8-1-2006

GeoExpert - An Expert System Based Framework for Data Quality in Spatial Databases

Aditya Kumar
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses

 Part of the Databases and Information Systems Commons

GEOEXPERT – AN EXPERT SYSTEM BASED FRAMEWORK FOR DATA
QUALITY IN SPATIAL DATABASES

A Thesis
Presented to
The Faculty of the Department of Computer Science
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Computer Science

By
Tadakaluru Venkat Raja Aditya Kumar

August 2006.

GEOEXPERT – AN EXPERT SYSTEM BASED FRAMEWORK FOR DATA QUALITY IN SPATIAL DATABASES

Date Recommended _July 12-2006_

Mostafa Mostafa

Director of Thesis

Richard G Bowker _8/1/2006_

Dean, Graduate Studies and Research   Date

Acknowledgement

First I thank my advisor Dr. Mostafa Mostafa, who helped me keep my morale up throughout the preparation of the thesis. Thank you for your excellent emotional and technical guidance throughout this thesis.

I wish to thank Dr. Andrew Ernest for giving me a job and chance to do this thesis at a time when I hit rock bottom during my student-life. Thank you for believing in me and my abilities.

I wish to thank Ms. Karla Andrew for pointing me in the right direction at all times when I got lost in this project. You have been ever so helpful. You are always my favorite boss.

Thank you, Dr. Ahmed Emam, for serving on my advisory committee. You, too, have been instrumental in the development of this thesis.

Last but not least, I wish to thank my parents, Mr. T. SivaKumar and Mrs.T.Rajasri, for being my pillars of support. I could not have made it without your faith in me. Thank you.

Table of Contents

List of Figures

GEOEXPERT – AN EXPERT SYSTEM BASED FRAMEWORK FOR DATA

QUALITY IN SPATIAL DATABASES

Aditya Kumar Tadakaluru Venkat Raja   July 12, 2005.   79 Pages

Directed by: Dr. Mostafa Mostafa

Department of Computer Science   Western Kentucky University

Usage of very large sets of historical spatial data in knowledge discovery process became a common trend, and in order to obtain better results from this knowledge discovery process the data should be of high quality. In this thesis we proposed a framework 'GeoExpert' for data quality assessment and cleansing tool for spatial data that integrates the spatial data visualization and analysis capabilities of the ARCGIS, the reason and inference capability of an expert system.

In this thesis we implemented the proposed framework both stand-alone and web versions using ArcGIS Engine and ArcGIS Server, respectively. We used JESS expert system shell for the expert system part of the GeoExpert. Use of expert system shell separates the application logic from the actual framework which makes the framework easily updatable and domain independent. In this thesis we implemented the GeoExpert on the spatially referenced water quality data.

Chapter 1: Introduction

Spatial data is one of the outcomes of the information revolution. Many commercial and research areas started looking at the data in more than one dimension. Looking at the data in relation to an underlying frame of reference is spatial data. Most popular form of spatial data is geographical data where earth is the underlying frame of reference. An example of other potential underlying frame of reference is an IC chip.

Spatial data mining became an answer for this tremendous production and usage of geographical data. Spatial databases due to its multidimensional structure will have crucial information hiding in it which is far away from what we can extract in traditional operational databases. Finding hidden patterns, mining association rules, and carrying other spatial data mining operations results in accurate results, only if the data that is being used is error free. Data quality is one of the major issues that need to be addressed before going to any knowledge discovery processes.

In this thesis we proposed and implemented a framework called 'GeoExpert – A framework for data quality in spatial databases' which addresses the data quality issue in spatially referenced data. In order to assess the data quality of spatial data, the framework has to have the capability of analyzing the spatial data such as performing spatial queries, spatial data visualization.

In this framework we dealt mainly with geographical data, but its robust and generalized architecture makes it capable of dealing with any kind of spatial data. We introduced ArcGIS Engine and ArcGIS Server into our framework to make it capable of analyzing spatial data.

Data quality is often compromised by calibration errors, threshold errors, missing data, extra data, missing records, garbling errors, duplications. The ways to deal with these error types are often different in different domains and depend upon that particular circumstance. For example, the dataset may be an anomaly in one geographical domain where as at the same time a correct dataset in other domain. If one uses traditional approach of developing programs that implement the relevant geographical domain expertise into application logic, a framework has to be developed for each geographical domain. Even if all the geographical domain expertise is incorporated in one single program, it will not be extensible when it comes to updating the logic or modification of the program. For this reason a rule based expert system is used in our framework. A rule based expert system separates the application logic from the actual framework, which makes the framework more generalized and easily updatable.

Its not only the erroneous data that can be assessed and cleansed with this framework, both spatial and non-spatial outliers in the data can be detected based on the rules that are fed to the expert system. An example for a spatial outlier is a departmental store in a residential area.

In this thesis the spatial data visualization and analysis capabilities of ArcGIS Server and ArcGIS Engine are integrated with the reason and inference capabilities of an expert system shell for the data assessment and cleaning of spatial data. Use of ArcGIS provides a rich user interface for analyzing spatial data. Our proposed framework GeoExpert is implemented both as a stand-alone version and also on-line.

Integrating the ArcGIS framework with the expert system shell resulting in a framework for assessing the spatial data quality is the major contribution in this thesis. GeoExpert is applied to spatially referenced water quality data which is another major contribution in this thesis.

This thesis is comprised of five chapters. In chapter 2 a detailed introduction to Geographical Information Systems, Expert Systems, and Data Quality is given and also the previous work with similar ideas and their limitations are explained. Chapter 3 focuses on the design criteria and framework architechture. Chapter 4 emphasizes the implementation of stand-alone and on-line versions of our proposed framework, GeoExpert. The thesis is concluded in the chapter 5.

Chapter 2: Background

2.1 Geographical Information Systems

Geographical Information Systems (GIS) is the key technology actuating the spatial database management system [2]. GIS not only includes the computer hardware, software but also geographical data for capturing, analyzing, managing and visualization of geographically referenced data [1]. The underlying frame of reference for geographical data is earth's surface. So each geographical dataset is viewed spatially in terms of latitude and longitude. In GIS the non-spatial data can be linked to a location like a particular combination of latitude and longitude e.g. street address, and store it as a separate layer. This new layer can be analyzed in terms of the location. An example of a spatial query will be "List all the auto servicing shops within fifty miles of the automobile parts shop." The following example gives an idea of how GIS stores and manages the data. Figure 1 is a screen capture of the geographic data view of National Hydrography Dataset (NHD) [25] stream coverage in the ArcMap. ArcMap is an ArcGIS desktop application for map analysis, editing and cartography [3]. In figure 1 there are two layers on the left most side. Greens100 is the layer that stores the NHD Stream data in the green county. Sampling point events is the layer that contains the data pertaining to water sampling points. Both these layers are associated to the earth's latitude and longitude.
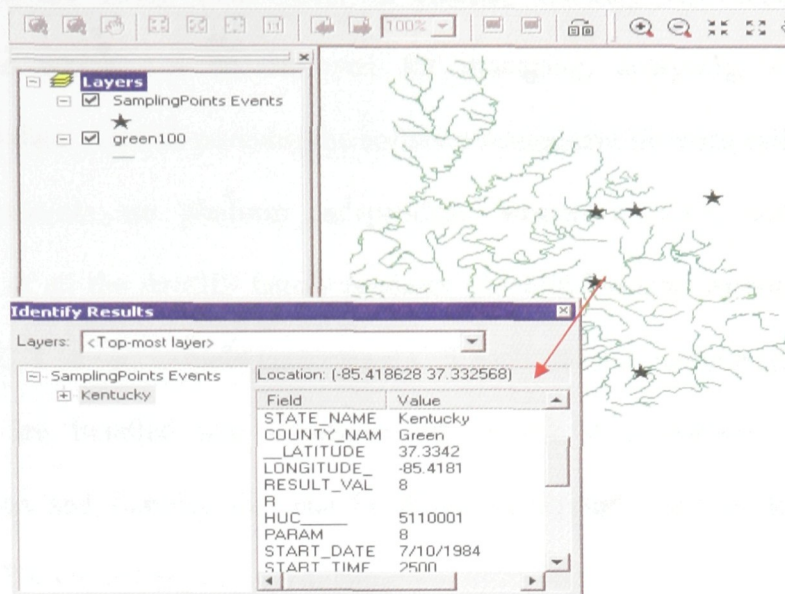
Figure 1: NHD Stream Coverage of Green Country, KY, USA

The actual data that is stored in one of the sampling points is shown above in the bottom left corner pointed by a red arrow. This is an example for the spatial data visualization capability of ESRI ArcGIS.
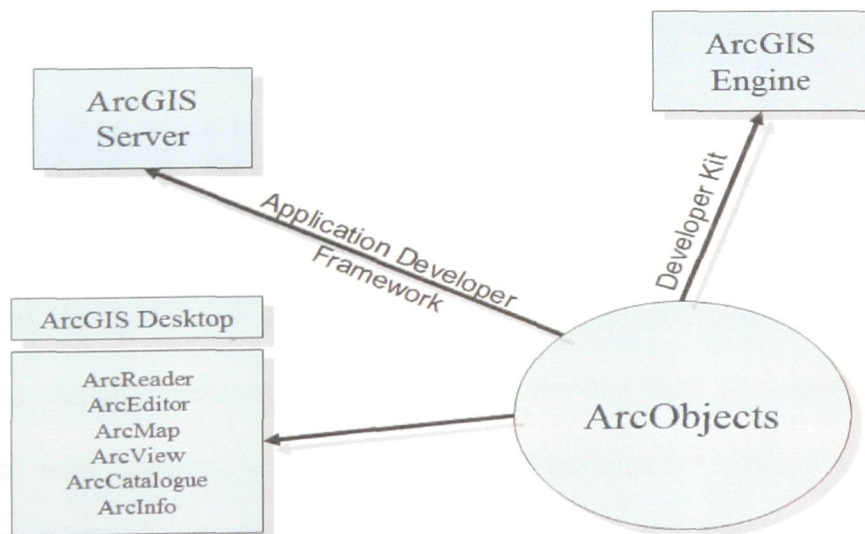
Figure 2: ESRI's ArcGIS products family [4]

There are many other tools in ArcGIS Desktop Arc Reader, ArcView, ArcCatalogue which can be deployed for managing, analyzing, visualization of geographical data. ArcGIS provides the software component libraries called ArcObjects. Those components are platform independent, written in C++, and the building components of all the ArcGIS family products (ArcGIS Desktop Applications, ArcGIS Engine, ArcGIS server) to build customized GIS application or functionality. Appropriate ArcObjects are bundled into developer kit, which is a software program with documentation and samples that can be deployed through ArcGIS desktop, ArcGIS engine, ArcGIS server based on the application nature [4].

ArcGIS Engine is a developer kit to implement GIS functionality in a stand-alone application. It provides libraries of the software components, ArcObjects the building components of ArcGIS desktop products. ArcObjects can be deployed in wide variety of operating systems using wide variety of software languages [6]. In the stand-alone version of GeoExpert, ArcObjects are deployed using Java programming language.

ArcGIS server provides a platform to develop enterprise GIS applications which are centrally managed and can be exploited by multiple users. It manages the geographical data as server objects which can be deployed over the web by web applications. ArcGIS framework can be used to develop both web services and web applications with advanced and customized GIS functionality. ArcGIS server can be divided into two parts: GIS Server and an Application Developer Framework (ADF) for .Net and Java.

ArcObjects are hosted on a GIS server which can be used by web and desktop applications. GIS Server includes core ArcObjects library from which the ArcGIS desktop applications like ArcMap, ArcView, and ArcInfo are built and thus have the capability of running ArcObjects in the server. Developers can use ADF to develop and deploy web applications and desktop applications in java and .Net that deploys ArcObjects running in the GIS server [7].

2.2 Expert Systems

An expert system is a computer program capable of solving complex problems by simulating the human reasoning process. It utilizes specific domain knowledge and inferences for simulating the human expertise [8].
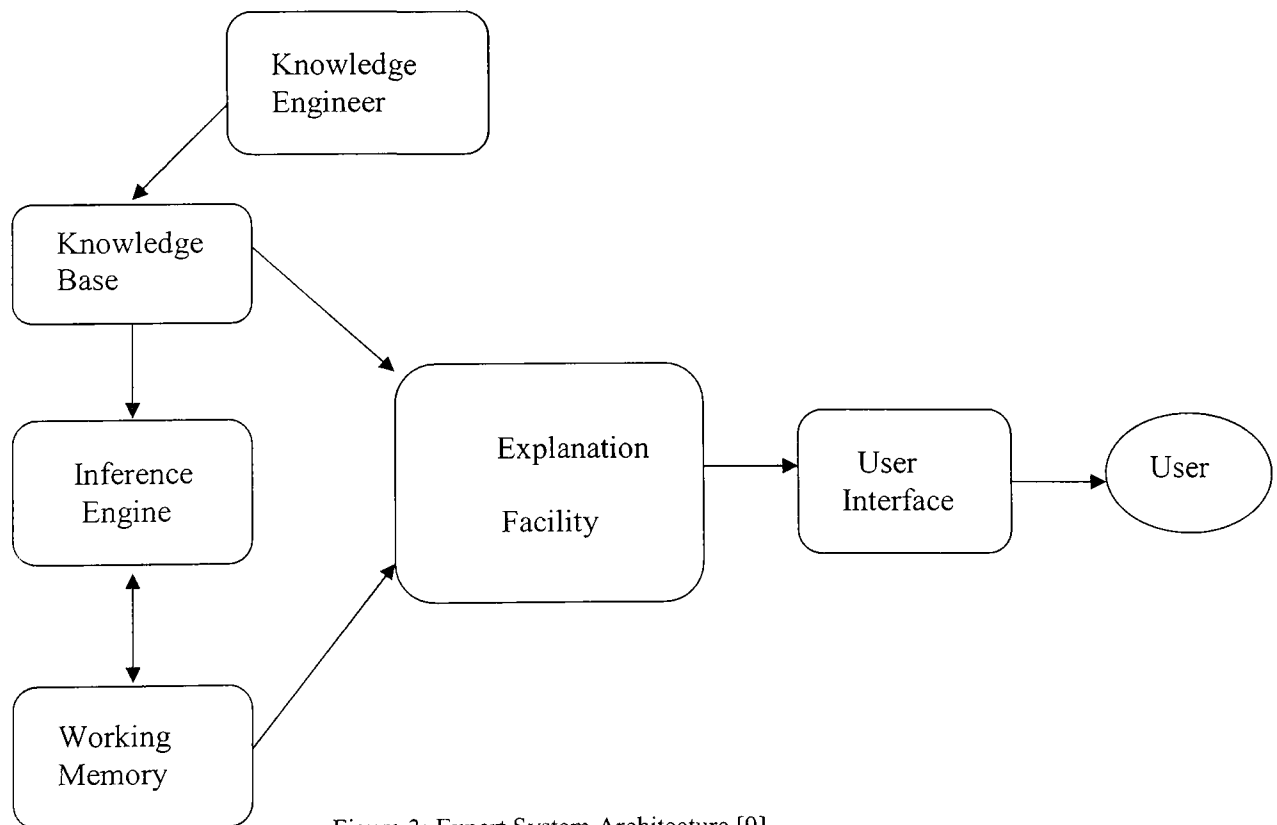
Figure 3: Expert System Architecture [9]

We can rather call an expert system as a system more than a program because it comprises different components like Knowledge base, inference engine and working memory. An expert system is a combination of four main components knowledge base, inference engine, working memory, explanation facility and GUI.

a) **Knowledge base :**The knowledge base is the place where an expert system stores the domain specific knowledge required to puzzle out a complex problem. Usually a knowledge engineer consults the experts in that particular domain and coordinates the knowledge into a format understood by the inference engine [9].

