



BASES DE DATOS

Tema 2

Modelos de Datos

UCLM- E.S. de Informática

Coral Calero, Marcela Genero, Francisco Ruiz



Objetivos

- Dar a conocer los **elementos** que forman parte de un **modelo de datos** (MD).
- Distinguir los principales **mecanismos de abstracción**.
- Analizar los distintos tipos de **restricciones** en un MD.
- Presentar una **clasificación** de los MD más utilizados.



Contenido

- Conceptos Básicos
 - Modelo, esquema y ejemplar
- Definición de modelo de datos
 - Estática
 - Dinámica
- Principales mecanismos de abstracción
 - Clasificación.
 - Agregación.
 - Generalización.
 - Asociación.
 - Jerarquías de abstracciones
- Restricciones de integridad
 - Componentes de una restricción
 - Clases de restricciones
- Taxonomía de los modelos de datos
- Los modelo de datos en el proceso de diseño de una BD



Bibliografía

- Básica
 - Piattini et al. (2006)
 - Cap. 4
- Complementaria
 - Tsichritzis y Lochovsky (1982)
 - Caps. 1-3



Conceptos Básicos

Modelo

- *"Construcción mental a partir de la realidad en la que se reproducen los principales componentes y relaciones del segmento de la realidad analizada".*
- Dos acepciones:
 - **Reproducción simplificada** de la realidad (ciencias empíricas)
 - **Realidad** propiamente dicha (pintor)



Conceptos Básicos

Modelo de Datos

- *"Dispositivo de **abstracción** que nos permite ver el bosque (esto es, la información contenida en los datos) en oposición a los árboles (valores individuales de los datos)"; Tsichritzis y Lochovsky (1982).*
- *"Conjunto de **herramientas conceptuales para describir** la representación de la información en términos de datos. Los modelos de datos comprenden aspectos relacionados con: estructuras y tipos de datos, operaciones y restricciones"; Dittrich (1994).*
- *"Conjunto de **conceptos, reglas y convenciones** que permiten describir y manipular los datos de la parcela de un cierto mundo real que deseamos almacenar en la base de datos"; De Miguel et al. (1999).*



$$\mathbf{LD = MD + Sintaxis}$$

Ejemplos:

$$\text{SQL} = \text{MDR} + \text{Sintaxis}$$

$$\text{QBE} = \text{MDR} + \text{Sintaxis (distinta)}$$

$$\text{OQL} = \text{MO} + \text{Sintaxis}$$

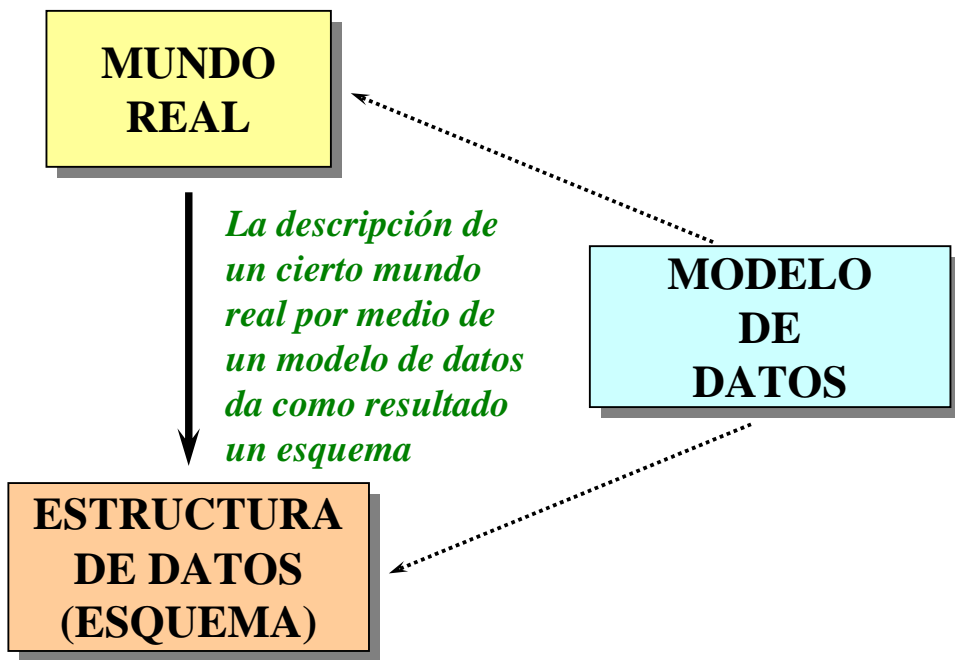


- "*La **descripción** específica de un determinado **mini-mundo** en términos de un modelo de datos se denomina esquema (o esquema de datos) del mini-mundo. La colección de datos que representan la información a cerca del mini-mundo constituye la base de datos*", Dittrich (1994).
- "**Representación** de un determinado mundo real (universo del discurso) en términos de un modelo de datos"; de Miguel, Piattini y Marcos (1999).



