
Transformación de Esquemas

Heterogeneidades en BDs

- **Modelo de Datos y Lenguajes**
- **Esquemas**
 - Semántica
 - Esquemática
 - Restricciones de Integridad
- **Optimización de consultas**
 - heurísticas, basadas en costos, combinación
 - basadas en la organización física: clustering sobre el disco, cache, etc.
- **Manejo de transacciones**
 - protocolos de control de concurrencia,
 - protocolos de commit
 - protocolos de recuperación

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Modelos de Datos y Lenguajes

- **Modelo de Datos**
 - Jerárquicos, de Redes, Relacionales, Funcionales, Semánticos, Orientados a Objetos...
- ⇒ pasar a un **Modelo de Datos Común** con un poder de modelaje mayor al de los modelos componentes.
- Funcional [Hurson et al. 94]
 - Relacional [Rusinkiewicz et al 89]
 - Entity Relational [Johanneson 94]
 - Object Oriented [Bukes and Elmagaramid 96]
- **Lenguajes de Consultas**
 - SQL(dialectos), QUEL, QBE, etc
- ⇒ pasar a un **Lenguaje de Consulta Global** con mayor poder expresivo.

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Transformación de Esquemas

Esquema :

- **Un conjunto de estructuras**
- **Un conjunto de restricciones sobre las estructuras**
(capturan las propiedades comunes a todas las instancias del esquema)
- **Un conjunto de operadores sobre las estructuras**
(proveen el vocabulario para la formulación de consultas sobre el esquema)

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Transformación de Esquemas

Como comparar esquemas?

⇒ a través de la **Capacidad de Información** del esquema.
[MIR93]

Sea $I(S)$ el conjunto de todas las posibles instancias válidas del esquema S .

Desde otra perspectiva la **Capacidad de Información** de S determina el conjunto $I(S)$.

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Definiciones previas

Sean dos conjuntos A y B .

- $f : A \rightarrow B$ es una **función** si para cada $a \in A$, existe como máximo un $b \in B$ tal que $f(a) = b$.
- f es **inyectiva** si su inversa es una función.
- f es **total** si está definida para todo elemento de A .
- f es **sobreyectiva** si su inversa es total.
- f es **biyectiva** si es una función inyectiva, total y subyectiva.

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Capacidad de Información (Dominancia)

Intuitivamente:

Un esquema **S2** tiene mayor capacidad de información que un esquema **S1** si cada instancia de **S1** puede ser mapeada a una instancia de **S2** sin pérdida de información.

Definición 1:

Un mapeo que preserve la capacidad de información entre las instancias de dos esquemas **S1** y **S2** es una función f total e inyectiva: $I(S1) \rightarrow I(S2)$

Definición 1.1:

Si $f: I(S1) \rightarrow I(S2)$ es un mapeo que preserve la capacidad de información entre las instancias de dos esquemas **S1** y **S2**, entonces **S2** **domina a S1** vía f , y se denota: $S1 \leq S2$.

Regina Motz - InCo INTEROPERABILIDAD Transf. de Esquema y Lenguajes

Capacidad de Información (Equivalencia)

Definición 2:

Un mapeo que preserve la **equivalencia** entre las instancias de dos esquemas **S1** y **S2** es una función $f: I(S1) \rightarrow I(S2)$ **biyectiva**.

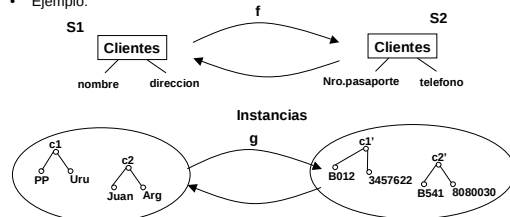
Definición 2.1:

Si $f: I(S1) \rightarrow I(S2)$ es un mapeo que preserve la equivalencia entre las instancias de dos esquemas **S1** y **S2**, entonces **S1** y **S2** **son equivalentes** vía f , y se denota: $S1 \equiv S2$.

Regina Motz - InCo INTEROPERABILIDAD Transf. de Esquema y Lenguajes

Comparar Capacidad de Información

- A través de la **intencion** del esquema
- Equivalencia: si puedo encontrar una función f **biyectiva** entre sus conjuntos de instancias posibles $I(S)$.
- Ejemplo:



Regina Motz - InCo INTEROPERABILIDAD Transf. de Esquema y Lenguajes

Ejemplos (Capacidad de Información)

Ejemplo 1:

S1: Emp1(E-nro, E-nombre, Sueldo)

S2: Emp2(E-nro, E-Nombre, Sueldo, Edad)

S1 es equivalente a S2?

S1 domina a S2?

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplos (Capacidad de Información)

Ejemplo 2:



S1 es equivalente a S2?