Four common methods of representing the knowledge are – predicate logic, production rules, frames (objects) and semantic network and conventional programs [9]. Proposed GeoExpert uses rule based expert system in which knowledge is represented in the knowledge base using the production rules.

A production rule is a powerful representation of knowledge with a set of antecedents and a set of consequents. The antecedent defines a set of conditions and consequent defines a set of actions [9]. The other word we can use for production rule is IF-Then Rule. An example for an If-Then rule will be **If** fecal coliform bacteria is over 200 colonies/100 ml of water sample **Then** there is a possibility of presence of pathogenic (disease causing) organisms in water [10].

b) Working memory:

The working memory stores the set of established facts for further use. The working memory gets constantly updated with feedback of the inference engine. The inference engine consists of one or more inference mechanisms that are capable of simulating the human reasoning process. There are two inference mechanisms that are used widely in expert systems development – forward chaining and backward chaining.

In forward chaining process, the inference engine searches the knowledge base for the rules that match the facts in working memory; this process will stop when either the goal is reached or there are no more rules to fire in the knowledge base [11]. Finally, forward chaining is a data driven process [9].

Backward chaining starts with a hypothesis or conclusion, and the inference engine searches the knowledge base for the rules that contains these conclusions and verifies whether the current working memory facts satisfies these rules. Finally backward chaining is a goal driven process [11].

c) Inference Engine:

The inference engine in the forward chaining as in the example below starts with a fact of passing the high school then the conclusion is whether he/she is eligible to enter bachelors program in any of the above listed majors – biology, economics, physics etc.

Bachelors

High School → Biology

Economics

Physics

Chemistry

Computer
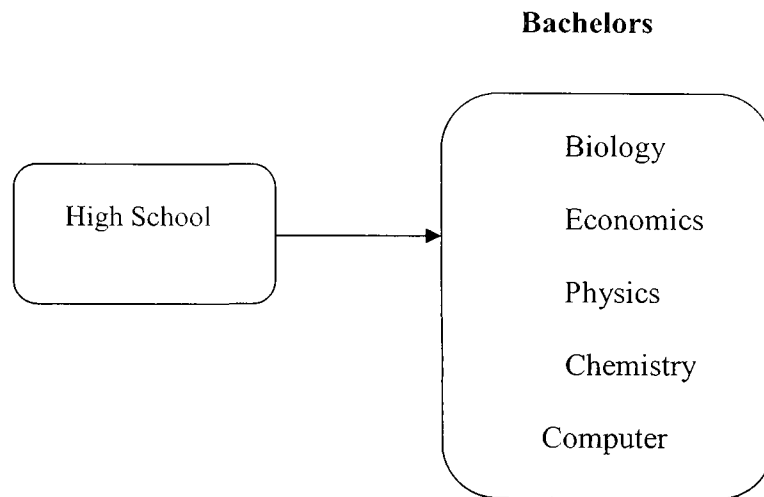
Figure 4: An example to illustrate forward chaining and backward chaining

In the backward chaining the inference engine starts with the conclusion 'enter bachelors program in any of the above listed majors" say conclusion 'X' and it checks the knowledge base to see if there is any rule with this conclusion 'X'.

The inference then verifies the facts in current working memory against the conditions of the rules with conclusion 'X' in the knowledge base [11].

d) GUI & Explanation facility

GUI & Explanation facility is the user interface which explains to the user how a particular decision is made by the expert system, what are the facts that are used, which rules are fired. An explanation facility can be described as the log file of the inference engine.

2.3 Data Quality

Data Quality is one of the important factors that affect the quality of the patterns, association rules and any other output achieved through data mining process. Poor data quality may lead to faulty results which in turn may lead to faulty decisions. The most common reasons behind poor data quality are incompleteness, noise, Inconsistency [13]. It's often difficult to assess the quality of the data and cleanse the data manually if the amount of data is very huge. The most commonly seen error types are [12]

- Calibration errors – calibration errors appear when the same additive or multiplicative error takes place in the performed measurement or observation.

- Threshold error – Threshold errors results when a particular data element exceeds the threshold value and the program is not able to notice or store that value.

- Missing Data – Missing data are the results of system malfunction or shutdown. They even occur due to human mistake of forgetting to enter a value.

- Extra data: Extra data results when two databases are merged or because of a poor database design.

- Missing records: Missing records occurs when a particular record is missing in the given sequence of the records.

STATE_NAME	COUNTY_NAM	_LATITUDE	LONGITUDE_	RESULT_VAL	R	HUC___	PARAM	START_DATE
Kentucky	Green	37.2539	-85.5031	7.5		5110001		2 6/12/1979
Kentucky	Green	37.2539	-85.5031	2		5110001		2 7/25/1979
Kentucky	Green	37.2539	-85.5031	4		5110001		2 9/25/1979
Kentucky	Green	37.2539	-85.5031	1		5110001		2 10/25/1979

Figure 5: Missing Records

In the above figure a record is missing with the START_DATE in the month of August 08, 1979.

- Garbling Errors: Garbling error occurs when a value is recorded incorrect.

- Duplication: Duplicate values result when the same value is recorded or stored more than once.

- Combination errors: Combination errors occur when a single event is recorded as two separate events or when two different events are combined into one single event [12].

Data Quality is multidimensional. For example if a customer is trying to purchase a new car where his main concern is about the price, mileage, automatic transmission, engine power, safety features. For this customer 'quality of the car' constitutes above attributes as dimensions. If there is another customer who is solely interested in price and mileage then the 'quality of the car' in his context will be decided on the dimensions price and mileage [29].

So to measure the quality of the data, certain dimensions or characteristics of the data have to be chosen based on the ultimate objectives of the user. The most commonly used dimensions in measuring the data quality are [30] [31]

- Accessibility: how quickly and easily the data can be retrieved or extracted

- Amount of Data: Is the amount of data facts available sufficient?

- Believability: Is the data true and credible?

- Completeness: Is the data available is complete without having any missing data facts?

- Concise representation: Is the data represented in a concise format?

●Consistent Representation: Is all the available data represented in same format?

●Ease of manipulation: How far it's easy to manipulate the data so that it can be used for different tasks at the same time?

●Free of error: Is the data free from errors and reliable?

●Objectivity: How far the data is unbiased or not prejudiced or not based on ones own interpretations

●Relevancy: Is the data available is relevant and can be used for a particular task?

●Reputation: How far the data is valued in terms of its source or content?

●Security: Is the data secure with needed access control?

●Timeliness: Is the data updated enough to use for a particular task ?

●Understandability: Is the data clear and understandable?

●Usability: How far the data is useful in particular task?

The user has to deicide upon the above dimensions depending upon the final objective or final task that will be conducted with the data. Once the user knows what characteristics to look for in the data, appropriate data quality metrics for those dimensions are formed to measure the quality of the data. The most common methods used for forming data quality metrics are [30]

●Simple Ratio: Simple ratio is the ratio of number of desired outcomes to the total number of outcomes. Using this method the *correctness* of the data can be evaluated by the ration of the number of data facts with errors to total number of data facts.

The other dimensions that can be evaluated by this method are concise representation, relevancy and ease of manipulation

●Min or Max operation: When the data has to be evaluated up on the aggregation of the various factors or dimensions of the data, min or max operation is used. The dimensions that can be evaluated using this Min or Max operation are believability or appropriate amount of data. The believability of the data has to evaluated based on the factors credibility, previous experience and ones own conception of the data. Hence the believability of the data will be evaluated with aggregation of the measurements of the above three factors with some minimum value.

●Weighted average: Weighted alternative can be used as an alternative to the Min operation. If the user has clear understanding of how important each factor is to the particular dimension of the data, then the weighted average of all the factors is a good choice.

Once the data quality is assessed in appropriate dimensions and metrics, an appropriate data cleansing technique has to be applied to the dataset to make it error free. Ultimate use of data plays an important role in deciding which data cleansing technique to use. Most commonly used techniques are: do nothing, correct, replace, insert, remove/split, and combine. Different algorithms are available for each of the data cleansing techniques [12].

2.4 Previous Work

FIELDS – an expert system for analyzing the contaminated sediments has been developed by Yichun Xie and George D. Graettinger [15]. The FIELDS system primarily focuses on finding pollutant hot spots and analyzing the contaminated sediments, which fits in the category of domain specific tools. A Query Knowledge Database, module Knowledge Acquisition has been developed as a part of FIELDS expert system framework. Arc View 2.1 [16] has been used for the spatial data visualization and the analysis part of the framework.

A rule based expert system in United States Environmental Protection Agency (USEPA) rules domain has been developed by Suresh Jayanty, Uta Ziegler using forward chaining and backward chaining engines [14]. Our proposed framework addresses all the major issues of data quality in spatial databases calibration, threshold, missing data, extra data and outliers. A forward chaining engine constitutes the expert system part of our proposed framework. Using the rule based expert system shell gives enough control on the behavior of expert system. Easy updating of knowledge is possible with the use of rule based expert system shell because it separates the rules logic from the GIS desktop application, which is another key point in our framework. The proposed framework supports any kind of spatial data irrespective of their domain, which is the major contribution in our framework. Using ARCGIS 9.1 Engine in the proposed framework can add GIS functionality to the native java framework. The power of the native java language is also harnessed in this framework. The Expert systems approach developed by Beltsville Human Nutrition Research Center for assessing the data quality of the food composition data was reviewed [26].

The Data quality fuzzy expert system developed by F. Moura-Pires, R. A. Ribeiro, A. Pereira, F. J. Varas, G. Mantovani, A. Donati [27] deploys a rules based expert system for assessing the data quality of telemetry data. The same IF-Then rules that are used in GeoExpert were also used in this Data quality fuzzy expert system.

This data quality fuzzy expert system takes the input measures as the input to the data quality fuzzy inference system and then normalizes the output to the interval [0, 1]. This system handles the real time data which is more dynamic and small in quantity [27]. Another expert system framework integrated with GIS developed by Kanjana Chuenpichai, Sutisak Chuenpichai and Ard Somrang [28] was reviewed. In this framework ArcView is deployed for GIS part and it is mainly a decision framework that does the land evaluation by recommending the physical suitability of the land for the given location and activity [28]. This expert system framework uses ArcView for spatial data visualization and analysis. This framework is a stand-alone program and is not generic enough to work with any kind of spatial data. In GeoExpert we solved this problem by implementing the Model-View-Controller (MVC) architecture which enables the GeoExpert to work with any kind of spatial data i.e. domain independent. The paper *A Software Engineering View of Data Quality* -Mónica Bobrowski, Martina Marré, Daniel Yankelevich [29] was reviewed for data quality metrics and dimensions. Another paper *Software Quality Metrics for Object-Oriented Environments* by Dr. Linda H. Rosenberg, Lawrence E. Hyatt [31] was also reviewed for data quality metrics and dimensions. Finally the paper *Data Quality Assessment* by Leo L. Pipino, Yang W. Lee, and Richard Y. Wang [30] was reviewed for the subjective and objective dimensions of the data quality.

Chapter 3: Framework and Prototype

3.1 Framework Objectives:

When we started designing this framework, we defined some objectives that need to be satisfied by this framework

a) **Independent GUI (Graphical User Interface)**

The GUI of this framework should be independent from the application logic. The framework followed the MVC architecture where model is to know what to-do i.e. it contains all the business logic, view is what user sees on the screen and controller is like an interface between the view and the model. The controller in turn applies the changes made by the model to the view. Implementing the MVC architecture gives the application greater flexibility and easy updatability.
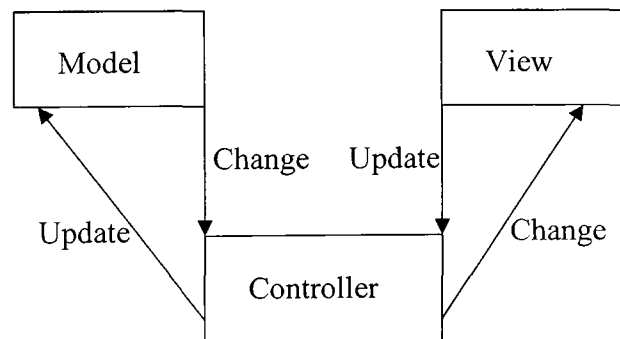
Figure 6: MVC Architecture

b) Domain Independent

There are vast numbers of sub domains in geography that use the Geographical Information Systems. Developing a tool for each domain, like developing a data quality tool for climate data or developing a data quality tool for water quality data is really a cumbersome and tedious process. The tool that is developed for handling data quality in spatial databases has to be a generic tool and it should be able to handle any kind of spatial data, domain independent [17].

3.2 Framework Architecture

A framework is a combination of available software with software components developed by the developer. As already said some of the components used in the implementation of this framework are off the shelf components. Establishing a co-ordination between different components and making them provide a completely new functionality in much more efficient way is a major contribution in this thesis.
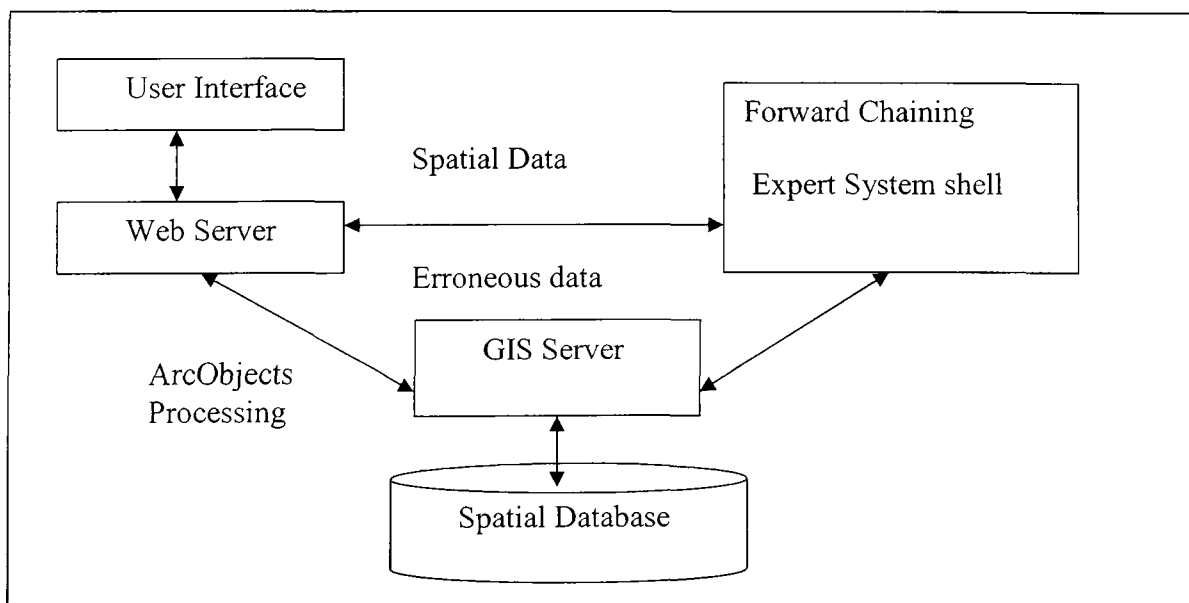
Figure 7: Conceptual architecture of the GeoExpert

Figure 7 is the conceptual framework of the GeoExpert. The internal communication in between the different components in the GeoExpert framework is explained. Initially, when the user selects the features in the user interface the web server extracts the spatial data of the selected features from the spatial database with the help of the GIS Server. GIS Server is having the capability to serve the spatial data remotely as server objects. GIS server is also used to deploy ArcObjects by the webserver. Once the extraction of the data is finished, the spatial data is submitted to the JESS. After the Jess process the data, it sends the erroneous data back to the webserver. Here the web server creates a new shape file with that erroneous data using ArcObjects on the GIS server and adds this new shape file as a new layer to the web user interface.