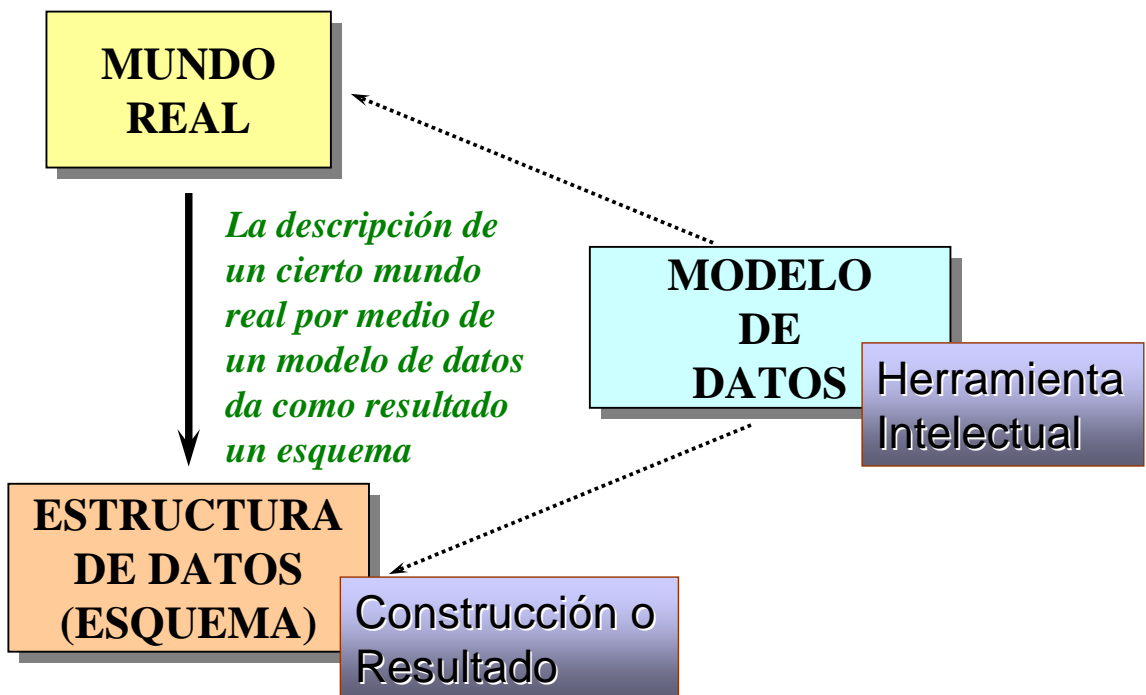
Conceptos Básicos

Modelo de Datos vs Esquema



Conceptos Básicos

Modelo de Datos vs Esquema





Conceptos Básicos

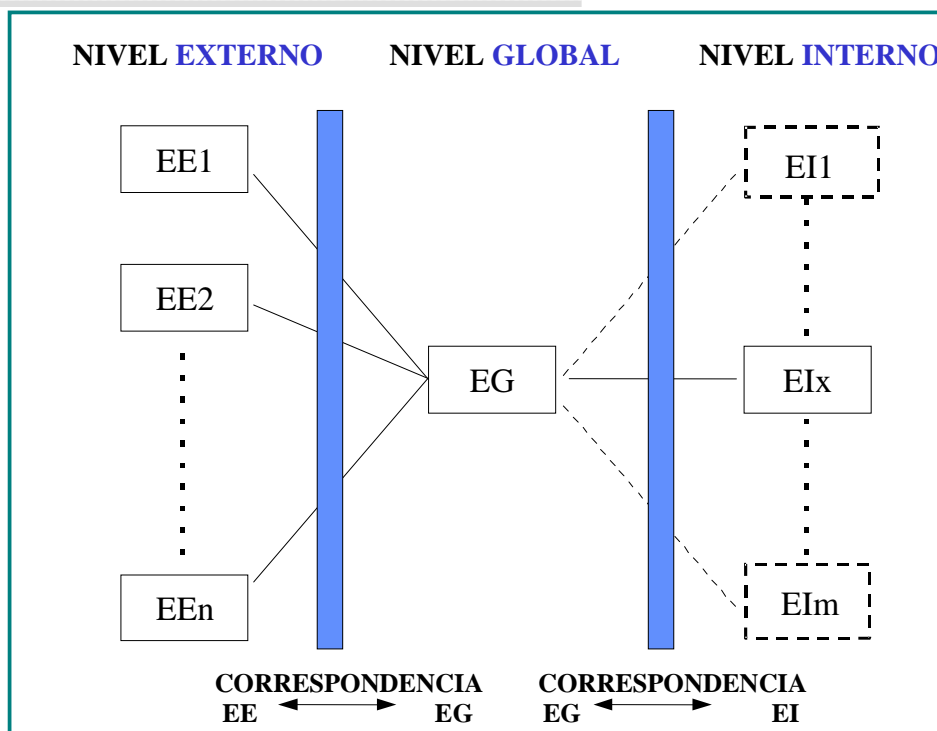
Esquemas y Ejemplares

- Un **ejemplar de un elemento** de un esquema son los **datos** que en un determinado momento se encuentran almacenados en el citado elemento del esquema.
- La colección de ejemplares de todos los elementos de un esquema en un momento determinado constituyen un **ejemplar del esquema**.
- Al igual que en los lenguajes de programación existen variables (constituidas por un tipo y un contenido), las cuales tienen en un momento determinado un cierto valor; en las bases de datos se debería hablar de "**variables de base de datos**", cuyo tipo sería el esquema y su contenido todos los posibles valores del esquema; su valor, en un momento determinado, sería un ejemplar del esquema.
- Nosotros utilizaremos la expresión "base de datos" en el sentido abstracto de todos los posibles ejemplares, y cuando queramos referirnos a su contenido en un cierto momento hablaremos de un ejemplar o bien de "la base de datos en el instante i" (BDi).

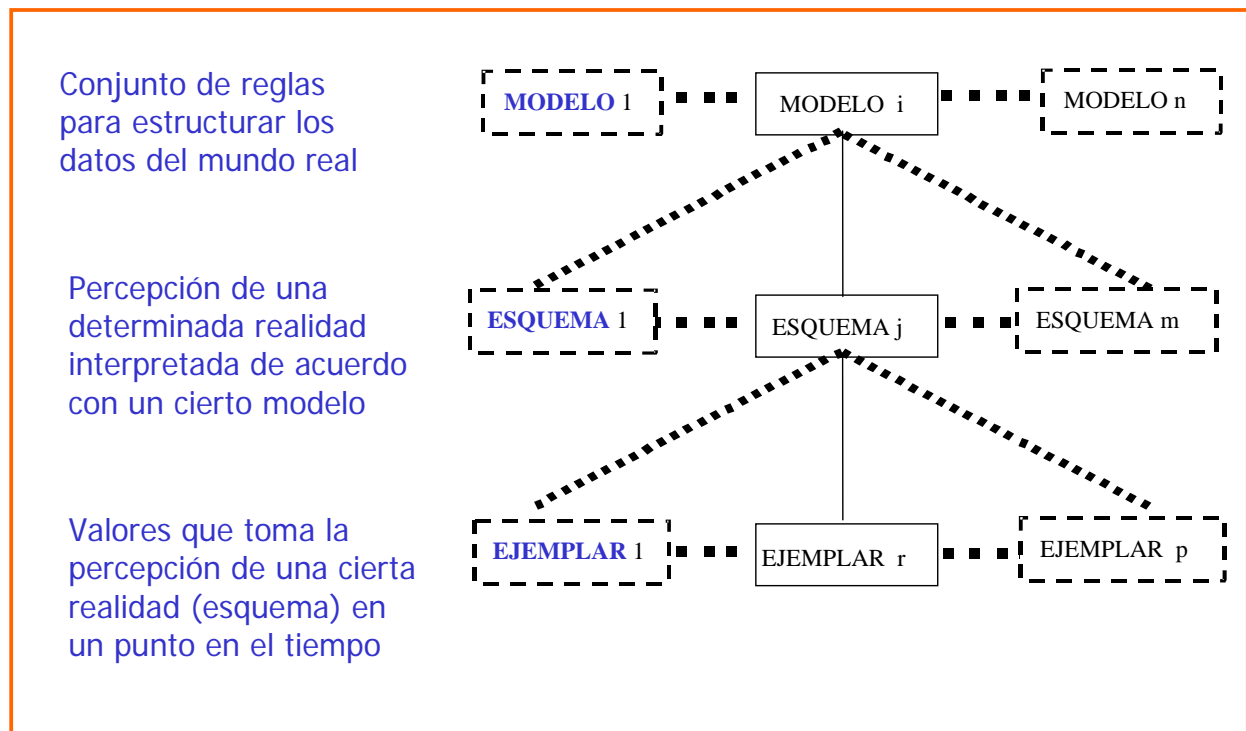


Conceptos Básicos

Tipos de Esquemas



Esquemas para los tres niveles de la arquitectura ANSI



Definición de Modelo de Datos

- Aunque existen muchos MD es posible abstraer una serie de características comunes a todos ellos, definiendo así el concepto de modelo de datos en general, de forma más precisa:
 - ***"un conjunto de conceptos, reglas y convenciones bien definidos que nos permiten aplicar una serie de abstracciones a fin de describir y manipular los datos de un cierto mundo real que deseamos almacenar en una base de datos"***.
- Los MD facilitan la creación de categorías mediante la aplicación de mecanismos de abstracción (clasificación, agregación, etc.). Esto lleva a diferenciar dos tipos de modelos (de forma similar a como ocurre con los lenguajes de programación):
 - Fuertemente tipados, ← (son los utilizados en BD).
 - Débilmente tipados.



Definición de Modelo de Datos

Propiedades

- **ESTÁTICAS**
 - Elementos permitidos
 - Objetos
 - Asociaciones
 - Características de los objetos
 - Dominios
 - Elementos no permitidos (restricciones)
 - Inherentes
 - De integridad o semánticas
- **DINÁMICAS**
 - Conjunto de operadores.
 - Cada operador tiene dos componentes:
 - Localización
 - Acción



Definición de Modelo de Datos

Formalización

- Un MD ha de proporcionar facilidades para recoger los aspectos estáticos y dinámicos de la realidad, por lo que se define formalmente como el par:

$$MD = \langle G, O \rangle$$

- donde **G** es el conjunto de reglas de generación que permiten representar la componente **estática**, es decir, describir las **estructuras** de nuestro universo del discurso, y
- **O** es el conjunto de **operaciones** autorizadas sobre la correspondiente estructura, operaciones que permiten representar la componente **dinámica**.
- La componente estática de un determinado MD expresado con una sintaxis determinada es el **Lenguaje de Definición de Datos (LDD)**, y la componente dinámica el **Lenguaje de Manipulación de Datos (LMD)**; ambos juntos constituyen el **Lenguaje de Datos (LD)**.
- Los SGBD suelen tener además un **Lenguaje de Consulta** (Query Language -QL-) y un **Lenguaje de Control** (Control Language).



