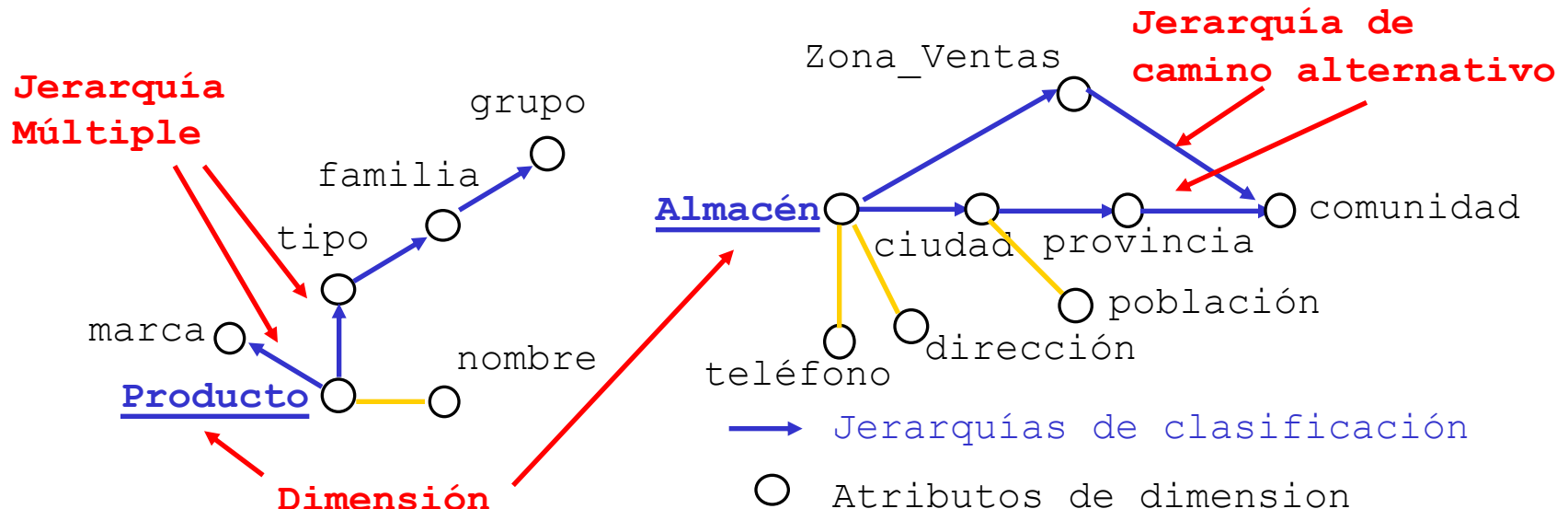
Parte estructural. Dimensiones. Jerarquías de clasificación

- Jerarquías múltiples y de camino alternativo
 - Ej. Ciudad se puede clasificar en comunidad y,
 - Ciudad también se puede clasificar en zona ventas
- Jerarquías paralelas → Más de una jerarquía definida para la misma dimensión
 - Independiente → Las distintas jerarquías no comparten niveles
 - Dependientes → Las distintas jerarquías comparten algún nivel

Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación

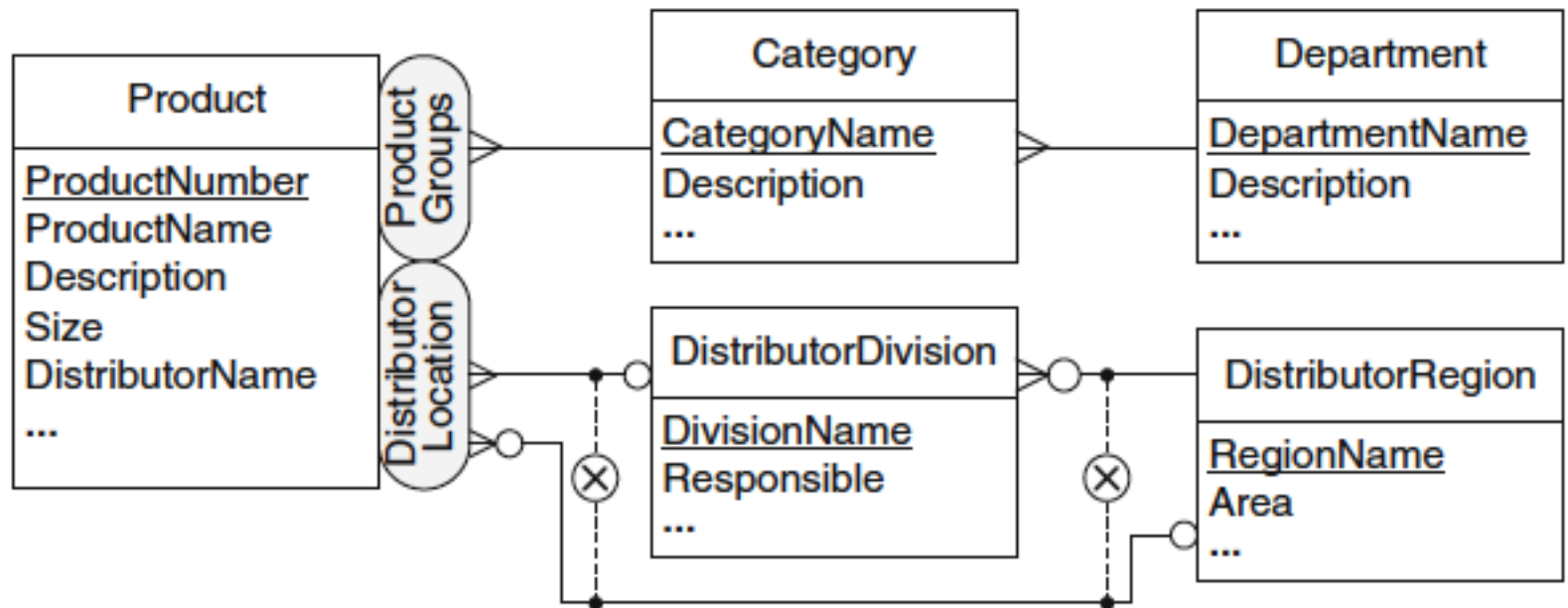
- Multiples → Normalmente se representan mediante Grafos Acíclicos dirigidos (G.A.D.)



Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación

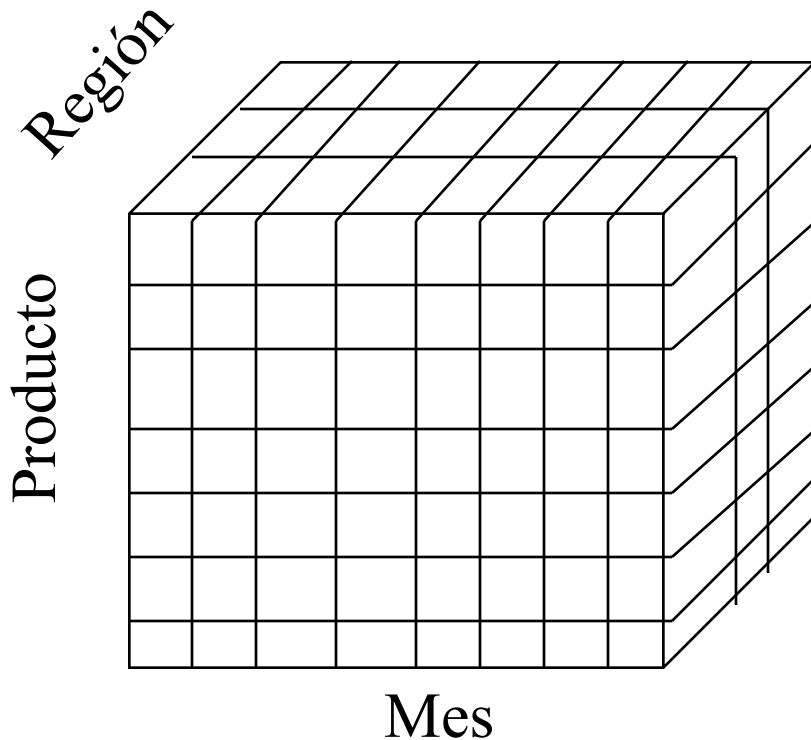
- Jerarquías paralelas independientes



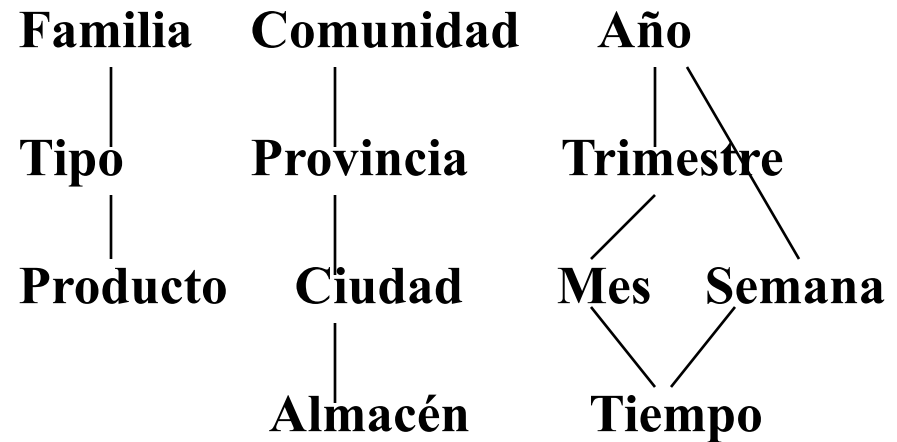
Modelado multidimensional

Parte estructural. Dimensiones. Jerarquías de clasificación y cubos

- Si utilizamos un cierto nivel de agregación
 - Ventas es una función del producto, mes y región