The off-the shelf components used in this actual implementation of this framework are

- **ARCGIS Server** – A standard GIS framework for designing and deploying Enterprise GIS applications on web, centrally manages ArcObjects where multiple users can use them at the same time. ARCGIS Engine is a developer kit to implement GIS functionality in a stand-alone application

- **JESS** – A forward chaining expert system shell

- **JSR-94** - Java Rule Engine API (JSR-94) [19] that provides access to the expert system shell from any java application.
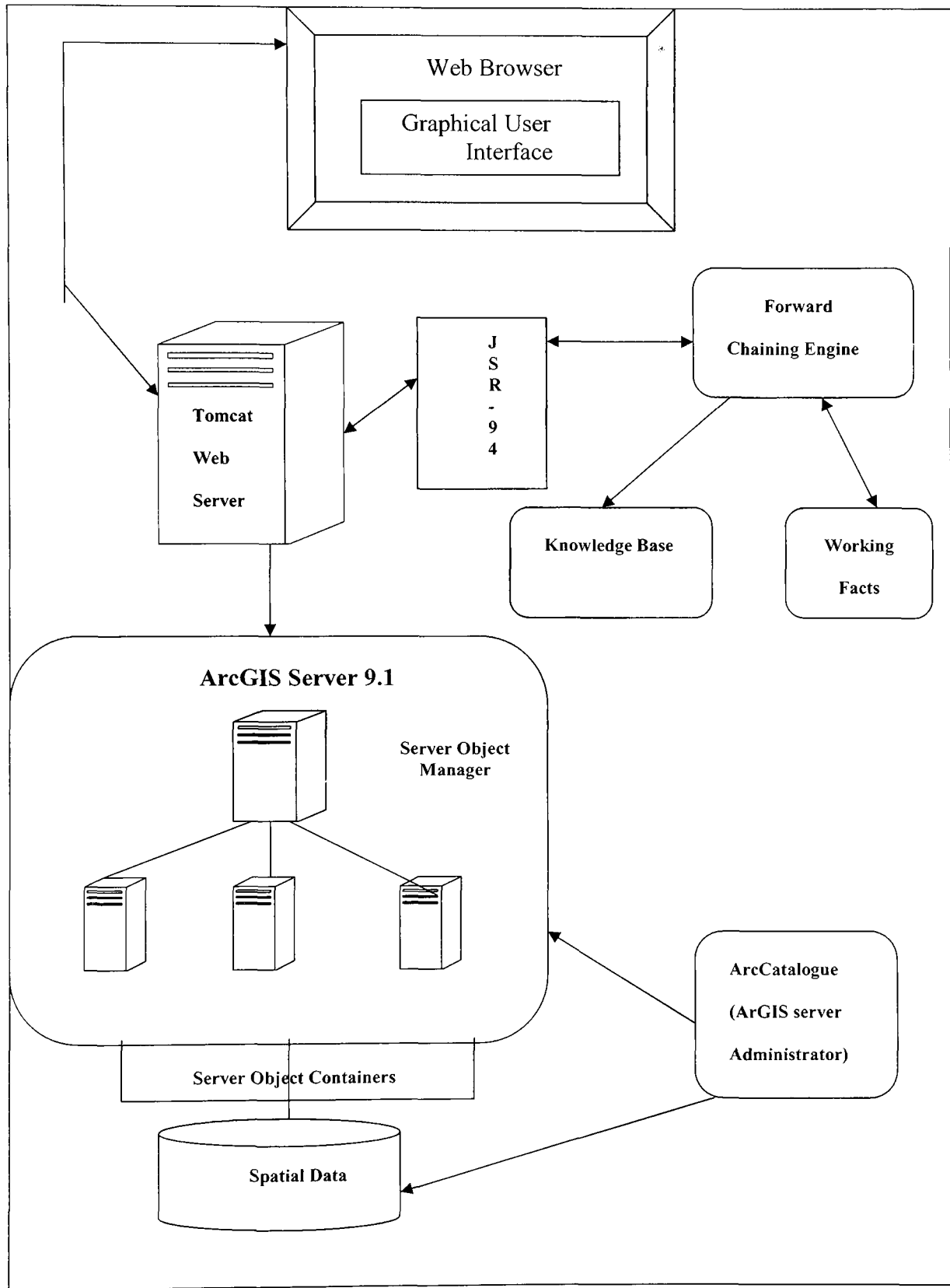
Figure 8: GeoExpert framework implementation architecture

3.2.1 Graphical User Interface

The user can use any of the current existing browsers to access web version of the GeoExpert. The graphical user interface contains three important features

   a) **Map Control:** The map control offers the ArcMap data view's display functionality for a map document served as a server object on the ArcGIS Server [7] and displays single data frame of the map document. We can build procedures to perform many functions on this map document like zoom in, zoom out, identify etc. The map control is a Java Server Faces (JSF) Component. Java Server faces technology is framework for developing sophisticated and flexible user interfaces for web applications. It provides a set of Application Programming Interface (API) for the creation and management of UI components, event handling, and validation of inputs [21]. It also provides custom Java Server Pages (JSP) tag library for deploying JSF components in Java Server Pages. It enables the users to develop the backend logic independent of its integration with user interface. Once the development of backend application logic is finished it can be assembled with the user interface components very easily. The Java Server Faces Component follows the MVC architecture and so are all the ArcGIS server java ADF web controls. An example JSF tag for creating a command button in a JSP page looks like

```
<jsfh:commandButton value="Run"
```

In the above JSF tag , rulesSelect is a java back bean. When the user clicks the command button, the 'save' method in the rulesSelect bean will be triggered.

Now returning to the map control, map control follows Model View Controller architecture. The following figure explains the response of the map control to tool action, listeners and how it accesses the application logic and renders the map [7].
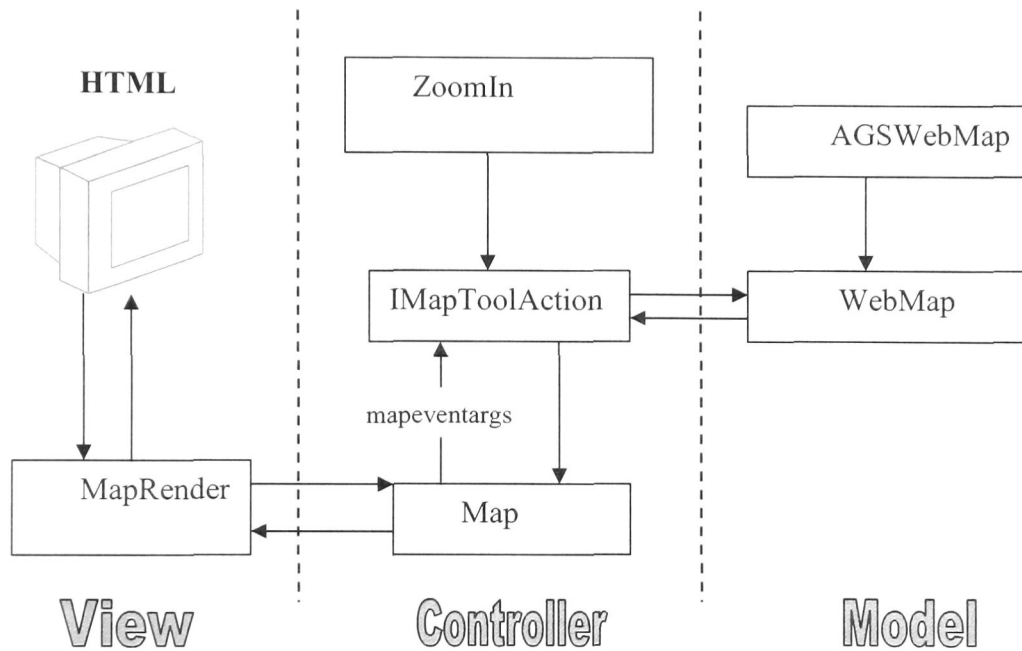
Figure 9: MVC architecture of the map control [7]

In a web application when a map control is created using the following map tag

```
<ags:context id="GeoExpertContext" resource="green@geoserver">
<ags:map id="Map0" left ="0" top = "0" width="500" height="405" />
```

The map tag uses the MapRenderer to generate and display the map. There are many operations like DragRectangle and Create point that can be preformed on the map.

Once the client-side action is completed the MapRenderer will capture and decodes the request and sends it to the map control. The map control serves the request by creating an event argument object called MapEventArgs , a representation of the event in the form of an object. Depending on the operation performed by the client, the map control sends the MapEventArgs to appropriate tool action. The tool action then uses the MapEventArgs to perform the action using Model business objects WebMap and AGSWebMap. This WebMap and AGSWebMap interact with the ArcGIS Server and other ADF objects to perfom the tool action. When the AGSWebContext is initialized at the time of the map control creation, a new AGSWebMap is created. AGSWebMap is a subtype of WebMap and map control associates itself with the WebMap Object [7].

b) Toc Control: The Table of Contents control is used to display the contents of the map in a hierarchical fashion and enables the user to turn on and off the layers. It provides the functionality similar to the table of contents in ArcMap

3.2.2 Apache Tomcat Web server

The Apache Tomcat web server is one of the popular web servers that are used to deploy Servlets and Java server pages on the web [22]. In GeoExpert web framework the ArcGIS server 9.1 Java ADF is deployed on the Tomcat web server. Jess – forward chaining expert system shell is deployed on Tomcat web server.

In our framework Tomcat plays an important role because it co-ordinates the ArcGIS server and Jess – Forward chaining expert system shell to work together and provide the framework both spatial data analysis and reasoning capabilities.

3.2.3 ArcGIS Server

The two major components in the ArcGIS server are server object manager (SOM) and server object container (SOC). ArcGIS server is a server object provider that hosts the GIS resources as server objects. For a GIS resource to be hosted and deployed on web or LAN, a server object which serves that GIS resource has to be created on the ArcGIS server using the ArcCatalogue [7].
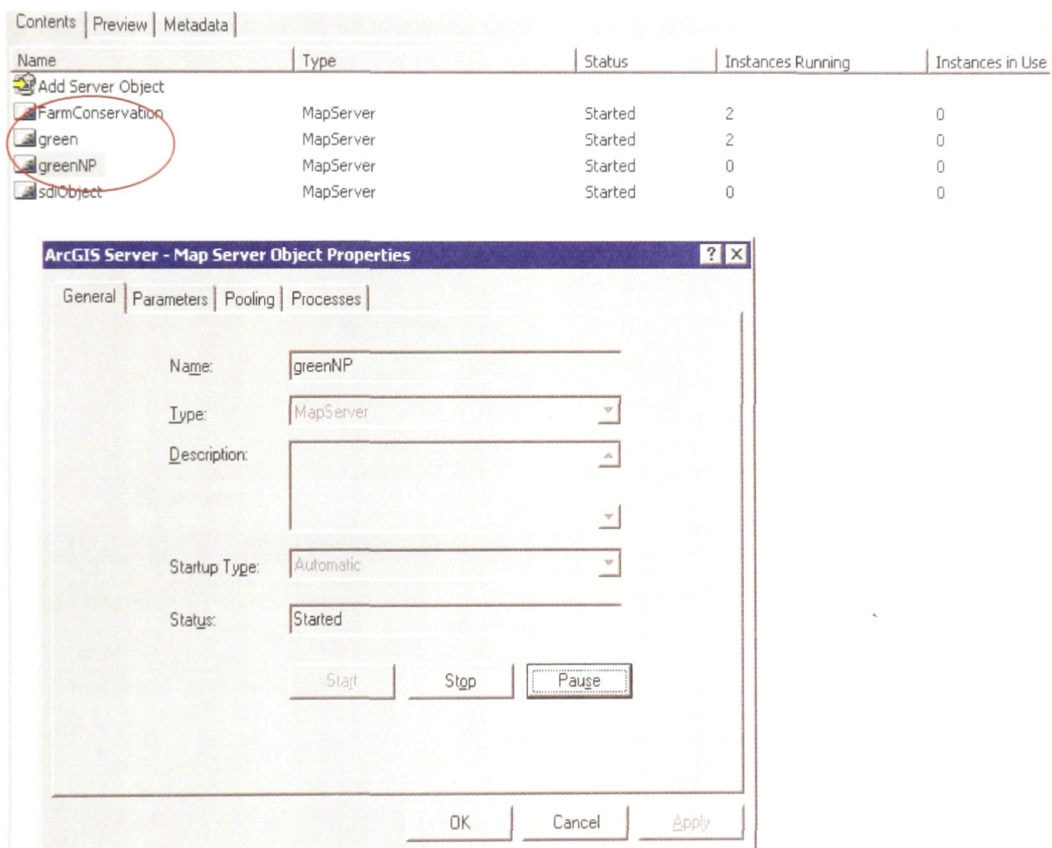
Figure 10: Creating a server object on ArcGIS server

- The server object manager is a Windows service running on a single computer that manages the server objects distributed across several server object container processes running on multiple hosts.

• The server object container is a process that contains one or more server objects running in it. There can be more than one server object container process running on a computer. In turn the server object containers have access to the GIS data resources [7].

