

Integración de Datos

Integración: *read-only*

- **Características:** necesidad de *fusión* de datos
- **Abordajes:**
 - **virtual**
(basado en modelo de descomposición de consultas, enviando las subconsultas a las BD fuentes y mezclando las respuestas.)
 - Close Word Assumption (CWA) La vista integrada posee exactamente la union de los datos en las BD fuentes.
Vista construida "bottom-up"
 - Open Word Assumption (OWA) La respuesta provista a través de una vista integrada tendrá un subconjunto de las respuestas completas que implican cada una de las bases de datos fuentes.
Vista construida "top-down"

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de

Integración: *read-only (II)*

- **Abordajes:**
 - **materializado**
 - Vista construida "bottom-up" en un abordaje CWA. ((DW))
 - Aspecto crítico: mantenimiento actualizado de la vista.
 - "Refresh" periódico
 - Mantenimiento incremental [GL95], [LMSS95]
 - "Self-maintainability" [GJM96] [ZHK96]
(deltas: representación, álgebra, aplicación inmediata o diferida)
 - Ventaja sobre el virtual: tiempo de respuesta.
 - **híbridos**

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de

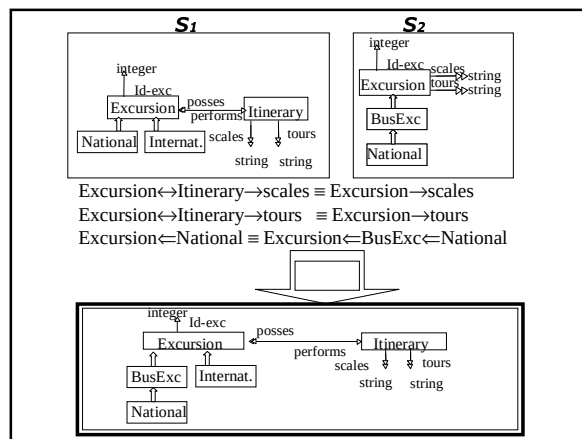
Metodologías

- Correspondencias al nivel de los esquemas.
Parten de la resolución de heterogeneidades entre los esquemas.
(Sistemas fuertemente acoplados)
[SpaccapietraParent 94], [Kosky 95], [Fankhauser 97]
- Correspondencias al nivel de las instancias .
(Sistemas débilmente acoplados)
[Fang et al. 93], [Scholl et al. 94] , [VermeeerApers 96]

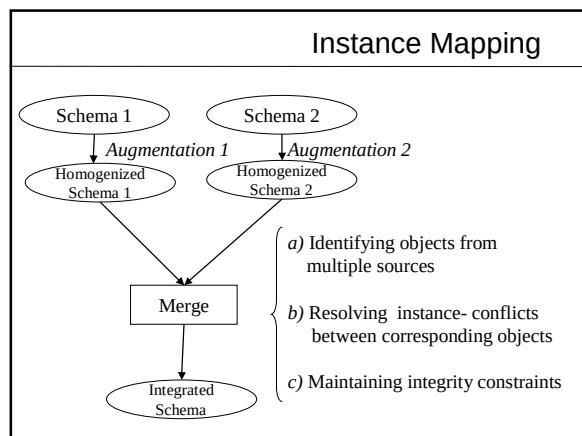
Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de



Instance Mapping



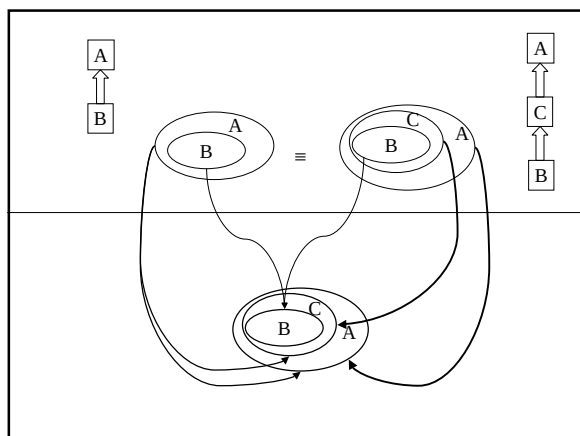
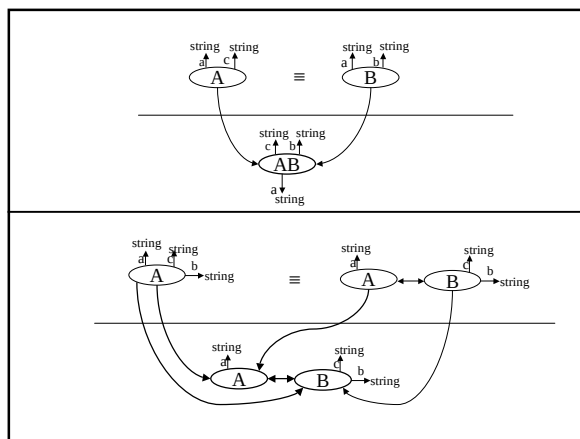
Generating an Integrated Schema