Estática – elementos permitidos

- Los **elementos permitidos** no son los mismos para todos los MD (varían especialmente en terminología), pero en general son:
 - **Objetos** (entidades, relaciones, registros, etc.)
 - **Asociaciones** entre objetos (interrelaciones, "set", etc.)
 - **Propiedades** o características de los objetos o asociaciones (atributos, campos, elementos de datos, etc.)
 - **Dominios**, que son conjuntos nominados de valores homogéneos sobre los que se definen las propiedades.
- A estos elementos permitidos se les podrán aplicar aquellas abstracciones reconocidas por el modelo.
- La **representación** de estos elementos depende de cada modelo de datos, pudiendo hacerse en forma de **grafos** (E/R, UML) o de **tablas** (Relacional).



Estática – elementos no permitidos

- Los elementos no permitidos se conocen como **restricciones**.
- Tipos:
 - Restricciones **inherentes** (del modelo):
 - Impuestas por la misma naturaleza del modelo de datos, el cual no admite ciertas estructuras.
 - Restricciones **de integridad o semánticas** (de usuario):
 - Permiten captar la semántica del universo del discurso que se quiere modelar y verificar la corrección de los datos almacenados en la BD.
 - Según los instrumentos que proporcione el modelo de datos para definir y gestionar las restricciones, éstas pueden ser:
 - **Propias** al MD: su definición le corresponde al diseñador, pero su gestión es responsabilidad del modelo de datos, el cual las reconoce y recoge en el esquema.
 - **Ajenas** al MD: son, por completo, responsabilidad del diseñador, ya que el modelo de datos no las reconoce ni proporciona instrumentos para manejarlas.



Definición de Modelo de Datos Estática – formalización

- Podemos definir la componente estática del modelo de datos como:

$$G = \langle Ge, Gr \rangle$$

- donde **Ge** es el conjunto de reglas de **generación de estructuras** (objetos del modelo y restricciones inherentes) y **Gr** es el conjunto de **restricciones de usuario**.
- La aplicación de la componente estática de un MD a un determinado Universo del Discurso (UD) nos da como resultado un esquema (E):

$$G[UD] = E$$

- E es la estructura de datos que describe, en el correspondiente modelo MD, las categorías que han resultado de las abstracciones aplicadas al mundo real (UD) que se trata de modelar.



Definición de Modelo de Datos Dinámica – formalización

- La componente **dinámica** consta de un conjunto de **operadores** que se definen sobre la estructura del correspondiente MD, ya que no todas las estructuras admiten el mismo tipo de operaciones.
- La aplicación de un operador a un ejemplar de un esquema transforma éste en otro ejemplar:

$$O [BDi] = BDj$$

- Tanto BDi como BDj deben ser ejemplares válidos de la BD, es decir, los valores de ambos deben pertenecer a alguna de las categorías definidas en el esquema y cumplir las restricciones de integridad (también deben cumplir, en caso de que existan, las posibles restricciones asociadas al cambio de estado).



Definición de Modelo de Datos Dinámica – LMD

- En un plano conceptual, sin seguir una sintaxis concreta, podemos expresar una sentencia del LMD de la siguiente forma:

LOCALIZACIÓN <condición>
ACCIÓN <objetivo>

```
SELECT DNI, Nombre
FROM Alumnos
WHERE Curso=3
```

- donde
 - La **Localización** o “enfoque” consiste en localizar un ejemplar de un objeto indicando un camino (lenguaje navegacional) o un conjunto de ejemplares especificando una condición (lenguaje de especificación).
 - La **Acción** se realiza sobre el(los) ejemplar(es) previamente localizado(s). Puede consistir en una recuperación o en una actualización (inserción, borrado o modificación).
 - <condición> es una expresión lógica que deben cumplir los objetos que se desea localizar o señala el camino que permite llegar a esos objetos.
 - <objetivo> indica los objetos (o las propiedades de éstos) sobre los que se aplica la acción.



Mecanismos de Abstracción

- El proceso de **abstracción** nos ayuda a modelar los datos al hacer que nos centremos en lo **esencial**, pasando por alto aspectos que no consideramos relevantes para nuestros objetivos en la representación del mundo real.

Ejemplo:

El concepto de **ambulancia** como una abstracción en la que únicamente recogemos aquellas características (chasis, ruedas, sirena, etc.), comunes a todas las ambulancias y que la distinguen de otros vehículos, que son de interés para nuestros fines.





Mecanismos de Abstracción

- Los MD ofrecen distintos mecanismos de abstracción a fin de facilitar la representación de los datos; siendo el esquema el resultado de aplicar un proceso de abstracción a un determinado mundo real. Los principales son:
 - **Clasificación**,
 - **Agregación**,
 - **Generalización** y
 - **Asociación**. ← (algunos autores piensan que es un tipo especial de agregación)
- **Pueden combinarse entre sí** ofreciendo interesantes mecanismos semánticos para estructurar los datos.
- Permiten establecer vinculaciones entre los elementos de un modelo.
- La clasificación establece una vinculación entre una categoría de objetos y cada objeto particular (ejemplar) de dicha categoría, mientras que en las otras tres el vínculo se establece entre categorías de objetos y, por tanto, también entre los correspondientes ejemplares de dichas categorías.



Mecanismos de Abstracción Uso Cotidiano

Los mecanismos de abstracción los utilizamos - consciente o inconscientemente - de manera continua:

- El vehículo de matrícula CR-0978-Z es de la clase ambulancia.
- Está formada por cuatro ruedas, un chasis, un motor, ...
- Una ambulancia es un vehículo para recoger y transportar enfermos.
- Su propietario es la empresa CUASER; su conductor es Fernández, ...





Mecanismos de Abstracción

Uso Cotidiano

Los mecanismos de abstracción los utilizamos - consciente o inconscientemente - de manera continua:

- **Clasificación:**
 - El vehículo de matrícula CR-0978-Z es de la clase ambulancia.
- **Agregación:**
 - Está formada por cuatro ruedas, un chasis, un motor, ...
- **Generalización:**
 - Una ambulancia es un vehículo para recoger y transportar enfermos.
- **Asociación:**
 - Su propietario es la empresa CUASER; su conductor es Fernández, ...



Mecanismos de Abstracción

Clasificación

- La **Clasificación** es la acción de abstraer las características comunes a un conjunto de ejemplares para crear una categoría a la cual pertenecen dichos ejemplares.
- El mecanismo contrario se llama **Particularización**.
- BRODIE (1984) define la clasificación como:
 - *Una forma de abstracción en la que una colección de objetos se considera como una clase de objetos de más alto nivel.*
 - *Una clase de objetos es una caracterización precisa de todas las propiedades compartidas por todos los objetos en la colección.*
 - *Un objeto es un ejemplar de una clase de objetos si tiene las propiedades definidas en la clase.*
- Ejemplo: Clasificamos como **Vehículos** a las máquinas, animales o cosas, con medios de propulsión propios, que sirven para desplazar seres u objetos desde una posición a otra.
 - **Ambulancia** => SI es un vehículo
 - **Grúa** => NO es un vehículo (incumple la autopropulsión).