When the user opens a web application in the internet browser, the web application in turn makes requests to the server object manager. The server object manager after validating the authentication credentials sends back a proxy of the server object running on one of the server object containers as requested by the web application. The web application works on the proxy GIS server object as if it exists in web application
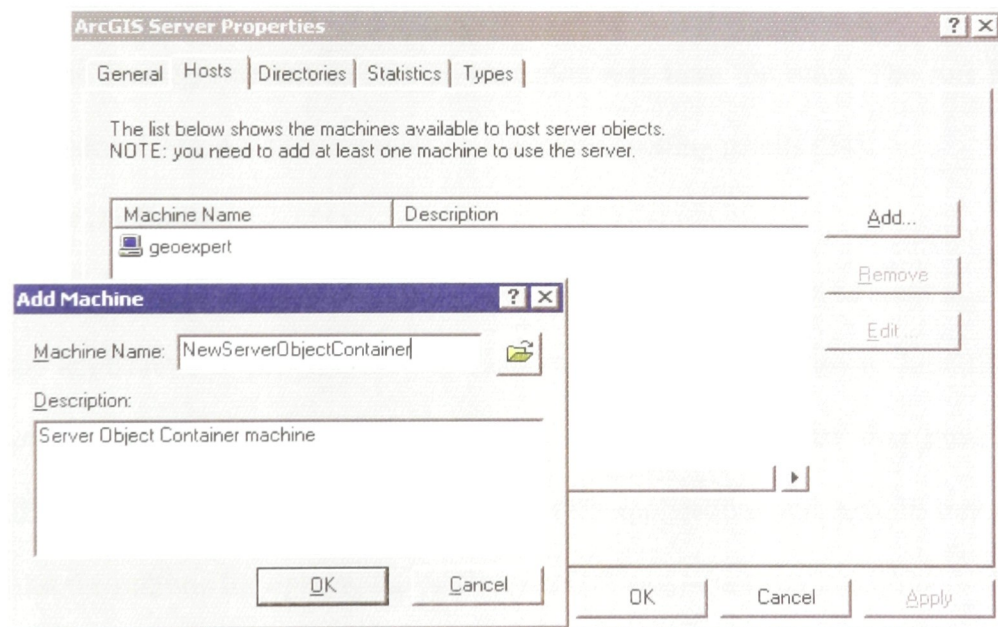
Figure 11: Adding a new server object container machine to ArcGIS server

. In reality all the execution takes place on the GIS server [7]. We will make a server object container machine accessible from GIS server by adding into the list server object container hosts available in ArcGIS server using ArcCatalogue.

3.2.4 Java Expert System Shell (JESS) and JSR - 94

JESS (Java Expert System Shell) is developed by Ernest Friedman-Hill. Jess is an expert system shell with empty knowledge base. The native language for encoding the domain expertise in to production rules is Jess language. Jess uses RETE algorithm resulting in efficient and fast pattern matching by inference engine [23]. JSR – 94 is java rule engine API that provides java runtime API for a large set of rule engines, provides access to the rule engines from a java application. JSR–94 contains two different sets of APIs: the rule administrator API and the rule runtime API. The rule administrator API provides classes for loading production rules that are written in Jess language in XML format and loading relevant actions as execution sets from the rules. The rule runtime API provides classes for running rules on clients and getting results [24].

3.2.5 ArcGIS Engine

ArcGIS engine is a package of ArcGIS components that can be used in a custom desktop application to provide the GIS functionality to that application. ArcGIS engine consists of two important modules, ArcGIS engine developer kit that provides the ArcGIS components and tools to build custom GIS applications and ArcGIS runtime that provides the runtime for running the custom ArcGIS engine applications [6].
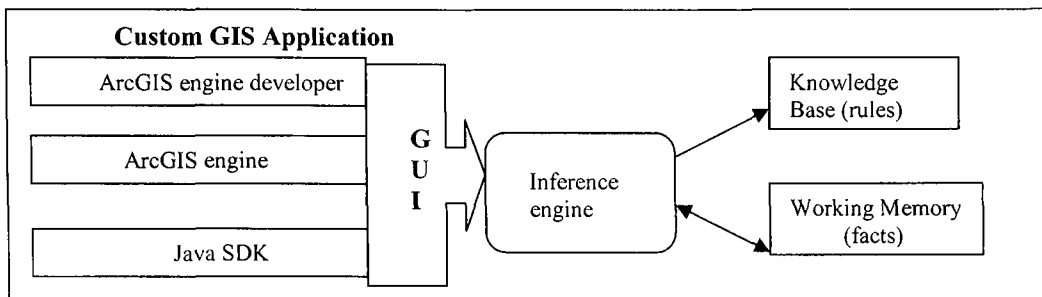
Figure 12: Architecture - GeoExpert Stand-alone version

Chapter 4: Implementation of GeoExpert

4.1 GeoExpert – Web Framework

In this section we will explains the user interaction with the GeoExpert and the backend processing that takes place while assessing the data quality. The GUI is a Java server page with map control and a table of contents (TOC) control and some other tools that can performs actions like zoom in, zoom out, select features and others on the map control. In the figure below have created a server object called 'greenNP' that serves the map document NHD Stream Coverage of Green Country, Kentucky, USA.
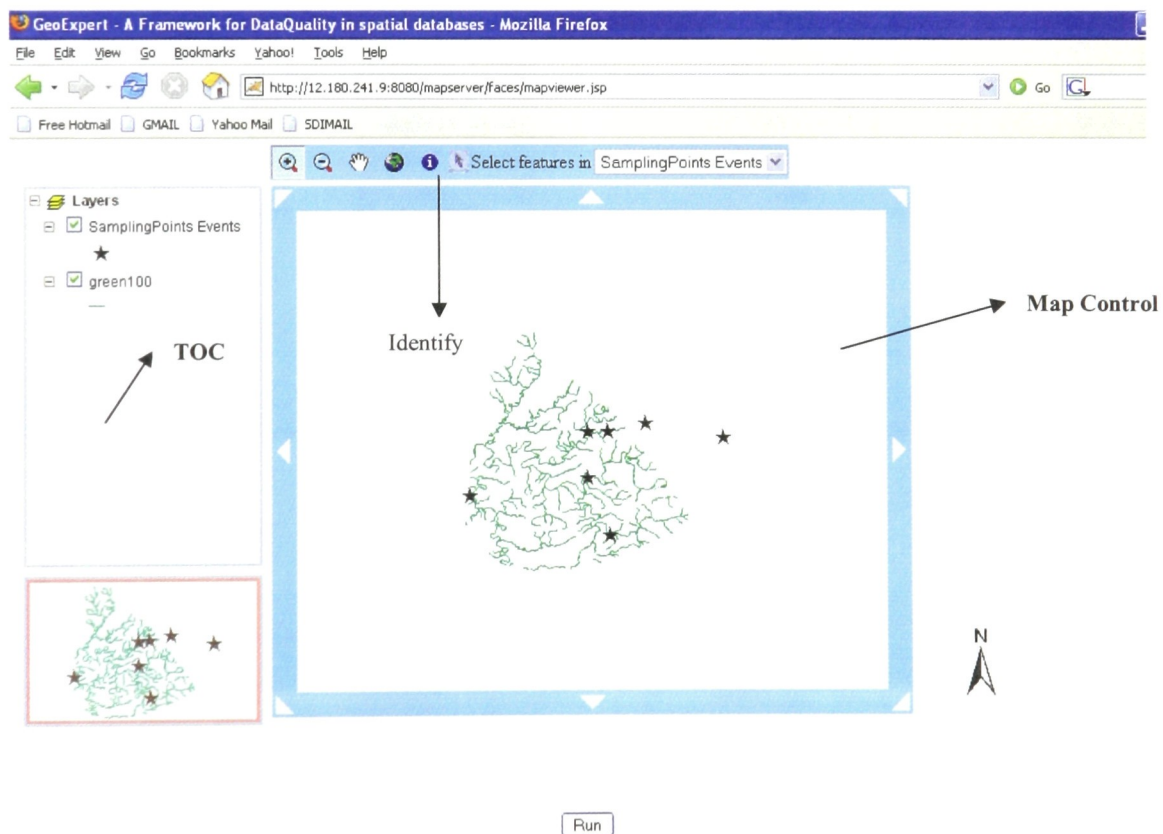
Figure 13: User Interface of the GeoExpert framework

The green100 layer is represented by the layer with green lines in the map control which representing water streams.

The sampling point's layer is represented with star symbols where each star represents a water sampling point at that latitude and longitude. Now we will assess data quality of the water sampling data which is stored in the sampling points layer. The data can be viewed in each sample point by using the identify tool.

Identify Results

Find features in: Top-most Layer

Features	Field	Value
SamplingPoints Events	OID	7
Kentucky	AGENCY__	21KY
	STATION___	SUD073
	STATION_NA	MIDDLE PI
	AGENCY_NAM	KY DEPT N
	STATE_NAME	Kentucky
	COUNTY_NAM	Green
	__LATITUDE	37.3342
	LONGITUDE_	-85.4181
	RESULT_VAL	8
	R	
	HUC_____	5110001
	PARAM	8
	START_DATE	7/10/1984
	START_TIME	2500
	END_DATE__	
	END_TIME	

Close
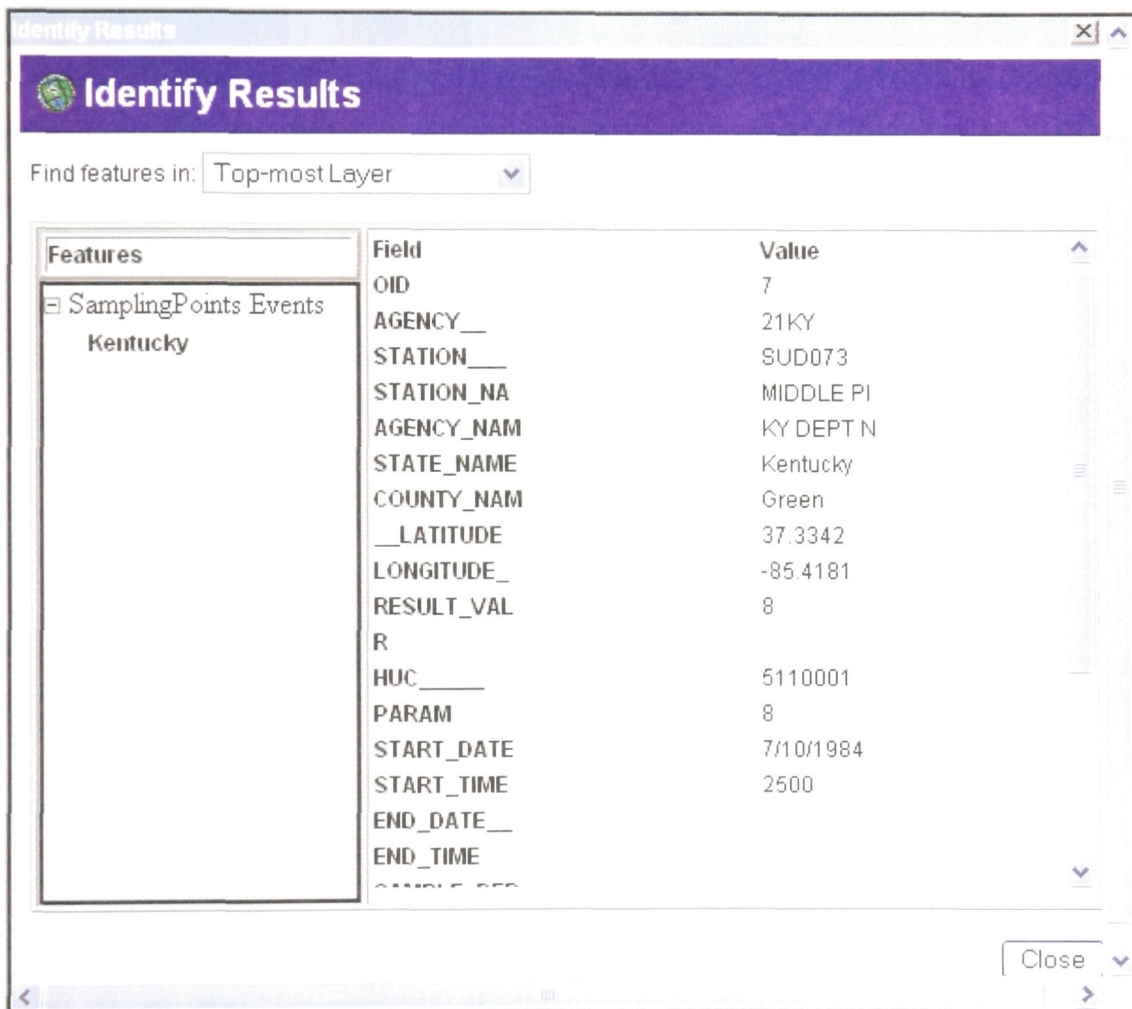
Figure 13: Geographical data stored in one of sample points displayed using identify tool

One of the important reasons for using the ArcGIS in our framework is spatial data visualization. Spatial data visualization provides rich GUI where user can directly interact with data visually.

In the following figure, the user selects the sampling points which should be assessed for data quality using the selection tool. Once after the user selects the features he/she has to click the button 'Run' to start data quality assessment with the help of Jess.
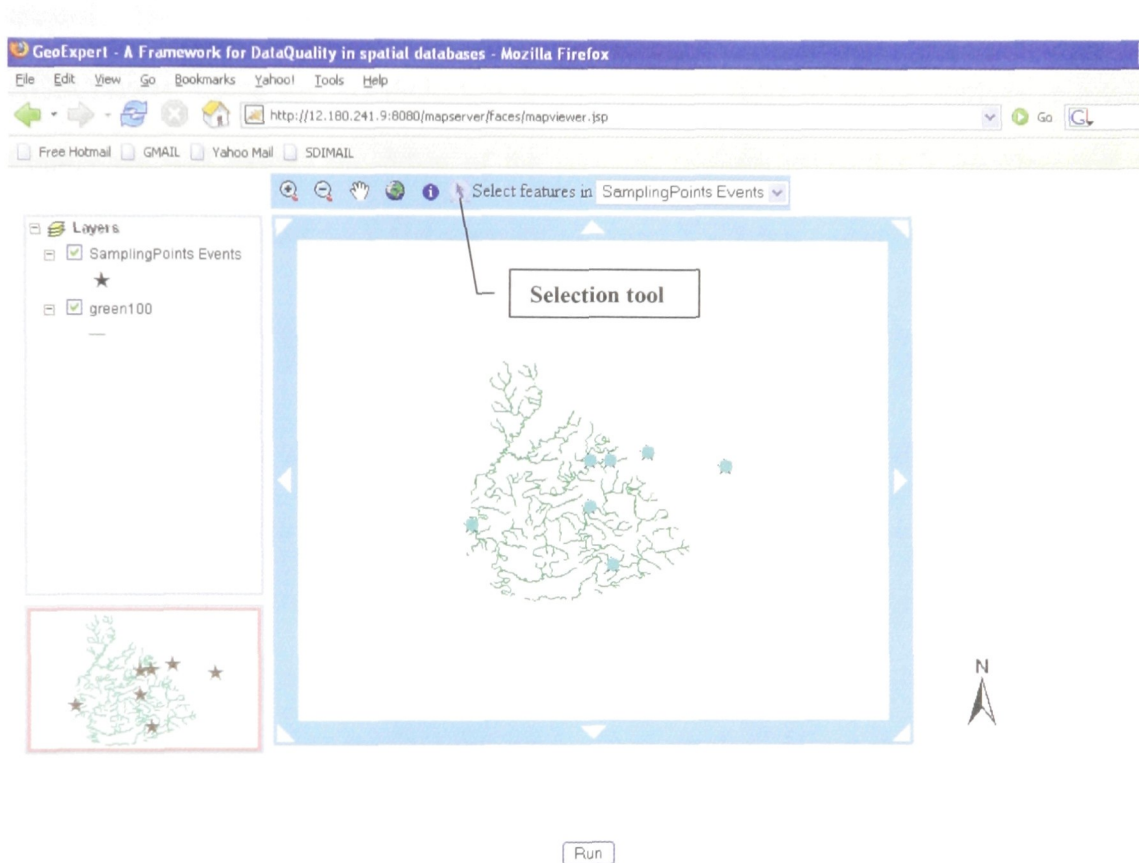
Figure 15: selecting features using select tool

```
AGSWebContext agscxt = (AGSWebContext) session.getAttribute("bufferContext");

AGSSelection selcxt = (AGSSelection)  agscxt.getAttribute(

                              AGSSelection.WEB_CONTEXT_ATTRIBUTE_NAME);

AGSWebMap wMap = (AGSWebMap)agscxt.getWebMap();

        layerID = selcxt.getSelectedLayerId();

        ILayerDescription lyDesc=agscxt.getLayerDescription (wMap.getFocusMapName(),layerID);
```

Figure 16: Extracting the selected features from the AGSWebContext

Once the user clicks the run button the selected features are extracted using the ArcGIS server API as shown in Figure 15. In this process, first the AGSWebContext object is extracted from which the selection context AGSSelection context object is created. The AGSWebMap object is created from the getWebMapMethod of the AGSContext. Using the AGSSelection we will get the layer ID of the features selected. Finally we will extract the fidset of the selected features using the ILayerDescription object. After extracting the fidset we iterate through the fidset and submit each feature as a fact to the Jess.