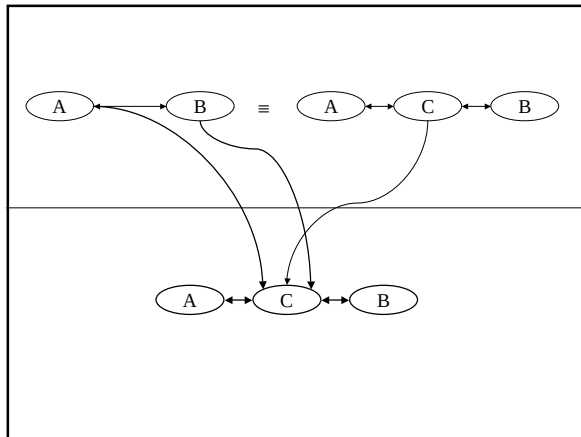
- **Identifying Instances:** keys values approach
- **Data Representation Conflicts:** conversion mappings or routines
- **Data Precision Conflicts:** conversion function (when possible)
- **Data Scaling Conflicts:** conversion functions
- **Default Value Conflicts:** preference criterion
- **Attribute Integrity Constraint Conflicts:**
generalization of constraint (when possible)

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de





Mapping Specification Syntax

```

interface <class name>: {<superclass name>}{
  extent    <extent name>
  keys      [<attribute name>]*;
  attribute <attribute type> <attribute name>
  relationship <class name> <relationship name>
               inverse <class name>: <relationship name>}

```

```

mapping <class name>{
  origins    <attribute name> [<attribute name>]*;
  def_ext    <oql expression>
  def_att    <attribute name> as <oql expression>
  def_rel    <traversal path name> as <oql expression>; }

```

Regina Motz - InCo
Data

INTEROPERABILIDAD

Integración de

```

mapping h_Excursion1{
  origins eorig; // eorig from Excursion1
  def_ext select h_Excursion1(eorig: e_inst)
           from e_inst in e_excursions;
  def_att h_exc_id as this.eorig.exc_id;
  def_rel posses as select e from e in posses
           where (this.eorig = e.eorig.tours) and
                 (this.eorig = e.eorig.scales); }

mapping h_Itinerary2{
  origins torig, eorig1, eorig2; // torig from Excursion2 ,
                                // eorig1 from Itinerary.tours,
                                // eorig2 from Itinerary.scales
  def_ext select h_Itinerary2(torig: inst, eorig1: itinerary.tours, eorig2: itinerary.scales)
           from inst in e_excursions,
           inst.tours in itinerary.tours,
           inst.scales in itinerary.scales;
  def_rel posses as select p from p in excursion
           where (this.torig.excursion.tours = p.eorig1) and
                 (this.torig.excursion.scales= p.eorig2) ;)

```

Integración basada en relaciones semánticas al nivel de las instancias

OBJETOS:

Cada *objeto* representa cierto objeto del mundo real.

Los objetos de la BD están agrupados en *clases*.

Cada clase posee un conjunto de *propiedades* por las cuales los objetos de esa clase son descriptos.

El conjunto de propiedades determina la *estructura* de un objeto.

Cada propiedad tiene un *dominio* del cual toma valores.

Para propiedades *referenciales* el dominio es una clase.

Cada objeto de la BD tiene *valores* para sus propiedades.

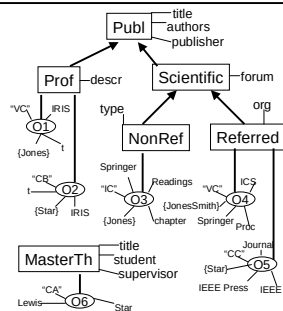
Este conjunto de valores determina el estado del objeto de la BD.

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de

Ejemplo [VermeerApers96]

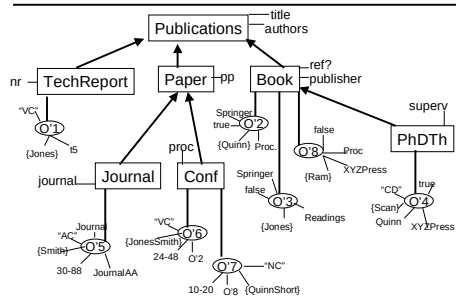


Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de

Ejemplo [VermeerApers96]



Regina Motz - InCo
Datos

INTEROPERABILIDAD

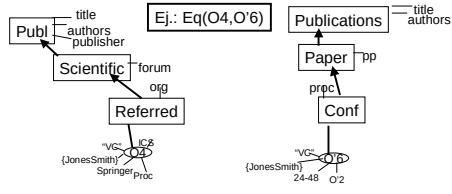
Integración de

Relacionamientos entre objetos [VermeerApers96]

- **Equal:** corresponden al mismo objeto del mundo real

$Eq(A,B)$ vale sii el objeto A es el mismo que el objeto modelado por la abstracción B.

a) B puede ser un objeto.