Mecanismos de Abstracción

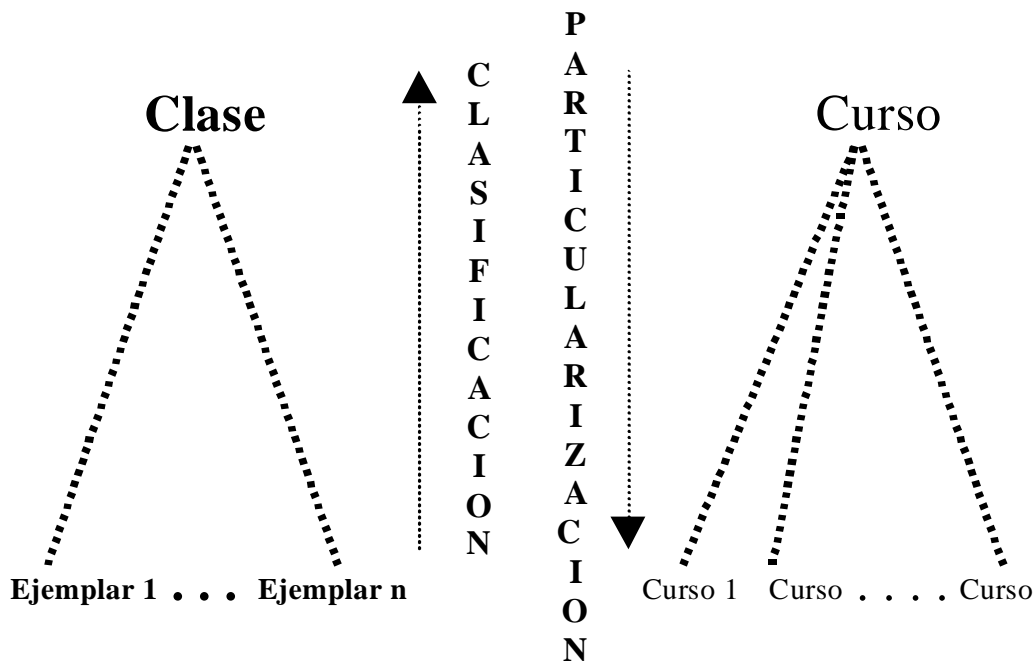
Clasificación

- La clasificación se corresponde con el concepto de **pertenencia** a un conjunto:
 - entre el elemento clase y los elementos miembros se establece una relación **ES_MIEMBRO_DE**.
- Los ejemplares de una clase tienen características similares por medio de las cuales describimos la correspondiente clase; estas características toman valores concretos para cada uno de los ejemplares de la clase.
- Los mismos objetos admiten clasificaciones distintas. Por ejemplo, podemos clasificar las asignaturas de varias maneras:
 - obligatorias / optativas,
 - anuales / semestrales,
 - de primer curso, segundo curso, etc,
 - teóricas / aplicadas, etc.
- Todos los MD de las bases de datos admiten la abstracción de clasificación.



Mecanismos de Abstracción

Clasificación - representación



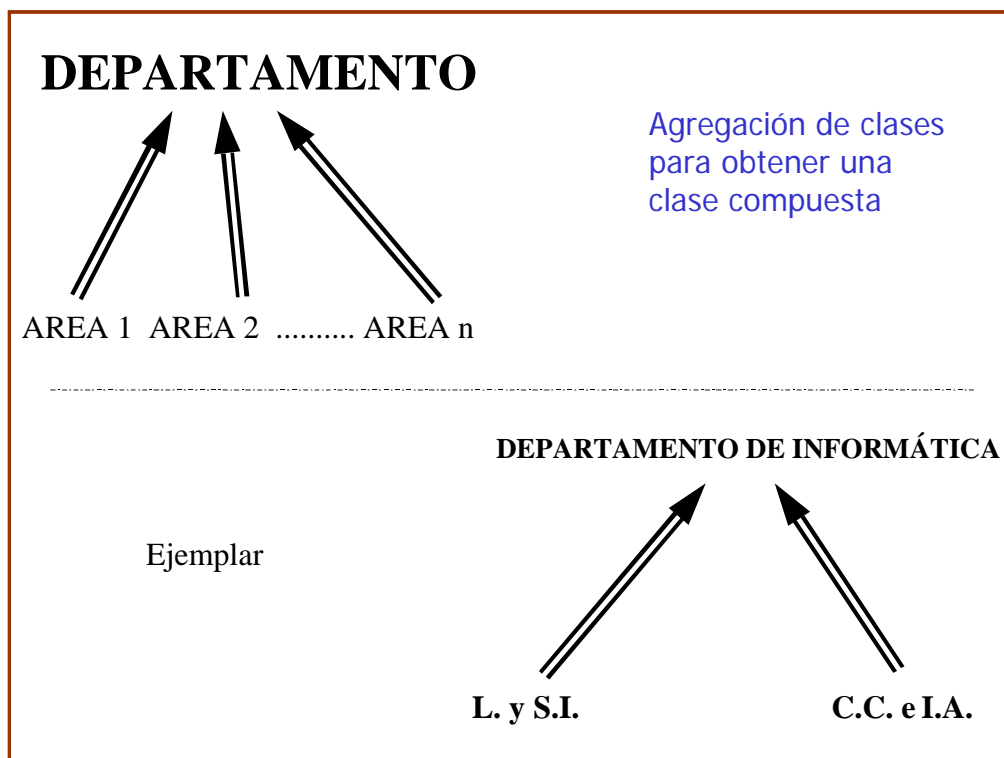


Agregación

- La abstracción de **Agregación** consiste en construir un nuevo elemento del modelo como **compuesto** de otros elementos (**componentes**):
 - Se establece una relación **ES_PARTE_DE** entre los elementos componentes y el elemento compuesto.
- El mecanismo contrario se llama **Desagregación**.
- Se pueden considerar tres tipos distintos de agregación:
 - Agregación de **clases** para obtener una **clase compuesta**,
(incluida en los MD semánticos: ER, UML)
 - Agregación de **propiedades** para obtener una **clase**, y
(admitida explícita o implícitamente por todos los MD)
 - Agregación de **propiedades** para obtener una **propiedad compuesta**.
(admitida por algunos MD: Codasyl Sí, Relacional No)