S2 domina a S1?

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Transformación de esquemas

Definición 3:

Sean S1 y S2 familias (conjuntos) de esquemas.

Una **transformación de esquemas** es una función total $F: S1 \rightarrow S2$. Si S1 y S2 están especificados en diferentes modelos de datos, la transformación F es llamada de **translation**.

$\Rightarrow$ Una **transformación de esquemas es deseable que sea un mapeo que preserve la equivalencia**.

Definición 3.1:

Una transformación $F: S1 \rightarrow S2$ es una transformación que preserva la equivalencia si para todo $S1 \in S1$, $S1 \equiv F(S1)$.

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Objetivos Operacionales y Capacidad de Información

Transformación de esquemas se ve en distintos contextos con diversos objetivos operacionales. De acuerdo al requerimiento operacional las transformaciones deben cumplir en diverso grado restricciones basadas en la capacidad de información.

Objetivo 1:

Consultar a través de S1 los datos almacenados en S2.

(S1 es una vista de S2)

Toda consulta de S1 debe corresponder a una única consulta en S2.

$\Rightarrow f: I(S2) \rightarrow I(S1)$ **es una función total.**

f no necesita preservar la información

Regina Motz - InCo

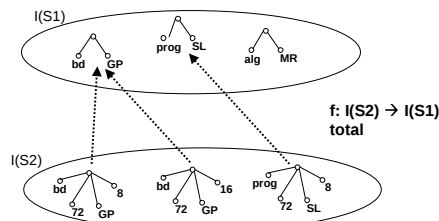
INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplo – Objetivo 1

S1: Cursos-Ing (curso, docente)

S2: Cursos-IngComp (curso, carrera, docente, horario)



Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Objetivos Operacionales y Capacidad de Información(II)

Objetivo 2:

El objetivo 1 más viendo desde S1 enteramente a S2.

$\Rightarrow$ necesitamos **f total** como en el objetivo 1, pero también **f** debe ser **sin pérdida de información**.

Una instancia de S1 debe únicamente determinar una instancia de S2.

$\Rightarrow f$ debe ser también **inyectiva** para que exista su inversa.

$$f^{-1}(i1) = f^{-1}(f(i2)) = i2$$

$\Rightarrow f$ debe ser un mapeo que preserve la información:

S2 $\leq$ S1

Regina Motz - InCo

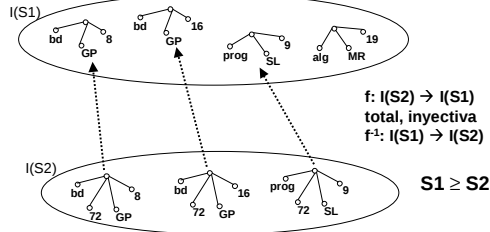
INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplo – Objetivo 2

S1: Cursos-Ing (curso, docente, horario)

S2: Cursos-IngComp (curso, carrera, docente, horario)



Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Objetivos Operacionales y Capacidad de Información)(III)

Objetivo 3:

El objetivo 2 más actualizar a través de S1 los datos de S2.

$\Rightarrow$ **f total, inyectiva** por objetivo 2..

Sea la actualización $u(i1) = i1'$

Para realizar la actualización en S2 todo $i1'$ debe determinar una única instancia de S2, $i2'$.

$\Rightarrow$ **f sobreyectiva** en S1 $\Rightarrow$ **f⁻¹** también está únicamente definida:

$u(i1) = i1' \Leftrightarrow u(f(i2)) = f(i2') \Leftrightarrow f^{-1}(u(f(i2))) = f^{-1}(f(i2')) = i2'$

Actualización corresponde a: $f^{-1} \circ u \circ f$

f y f^{-1} deben preservar al capacidad de información

$\Rightarrow$ **S1 $\equiv$ S2**

Regina Motz - InCo

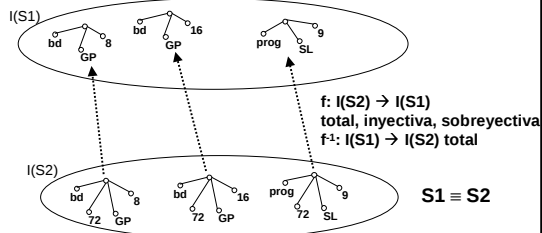
INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplo – Objetivo 3

S1: Cursos-IngComp (curso, docente, horario)

S2: Cursos-IngComp (curso, carrera, docente, horario)



Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Objetivos Operacionales y Capacidad de Información(IV)

Objetivo 4:

Consultando a traves de S1 los datos de S2 y viceversa.

Este objetivo es equivalente al objetivo 1 para ambas direcciones.

⇒ **f total** para una direccion y función **sobreyectiva** para la otra dirección,
o sea **bijectiva**.