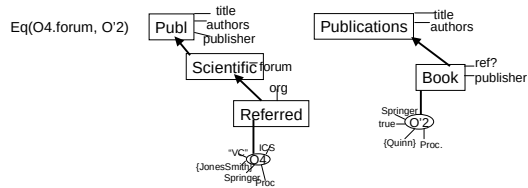
Regina Motz - InCo
Datos INTEROPERABILIDAD Integración de

Relacionamientos entre objetos [VermeerApers96] (II)

b) B puede ser un conjunto de valores de propiedades.

Ejemplo:

Publicaciones científicas de un forum son vistas como objetos en DB2 mientras que son vistas como valores describiendo publicaciones en libros en DB1.



Regina Motz - InCo
Datos INTEROPERABILIDAD Integración de

Relacionamientos entre objetos [VermeerApers96] (III)

- **Similar**

Un objeto puede ser similar a un conjunto de objetos de una clase C

- Strict Similarity

Sim(A,B) es valido sii objeto A es similar a los objetos representados por la abstracción B.

Ejemplo: Una PhDTh es con referato entonces:

$SIM(O'4, Referred)$

- Approximate Similarity

Sim(A,B, NewC) Ocurre cuando O' y los objetos de C son suficientemente similares como para agruparlos en una nueva clase NewC.

Ejemplo: $Sim(O'4, MasterTh, GradTh)$

Regina Motz - InCo
Datos INTEROPERABILIDAD Integración de

Relacionamientos entre objetos [VermeerApers96] (IV)

- **Aggregate**
Es usada para la composición de objetos dentro de un objeto mayor.

Aggr(A, B [, Role]) Vale sii el objeto modelado por la abstracción A es una agregación del objeto representado por la abstracción B.

Opcionalmente se puede especificar un papel (role) específico por B en A.

Ejemplo: El chapter de Jones esta contenido en su libro:
Aggr(O3, O'3)

Regina Motz - InCo INTEROPERABILIDAD Integración de Datos

Reglas de Comparación de Objetos

Ejemplo:

Sim(O':ConfPaper, refereed) $\leftarrow$ O'.proc.ref? = true

Eq(O':Prof, O:TechReport) $\leftarrow$ occurs(concat("TR", O.nr), O'.descr

- Resuelven heterogeneidades esquemáticas

- El conj. de objetos relacionados que definen no necesariamente coinciden con las extensiones de ninguna clase.

Regina Motz - InCo INTEROPERABILIDAD Integración de Datos

Proceso

Entrada:

Set of local objects LO1, LO2.
Clasificación de los objetos locales SC1, SC2.
Relacionamientos entre LO1 y LO2.

Salida:

- **Un nuevo conjunto de objetos integrados IO**
⇒ Reconciliación de objetos
Ejemplo: Eq(O, O'.prop)
como objeto de DB1 o como valor de DB2?: Estrategia fija
Diferentes dominios: Funciones de conversión
Diferentes valores: Funciones de decisión
- **Una nueva clasificación para estos objetos.**

Regina Motz - InCo INTEROPERABILIDAD Integración de Datos

Características

- Aplicable para sistemas débilmente acoplados.
- Aplicable en ambientes estables
- Carece de integración de métodos
- Carece de formas de detectar inconsistencias
- Usado para *read-only*

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de

Bibliografía

- [Fang et al. 93] Fang D., Ghandeharizadeh S., McLeod D. And Si A.
The design, implementation and evaluation of an object-based sharing mechanism for federated database systems.
In Proc. Ninth Internat. Conf. on Data Engineering, Vienna, Austria, April 19-23, 1993. Washington, DC.: IEEE Computer Society Press, pp 467-475,1993.
- [Scholl et al. 94] Scholl M. H., Scheck H-J. And Tresch M.
Object algebra and views for multiobjectbases.
In Distributed Object Management, M.T. Oszu, U.Dayal and P. Valduriez, Eds. San Mateo, CA: Morgan Kaufmann Publishers,pp353-374,1994.
- [VermeerApers96] Vermeer M.W.W. And Apers P.M.G.
On the applicability of schema integration techniques to database interoperation.
ER'96

Regina Motz - InCo
Datos

INTEROPERABILIDAD

Integración de