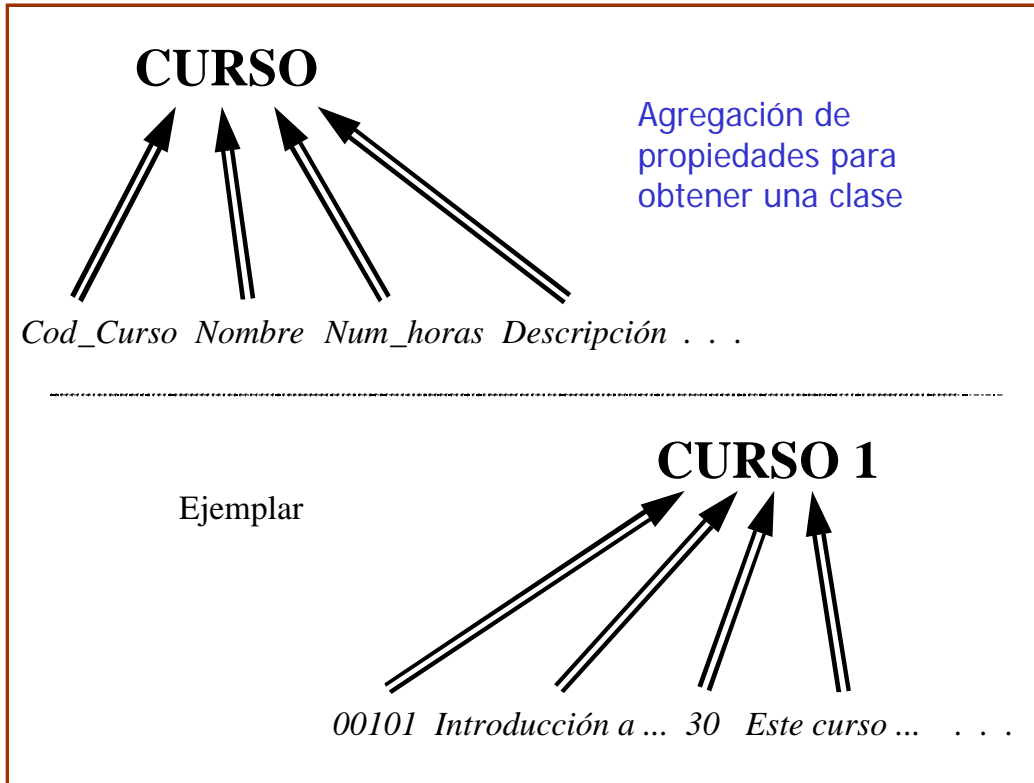


Agregación – de clases

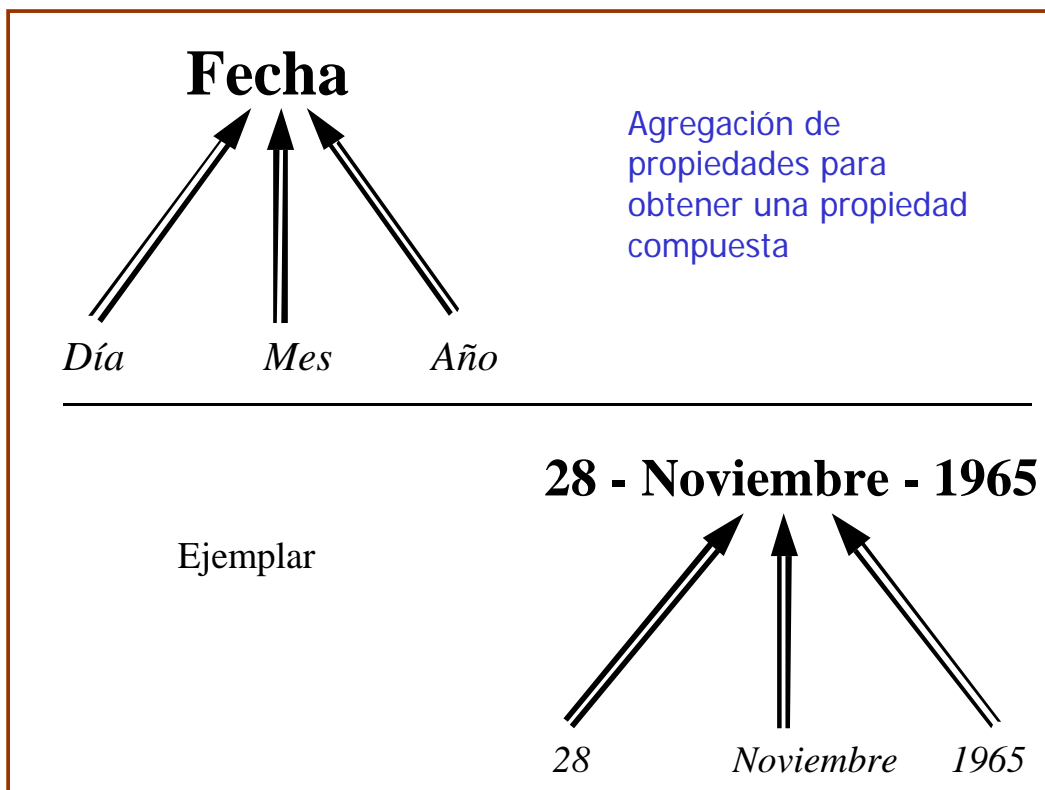




Agregación – de propiedades para obtener una propiedad compuesta



Agregación – de propiedades





Mecanismos de Abstracción

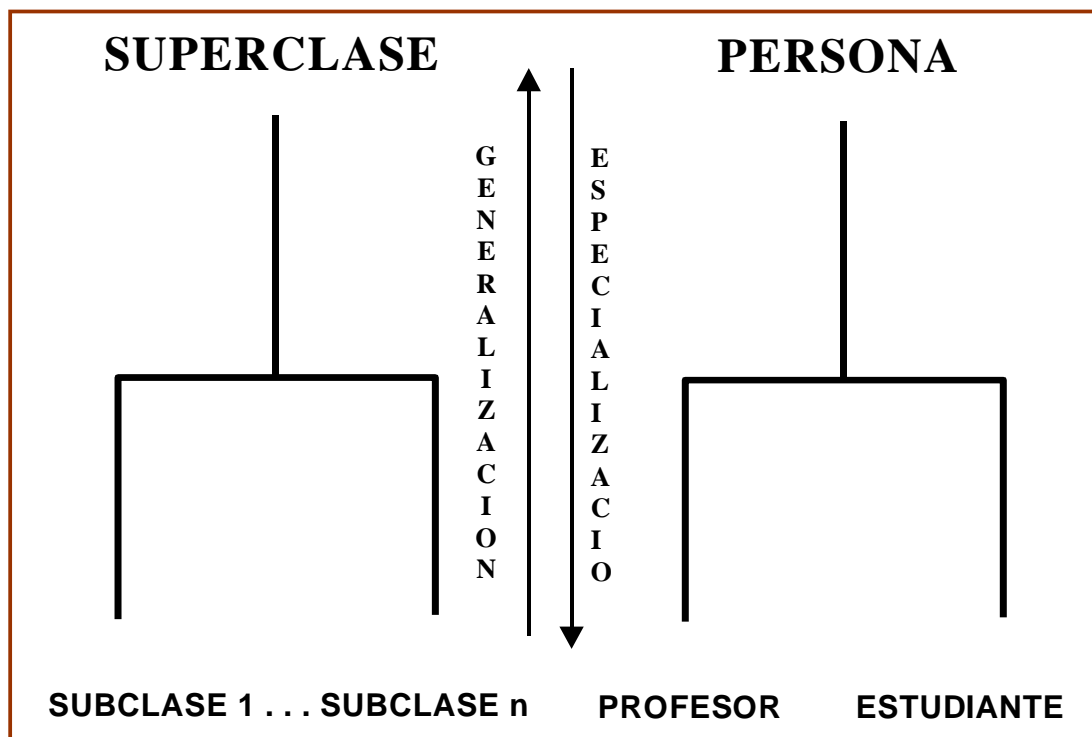
Generalización

- La Generalización es la acción de abstraer las características comunes a varias clases (**subclases**) para constituir una clase más general (**superclase**) que las comprenda:
 - El conjunto de ejemplares de una subclase “es un” subconjunto de los ejemplares de la correspondiente superclase.
 - Entre los elementos subclase y el elemento superclase se establece una relación del tipo **ES_UN**.
 - Ejemplo: La superclase PERSONA es una generalización de las subclases PROFESOR y ESTUDIANTE.
- El mecanismo inverso de la generalización es la **Especialización**.
- Todo ejemplar de una subclase es también ejemplar de la superclase y, además de poseer las características específicas de la subclase, **hereda** todas las correspondientes a la superclase.
- Aunque esta abstracción es muy intuitiva y muy útil, no se contempla en bastantes modelos de datos (p.e. Relacional).



