Spatial Validation
Academic References

Refractions RESEARCH INC.

Submitted To: Program Manager
GeoConnections
Victoria, BC, Canada

Submitted By: Jody Garnett
Brent Owens
Suite 400, 1207 Douglas Street
Victoria, BC, V8W-2E7
jgarnett@refractions.net
Phone: (250) 885-0632
Fax: (250) 383-2140

Table of Contents

1 INTRODUCTION

This document describes our findings in the academic literature on spatial validation. We undertook this research to discover the breadth of this topic, and references to existing tools.

Of interest:

- Classification of spatial validation:
  - Attribute Accuracy:
    Validates the measurements with respect to the real world
  - Completeness:
    Validates a geospatial database with respect to an application
  - Logical Consistency:
    Validates a geospatial database with respect to spatial reasoning
- Qualitative Spatial Reasoning
  - Spatial Relationships:
    Allows for the specification of relationships and constraints
  - Inference Engines:
    Allow for the specification of constraints and for their propagation.

Academic research has not provided us with many references to existing tools. For more information on existing tools refer to our document on online research.

2 VALIDATING WEB FEATURE SERVER REQUIREMENTS

The Validating Web Feature Server is an extension to the existing GeoServer reference implementation of the Open GIS Consortium's Web Feature Server Specification.

The extension consists of the addition of a Validation Processor to the Transaction operation.

The Validation Processor makes use of a series of Plug-ins to refuse operations that would otherwise leave the geospatial database in an inconsistent state.

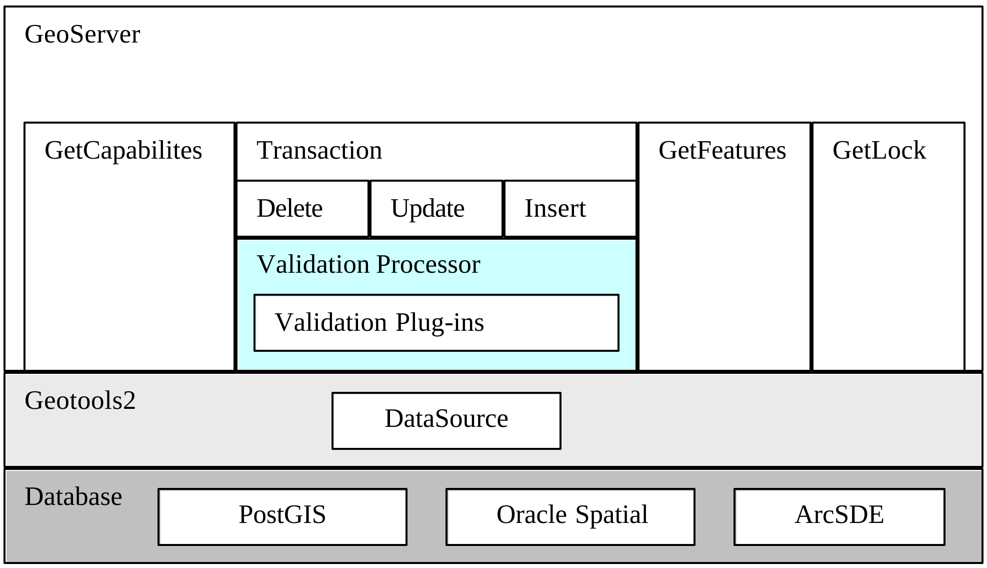

Figure 1 - VWFS Layer Diagram

The geotools2 library is used to provide Database access, high-level geospatial query operations and a programmatic interface to Features.

The VWFS requirements include:

- the specification of attribute constraints
- the specification of topology constraints
- verification of geospatial database consistency

The focus of this document is on science targeting the spatial nature of this problem. For database integrity issues, such as consistency across updates and delete operations, we will be using traditional database techniques.

3 CLASSIFICATION OF SPATIAL VALIDATION ISSUES

Issues in spatial validation can be broken down into several categories: Attribute Accuracy, Completeness and Logical Consistency.

3.1 Attribute Accuracy

Validating attribute accuracy of a geospatial database ensures the accuracy in which it's measurements reflects the real world.

Validation issues with attribute accuracy:

- Modeling data confidences over time
- Providing an interpolation between two temporal measurements
- Providing an interpolation between grid cells
- Using statistical measures of a dataset to identify outliners

To handle these issues the pedigree of a dataset can be tracked.

Metadata can be used to:

- Record initial sampling error:
  Measure of collection technique
- Record confidences for the dataset:
  A statistical model of accuracy, often with respect to time
- Record history for a dataset:
  Changes in accuracy over time
- Record accumulated error:
  Changes in accuracy over computation

Accuracy can be increased by intelligent use of other datasets. An example is the process of conflation or snapping, in which information is referenced against a more accurate or detailed dataset. As with interpolation, the data from this process will need a sophisticated confidence model.

