

UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
MAGISTER EN CIENCIAS DE LA COMPUTACIÓN

INFORME DE TESIS
MEDIDAS DE INCONSISTENCIA DE UNA BASE DE DATOS ESPACIAL



ALUMNA : JAZNA MEZA HIDALGO
PROFESOR PATROCINANTE : ANDREA RODRÍGUEZ

2011

INTRODUCCIÓN

La inconsistencia en una base de datos espacial ocurre con mayor frecuencia cuando las restricciones de integridad no forman parte de la definición del dominio de los datos definidos en el sistema administrador de la base de datos y se dejan en manos del conocimiento/experiencia del diseñador de la base de datos. Este es el caso de las restricciones de integridad topo-semánticas (equivalentes a las restricciones semánticas en el contexto relacional). Una restricción topo-semántica impone restricciones topológicas relacionadas con las consideraciones acerca de la semántica de los objetos geográficos. Por ejemplo, las siguientes sentencias muestran ejemplos clásicos de este tipo de restricciones:

- *dos parcelas no pueden estar superpuestas,*
- *una casa debe estar contenida dentro de una parcela.*

Las bases de datos espaciales presentan nuevos desafíos para el manejo de la consistencia debido a la naturaleza compleja de los atributos espaciales. Hasta este momento, la mayoría de los trabajos se relacionan con la definición de restricciones geométricas [3], modelos de consistencia para la representación a múltiples escalas de un mismo objeto [4] y la especificación de restricciones topológicas.

Existen diversas razones que dan origen a la inconsistencia de una base de datos. En muchos casos la inconsistencia en una base de datos se origina a partir de la naturaleza imprecisa de los datos espaciales, lo que genera diversas observaciones para un mismo fenómeno espacial, como por ejemplo, el límite de un lago varía y puede ser impreciso. Además, existen aquellos casos en que los datos se obtienen de distintas fuentes y se tiene que la instancia global de la base de datos es inconsistente respecto de un conjunto de restricciones de integridad, aunque los datos puedan ser consistentes en la instancia de la fuente de origen.

Asumiendo que la inconsistencia puede ocurrir, la calidad de los datos de una base de datos cobra relevancia, llegando a ser clave para determinar el grado de satisfacción de una instancia de base de datos respecto a un conjunto definido de restricciones de

integridad.

En este trabajo se aborda el tema de definir medidas para establecer el grado de inconsistencia de una base de datos espacial; grado que va más allá de la simple dicotomía de consistente versus inconsistente. Este grado de inconsistencia se va a medir tomando como base un conjunto de restricciones topológicas o topo-semánticas previamente definidas sobre una instancia de una base de datos. En base a nuestro conocimiento actual, no existe ningún otro trabajo en la literatura que aborde la evaluación del grado de inconsistencia de una base de datos espacial.

Este trabajo consiste en estudiar la cuantificación de la inconsistencia de una base de datos respecto de un tipo particular de restricciones topo-semánticas llamadas restricciones de dependencia topológica [3]. Se propone la definición de medidas cuantitativas acerca del grado de satisfacción de dos geometrías respecto de una relación topológica representada a través de una restricción de dependencia topológica. Se ha definido una medida de consistencia para cada conflicto entre una relación topológica T_1 y una relación topológica esperada T_2 , donde $T_1 \neq T_2$. Mientras T_2 es especificada por una restricción de dependencia topológica, T_1 es determinada a partir de dos geometrías existentes en la base de datos. La evaluación experimental compara los valores de las medidas definidas en bases de datos reales y sintéticas a dos medidas de base: (1) distancia semántica de las relaciones topológicas en un grafo de vecindad [5] [7] y (2) distancia entre los puntos del borde de las geometrías en una base de datos inconsistente versus una base de datos consistente.

1.1. Hipótesis

Existen distintos grados de consistencia de una base de datos espacial respecto a un conjunto de restricciones de integridad topo-semántica que dependen de una comparación o valoración de relaciones topológicas.

1.2. Objetivo General

Definir medidas para determinar el nivel de inconsistencia de una base de datos espacial con respecto a un conjunto de restricciones topo-semánticas.

1.3. Objetivos Específicos

- Estudiar restricciones de integridad de dependencia topológica definido en [3].
- Definir medidas de inconsistencia según la relación topológica especificada por una dependencia topológica.
- Definir algoritmos de cómputo de las medidas.