Mecanismos de Abstracción

Generalización - representación





Mecanismos de Abstracción

Asociación

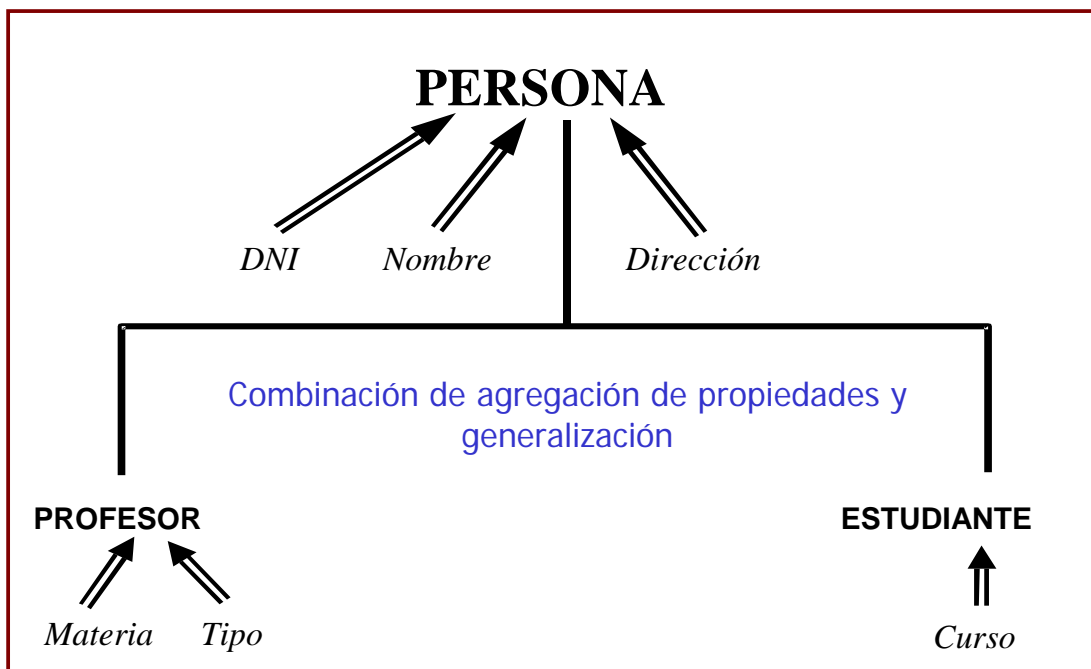
- La **Asociación** es una abstracción que se utiliza para relacionar dos o más clases (y, por tanto sus ejemplares), creándose un elemento de un tipo distinto.
- En algunos MD no aparece (p.e. Relacional).
- El mecanismo inverso es la **Disociación**.
- Algunos autores consideran que es un tipo especial de agregación, pero en De Miguel et al. (1999) se determinan las siguientes diferencias:
 - Cuando se asocian dos o más categorías, el nuevo elemento que aparece tiene determinadas características que lo distinguen de las categorías normales, por lo que, en general, los MD crean un nuevo concepto para representarlo.
 - El nuevo elemento no *está compuesto*, como en el caso de la agregación, por los elementos que asocia.
 - En la agregación puede existir herencia, y no así en la asociación.

- Ejemplo **PROFESOR** **Imparte** **CURSO**



Mecanismos de Abstracción

Jerarquías

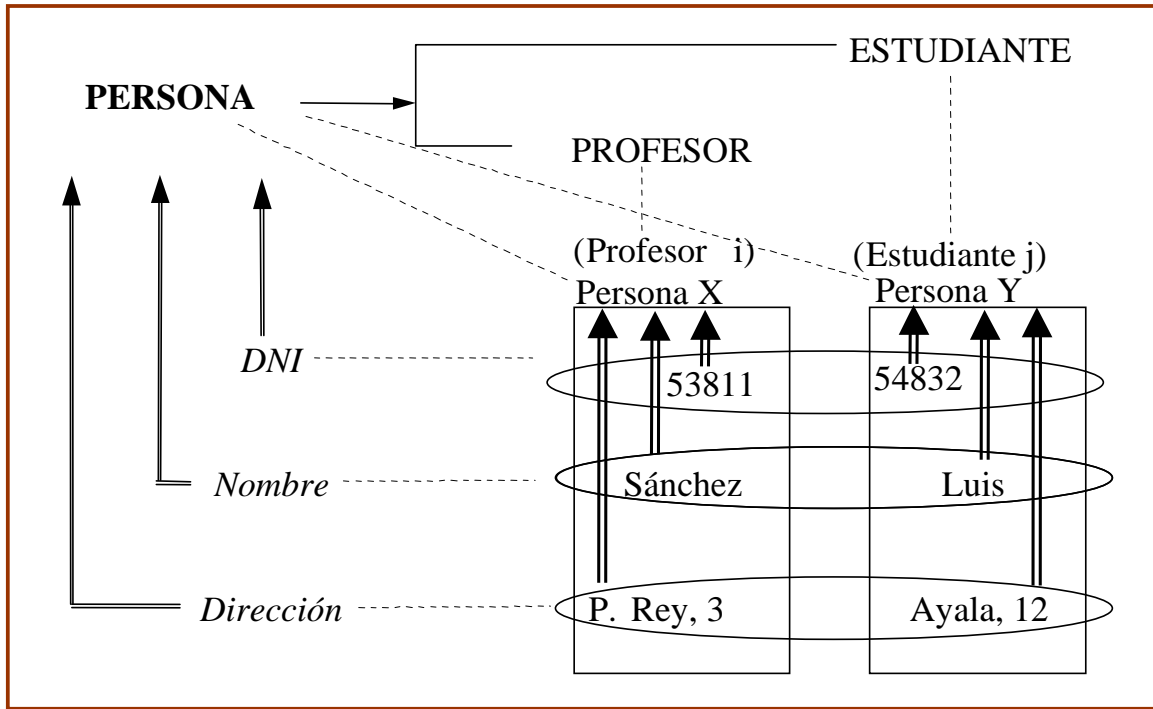


En el proceso de modelado de una determinada realidad es preciso, a menudo, combinar distintas abstracciones formando una jerarquía de abstracciones