In order to use the Jess for data assessment, the rules should be loaded into the knowledge base of the Jess, register the rule execution set with the rule administrator and then acquire the Jess rule runtime. After acquiring the jess rule runtime we should acquire the rule session for the rule runtime. The rules for the Jess are stored in a XML file which is read through the input stream by the rule administrator. Using JSR-94 will preserves the MVC architecture of the GeoExpert framework by enabling the same framework to work with any other rule based expert system shell in the place of Jess. Thus the user does not have to change the user interface nor either the logic of the program handling the Jess if any other rule based expert system is used in the place of Jess. JSR 94 acts like an interface between the application and the expert system shell. It converts the application requests so that they can be understood regardless of which expert system shell is used.

```
<?xml version="1.0" encoding="UTF-8"?>

<rule-execution-set>

  <name>executionset1</name>

  <code>

    (defclass RuleEngine.Compound RuleEngine.Compound)
    (defclass RuleEngine.Ulimit RuleEngine.Ulimit)
    (defclass RuleEngine.Llimit RuleEngine.Llimit)

    (defrule MAXLIMIT

?limobj&lt;- (RuleEngine.Ulimit   (ulimit ?ulim) (OBJECT ?D))
?com &lt;- (RuleEngine.Compound   (mcl ?amt&amp;:(&gt; ?amt ?ulim))
                                                    (OBJECT ?C))


    =&gt;

    (call ?C setmcl 9999.99)
    (printout t "The strange arsenic Value is " (get ?C mcl) crlf))


    (defrule MINLIMIT

?limobj1&lt;- (RuleEngine.Llimit    (llimit ?llim) (OBJECT ?E))
?com1 &lt;- (RuleEngine.Compound   (mcl ?amt&amp;:(&lt; ?amt ?llim))
                                                    (OBJECT ?B))
    =&gt;

    (call ?B setmcl 8888.88)
    (printout t "The strange arsenic Value is " (get ?B mcl) crlf))

  </code>
  </rule-execution-set>
```

Figure 17: Rules in the XML file for Jess

In figure 15 there are two attributes PARAM and RESULT_VAL. According to USEPA each PARAM represents a compound test and each compound will have a lower limit and upper limit beyond which that result value is treated as either anomaly or outlier. We supply this information to the GeoExpert using a data file which has the PARAM value, appropriate compound, upper limit and lower limit. So GeoExpert reads the current value from each selected feature and get the upper limit and lower limit of that PARAM from the data file and submits them to Jess for processing.

In figure 16

●<rule-execution-set> is the set of rules and defclass is a command to create a template out of the java class file.

● Defrule is a construct used to create the rule.

● In the above rule MaxLimit, ?ulim holds the value of member variable 'ulimit' of RuleEngine.Ulimit object.

● In the next line ?amt hold the value of member value 'mcl' of RuleEngine.Ulimit object.

● The current maximum contamination level value is compared against the compounds maximum contamination level limit and if it exceeds the maximum limit , a flag is raised.

Once after processing of all the selected features is finished, a new shape file is created using the flagged features. There is an explanation provided for each feature in the newly created layer. An attribute called RuleResult can have two values. '+' indicates the compound is over the maximum level and '-' denotes the compound is less than compounds minimum level. The name of the compound is stored in the attribute 'Compound'. Once the erroneous features are found, the user can use the appropriate data cleansing technique depending upon the ultimate use of data for data cleansing.
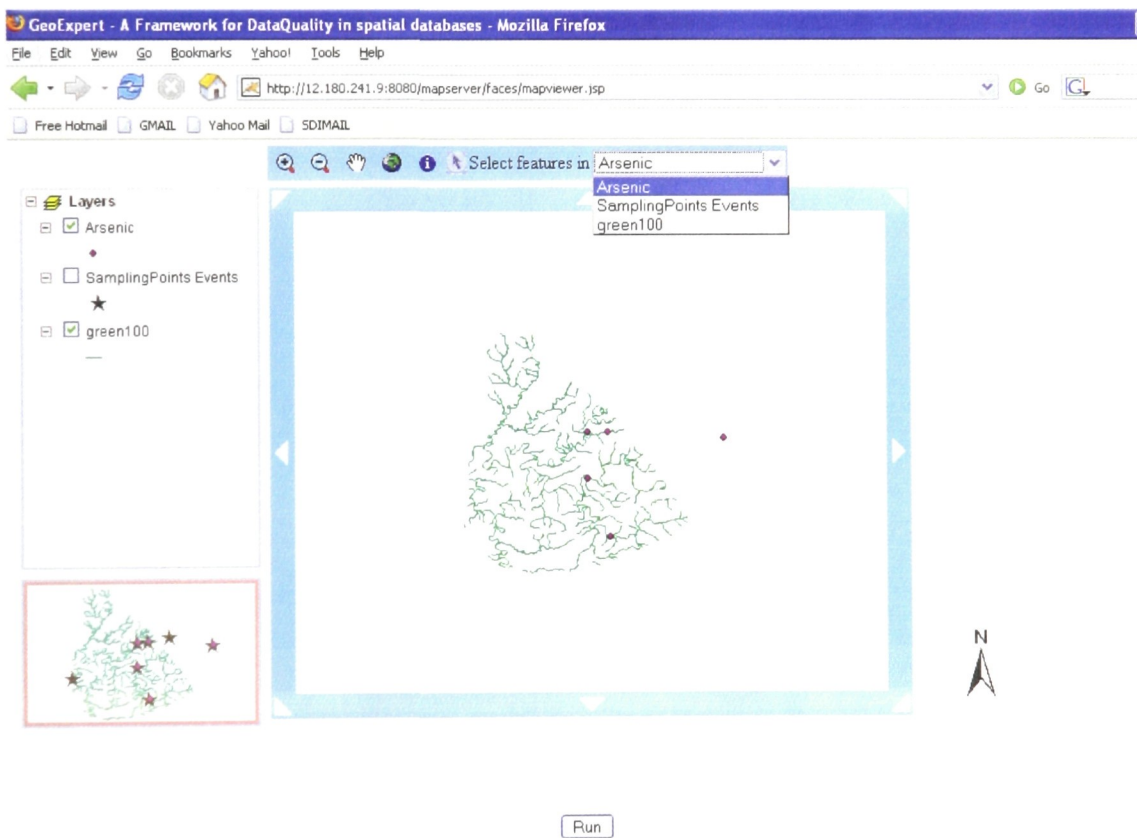
Figure 18: GeoExpert with newly created layer that stores the erroneous features

Figure 19: Erroneous feature identified by JESS

4.2 GeoExpert – Stand-alone Framework

A java stand-alone application capable of handling GIS data is developed for assessing the quality of the spatial data [17]. For this we used Java 2 Platform Standard Edition 5.0 and ArcGIS Engine 9.1 that provides access to the ArcObjects, and the Standard GIS Framework. We used Jess for forward chaining in our application. When the user selects the data, with the help of JSR-94 the application acquires a stateful rule runtime session of the Jess and submits the selected spatial datasets as facts to the Jess.
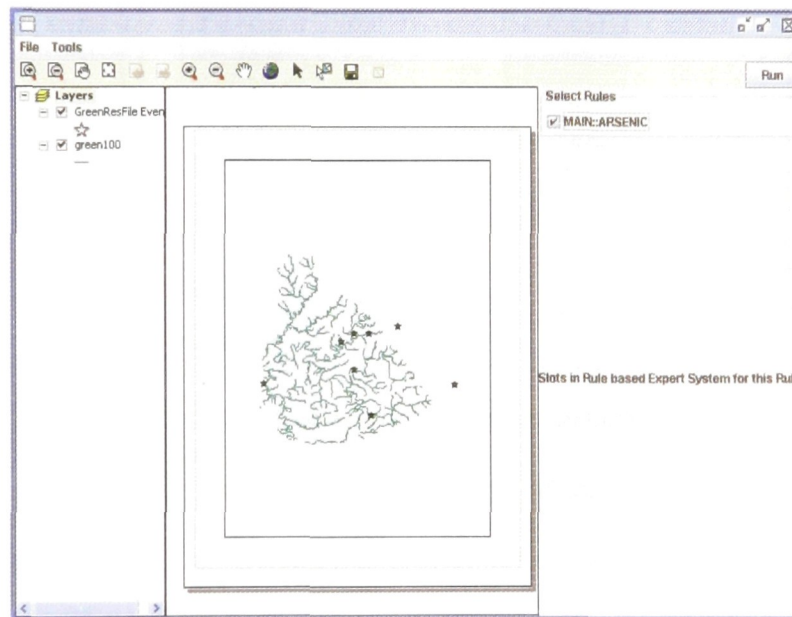
Figure 20: Application after loading the MXD document

Currently GeoExpert application can handle the spatial data in the format of ArcMap Coverages. Initially the user will load the data into the application in MXD format (.MXD extension). The above MXD document in Figure 18 represents the water quality data of Green County, Kentucky, USA.

The start points indicate the water sampling points in the green county and the underlying layer is NHD stream coverage of Green County, Kentucky, USA [25]. Initially the user loads the MXD document, using the "select features" tool, then, selects the sample sites that will be checked against the rules in the Jess expert system shell.
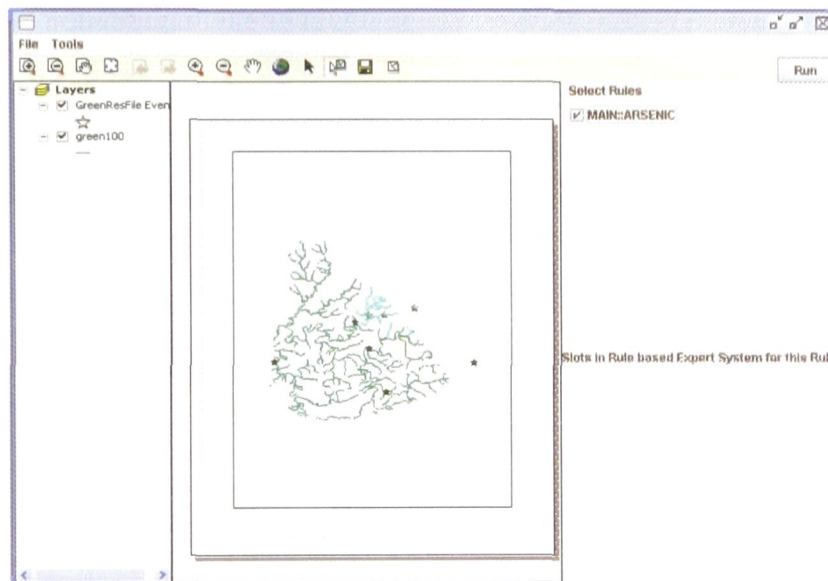
Figure 21: Application after selecting the features

After selecting the features, the user selects the rules to be executed against the selected features from the list of rules that are displayed on the right-most panel. As soon as the user clicks on the RUN button on the right most upper corner, the data of the selected features is retrieved using a spatial query. This application is connected to the Jess rule base expert system shell [23] using JSR-94 [19]. Now the retrieved data is passed to the Jess as facts.

In the above rule we are setting the error flag ON if the contaminant level exceeds the maximum level or if it is below the minimum level. Once the execution of a rule is completed using the errors detected, a separate shape file (with .shp extension) is

generated for each rule executed along with the errors detected by that rule. The newly

created shape file will be added as a new layer to the MXD document currently loaded. In

this case a new shape file named arsenic.shp is created with erroneous data.
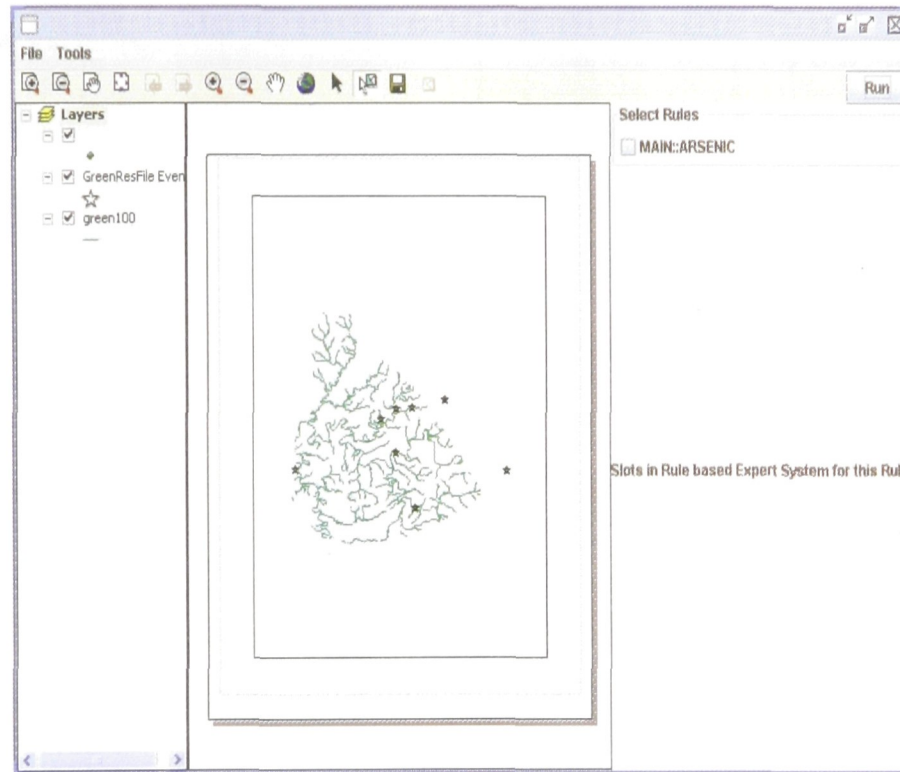
Figure 22: Application after the data validation process

In Figure 20, a new layer is added on the top of the tree structure. The new

layer holds the data that has been detected erroneous by Jess. Hence it is proved that that

spatial data visualization and analysis capabilities of ARCGIS can be integrated with the

reasoning and inference capabilities of the expert system shell for assessing the data

quality of the spatial databases.