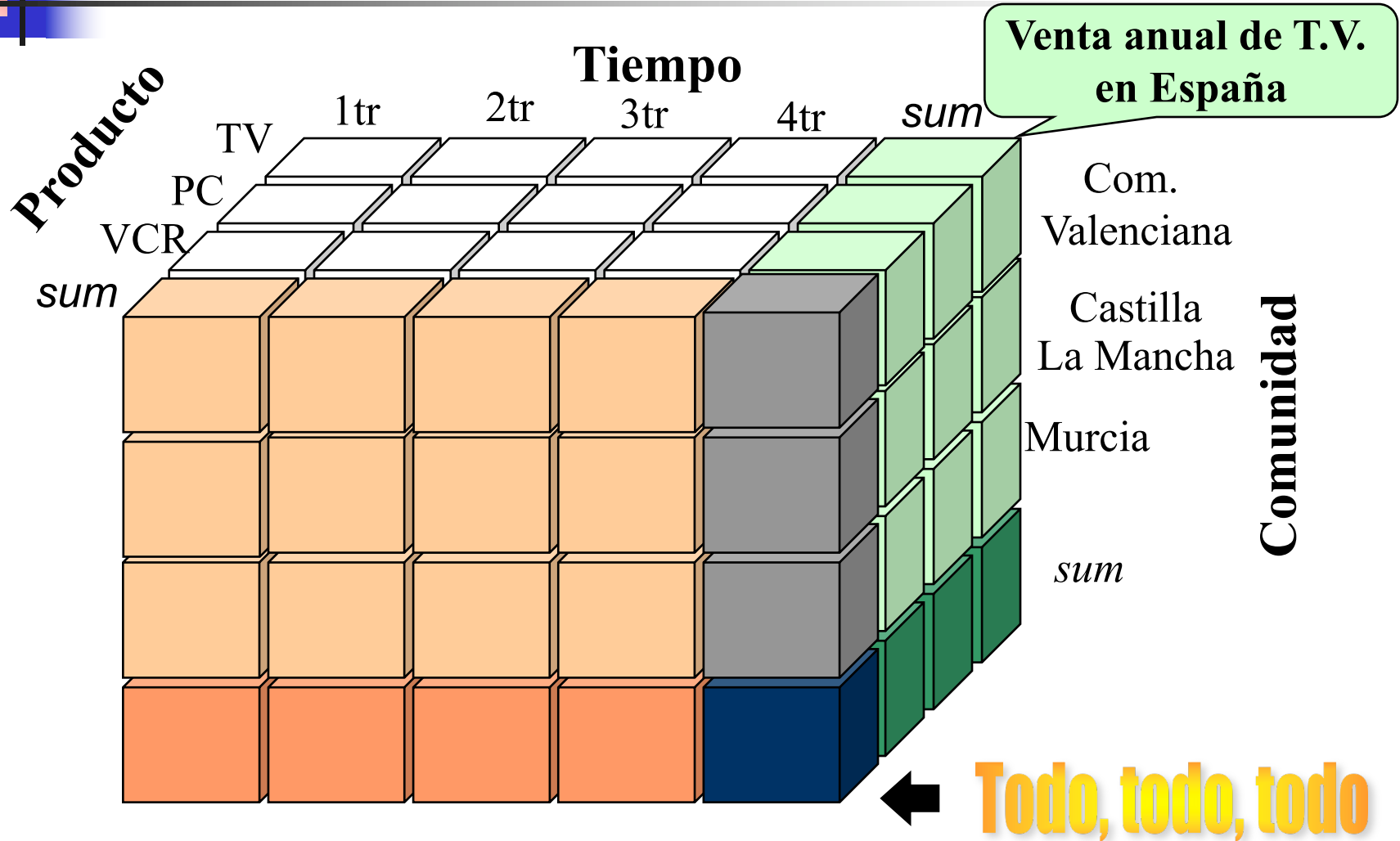


Dimensiones: Producto, Almacén, Tiempo
Caminos de jerarquía por los que agregar



Modelado multidimensional

Parte estructural. Un ejemplo de cubo de datos





Modelado multidimensional

Parte estructural. Hechos

- Atributos de hecho o medidas
 - Atómicos
 - Ej. Cantidad vendida, precio, etc.
 - Derivados
 - Utilizan una fórmula para calcularlos
 - Ej. $\text{Precio_total} = \text{precio} * \text{cantidad_vendida}$



Modelado multidimensional

Parte estructural. Hechos

■ Aditividad

- Conjunto de operadores de agregación (SUM, AVG, etc.) que se pueden aplicar para agregar los valores de medidas a lo largo de las jerarquías de clasificación (*Kimball, 1996*)
- Es aditiva → SUM sobre todas las dimensiones
- Semi-aditiva → SUM sólo sobre algunas dimensiones
- No aditiva → SUM sobre ninguna dimensión



Modelado multidimensional

Parte estructural. Hechos

- Si no aditiva → otros operadores pueden aplicarse (ej. AVG, MIN, etc.)
 - Ej. Atributos que miden niveles (ej. Inventarios) no son aditivos sobre la dimensión tiempo
 - Es aditivo sobre la dimensión producto
- Las medidas de temperatura no son aditivas
- Algunas son semánticamente incorrectas
 - Ej. Atributo número de clientes que cuenta el número de tickets emitidos no es aditiva sobre la dimensión producto



Modelado multidimensional

- En aplicaciones OLTP...
 - Modelado conceptual → Entidad-Relación (EER)
 - ¿ Podría reflejar la multidimensionalidad de los datos ?
 - Hechos,...
 - Dimensiones,...
 - ¿ Podría ser interrogado por un analista de la información ?

Modelado multidimensional

Parte dinámica.

- BD Multidimensionales parecen más naturales y
⇒ Queries son también más naturales
- Ejemplo: llamadas de tfno. por producto y región
 - En un cubo:
 - ROLLUP*(calls, producto->producto, localizacion->comunidad, SUM(numLlamadas))

Producto/Comun.	C. Valenciana	Cast. Mancha
Fax	44	28
E-mail	27	51
Mobiles	46	11

- En tablas relacionales:

```
Select producto, comunidad, sum(llamadas)
from Llamadas, Comunidad
where Llamadas.comunidad= Comunidad.comunidad
Group by producto, comunidad
Order by producto, comunidad;
```

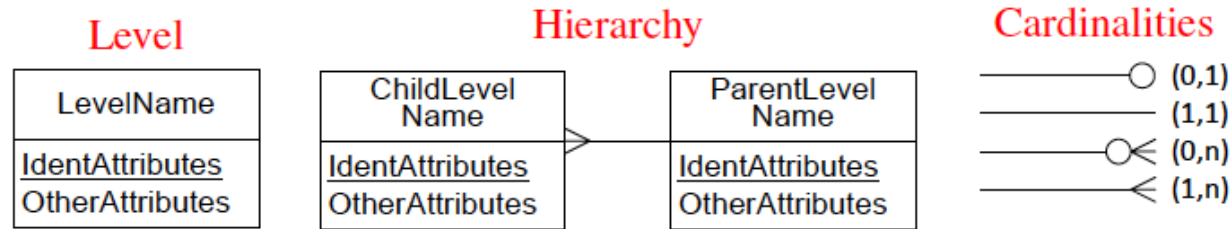
Product	Region	Sum(calls)
Fax	Vaud	44
Fax	Valais	28
Mobile	Vaud	27
Mobile	Valais	51
Standard	Vaud	46
Standard	Valais	11