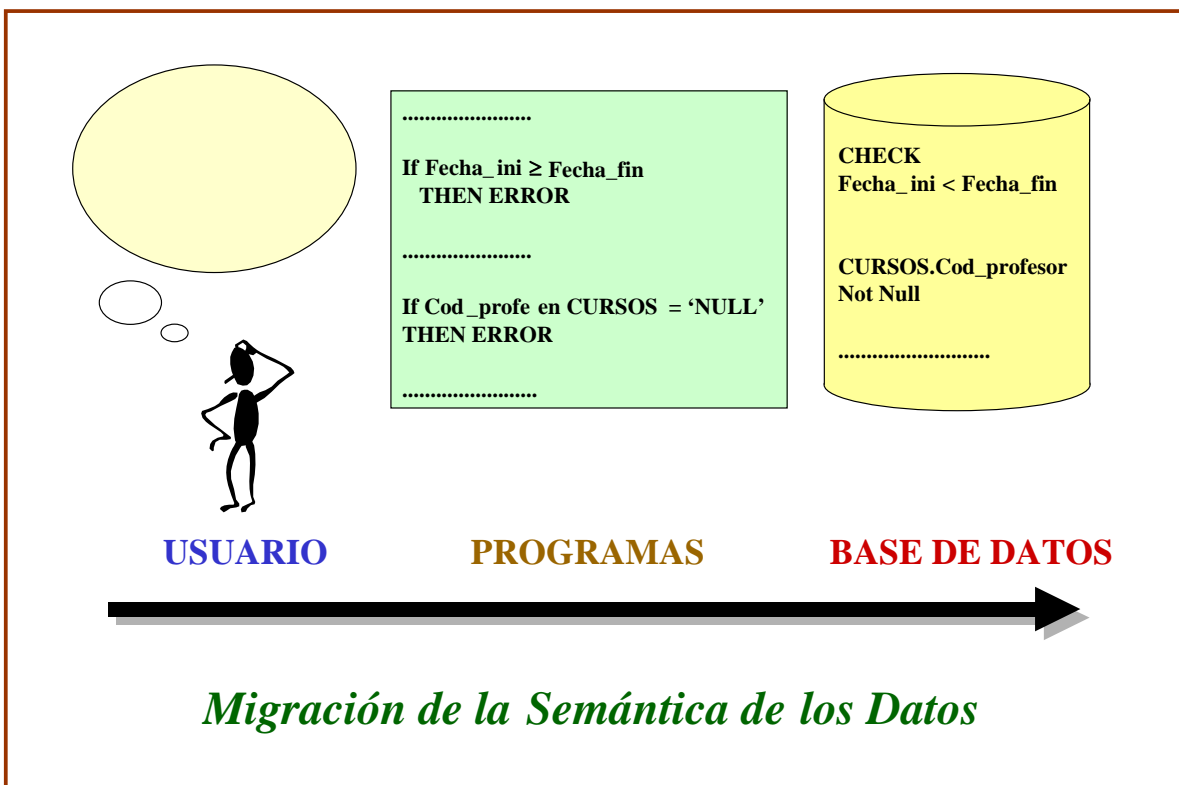


Mecanismos de Abstracción Jerarquías

Ejemplo de abstracciones de clasificación, agregación y generalización

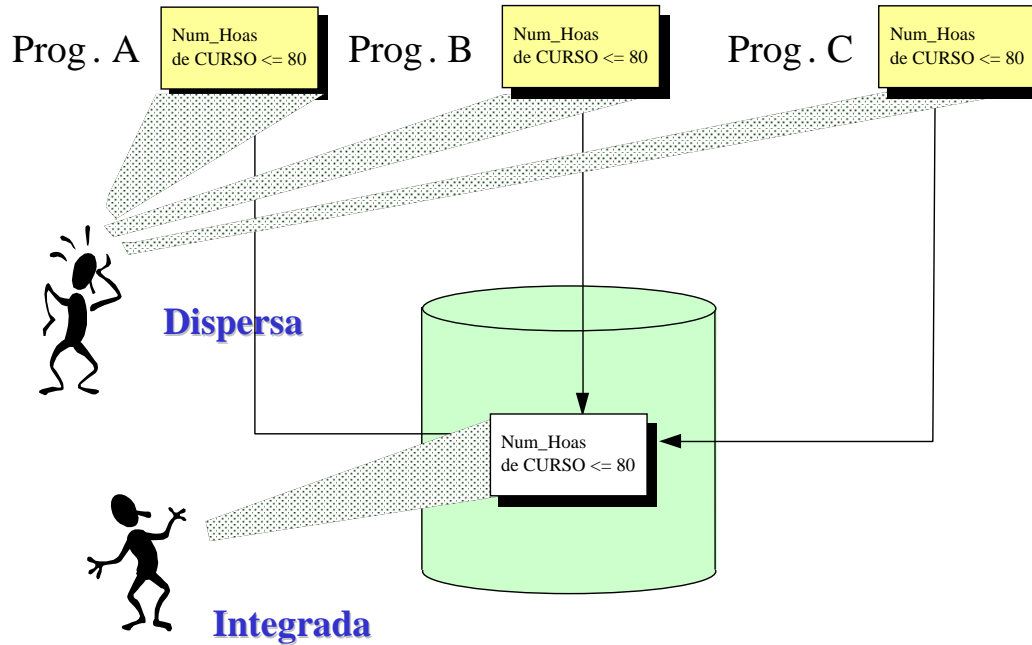


Restricciones de Integridad





Restricciones de Integridad



Semántica de los datos “dispersa” vs “integrada”

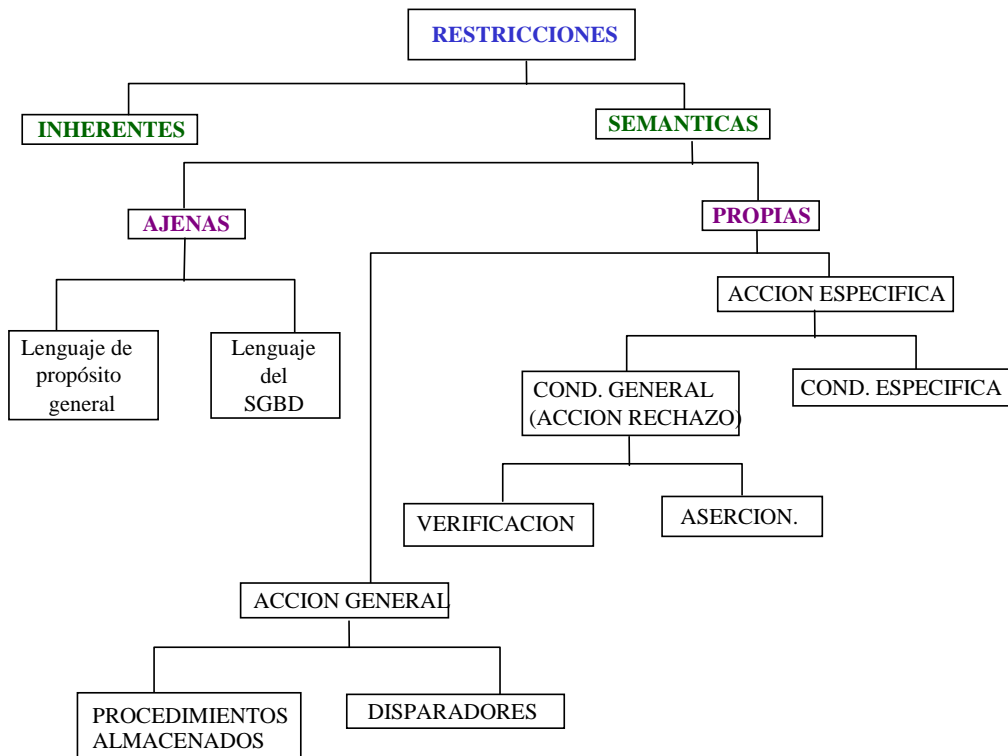


Restricciones de Integridad Componentes

- Una restricción de integridad tiene los siguientes componentes:
 - La **operación** de actualización (inserción, borrado o modificación) cuya ejecución ha de dar lugar a la comprobación del cumplimiento de la restricción.
 - La **condición** que debe cumplirse, la cual es en general una proposición lógica, definida sobre uno o varios elementos del esquema, que puede tomar uno de los valores de verdad (cierto o falso).
 - La **acción** que debe llevarse a cabo dependiendo del resultado de evaluar la condición.
- Las restricciones de integridad se pueden considerar, en cierto modo, como **reglas ECA (Evento, Condición, Acción)**:
 - al ocurrir un evento (en este caso una actualización), se comprueba una condición y dependiendo de su resultado se pone en marcha una acción (rechazar la operación, informar al usuario, corregir el error, etc.).
- Además de estos elementos, también pueden tener un **nombre**, por medio del cual es posible identificarlas, y también puede indicarse el **momento** en el que ha de evaluarse la condición.



Restricciones de Integridad Clasificación



Restricciones de Integridad Clasificación – semánticas propias

- Las **restricciones propias** se especifican al definir el esquema mediante las facilidades que proporciona la función de definición de datos, almacenándose en la base de datos (no en los programas), por lo que no pueden ser violadas por ninguna aplicación, es decir, cualquier actualización está obligada a respetarlas.
- Según sea o no preciso definir la acción tenemos existen tipos:
 - De **acción general**: es preciso programar un procedimiento (en algún lenguaje) que determine la acción que hay que llevar a cabo. Se subdividen en:
 - **Procedimientos almacenados**: se definen totalmente de forma procedimental (tanto la acción como la condición).
 - **Restricciones de disparo**: se definen mediante **disparadores (triggers)**. En ellas se formula una condición de forma declarativa, mediante una proposición lógica; el *cumplimiento* de la misma "dispara" una acción especificada de forma procedimental.
 - De **acción específica**: la acción (en general rechazo, aunque puede ser otra, bien predeterminada bien elegida mediante opciones) está implícita en la misma restricción.



Clasificación – de acción específica

- En ellas se distinguen las siguientes clases:
 - De **condición general**: la condición se define mediante una **proposición lógica**. La operación será una actualización. No se declara la acción porque siempre lleva asociado el rechazo: el sistema evalúa la condición y si el resultado es cierto se actualiza y si no es cierto no se lleva a cabo la operación. En **SQL 92** se incluyen dos tipos:
 - De **verificación**: La expresión lógica mediante la cual se formula la condición está definida sobre uno o varios atributos de un mismo elemento. Por ejemplo, una cláusula "**CHECK**" dentro de un CREATE TABLE.
 - De **aserción**: son análogas a las anteriores pero pueden estar referidas a más de un elemento del esquema ya que tienen existencia por sí mismas (por tanto tienen un nombre). Ejemplo: **CREATE ASSERTION**.
 - De **condición específica**: también llamadas de "caso especial" ó "implícitas". Se refieren a las diversas opciones que facilitan los distintos MD cuando se definen los elementos de su esquema y que en realidad son restricciones. Por ejemplo, en el modelo Relacional: **PRIMARY KEY, FOREIGN KEY, NOT NULL ...**



Clasificación – otros criterios para clasificar

- A) Según la **dimensión temporal**:
 - De **Estado** o *Estáticas*: las restricciones se aplican a un determinado estado de una BD y no hay necesidad de conocer los estados anteriores para saber si se cumple o no la condición. Ejemplo: edad >= 0.
 - De **Transición** o *Dinámicas*: la restricción hay que aplicarla a la transición entre dos estados. Ejemplo: el salario de un empleado no puede disminuir.
- B) Según el **ámbito** o alcance:
 - Afectan a un **único ejemplar** de un tipo dado: edad > 14.
 - Afectan a **varios ejemplares** de un tipo dado. A su vez, pueden afectar sólo a algunos o a todos los ejemplares de un cierto tipo. (el sueldo de un empleado menor que el sueldo de su jefe)
- C) Según los **criterios de comparación**:
 - De **valor**: en la condición se comparan los valores que pueden tomar las propiedades.
 - **Estructurales**: imponen limitaciones a la estructura de los elementos del modelo. Por ejemplo, que un atributo no puede tomar más de un valor.