A Quality Assurance program, complete with field spot checks is the only real way to verify how accurate a dataset is.

3.2 Completeness

Validating the completeness of a geospatial database ensures that it is capable of meeting an intended use.

As an example a geospatial database for a road atlas is complete when it has all the roads required for publication.

```
Completeness tests for a road atlas:
Ensure that there are no gaps in routes
Ensure there are no holes in coverage information
Issue with Geospatial completeness:
Perform logical tests of the spatial and attribute data
Perform statistical tests of the spatial and attribute data, to identify
outliners
Summary data can be compiled and checked to be within known tolerances.
Spatial data can be automatically snapped correct for gaps and undershoots.
```

3.3 Logical Consistency

Validating that a geospatial database is logically consistent ensuring that it is complete for the purposes of logical and spatial reasoning.

Issues with Logical Consistency:

- Ensuring internal consistency:
  Maintain required dependencies between attributes in a feature
- Ensuring logical consistency:
  Maintain required relationships between features or feature types
- Limiting attributes to a specific range of valid values
- Requiring well formed spatial data

Specific applications will provide their own requirements for Spatial Reasoning.

```
Sample spatial inference requirements:
•  Lakes form closed polygons
•  One label for each feature
•  No duplicate arcs
•  No gaps in rivers or roads
•  An intersection relationship between roads and bridges and rivers
•  A network built from a river connection relationship be validated with
   respect to a height from a digital elevation map.
```

Logical and Spatial Reasoning are both very demanding applications that become brittle and indeterminate when used with inconsistent data. When interpolated values it is difficult to ensure the result is analytically stable.

4 GEOMETRIC ALGEBRA

Geometric Algebra is a different take on the mathematical problem of geometry.

The traditional Linear Algebra approach is popular, often offers hardware support and is widely understood. Geometric Algebra is a competitor, with several mathematical advantages.

Advantages of Geometric Algebra:

- Consistency of geometric operations, no special cases for points, lines, planes or time
- Scalability to higher order vector spaces.

Geometric Algebra achieves this success through two means.

- A unified approach to handling spatial constructs
- Use of a number of high value geometric operations

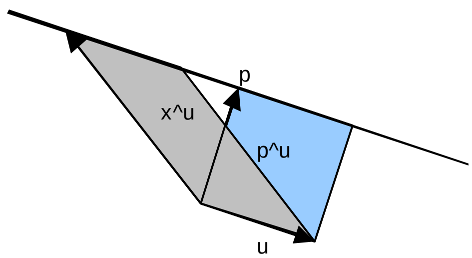

Figure 2 - Geometric Algebra Representation of a Line

Geometric Algebra is not yet widely supported in computer science. The advantages of working with an existing, tested, and public code base outweigh the benefits in simplification afforded by the new math.

5 GEOSTATISTICAL ANALYSIS

Statistical analysis with the geospatial data is currently an area of research, and opportunity for tool development.

Statistical analysis is used for producing quality assessments. The techniques are well suited to producing quality metrics for geospatial metadata.

For our purposes these statistical measures can be used to provide a reference point for identifying outliers.

6 QUALITATIVE SPATIAL REASONING

Artificial Intelligence research into Qualitative Reasoning provides a basis for constraint propagation, and constraint based logic systems. The application of Qualitative Reasoning to the spatial domain has been the focus of recent AI work into the guiding of robots.

Constraint propagation is of particular interest to the construction of a Validating Web Feature Server. Direct applications involve working with constraints on networks of roads or rivers and maintaining the stability of the network across modifications.

6.1 Qualitative Reasoning

Qualitative Reasoning is the practice of reasoning with qualities rather than exact numbers.

An example of Qualitative Reasoning is the use of predicate calculus to provide inference methods for objects and their relationships.

Example using Unification as an inference technique:

```
Given the following predicate assumptions:
  spot.isA(dog)
  dog.can(run)
  dog.can(bark)

We test the consistency of the following predicate:
  Spot.can(run)?

By using predicate unification of spot and dog we have:
  spot.isA(dog)
  dog.can(run)
  dog.can(bark)
  spot.can(run)
  spot.can(bark)

Since we now have the predicate spot.can(run) in our set know that
spot.can(run) is consistent.
```

By reducing the quantitative topological information to a series of qualitative spatial relationships we can make use of predicate-based inference.

The example of unification and predicates calculus is just one of many qualitative reasoning techniques available.

7 GEOSPATIAL RELATIONSHIPS

Describing spatial relationships between geographical objects is central to the success of Qualitative Spatial Reasoning. What follows are two techniques for establishing relationships between features.

