

## Capítulo 5. Integración de componentes locales

El presente capítulo describe el proceso operacional requerido para llevar a cabo la integración de componentes locales. La selección apropiada de esquemas locales, es uno de los conceptos que cobran importancia en este capítulo, pues de esto depende llevar a cabo satisfactoriamente la integración.

Las opciones de integración de información descriptiva o espacial demandan tratamientos diferentes, por lo que es importante identificar el tipo de información contenida en los esquemas locales. Finalmente, se presenta la forma en se solucionan los conflictos y la manera en que se registra la información necesaria en el esquema auxiliar para ofrecer un acceso transparente a los componentes locales.

### 5.1 SELECCIÓN DE COMPONENTES

Uno de los aspectos críticos que garantizan una buena integración, es la selección apropiada de los componentes locales; puesto que si no existe una relación lógica entre ellos, no es posible ofrecer una consistencia en el esquema global generado.

Los conceptos de equivalencia de entidades presentados en el capítulo 3 deben considerarse para una selección apropiada de los esquemas locales.

La herramienta de integración de esquemas permite integrar componentes basados en el modelo relacional, con información descriptiva y/o espacial. Para el caso de información espacial se considero el modelado de Quadtree, modelo descrito en la sección 5.1.2.

Antes de seleccionar los componentes a integrar es importante considerar los requerimientos del usuario del esquema global.

#### *5.1.1 Especificación de requerimientos de integración.*

La recopilación de la información pertinente a los requerimientos de consultas globales, permitirá plantear de manera formal las necesidades y la factibilidad para llevar a cabo la integración de los esquemas.

La especificación de requerimientos significa la traducción de las ideas en la mente de los usuarios (las entradas), en un documento formal (las salidas de los requerimientos). Así, la salida de este proceso es el conjunto de requerimientos especificados formalmente

El proceso de especificación de requerimientos consiste de dos actividades básicas: el análisis del problema y la especificación de requerimientos.

### **Análisis del problema**

Su propósito es obtener claramente las necesidades del usuario global del sistema, y conocer exactamente la información que desea acceder. Durante esta actividad se colecta toda la información relacionada, tal como: bases de datos, site de residencia, modelo de datos de las BD, esquemas de los componentes locales, propósito de la BD, tipo de información almacenada y privilegios de acceso. Toda la información recolectada necesita organizarse de alguna manera, se sugiere la utilización de una especificación estructurada y de un Diccionario de Datos para tal efecto.

Con la información organizada, es posible determinar que es lo que se desea integrar y la factibilidad de llevarlo a cabo. Una vez determinada la posibilidad de integrar debe ser especificada.

### **Especificación de Requerimientos**

La especificación formal de lo deseado debe representarse de una manera comprensible y clara. Para este caso se sugiere un modelo entidad-relación para ofrecer una visión clara de las entidades, relaciones y atributos que pueden estar disponibles a través de la federación.

#### **5.1.2 Componentes de Bases de Datos**

La información a considerar para la integración debe ser del tipo espacial o descriptiva. Dicha información con las restricciones que se describen a continuación.

#### **Componentes espaciales**

Uno de los tipos de información considerados para la integración, es la espacial. Información perteneciente a un Sistema de Información Geográfico (GIS), el cual describe datos tales como mapas, planos, terrenos, edificios, ríos, montañas, etc..

Los métodos de representación de la información geográfica, se basan en vectores, rastreo y en objetos [Briones 1998]. El modelo considerado para llevar a cabo la integración es el basado en rastreo (normalmente denominado raster).

El modelo de rastreo se caracteriza por la división del espacio en fragmentos, y donde es posible describir un objeto a partir de un conjunto de fragmentos. El modelado de un fragmento presenta la siguiente estructura en el modelo relacional:

```
FRAGMENTO(Identificador_objeto,Llave_Peano,  
[Nivel_presición])
```

donde el `Identificador_objeto`, hace referencia a un objeto del espacio. La `Llave_Peano` es la llave principal, determinada mediante el recorrido de la curva de peano, representada de manera binaria.. El Nivel, que puede ser opcional, es el nivel de resolución que tiene cada uno de los cuadrantes Para mayor información del modelado de fragmentos ver Apéndice A.

### Componentes descriptivos

Consideremos como información descriptiva aquella que describe un objeto (persona, cosa o evento) por medio de un conjunto de atributos. Es aquella información que no requiere de una interpretación especial para su entendimiento, tal como la espacial.

Este tipo de información es la tradicionalmente usada en aplicaciones de uso común. Nómina, contabilidad, inventarios y control escolar son algunos de los ejemplos que entran en esta categoría.

La única restricción para usar componentes descriptivos es que sean conceptualizados a través del modelo relacional.

## 5.2 IDENTIFICACIÓN Y SOLUCIÓN DE CONFLICTOS

Considere niveles de precisión 6 para el atributo *llave* y nivel 4 para *clave\_peano*, los cuales se determinan a través de un análisis de longitud de todas las llaves Peano de cada componente. Una vez determinado el nivel máximo de precisión de los objetos espaciales, se selecciona y se registra el menor de ellos como valor predeterminado para el esquema global. Para este caso en particular, el nivel de precisión menor es 4, el cual se registra en la tabla de *Funciones\_mapeo* para los atributos que almacenan llaves de Peano.