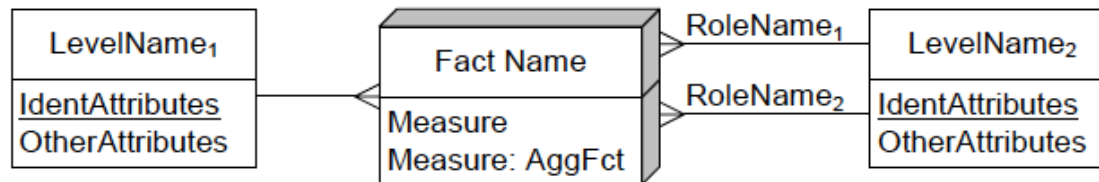
Diseño conceptual de almacenes de datos

- Es el primer paso para el diseño de un almacén de datos
- Se parte de la documentación relacionada con la base de datos integrada y consta de:
 1. Definición de hechos
 2. Para cada hecho:
 - definición del árbol de atributos
 - edición del árbol de atributos
 - definición de las dimensiones
 - definición de las medidas
 - definición de las jerarquías

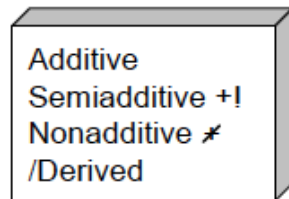
Representación Modelo multidimensional



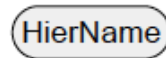
Fact with measures and associated levels



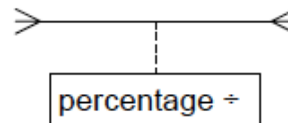
Types of measures



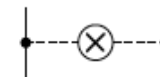
Analysis criterion



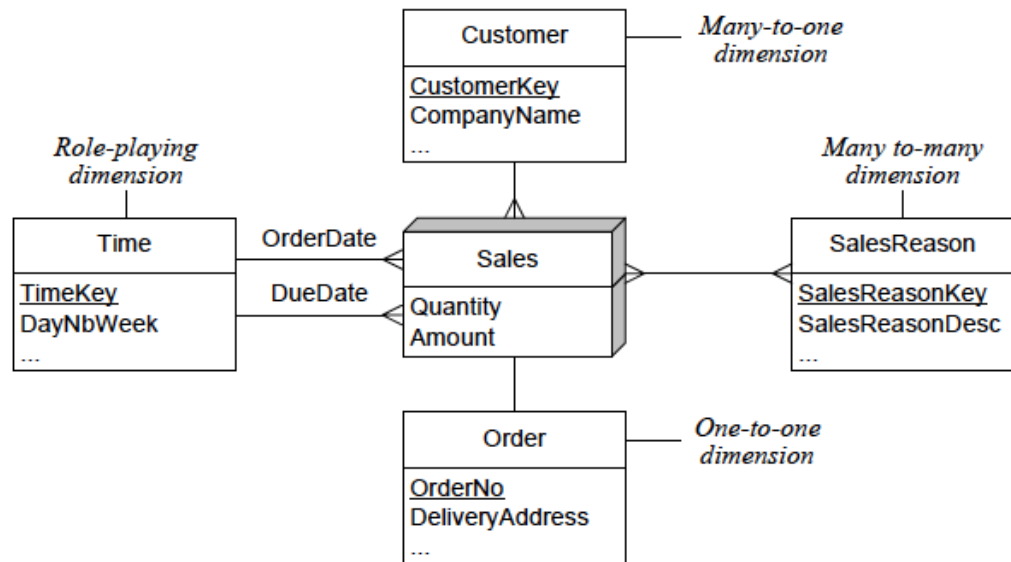
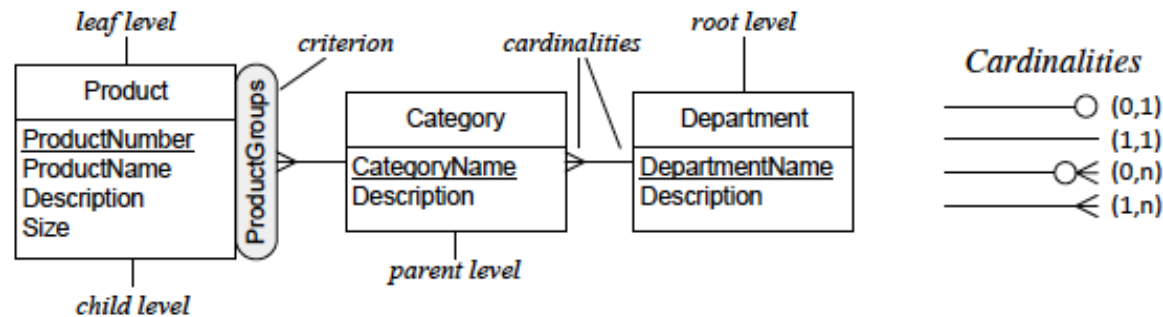
Distributing factor