7.1 Region Connection Calculus

The approach is based on the idea of primitive connection relationship between regions. This primitive is then used to describe other geospatial relationships.

C(x,y)	x and y are connected
~C(x,y)	x is disconnected from y
For all z[C(z,x) → C(z,y)]	x is part of y

This system does not handle relationships with points and lines and may only be applicable to a few domains (such as watersheads).

7.2 Point Set Theory

Point Set Theory includes concepts of lines and points in addition to Regions. These constructs are defined as the set of points contained in their Interior, Boundary and Exterior.

	Interior	Boundary	Exterior
A Region h	h^0	δh	^-h
A Point p	p^0	n/a	^-p
A line l	l^0	end points	^-l

Relationships between Regions are described as a matrix produced by comparing the intersection of the Interior, Boundary and Exterior properties of both regions. This comparison referred to as the Dimensionally Extended 9-Intersection Matrix or DE-9IM.

	Interior	Boundary	Exterior	
Interior	$x^0 \cap y^0$	$x^0 \cap \delta y$	$x^0 \cap {}^-y$	
Boundary	$\delta x \cap y^0$	$\delta x \cap \delta y$	$\delta x \cap {}^-y$	$\leftrightarrow x \cap y$
Exterior	$^-x \cap y^0$	$^-x \cap \delta y$	$^-x \cap {}^-y$	

Where x and y are a region, point or line defined by their Interior, Boundary and Exterior.

This adopted has been adopted by the OCG and will be used by the Validating Web Feature Server.

7.2.1 Dimensionally Extended 9-Intersection Matrix Example

Please consider the following comparison of two overlapping polygons x and y.

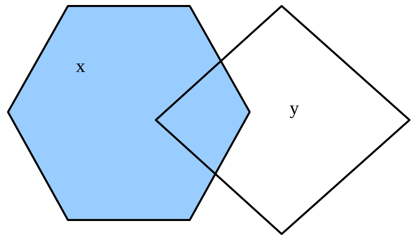

Figure 3 - Overlap Relationship

The above relationship between x and y can be described with the following DE-9IM in which only the dimensionality of the resulting set is considered.

	Interior	Boundary	Exterior	
Interior	2	1	2	
Boundary	1	0	1	$\leftrightarrow$ dim(x $\cap$ y)
Exterior	2	1	2	

This relationship can be represented as the string “212101212”.

7.2.2 Geospatial Relationships

Geospatial relationships can be described in terms of a mask, or pattern, which may be matched by the DE-9IM.

Consider the following definition of Area/Area overlap.

	Interior	Boundary	Exterior	
Interior	T	*	T	
Boundary	*	*	*	$\leftrightarrow (x^0 \cap y^0 \neq \emptyset) \wedge (x^0 \cap {}^-y \neq \emptyset) \wedge ({}^-x \cap y^0 \neq \emptyset)$
Exterior	T	*	*	

This is can be represented as the string “T*T***T**” where:

- T: value is not equal to zero
- F: value is equal to zero
- *: Don’t care what the value is
- 0: value is exactly zero
- 1: value is exactly one
- 2: value is exactly two

Using this technique the geospatial relationships can be defined as follows:

	Relationship	Pattern(s)	Limitation
1	x.Disjoint(y)	FF*FF****	
2	x.Touches(y)	FT*******	Area/Area, Line/Line, Line/Area, Point/Area
		F**T*****	Not Point/Point
		F***T****	
3	x.Crosses(y)	T*T******	Point/Line, Point/Area, Line/Area
	x.Crosses(y)	0********	Line/Line
4	x.Within(y)	TF*F*****	
5	x.Overlaps(y)	T*T***T**	Point/Point, Area/Area
	x.Overlaps(y)	1*T***T**	Line/Line

8 REFERENCES

Proposed Grid Timestamp Standard,
http://www-didc.lbl.gov/GGF-PERF/GMA-WG/papers/GWD-GP-4-1.pdf

Tiers of Ontology,
http://citeseer.nj.nec.com/frank98tiers.html

Product quality assurance for GIS life-cycle,
http://www.gisdevelopment.net/technology/gis/techgi0042c.htm

GIS Data Reviewer,
http://www.esri.com/library/fliers/pdfs/gis-data-reviewer.pdf

L. Dorst, Honing geometric algebra for its use in the computer sciences,
http://citeseer.nj.nec.com/dorst01honing.html

Krivoruchko, Bridging the Gap Between GIS and Solid Spatial Statistics,
www.dpi.inpe.br/gilberto/csiss/papers/krivoruchko.pdf

OCG, OpenGIS Simple Features Specification
http://www.opengis.org/techno/specs/99-050.pdf

Qualitative Spatial Representation and Reasoning: An Overview (2001),
http://citeseer.nj.nec.com/cohn01qualitative.html