Taxonomía de los Modelos de Datos

- Según el **nivel de abstracción** que consideremos en la arquitectura ANSI, tendremos modelos **internos**, **globales** y **externos**:

MODELO DE DATOS

EXTERNO

* (punto de vista de cada usuario en particular)

GLOBAL

* (punto de vista del conjunto de usuarios -empresa-)

INTERNO

* (punto de vista de la máquina)

- También se utiliza la expresión “modelos **lógicos**” para hacer referencia tanto a los globales como a los externos, ya que ambos describen aspectos lógicos de los datos -frente a los modelos internos que describen aspectos **físicos**.



Taxonomía de los Modelos de Datos Modelos Globales

MD GLOBALES

CONCEPTUALES O SEMÁNTICOS

- enfocados a describir el mundo real con independencia de la máquina -

Entidad/Interrelación (E/R)
Objetos (UML)

CONVENCIONALES O LÓGICOS

- orientados a su implementación en SGBD -

Jerárquico
Red (Codasyl)
Relacional



Modelos Globales – convencionales vs conceptuales

CONVENCIONALES

- Implementados en SGBD comerciales
- Dependen del SGBD
- Más próximos al ordenador
- Poca capacidad semántica
- Más enfocados a la implementación
- Interfaz informático/sistema
- Nivel de “mediación” entre el nivel externo e interno

CONCEPTUALES

- No suelen estar implementados en SGBD
- Independientes del SGBD
- Mayor nivel de abstracción
- Mayor capacidad semántica
- Más enfocados al diseño de alto nivel (modelado conceptual)
- Interfaz usuario/informático



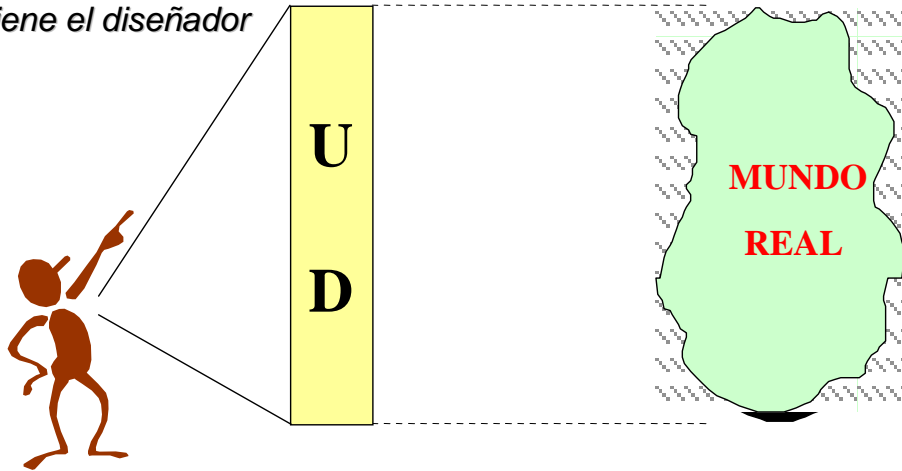
Los MD en el proceso de diseño de una BD

- Se conoce como proceso de diseño de una BD al conjunto de tareas necesarias para pasar de una determinada realidad (Universo del Discurso) a la BD que la representa. Los MD juegan un importante papel en el proceso de diseño de una BD al ofrecernos facilidades de abstracción que nos ayudan a representar la realidad.
- Los **objetivos** que persigue todo MD son de dos tipos:
 - a) De **Formalización**: el MD permite definir formalmente las estructuras permitidas y las restricciones; también establece la base para la definición de un lenguaje de datos.
 - b) De **Diseño**: el MD es un elemento fundamental en el desarrollo de una metodología de diseño de BD, en el cual se basan los otros componentes de la metodología (lenguajes, documentación y otras herramientas).



Los MD en el proceso de diseño de una BD

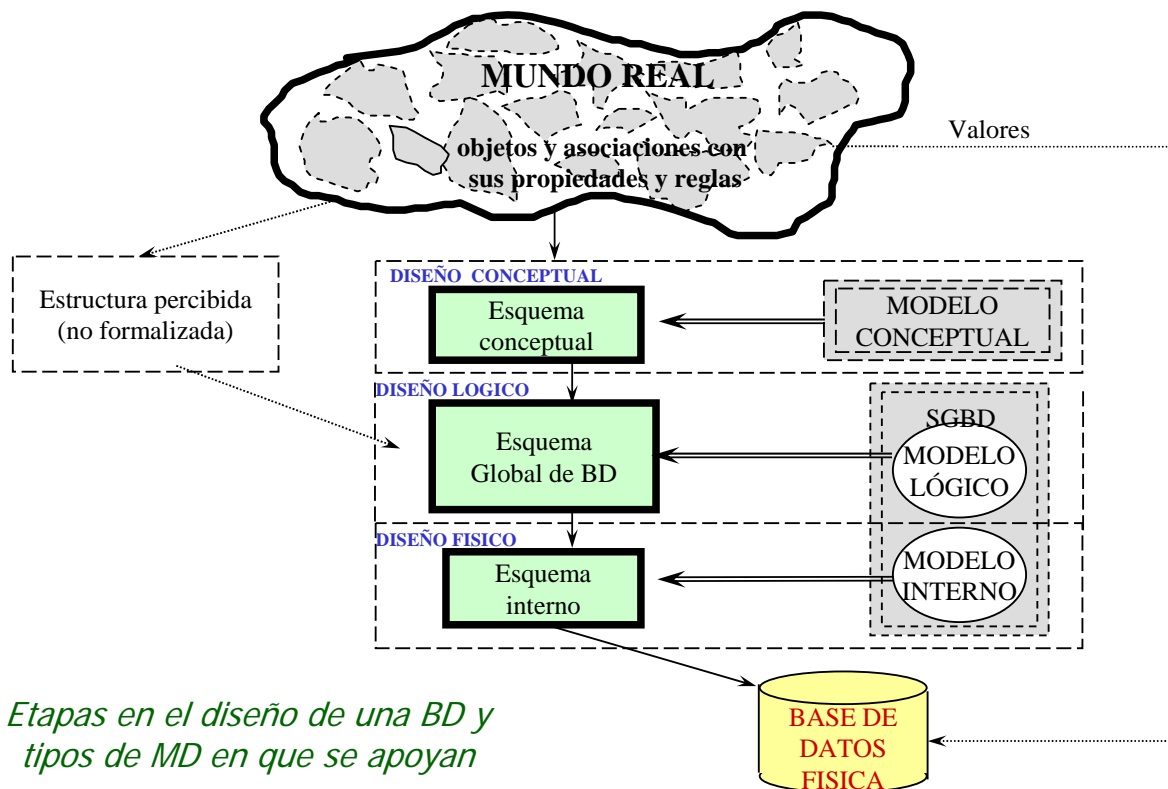
visión del mundo real que tiene el diseñador



Universo del Discurso y Mundo Real



Los MD en el proceso de diseño de una BD



Etapas en el diseño de una BD y tipos de MD en que se apoyan