Report on MSRM Tools and Required Feature Set

Submitted To: Program Manager
GeoConnections
Victoria, BC, Canada

Submitted By: Jody Garnett
Brent Owens
Refractions Research Inc.
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

Upon researching various tools used by the Ministry of Sustainable Resource Management (MSRM) we came across some relevant validation software. One such program is called DRAQA (Digital Road Atlas Quality Assurance). The DRAQA tool is a Java application that performs various consistency checks on a road data set, specifically the DRA data set. We can use the tests that the software performs on the DRA data for our data consistency checks. Although some tests are far too specific for our needs, many can be generalized or used directly for various data sets. We took a look at the tests it performs and extracted the ones that would be useful to us.

Another quality assurance tool that was mentioned was FME (Feature Manipulation Engine). FME provides some functions capable of testing and fixing data. Some of which will be quite useful to us.

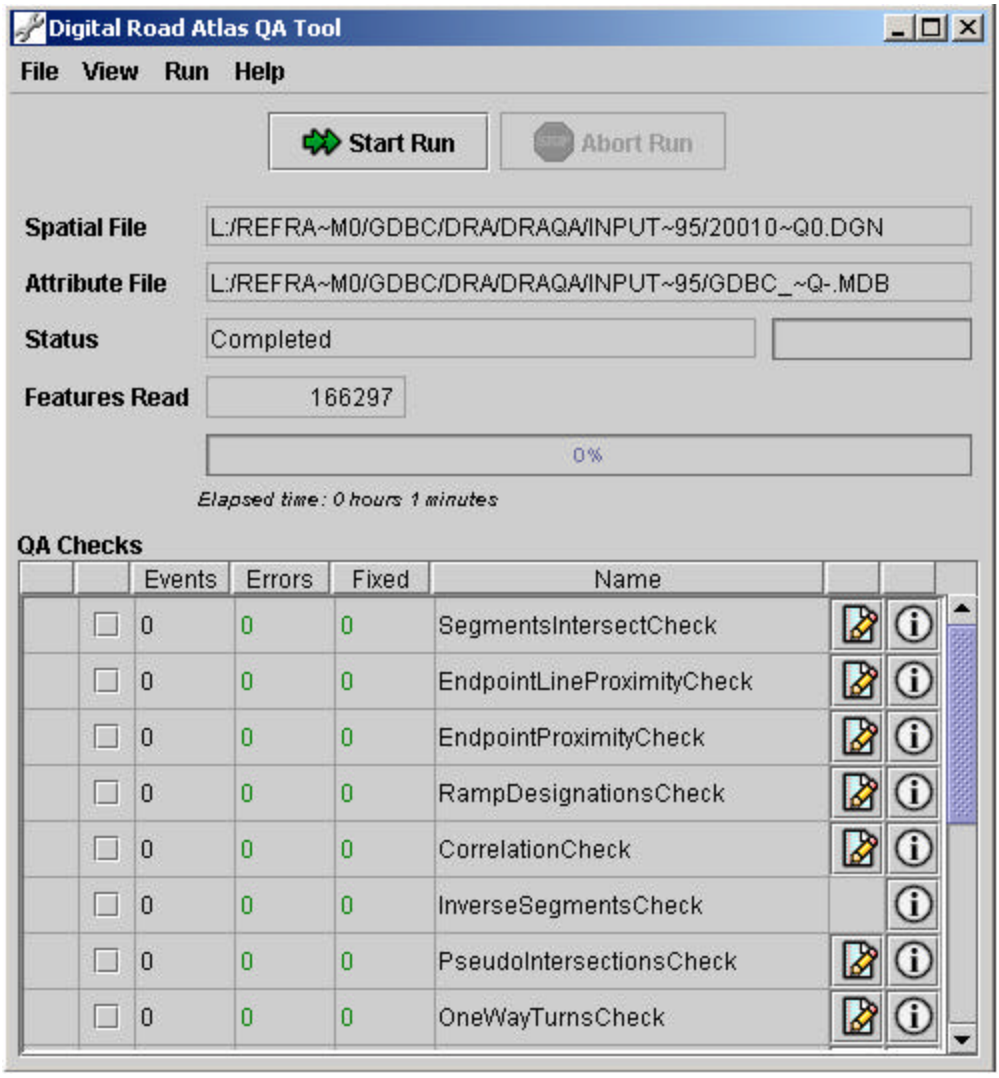
Further, this paper discusses GeoConnection's use of GML for useful validation techniques such as a gazetteer.

2 DRA QUALITY ASSURANCE (DRAQA)

2.1 Feature List

The DRA Quality Assurance program is used to validate DRA road data to make sure it conforms to specified rules. DRAQA has a list of tests that are performed against the supplied road data. The program tells the user whether the data was valid or not. The list that follows contains the various consistency tests that DRAQA performs. Each test is accompanied by a short description and whether or not it can be generalized to be performed on different data sets for our needs.