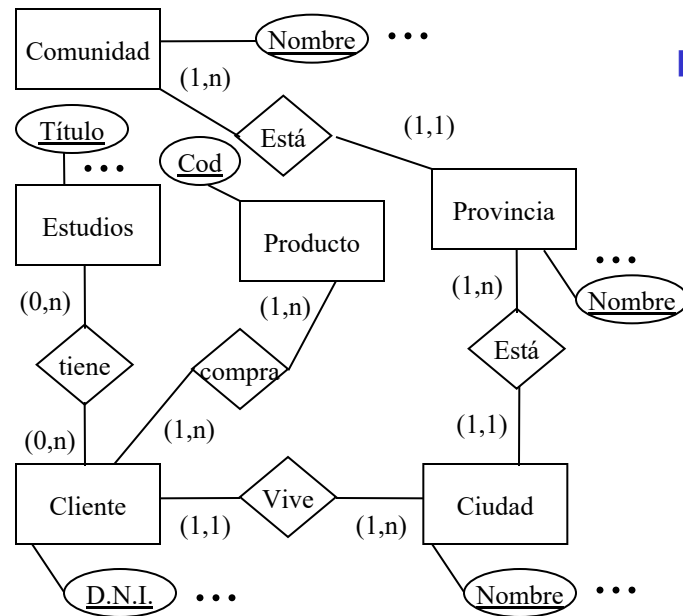
Exclusive relationships



Representación Modelo multidimensional



Modelado multidimensional



■ ¿Hechos?

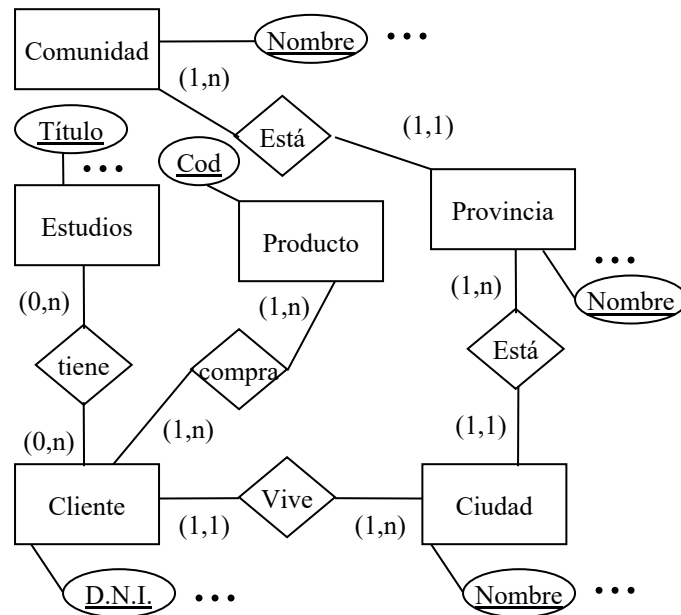
■ ¿Dimensiones?

■ ¿Jerarquías?

↓
Transformar

■ ¿Podría un analista interrogarlo?

Diseño conceptual de almacenes de datos



Ejemplo 1.
Determina el
diseño
conceptual de un
almacén a partir
de diagrama ER



Ejercicio 1

- Diseña un esquema multidimensional para la tarea del proveedor de teléfono (ver tema anterior) utilizando el modelo de representación de las transparencias anteriores