⇒ **S1 = S2**

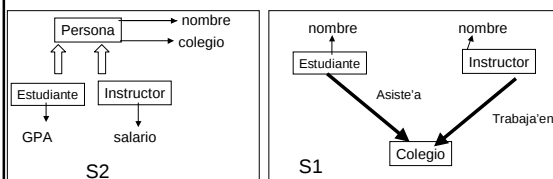
Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplo de mapeamento de consultas

- Para que el objetivo 2 se verifique se debe cumplir que f tiene que ser un mapeamiento que preserva informacion.
- Como transformar en la practica las consultas sobre la vista S2 hacia el esquema S1.



Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Ejemplo de mapeamento de consultas

- **En S2:**
Q1: Retrieve P where nombre(P) = "A. Walker"
- **En S1:**
Q2: Retrieve E
where Estudiante.nombre (E) = "A.Walker"
union
Retrieve I
where Instructor.nombre (I) = "A.Walker"

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Contextos

- Integración de vistas
 - Mismo modelo
 - Puede interesar: ver, actualizar
- Integración de BD heterogeneas
 - Transformación de esquemas de un modelo a otro
 - Dificultad importante: como definir la función g que implica identificar objetos comunes
- Evolución de esquemas
 - Mismo modelo
 - Mismas instancias

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Correspondencias entre esquemas

- Ej.
 
- *Equivalencia*. $S1 \equiv S2$
 - La intención de los 2 esquemas es la misma (semántica)
 - Existe f función biyectiva
- *Inclusion*. $S1 \subset S2$
 - Ej.: En S1 son solo las reservaciones que ya están pagas
 - Dominancia de S2 a S1 (f inyectiva)

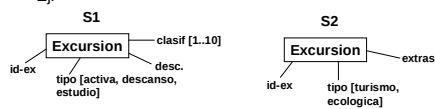
Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Correspondencias entre esquemas

- *Solapamiento* (superposición).
 - No hay equivalencia ni inclusión entre S1 y S2
 - Ej.



Solo las excursiones de S1 que tienen tipo=activa y clasif=4 se corresponden con las tipo=turismo de S2

$f: \{(S1.Excursion) / clasif=4 \wedge tipo=activa\} \rightarrow \{(S2.Excursion) / tipo=turismo\}$
f biyectiva

Regina Motz - InCo

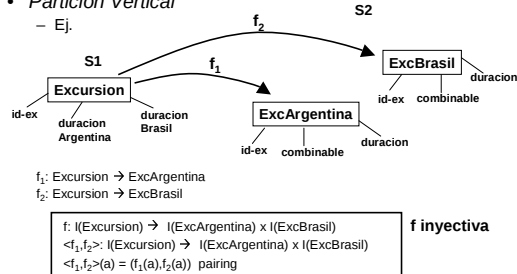
INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Correspondencias entre esquemas

• Particion Vertical

– Ej.



Regina Motz - InCo

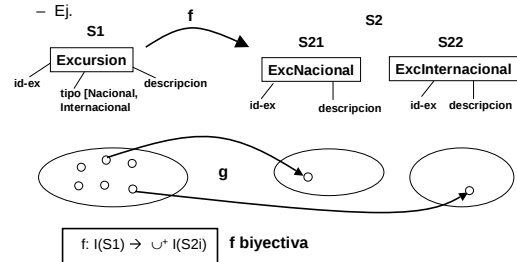
INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Correspondencias entre esquemas

• Particion Horizontal

– Ej.



Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Bibliografía

- [Bukes and Elmagaramid 96] Bukes O. and Elmagaramid A.K.
Object Oriented Multidatabase Systems: A solution for advanced applications. Prentice Hall, Englewood Cliffs, new jersey, 1996.
- [Hurson et al 94] Hurson, A.R., Bright, M. W and Packzad, H.
Multidatabase Systems: An advanced solution for Global Information Sharing. IEEE Computer Society Press, Los Alamitos, CA, 1994.
- [Johannesson 94] Johannesson P.
A metod to translating relational schemas into conceptual schemas. In proceedings of the Tenth international conference on data Engineering, 1994.

Regina Motz - InCo

INTEROPERABILIDAD

Transf. de Esquema y Lenguajes

Bibliografía

[MIR93] Miller,R.J., Ioannidis, Y.E and Ramakrishnan R.

The use of Information Capacity in Schema Integration and translation. VLDB Conference, 1993.

[Rusinkiewicz et al 94] Rusinkiewicz M. , El-Masri, R., Czejdo B., Georgakopoulos D. , Karabatis G. , Jamousi, A. , Loa, K. and Li, Y.
Query processing in a heterogeneous multidatabase environment.
In proceedings of the first annual symposium on Parallel and Distributed Processing, 1989.