Chapter 5: Future work and Conclusion

5.1 Future Work

A backward engine can be used for data cleansing phase in this framework. Once the errors in the dataset are identified, simultaneously a backward engine asks the user the purpose of the data. The backward engine proposes to the user the appropriate cleansing technique according to his ultimate purpose of the data. Once the user selected a particular procedure, that procedure will be applied to the erroneous dataset [17].

To make GeoExpert to work for complex rules, the user has to encode the complex rule in native jess language and supply the Jess expert system shell with the necessary data. The MVC architecture of this framework makes it easy for the user to update the application logic without changing the user interface. Once the rules are there in knowledge base of the expert system shell and required data is supplied, the data will be assessed by the expert system shell.

5.2 Conclusion

The use of a rule based expert system for data quality assessment automates the data cleansing and use of ArcGIS framework in the application allows us to provide the user with a user – friendly interface that supports the spatial data visualization, which is one of the major advantages of the framework. In the future users may integrate this framework in their own application to automate the data quality assessment and cleansing phase.

Deploying Jess on the web is the first step towards harnessing the reasoning and inference capability of expert systems remotely on the internet. The use of expert systems in assessing the data quality of the spatial data with the help of ArcGIS framework opens a new way of deploying expert systems. Also Implementing the MVC architecture gives the application greater flexibility and easy updatability.

Appendix – I

GeoExpert – Web Framework

mapviewer.jsp (Some of the code in mapviewer.jsp is copied from mapviewer template of ArcGIS server Java ADF Framework, sample codes in http://edn.esri.com and JSR-94 sample code in http://java.sun.com/developer/technicalArticles/J2SE/JavaRule.html)

In the following source code we used Java Server Faces, Java SDK, servlets and JSP Tags. In initially in this code a server connection is established with the ArcGIS server with a particular server object. After that as explained in the thesis a map control was created which displays the map of the server object. At the same time there are map navigation tools like Zoom In, Zoom Out , Pan, Next Extent, Previous Extent etc in this page.

```
<%--
   The Java ADF was developed with Java Server Faces (JSF) version 1.0.
   For the latest information on supported servlet engines and application servers as
   well as software and documentation updates, visit the ESRI Support Web site at
   http://support.esri.com. For more information about JSF, visit
   http://java.sun.com/j2ee/Java Serverfaces.
--%>

<%@ page contentType="text/html; charset=ISO-8859-1" %>
<%@page session="true"%>
<%@ taglib uri="http://java.sun.com/jsf/core" prefix="jsfc" %>
<%@ taglib uri="http://java.sun.com/jsf/html" prefix="jsfh" %>
<%@ taglib uri="http://java.sun.com/jstl/core" prefix="c" %>
<%@ taglib uri="http://www.esri.com/arcgis/webcontrols" prefix="ags" %>

<ags:noCache />
<jsfc:loadBundle basename="res.ResTemplates" var="resource" />
<jsfc:view>
<html>
 <head>
  <title><c:out value="${resource['MapViewer.title']}" /></title>
    <ags:htmlBase />
  <link href="css/mapviewer.css" rel="stylesheet" type="text/css">
  <script language="Javascript" src="js/templates.js"></script>
 </head>
<body topmargin="5">
```

```
<jsfh:form id="frmMap"> <ags:context id="bufferContext"
                                        resource="greenNP@GeoExpert">
        <!-- Title and Logo -->
<%-- Selection Tool--%>
<c:setvar="selection"
value="${sessionScope['bufferContext'].attributes['esriAGSBuffer'].selection}"/>
<%-- Number of features selected are stored the esriAGSBuffer Context--%>
<c:set var="selectionCount"
value="${sessionScope['bufferContext'].attributes['esriAGSBuffer'].selection.selectionCo
unt}"/>
 <c:set var="resultCount"
value="${sessionScope['bufferContext'].attributes['esriAGSBufferResults'].count}"/>
<c:set var="buffer"
value="${sessionScope['bufferContext'].attributes['esriAGSBuffer'].selection.selectedLay
erId}"/>
<c:set var="bufferResults"
value="${sessionScope['bufferContext'].attributes['esriAGSBufferResults']}"/>
  <table cellspacing="0">
        <tr class="lightBlue">
        <td width="210" style="background-color:#FFFFFF;"> </td>
        <td>
          <%-- ZoomIn tool --%>
          <img id="imgZoomIn" name="imgZoomIn" src="images/zoomin.gif"
              alt="<c:out value="${resource['Templates.ZoomIn']}" />"
              title="<c:out value="${resource['Templates.ZoomIn']}" />"
              onmouseover="this.src='images/zoominU.gif"
onmousedown="this.src='images/zoominD.gif';MapDragRectangle('Map0', 'ZoomIn');
HighlightTool('ZoomIn');"
onmouseout="ButtonOut('imgZoomIn', 'Map0', 'ZoomIn', 'images/zoomin.gif',
'images/zoominD.gif')">
        </td> <td>
<%-- ZoomOut tool --%>
<img id="imgZoomOut" name="imgZoomOut" src="images/zoomout.gif"
 alt="<c:out value="${resource['Templates.ZoomOut']}" />"
title="<c:out value="${resource['Templates.ZoomOut']}" />
onmouseover="this.src='images/zoomoutU.gif"
onmousedown="this.src='images/zoomoutD.gif';MapDragRectangle('Map0',
'ZoomOut');HighlightTool('ZoomOut');
onmouseout="ButtonOut('imgZoomOut', 'Map0', 'ZoomOut', 'images/zoomout.gif',
'images/zoomoutD.gif')">
</td><td>
<%-- Pan tool --%>
<img id="imgPan"  name="imgPan" src="images/pan.gif"
              alt="<c:out value="${resource['Templates.Pan']}" />"
              title="<c:out value="${resource['Templates.Pan']}" />"
              onmouseover="this.src='images/panU.gif"
```

```
                    onmousedown="this.src='images/panD.gif';MapDragImage('Map0',
'Pan');HighlightTool('Pan');"
                    onmouseout="ButtonOut('imgPan', 'Map0', 'Pan', 'images/pan.gif',
'images/panD.gif')">
            </td>
            <td>
            <%-- Zoom to Full Extent button --%>
            <jsfh:commandButton id="cmdFullExt" image="images/fullext.gif"
                    title="#{resource['Templates.FullExtent']}"
                    alt="#{resource['Templates.FullExtent']}"
                    onmousedown="this.src='images/fullextD.gif"
                    onmouseover="this.src='images/fullextU.gif"
                    onmouseout="this.src='images/fullext.gif">
        <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.ZoomFullExtentListener" />
            </jsfh:commandButton>
            </td>
            <%-- Load Extent Stack --%>
            <c:set var="initStack"
value="S{sessionScope['mapContext'].attributes['esriAGSExtentStack']}" />
            <%-- <td>
            <%-- Zoom to Previous Extent button
            <jsfh:commandButton id="cmdBackExt"
action="#{sessionScope['mapContext'].attributes['esriAGSExtentStack'].doPrevious}"
image="images/back.gif"
                    title="#{resource['Templates.PreviousExtent']}"
                    alt="#{resource['Templates.PreviousExtent']}"
                    onmousedown="this.src='images/backD.gif"
                    onmouseover="this.src='images/backU.gif"
                    onmouseout="this.src='images/back.gif" />
            </td>  <td>


            <%-- Zoom to Next Extent button
            <jsfh:commandButton id="cmdForwardExt"
action="#{sessionScope['mapContext'].attributes['esriAGSExtentStack'].doNext}"
image="images/forward.gif"
                    title="#{resource['Templates.NextExtent']}"
                    alt="#{resource['Templates.NextExtent']}"
                    onmousedown="this.src='images/forwardD.gif"
                    onmouseover="this.src='images/forwardU.gif"
                    onmouseout="this.src='images/forward.gif" />
            </td>--%>
            <td>
            <%-- Identify tool --%>
            <IMG id="imgIdentify" name="imgIdentify" src="images/identify.gif"
                    alt="<c:out value="${resource['Templates.Identify']}" />"
```

```
                title="<c:out value="${resource['Templates.Identify']}" />"
                onmouseover="this.src='images/identifyU.gif'"
                onmousedown="this.src='images/identifyD.gif';MapPoint('Map0',
'Identify');HighlightTool('Identify');"
                onmouseout="ButtonOut('imgIdentify', 'Map0', 'Identify',
'images/identify.gif', 'images/identifyD.gif')" .
        </td>
        <td>
            <img id="imgSelection" name="imgSelection" src="images/sel.gif"
                alt="<c:out value="${resource['Buffer.SelectFeatures']}" />"
                title="<c:out value="${resource['Buffer.SelectFeatures']}" />"
                onmouseover="if(!this.disabled) {this.src='images/sel.gif'; }"
                onclick="if(!this.disabled)
    {
    this.src='images/sel_on.gif';
    MapDragRectangle('Map0', 'Selection');
    divVisible('frmBuffer:selSelectionType', true);
    }"
                onmouseout="if(!this.disabled) { ButtonOut('imgSelection', 'Map0',
'Selection', 'images/sel.gif', 'images/sel_on.gif'); }">
        </td>
        <td><c:out value="${resource['Buffer.Label1']}" /></td>
        <td>
        <jsfh:selectOneMenu id="selSelectedLayerID"
value="#{sessionScope['bufferContext'].attributes['esriAGSSelection'].selectedLayerId}"

disabled="#{sessionScope['bufferContext'].attributes['esriAGSBufferResults'].resultsFetc
hed}"
                style="width:165px">
            <jsfc:selectItems id="selSelLayerOption"
value="#{sessionScope['bufferContext'].webMap.featureLayers}"/>
        </jsfh:selectOneMenu>
        </td>
    </tr>
    </table>
    <table cellpadding="3">
      <tr valign="TOP">
        <td><ags:toc height="320" width="200" cssClass="tocClass" borderWidth="2"
/></td>
        <td rowspan="3">
          <table cellspacing="0">
            <tr class="lightBlue" align="CENTER">
            <td>
                <jsfh:commandButton id="cmdPanUL" image="images/upperleft.gif"
title="#{resource['Templates.PanUpperLeft']}"
alt="#{resource['Templates.PanUpperLeft']}">
```

```
            <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
            <jsfc:attribute name="direction" value="upperLeft" />
            <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
    </td>
    <td>
        <jsfh:commandButton id="cmdPanU" image="images/up.gif"
title="#{resource['Templates.PanUp']}" alt="#{resource['Templates.PanUp']}">
            <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
            <jsfc:attribute name="direction" value="up" />
            <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
    </td>
    <td>
        <jsfh:commandButton id="cmdPanUR" image="images/upperright.gif"
title="#{resource['Templates.PanUpperRight']}"
alt="#{resource['Templates.PanUpperRight']}" >
            <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
            <jsfc:attribute name="direction" value="upperRight" />
            <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
    </td>
</tr>
    <tr align="CENTER" class="lightBlue">
    <td>
        <jsfh:commandButton id="cmdPanL" image="images/left.gif"
title="#{resource['Templates.PanLeft']}" alt="#{resource['Templates.PanLeft']}">
            <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
            <jsfc:attribute name="direction" value="left" />
            <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
    </td>
    <td>
        <ags:map id="Map0" width="500" height="405" activeTool="ZoomIn" />
    </td>
    <td>
        <jsfh:commandButton id="cmdPanR" image="images/right.gif"
title="#{resource['Templates.PanRight']}" alt="#{resource['Templates.PanRight']}">
            <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
            <jsfc:attribute name="direction" value="right" />
            <jsfc:attribute name="factor" value=".5" />
```

```
        </jsfh:commandButton>
      </td>
    </tr>
    <tr class="lightBlue" align="CENTER">
      <td>
        <jsfh:commandButton id="cmdPanLL" image="images/lowerleft.gif"
title="#{resource['Templates.PanLowerLeft']}"
alt="#{resource['Templates.PanLowerLeft']}">
          <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
          <jsfc:attribute name="direction" value="lowerLeft" />
          <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
      </td>
      <td>
        <jsfh:commandButton id="cmdPanD" image="images/down.gif"
title="#{resource['Templates.PanDown']}" alt="#{resource['Templates.PanDown']}">
          <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
          <jsfc:attribute name="direction" value="down" />
          <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
      </td>
      <td>
        <jsfh:commandButton id="cmdPanLR" image="images/lowerright.gif"
title="#{resource['Templates.PanLowerRight']}"
alt="#{resource['Templates.PanLowerRight']}">
          <jsfc:actionListener
type="com.esri.arcgis.webcontrols.ags.faces.event.DirectionalPanListener" />
          <jsfc:attribute name="direction" value="lowerRight" />
          <jsfc:attribute name="factor" value=".5" />
        </jsfh:commandButton>
      </td>
    </tr>
  </table>


      </td>
      <td rowspan="2" valign="BOTTOM"><ags:northArrow /></td>

    </tr>
    <tr>
      <td valign="top">
        <ags:overview id="Overview0" resource="greenNP@GeoExpert" width="200"
height="125" borderWidth="2" cssClass="overviewClass" />
      </td>
```

```
</tr>

</table>
```

<%-- Identify Results control --%>
```
<ags:identifyResults id="IdentifyResults" left="40" top="40" />
<%-- Number of feature(s) selected on the buffer layer --%>
        <c:if test="${rulesSelect.cmdClick}">
        <%@ page import="com.esri.arcgis.webcontrols.ags.data.* " %>
        <%@ page import="com.esri.arcgis.webcontrols.util.*" %>
        <%@ page import="javax.faces.*" %>
        <%@ page import = "com.esri.arcgis.geodatabase.* "%>
        <%@ page import="com.esri.arcgis.geometry.*"%>
        <%@ page import = "com.esri.arcgis.carto.*"%>
        <%@ page import="javax.servlet.*" %>
         <%@ page import="java.lang.* " %>
         <%@ page import="java.util.*" %>
        <%@ page import="java.io.*" %>
        <%@ page import="com.esri.arcgis.webcontrols.data.*"%>
         <%@ page import="com.esri.arcgis.server.* "%>
        <%@ page import="javax.rules.admin.*"%>
      <%@page import="javax.rules.*" %>
      <%@page import="javax.faces.*" %>
      <%@ page  import="RuleEngine.*" %>
      <%@ page import ="com.esri.arcgis.datasourcesfile.*"%>



        <%
        javax.faces.context.FacesContext fc =
javax.faces.context.FacesContext.getCurrentInstance();
        String[] ruleChoice =
(String[])fc.getApplication().createValueBinding("#{rulesSelect.JessRules}").getValue(f
c);
        IFeatureCursor pFeatureCursor;
    IFeature pFeature;
    IFields sFields;
//Rule service provider for jess in JSR-94
    String RULE_SERVICE_PROVIDER = "org.jcp.jsr94.Jess";
//Rule execution set
    RuleExecutionSet res1;
//Rule Service Provider
    RuleServiceProvider serviceProvider;
    RuleAdministrator ruleAdministrator;
    CreateShapeFromText shapeFromText;
    IFeatureClass fClass;
```

```
IFeatureClass pFeatClass = null;
Hashtable paramHash = new Hashtable();
RuleEngine.dataElement dt;
String iostring;

String ruleValues = new String();
boolean ruleViolate = false;

try
{
```
//Initially reading the parameters.dat file for the maximum and minimum levels
//and the compound name for the parameters into the Jess

```
        FileReader fr = new FileReader("C:\\Program Files\\Apache Software
Foundation\\Tomcat 5.5\\webapps\\mapserver\\WEB-INF\\classes\\parameters.dat");

        BufferedReader br = new BufferedReader(fr);
```
//Storing the Contents in the parameters.dat into the hash table
```
        while((iostring = br.readLine())!=null)
                {

        String[] result = iostring.split("\\s");
dt = new dataElement(Integer.parseInt(result[0]),result[1],
                Double.parseDouble(result[2]),Double.parseDouble(result[3]));

                paramHash.put(Integer.parseInt(result[0]),dt);

                }

}
catch(Exception e)
{
    out.println(e+"Exception in reading parameter file");
}
```