Test Name	Functionality	Generalizable
AddressGapCheck	This test checks if adjoining segments of a road have a gap in their address ranges.	**yes**
AddressOverlapCheck	This test checks if the address ranges of adjoining segments of a road overlap.	**yes**
AngleSizeCheck	This test checks if two segments of a road meet at an unexpectedly sharp angle.	**yes**
AttributeSpecificationCheck	This test checks if a value of a table is not on the specified list of allowed values.	**yes**
CorrelationCheck	This test verifies that there are no disparities between two attributes.	**yes**
FeatureIDCheck	This test checks to see if two segments of a road have different feature ID's.	**yes**
InverseSegmentsCheck	This test checks to see if any segments are inversely oriented.	no
NameChangeCheck	This test checks to see if a road changes its name unexpectedly.	no
OneWayTurnsCheck	This test checks to see if the indicated turn restriction allows turning the wrong way onto a one-way street.	no
PseudoIntersectionsCheck	This test checks that no turns are allowed on a pseudointersection.	no
RampDesignationCheck	This test checks to make sure that an on-ramp connects from a lower class to a higher class, and that an off-ramp connects from a higher class to a lower class.	no
RoadchangesCheck	This test checks to see if two adjoining segments of a road have different road classes.	no
StarNodeCheck	This test checks to see if three or more road segments with the same name meet at a point.	**yes**



2.2 Potential Generalization Tests

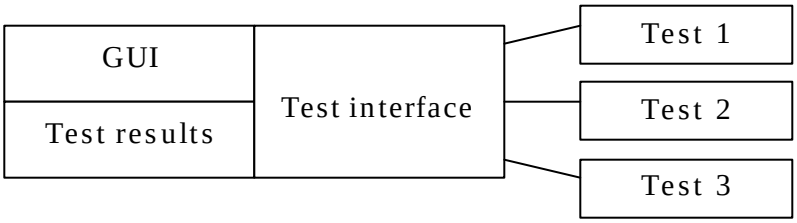
Many of the tests that are present in DRAQA can be extended and generalized for use on different data sets, not just road structures. The tests listed below will be the most useful for generalized spatial consistency validation.

2.3 Validation of Generalized DRAQA Tests

We can use these generalized DRAQA tests and compare their results against the actual tests in DRAQA, using the same test data. This will show that we can use the generalized tests to perform specific checks on data that is required by other applications.

2.4 Architectural Review

DRAQA is a modular design that allows for expansion of tests. The application consists of a GUI, a list of available tests that can be used, and output from the tests. The GUI doesn't have to know anything about the tests or the data that it is testing. It is simply passed information, defined by a specification, that it knows how to display.



3 FME

3.1 Overview

The Feature Manipulation Engine (FME) provides tools to transform and run tests on data. Although most of the functions it provides are straight transforms, tests can be applied to the transforms to build valuable validation tools. There are two types of tools that can be used for validation and transformations, they are functions and factories. The difference between the two is not relevant and beyond the scope of the document. Therefore they will both be referred to as just functions.

3.2 Validation Functions

There are several useful FME functions that can be used alone or in conjunction with each other to perform validation and consistency checks. More information and documentation about FME can be found on their web site at www.safe.com.

Test Name	Functionality
AttributeClassifier	Tests if the contents of the source attribute are entirely of a particular character classification.
AttributeFilter	Routes features to different output ports depending on the value of an attribute.
ChangeDetector	Detects changes between two sets of input features.
CommonSegmentFinder	Tests to see which of the candidate features have even one segment in common with any base feature.
DuplicateRemover	Detects duplicate features based on the value of a key attribute.
EnvelopeFilter	Determines whether or not features intersect the bounding box of some envelope feature.
GeometryFilter	Routes a feature based on its geometry type.
TCLCaller	Runs a Tool Command Language (Tcl) command and assigns its return value to an attribute.

4 GEOCONNECTIONS USE OF GML

4.1 Postal Code Lookup Service

This tool provides point geometry for Postal Codes by 3-digit Forward Sortation Areas (FSA) in GML 3.0.0.

GetCapabilities

```
http://geoservices.cgdi.ca/cgi-
bin/postalcode/postalcode.cgi?version=1.0.0&request=GetCapabilities
```

GetPostalCode

```
http://geoservices.cgdi.ca/cgi-
bin/postalcode/postalcode.cgi?version=1.0.0&requet=GetPostalCode&sortarea=fsa&
code=V3W
```

4.2 Gazetteer Service

The gazetteer service: provides complex geometry for toponymic information in GML 2.1.2; provides virtual linking of placenames and features via WFS; combines spatial and aspatial searching; and provides alternative service to now operational Canadian Geographical Names Service (CGNS) WFS (authoritative national view of placenames in Canada).

GetCapabilities

```
http:// cgdi-dev.geoconnections.org/cgi-
bin/prototypes/cgdigaz/cgdigaz.cgi?version=1.0&request= GetCapabilities
```

GetPlacenameGeometry

```
http://cgdi-dev.geoconnections.org/cgi-
bin/prototypes/cgdigaz/cgdigaz.cgi?version=1.0&request=GetPlacenameGeometry&wi
ldcards=true&geomtype=bbox&placename=vancouver&provterr=59
```

5 ARC/INFO ARC MACRO LANGUAGE (AML)

Arc Macro Language files are used to run Arc/Info commands. These can be used to perform validation on data and have been used in the Corporate Watershed Base project for data quality assurance. Several of these tests a similar to what we will want to implement, but on a more generalized level. The following are a list of tests that are performed on the watershed data:

- Distinguish stream “main flows” from “secondary flows”.
- Consistent flow direction of streams.
- Stream connectivity.
- Closed polygonal figures
- Connect all stream features to the network.
- Connect all large lake/marsh/swamp features to the network.
- Corrected incorrectly coded TRIM features.

6 REFERENCES

“GML Usage in the CGDI”: from GML Dev Days conference.

MSRM: http://srmwww.gov.bc.ca/gis/