Este capítulo presento la forma en que se lleva a cabo la integración de esquemas con auxilio de la herramienta de integración. La idea de ejemplificar la identificación y solución de conflictos, es para mostrar claramente como generar una federación. La identificación de requerimientos y la selección de componentes determinan, en conjunto, uno de los factores de éxito del proceso de integración, pues se garantiza cumplir con las expectativas del usuario de la federación.

Con esta última parte es posible determinar las restricciones y/o limitaciones de la herramienta. Tales restricciones pueden establecer la pauta para trabajos futuros sobre la misma línea de investigación.

**Tabla 5.1 Requerimientos de almacenamiento para conflictos semánticos y de datos**

Conflicto	Tipo Conflicto		Tablas requeridas			ID Confl
	Esquema	Datos	Equivalencia	Tabla_auxiliar	claves_primarias	
Nombrado de tablas						
Sinónimos	X		X			
Nombrado de atributos						
Sinónimos	X		X			
Restricciones de integridad						
Clave primaria	X				X	
Grado de atomicidad	X		X			CMP (Compos
Fragmentación						
Horizontal						
Vertical	X		X			
	X		X			
Identidad	X		X			H (Horiz
Estructura de las tablas						
Atributos	X		X			FA (Falta Atributo)

faltantes						
Diferentes representaciones						UD
Unidades distintas		X	X	X		(Unidade Distintas
Diferentes niveles de precisión						NP
Nivel de precisión		X	X	X		(Nivel de Precisión

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tip</b>
municipios	id_municipio	municipios	id_municipio	1	
municipios	nombre	municipios	nombre	1	
municipios	num_hab	muni	num_hab	2	
municipios	superficie	muni	superficie	2	

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tip</b>
municipios	id_municipio	municipios	id_municipio	1	
municipios	nombre	municipios	nombre	1	

municipios	num_hab	muni	num_hab	2	
municipios	superficie	muni	superficie	2	
presidentes	rfc	presidentes	rfc	1	
presidentes	nombre	presidentes	nombre	1	
presidentes	domicilio	presidentes	domicilio	1	

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tipo_</b>
municipios	objeto	mun_tlx	objeto	1	H
municipios	objeto	muni_pue	num_objeto	2	H
municipios	clave_peano	mun_tlx	llave	1	H
municipios	clave_peano	muni_pue	clave_peano	2	H

<b>id_componente</b>	<b>tabla_local</b>	<b>cve_primaria</b>
1	municipios	id_municipio
2	muni	id_mpio

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tip</b>
universidad	id_municipio	universidad	id_municipio	1	H
universidad	id_municipio	universidad	id_mpio	2	H
universidad	nom_muni	universidad	nombre	1	H
universidad	nom_muni	universidad	nom_unii	2	H
universidad	domicilio	universidad	domicilio	1	H
universidad	domicilio	universidad	calle	2	H:(
universidad	domicilio	universidad	localidad	2	H:(

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tip</b>
universidad	id_municipio	universidad	id_municipio	1	H
universidad	id_municipio	universidad	id_mpio	2	H
universidad	nom_muni	universidad	nombre	1	H
universidad	nom_muni	universidad	nom_muni	2	H
universidad	domicilio	universidad	domicilio	1	H
universidad	domicilio	universidad	calle	2	H:(

universidad	domicilio	universidad	localidad	2	H:(
universidad	tel_uni	universidad	tel_uni	1	H
universidad	tel_uni	universidad	tel_uni	2	H:f

<b>tabla_global</b>	<b>atributo_g</b>	<b>tabla_local</b>	<b>atributo_l</b>	<b>id_componente</b>	<b>tip</b>
universidad	id_municipio	universidad	id_municipio	1	H
universidad	id_municipio	universidad	id_mpio	2	H
universidad	nom_muni	universidad	nombre	1	H
universidad	nom_muni	universidad	nom_muni	2	H
universidad	domicilio	universidad	domicilio	1	H
universidad	domicilio	universidad	calle	2	H:(
universidad	domicilio	universidad	localidad	2	H:(
universidad	tel_uni	universidad	tel_uni	1	H
universidad	tel_uni	universidad	tel_uni	2	H:f
universidad	calif_prom	universidad	calif_prom	1	H
universidad	calif_prom	universidad	promedio	2	H:l



id_componente	tabla_local	atributo_l	expresión
1	universidad	promedio	promedio*10

tabla_global	atributo_g	tabla_local	atributo_l	id_componente	tipo_
municipios	objeto	mun_tlx	objeto	1	H
municipios	objeto	muni_pue	num_objeto	2	H
municipios	clave_peano	mun_tlx	llave	1	H:NP
municipios	clave_peano	muni_pue	clave_peano	2	H:NP

id_componente	tabla_local	atributo_l	expresión
1	muni_tlx	llave	4
1	muni_pue	clave_peano	4

Alvarez Carrión, G. 1999. **Integración de esquemas en bases de datos heterogéneas fuertemente acopladas**. Tesis Maestría. Ciencias con Especialidad en Ingeniería en Sistemas Computacionales. Departamento de Ingeniería en Sistemas Computacionales, Escuela de Ingeniería, Universidad de las Américas Puebla. Mayo. Derechos Reservados © 1999.