// **Getting the current AGS webcontext object**
```
        AGSWebContext agscxt =
(AGSWebContext)session.getAttribute("bufferContext");
```
// **Getting the current AGS selection context**
```
        AGSSelection selcxt = (AGSSelection) agscxt.getAttribute(
AGSSelection.WEB_CONTEXT_ATTRIBUTE_NAME);
```
//**Getting the current AGSBufferContext**

```
            AGSBuffer bufcxt = (AGSBuffer) agscxt.getAttribute(
AGSBuffer.WEB_CONTEXT_ATTRIBUTE_NAME);
            //AGSWebMap wMap = selcxt.getAGSWebMap ();
            //agscxt.applyDescriptions();
    // Getting the current AGSWebMap
            AGSWebMap wMap = (AGSWcbMap)agscxt.getWebMap();
    // Getting the current AGSWebToc
            AGSWebToc wToc = (AGSWebToc)agscxt.getWebToc();


    //Getting the server context object
            IServerContext serverContext = agscxt.getServerContext();
    //Getting the layered of the layer in which the features are selected using select
    //features tool
            int layerID = selcxt.getSelectedLayerId();
    //Getting the Ilayerdescription object
            ILayerDescription lyDesc=agscxt.getLayerDescription
(wMap.getFocusMapName(),layerID);
    //Getting the FIDSet of the selected featues and iterating thru that features using
    //IEnumIDs object
            IFIDSet fidSet=lyDesc.getSelectionFeatures();
            IEnumIDs enumids = new IFIDSet2Proxy(fidSet).getIDs();
            IMapServer mapServer = new IMapServerProxy(
serverContext.getServerObject());


    // IMapServerObjects mapServerObjects = new
IMapServerObjectsProxy(mapServer);
            // IMap map = mapServerObjects.getMap(mapServer.getDefaultMapName());
            MapServer mapServerObjects = new MapServer(agscxt.getServer());
        IMap map = mapServerObjects.getMap(wMap.getFocusMapName());


            // ILayer il = new ILayerProxy(map.getLayer(layerID));
            IFeatureLayer ifl = new IFeatureLayerProxy(map.getLayer(layerID));
            //Getting the featureclass of the layer in which the features are selected
            IFeatureClass pfClass = ifl.getFeatureClass ();


            IFeatureSelection pFeatureSelection = new IFeatureSelectionProxy( ifl);
            //ISelectionSet pSelectionSet = pFeatureSelection.getSelectionSet();
        com.esri.arcgis.geodatabase.IQueryFilter pQueryFilter;
    // Getting the printstream for the XY.txt file where the information about the
    //outliers is stored
            PrintStream LatLong = new PrintStream( new FileOutputStream("C:\\Program
Files\\Apache Software Foundation\\Tomcat 5.5\\webapps\\mapserver\\WEB-
INF\\classes\\RuleEngine\\XY.txt"));


            FeatureLayer arsenic_Layer = new
FeatureLayer(serverContext.createObject(FeatureLayer.getClsid()));
```

//Iterating thru the each FID and extracting the relevant feature of the fid from
//feature class and passing the required water sampling data into the Jess

```
        for(int i=0;i<fidSet.count();i++)
        {
        try
          {
          pQueryFilter = new
com.esri.arcgis.geodatabase.QueryFilter(serverContext.createObject(QueryFilter.getClsid
()));

          pQueryFilter.setWhereClause ("OID = "+enumids.next());

          //out.println(pQueryFilter) ;
          //pFeatureSelection.selectFeatures(pQueryFilter,esriSelectionResultEnum.esriSelectio
nResultNew,false);
          //pSelectionSet.search(pQueryFilter,false,pFeatureCursor);
          pFeatureCursor = ifl.search(pQueryFilter,false);

          pFeature = pFeatureCursor.nextFeature();
          while(pFeature!=null)
          {
          sFields = pFeature.getFields();
                Class.forName ("org.jcp.jsr94.Jess.RuleServiceProviderImpl");




                serviceProvider =



RuleServiceProviderManager.getRuleServiceProvider(RULE_SERVICE_PROVIDER);
```

//Acquiring the jess stateful session
```
          ruleAdministrator = serviceProvider.getRuleAdministrator();
```

//Reading the rules into the expert system of the jess
```
          InputStream inStream = this.getClass().getResourceAsStream("/"+"tck.xml");
```

//Loading the rules from the tck.xml into the knowledge base of the JESS
```
        res1 =
ruleAdministrator.getLocalRuleExecutionSetProvider(null).createRuleExecutionSet(
inStream, null );

        inStream.close();



        String uri = res1.getName();

        ruleAdministrator.registerRuleExecutionSet(uri, res1, null);



        RuleRuntime ruleRuntime = serviceProvider.getRuleRuntime();
```

//Acquiring the stateful session
```
        StatefulRuleSession statefulRuleSession =

            (StatefulRuleSession) ruleRuntime.createRuleSession(uri,new HashMap(),
RuleRuntime.STATEFUL_SESSION_TYPE);
            java.util.List JessInput = new ArrayList();
            dataElement obj =(dataElement)
paramHash.get(Integer.parseInt(pFeature.getValue(sFields.findField("PARAM")).toStrin
g()));

            Compound val = new Compound(Double.parseDouble
(pFeature.getValue(sFields.findField("RESULT_VAL")).toString()));
            Ulimit ulimitObj = new Ulimit(obj.Ulimit);
            Llimit llimitobj = new Llimit(obj.Llimit);
```

//Creating the objects which contains the spatial data , that has to be sent into
//jess for processing
```
        JessInput.add(val);
        JessInput.add(ulimitObj);
        JessInput.add(llimitobj);
        statefulRuleSession.addObjects (JessInput);

        statefulRuleSession.executeRules();

        ruleValues = null;
        ruleViolate = false;
```

//Once the processing is finished, the objects returned by the jess is checked
//whether the flag is on or not.

```
if(val.getmcl() == 9999.99)
{
ruleValues = "+";
ruleViolate = true;

}
else if(val.getmcl() == 8888.88)
{
 ruleValues = "-";
 ruleViolate = true;
}
else
{
 ruleValues = "0";
}


if(ruleViolate)
{
```

//If the object is an outlier write it into the XY.txt

```
        LatLong.print
(Double.parseDouble(pFeature.getValue(sFields.findField("LONGITUDE_")).toString())
+" ");


        LatLong.print( Double.parseDouble(pFeature.getValue
(sFields.findField("__LATITUDE")).toString())+" ");


    // LatLong.print(obj.Name+" ");


    LatLong.print(ruleValues+" ");

    LatLong.println(obj.Name+" ");


    out.println("FLAGGED -- "+ruleValues);

}
```

```
            statefulRuleSession.release();


        pFeature = pFeatureCursor.nextFeature();
        }
    }
    catch(Exception e)
    {
        out.println("Exception is"+e);
    }}
        try {
            //Creating a new shape file using the data stored in XY.txt
            String strFolder = new String("C:\\Program Files\\Apache Software
Foundation\\Tomcat 5.5\\webapps\\mapserver\\arsenic\\");
            String sessionid =  (String) session.getId();
            String strName = new String("ARSENIC"+sessionid); // Dont include .shp
extension
            String strShapeFieldName = new String("Shape");


            //Open the folder to contain the shapefile as a workspace
            out.println("Open " + strFolder +
                        " to contain the shapefile as a workspace");


            //Set up a simple fields collection
            IWorkspaceFactory pWorkspaceFactory = new
IWorkspaceFactoryProxy(serverContext.createObject(ShapefileWorkspaceFactory.getCls
id()));
            IFeatureWorkspace pFWS = new
IFeatureWorkspaceProxy(pWorkspaceFactory.openFromFile(strFolder, 0));
            IFields pFields = null;
            IFieldsEdit pFieldsEdit = null;
            pFields = new Fields(serverContext.createObject(Fields.getClsid()));

            pFieldsEdit = (IFieldsEdit) pFields;



            IField pField = null;
            IFieldEdit pFieldEdit = null;


            //Make the shape field it will need a geometry definition, with a spatial
//reference
            pField = new Field(serverContext.createObject(Field.getClsid()));
            pFieldEdit = new IFieldEditProxy(pField);
            pFieldEdit.setName(strShapeFieldName);
            pFieldEdit.setType(esriFieldType.esriFieldTypeGeometry);
```

```
        IGeometryDef pGeomDef = null;
        IGeometryDefEdit pGeomDefEdit = null;
        pGeomDef = new
GeometryDef(serverContext.createObject(GeometryDef.getClsid()));
        pGeomDefEdit = new IGeometryDefEditProxy(pGeomDef);
        pGeomDefEdit.setGeometryType(esriGeometryType.esriGeometryPoint);

        ISpatialReferenceFactory2 spaRefFact;
        spaRefFact = new
SpatialReferenceEnvironment(serverContext.createObject(SpatialReferenceEnvironment.
getClsid()));
        IGeographicCoordinateSystem geoCoordSys;
        geoCoordSys = spaRefFact.createGeographicCoordinateSystem(
            esriSRGeoCSType.esriSRGeoCS_NAD1983);
        ISpatialReference sRef;
        sRef = new ISpatialReferenceProxy(geoCoordSys);
        pGeomDefEdit.setSpatialReferenceByRef(sRef);

        pFieldEdit.setGeometryDefByRef(pGeomDef);

        pFieldsEdit.addField(pField);

// Add X, Y, and Value fields
// NOTE: We do not have to explicitly add fields for Lat. and Long. b/c the
//shape field holds this type of data.  It is sometimes useful to have the Lat.
//and Long. fields
// explicitly in the table, however, and we do this here.
        IField ipX = new Field(serverContext.createObject(Field.getClsid()));
        IField ipY = new Field(serverContext.createObject(Field.getClsid()));
        IField ipValue = new Field(serverContext.createObject(Field.getClsid()));
        IField ipCompound = new Field(serverContext.createObject(Field.getClsid()));


        int len = 30;
        pFieldEdit = new IFieldEditProxy(ipX);
        pFieldEdit.setLength(len);
        pFieldEdit.setName("Longitude");
        pFieldEdit.setType(esriFieldType.esriFieldTypeDouble);
        pFieldsEdit.addField(ipX);

        pFieldEdit = new IFieldEditProxy(ipY);
        pFieldEdit.setLength(len);
        pFieldEdit.setName("Latitude");
        pFieldEdit.setType(esriFieldType.esriFieldTypeDouble);
        pFieldsEdit.addField(ipY);
```

```
            pFieldEdit = new IFieldEditProxy(ipValue);
            pFieldEdit.setLength(30);
            pFieldEdit.setName("RuleResults");
            pFieldEdit.setType(esriFieldType.esriFieldTypeString);
            pFieldsEdit.addField(ipValue);

            pFieldEdit = new IFieldEditProxy(ipCompound);
            pFieldEdit.setLength(30);
            pFieldEdit.setName("Compound");
            pFieldEdit.setType(esriFieldType.esriFieldTypeString);
            pFieldsEdit.addField(ipCompound);

        //Create the shapefile (some parameters apply to geodatabase options and can be
defaulted as Nothing)
            out.println("Create " + strName);

            pFeatClass = pFWS.createFeatureClass(strName, pFields,null, null,
                                esriFeatureType.esriFTSimple,
                                strShapeFieldName, "");

        }
        catch (java.lang.Exception e) {
          out.println("Exception :" + e.getMessage());
            }

/***************************************************************
***********************************************/
        try {

        // use CreateFeature/Store
            // Get the column index in table for X, Y, Value fields
            IFields ipFieldsNew;
            ipFieldsNew = pFeatClass.getFields();
            int indX, indY, indV;
            indX = ipFieldsNew.findField("Longitude");
            indY = ipFieldsNew.findField("Latitude");
            indV = ipFieldsNew.findField("RuleResults");

        BufferedReader br = new BufferedReader(new java.io.FileReader("C:\\Program
Files\ Apache Software Foundation\\Tomcat 5.5\\webapps\\mapserver\\WEB-
INF\\classes\\RuleEngine\\XY.txt"));
            String str;
            str = br.readLine();
            // Loop through the X, Y, Value tuples in from text file and update the feature
buffer
            double xx = 0;
```

```
double yy = 0;
String vv = null;
String uu = null;
while (str != null) {
  java.util.StringTokenizer stk = new java.util.StringTokenizer(str,
    " ");
  while (stk.hasMoreElements() && stk.countTokens() == 4) {
    xx = Double.parseDouble( (String) stk.nextElement());
    yy = Double.parseDouble( (String) stk.nextElement());
    vv = (String)stk.nextElement();
    uu = (String)stk.nextElement();


  out.println("Note"+vv);
  }

  out.println("Longitude: " + xx + " Latitude: " + yy +
          " RuleResults " +
          vv+"Compound"+uu);
  // Create a new feature
  IFeature ipFeat;
  ipFeat = pFeatClass.createFeature();

  // Create a new point and add it to the shape field
  IPoint ipPoint = new Point(serverContext.createObject(Point.getClsid()));
  ipPoint.setX(xx);
  ipPoint.setY(yy);
  IGeometry ipGeom = new IGeometryProxy(ipPoint);
  ipFeat.setShapeByRef(ipGeom);
```

// NOTE: We do not have to explicitly add the data for Lat. and Long. b/c
//the shape field holds this type of data. It is sometimes useful to have the Lat. and
//Long. data explicitly in the table, however, and we do this here.

```
  Object x = new Double(xx);
  Object y = new Double(yy);
  Object v = new String(vv);
  Object u = new String(uu);

  ipFeat.setValue(indX, x);
  ipFeat.setValue(indY, y);
  ipFeat.setValue(4, v);
  ipFeat.setValue(5, u);
```

// Persist your changes by calling Store()

```
        ipFeat.store();
          str = br.readLine();
          }


    }
    catch (IOException ex) {
      out.println("Exception" + ex.getMessage());
    }
    catch (NumberFormatException ex) {
    }




/*****************************************************************
*******************************************/
        //IFeatureClass arClass =arsenic_Layer.getFeatureClass();

      arsenic_Layer.setFeatureClassByRef(pFeatClass);
      arsenic_Layer.setName("Arsenic");
    ILayer layer = arsenic_Layer;
        //The new layer with the outliers is added to the MapControl in the user
//interface
    map.addLayer(layer);
    out.println("The new map layer count - "+map.getLayerCount());
    wToc.reload();
    agscxt.reloadDescriptions();

    agscxt.refresh();
    agscxt.refresh(AGSRefreshId.LAYER_VISIBILITY_TOGGLED );




fc.getApplication().createValueBinding("#{rulesSelect.cmdClick}").setValue(fc,false);



        %>
          <c:out value="${selectionCount}"/>



          <script>
            var layerIdElem =
document.forms[docFormID].elements['frmBuffer:selSelectedLayerID'];
                document.write(layerIdElem.options [layerIdElem.selectedIndex].text);
```

```
        </script>
        <c:out value="${resource['Buffer.FeaturesSelected']}" />

      </c:if>

<%--------------------------------------------------%>
<table>
<br><br><br>
<td>

<td>
<jsfh:selectManyCheckbox value="#{rulesSelect.JessRules}">

< jsfh:selectManyCheckbox>

</td>
</tr>

</table>

<table>
<tr>

   &nbsp   &nbsp   

<jsfh:commandButton value="Run"
action="#{rulesSelect.save}" />

</tr>
</table>
</ags:context>

<%-- Hightlight the current active tool on the tool bar--%>
<script
language="JavaScript">HighlightTool(document.forms[docFormID].elements["Map0_m
ode"].value);</script>
</jsfh:form>
<%--<form name ="Jess" method="POST" action="faces/mapviewer.jsp">
<p 
<b> Rules in the Jess</b></p>
<p       <input type="checkbox"
name="arsenic_CB" value="arsenic_ON" /> 
   Arsenic</p>
<p   

<input type="checkbox" name="Arsenice0" value="ON" />  Bromide</p>
```

```
<INPUT TYPE="submit" NAME="InvokeRules" VALUE="Submit">

</form>--%>
</body>
</html></jsfc:view>
```

//**java backend bean for run button**
RuleSelect.java

```java
public class rulesSelect {
private String[] JessRules;
private boolean cmdClick;

public String[] getJessRules( ) {
return JessRules;
}
public void setJessRules (String[] JessRules) {
this.JessRules = JessRules;
}
public boolean getCmdClick()
{
   return cmdClick;
}
public void setCmdClick(boolean i)
{
    this.cmdClick = i;
}

public String save()
{
   cmdClick = true;
   return "success";
}
}
```

Appendix- II

GeoExpert – Stand-alone Framework

IFrame.java(*Some of the code in IFrame.java is copied from mapviewer template in java ADF , samples in http://edn.esri.com, JSR-94 sample code in http://java.sun.com/developer/technicalArticles/J2SE/JavaRule.html)*

```java
import java.io.*;
import java.awt.*;
import java.awt.event.*;
import javax.swing.*;
import javax.swing.border.Border;
import java.util.*;
import javax.swing.filechooser.*;
import com.esri.arcgis.beans.map.MapControl;
import com.esri.arcgis.beans.pagelayout.PageLayoutBean;
import com.esri.arcgis.system.AoInitialize;
import com.esri.arcgis.system.EngineInitializer;
import com.esri.arcgis.system.esriLicenseProductCode;
import com.esri.arcgis.beans.TOC.*;
import com.esri.arcgis.beans.toolbar.*;
import com.esri.arcgis.controlcommands.*;
import com.esri.arcgis.systemUI.esriCommandStyles;
import javax.rules.admin.*;
import javax.rules.RuleServiceProvider;
import javax.rules.RuleServiceProviderManager;
import javax.rules.admin.Rule;
import javax.rules.*;
import com.esri.arcgis.carto.*;
import com.esri.arcgis.geodatabase.*;
import com.linar.jintegra.AutomationException;
import com.esri.arcgis.carto.IMap;
import com.esri.arcgis.datasourcesfile.*;
public class IFrame extends JFrame implements ActionListener{


    JFileChooser fc;

    JMenuBar jmb;
    JMenu actionmenu,toolmenu;
    JMenuItem mi,mi_saveas,mi_exit,mi_DQ;
    String path = new String();
    PageLayoutBean pageLayout;
    JButton run =  new JButton("Run");;
```

```
JComboBox ruleset= new JComboBox();
JOptionPane DQ_Pane;
JScrollPane scrollPane;
TOCBean tocb;

ToolbarBean DQ_Toolbarbean;
JLabel Select_Rule;
JPanel innerPanel;
JPanel outerPanel;
JPanel ruleDisplay;
private static final String RULE_SERVICE_PROVIDER = "org.jcp.jsr94.Jess";
int numberofRules;
JCheckBox ruleBox;
Rule rule01;
JPanel pLayout_ruleDisplay = new JPanel(new BorderLayout());
JPanel ruleSelect = new JPanel(new GridLayout());
JPanel objectPanel = new JPanel(new GridLayout());
Border ruleBorder;
int LayerCount,dataCount,count;
Double Latitude,Longitude;
com.esri.arcgis.carto.Map map;
FeatureLayer FLayer;
IRow row;
IFeature feature;
IFields fields;
RuleExecutionSet ruleExecSet;
RuleServiceProvider serviceProvider;
RuleAdministrator ruleAdministrator;
PrintStream LatLong;
CreateShapeFromText shapeFromText;
 IFeatureClass fClass;
public void RuleBase()
{
    try {
```

//Declaring the rule service provider
```
        Class.forName("org.jcp.jsr94.Jess.RuleServiceProviderImpl");
        serviceProvider =
RuleServiceProviderManager.getRuleServiceProvider(RULE_SERVICE_PROVIDER);
        ruleAdministrator = serviceProvider.getRuleAdministrator();
```
//Creating the inputstream object for reading rules into the jess from XML file
```
        InputStream inStream = this.getClass().getResourceAsStream("tck.xml");
        System.out.println("Acquired InputStream to RI tck_res_1.xml: " + inStream);

        ruleExecSet =
```

//Creating the rule execution set
```
        ruleAdministrator.getLocalRuleExecutionSetProvider(

        null).createRuleExecutionSet( inStream, null );

    inStream.close();
```
//Storing the rules in the list object
```
        java.util.List results = ruleExecSet.getRules();

        Iterator itr = results.iterator();

        numberofRules = results.size();
        System.out.println(numberofRules);

        ruleDisplay = new JPanel(new GridLayout(0,1));
        ruleDisplay.setSize(300,200);
    // ruleDisplay.setBackground(Color.white);
        ruleBorder = BorderFactory.createTitledBorder("Select Rules");
        ruleDisplay.setBorder(ruleBorder);
```

//Creating the userinterface according to the rules in the jess, so that the
//user will have an option to select rules he wants to execute the facts against
```
        for(int i =0; i<numberofRules;i++)
        {
            rule01=(Rule)results.get(i);
            ruleBox = new JCheckBox((rule01.getName()));
            System.out.println(rule01.getName());
            ruleBox.setBackground(Color.white);
            ruleDisplay.add(ruleBox);
        }



        }
    catch(Exception E)
    {
        System.out.println("Error: "+E);
    }


            }

public boolean checkData(Double input)
{
    int flag_ArsenicObject = 0;
    try
```

```
{

// Acquiring rule runtime

   String uri = ruleExecSet.getName();

   ruleAdministrator.registerRuleExecutionSet(uri, ruleExecSet, null);


   RuleRuntime ruleRuntime = serviceProvider.getRuleRuntime();


// Acquiring the statefull rule session

StatefulRuleSession statefulRuleSession =

      (StatefulRuleSession) ruleRuntime.createRuleSession(uri,

      new HashMap(), RuleRuntime.STATEFUL_SESSION_TYPE);


//Creating the data object that holds the spatial data, which is sent into jess

   Arsenic ArsenicObject = new Arsenic(input);

   // Create an input list.

   java.util.List JessInput = new ArrayList();

   JessInput.add(ArsenicObject);



      statefulRuleSession.addObjects(JessInput);
      statefulRuleSession.executeRules();
      if(ArsenicObject.getmcl()==9999.99)
         flag_ArsenicObject=1;

      statefulRuleSession.release();


} catch (NoClassDefFoundError e) {

   if (e.getMessage().indexOf("JessException") != -1) {
```

```
            System.err.println("Error: The RI Jess could not be found.");

        } else {

            System.err.println("Error: " + e.getMessage());

        }

    } catch (Exception e) {

        e.printStackTrace();
        return false;
    }
    if(flag_ArsenicObject == 1)
return true;
    else
        return false;
}
```

//function to extract the selected features

```
public void afterSelection(com.esri.arcgis.carto.Map mo) throws Exception
{
    IFeatureSelection selectedFeature= new IFeatureSelectionProxy(mo);
    SelectionSet dataset = (SelectionSet)selectedFeature.getSelectionSet();
    int count = dataset.getCount();
    System.out.println(count);
    System.out.println("aditya");
}
```

//Constructor to render the user interface
```
public IFrame() {

    this.setName("GeoExpert");
    this.setLayout(new BorderLayout());
    pageLayout = new PageLayoutBean();
    //mxd=new MapControl();
    tocb = new TOCBean();
    Select_Rule = new JLabel("Select a Rule");
    fc = new JFileChooser();
    jmb = new JMenuBar();
    DQ_Toolbarbean = new ToolbarBean();
    toolmenu=new JMenu("Tools");
    actionmenu = new JMenu("File");
    actionmenu.add(mi = new JMenuItem("Open"));
    mi.setAccelerator(KeyStroke.getKeyStroke(
```

```
                KeyEvent.VK_O, ActionEvent.ALT_MASK));
mi_saveas = new JMenuItem("SaveAs",
                KeyEvent.VK_S);
mi_saveas.setAccelerator(KeyStroke.getKeyStroke(
     KeyEvent.VK_S, ActionEvent.ALT_MASK));
actionmenu.add(mi_saveas);
mi_exit = new JMenuItem("Exit",
                KeyEvent.VK_X);
mi_exit.setAccelerator(KeyStroke.getKeyStroke(
     KeyEvent.VK_X, ActionEvent.ALT_MASK));
actionmenu.add(mi_exit);


mi.setActionCommand("Open");
mi.addActionListener(this);

jmb.add(actionmenu);
//Creating the tool menu

mi_DQ = new JMenuItem("DataValidator",
                KeyEvent.VK_D);
mi_DQ.setAccelerator(KeyStroke.getKeyStroke(
     KeyEvent.VK_D, ActionEvent.ALT_MASK));
mi_DQ.setActionCommand("DQ");
mi_DQ.addActionListener(this);
toolmenu.add(mi_DQ);
jmb.add(toolmenu);
actionmenu.getPopupMenu().setLightWeightPopupEnabled(false);
toolmenu.getPopupMenu().setLightWeightPopupEnabled(false);
ruleset.setLightWeightPopupEnabled(false);
this.setJMenuBar(jmb);
     //ruleset = new JComboBox();
        innerPanel= new JPanel();
        innerPanel.setLayout(new BorderLayout());

        outerPanel = new JPanel();
        outerPanel.setLayout(new BorderLayout());
        // innerPanel.setLayout(new BorderLayout());
        //innerPanel.add(Select_Rule,BorderLayout.CENTER);
        // innerPanel.add(ruleset,BorderLayout.EAST);
        outerPanel.add(DQ_Toolbarbean,BorderLayout.CENTER);
        outerPanel.add(run,BorderLayout.EAST);
     this.getContentPane().add(outerPanel, BorderLayout.NORTH);


     try{
```

//Creating the map navigation tools like zoomin , zoomout etc

```
        DQ_Toolbarbean.addItem( new ControlsPageZoomInTool(), 0, -1, false,
0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsPageZoomOutTool(), 0, -1, false,
0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsPagePanTool(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new
ControlsPageZoomWholePageCommand(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new
ControlsPageZoomPageToLastExtentBackCommand(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new
ControlsPageZoomPageToLastExtentForwardCommand(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsMapZoomInTool(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsMapZoomOutTool(), 0, -1, false,
0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsMapPanTool(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsMapFullExtentCommand(), 0, -1,
false, 0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsSelectTool(), 0, -1, false, 0,
esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsSelectFeaturesTool(), 0, -1, false,
0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsSaveAsDocCommand(), 0, -1,
false, 0, esriCommandStyles.esriCommandStyleIconOnly );
        DQ_Toolbarbean.addItem( new ControlsClearSelectionCommand(), 0, -1,
false, 0, esriCommandStyles.esriCommandStyleIconOnly );




        }
        catch(IOException z)
        {
           System.out.println("ERROR"+z);
        }
```

```java
        this.pack();
        this.setVisible(true);

    }

public void actionPerformed(ActionEvent e){

String actionName = e.getActionCommand();


    if (actionName.equals("Open"))
    {
        RuleBase();
        int returnVal = fc.showOpenDialog(this);

        if (returnVal == JFileChooser.APPROVE_OPTION) {
            File file = fc.getSelectedFile();
            path = file.getAbsolutePath();
            try{
            //Function to open the MXD file into the map control
                if (pageLayout.checkMxFile(path))
                pageLayout.loadMxFile(path,null);
                // pageLayout.zoomToWholePage() ;

                tocb.setBuddyControl(pageLayout);
                JPanel contentPane = new JPanel(new BorderLayout());
                JPanel mappanel = new JPanel(new BorderLayout());
                // contentPane.add(tocb, BorderLayout.WEST);
                mappanel.add(pageLayout, BorderLayout.CENTER);
                DQ_Toolbarbean.setBuddyControl(pageLayout);
                Select_Rule.setText("Slots in Rule based Expert System for this Rule");
                ruleSelect.add(Select_Rule);
                contentPane.add(ruleDisplay,BorderLayout.NORTH);
                contentPane.add(ruleSelect,BorderLayout.CENTER);
                pLayout_ruleDisplay.add(mappanel,BorderLayout.CENTER);
                pLayout_ruleDisplay.add(contentPane,BorderLayout.EAST);

this.getContentPane().add(pLayout_ruleDisplay,BorderLayout.CENTER);
                this.getContentPane().add(tocb,BorderLayout.WEST);
                //fl=new FeatureLayer();


        run.addActionListener(new ActionListener() {
            public void actionPerformed(ActionEvent ae){
                try{
```

```
LatLong = new PrintStream( new FileOutputStream("XY.txt"));
//int LatLong_Flag = 0;
// afterSelection(mapObject);

map   = new
com.esri.arcgis.carto.Map(pageLayout.getActiveView().getFocusMap());
         LayerCount = map.getLayerCount();

         for(int i=0;i<LayerCount;i++)
         {

         FLayer = new FeatureLayer(map.getLayer(0));
         ISelectionSetProxy dataset =(ISelectionSetProxy)
FLayer.getSelectionSet();
         dataCount = dataset.getCount();
         //System.out.println(dataCount);
         ICursor cur[] = new ICursor[dataCount];
         dataset.search(null,true,cur);
         row = cur[0].nextRow();


         while(row !=null){
//Feature feature = new IFeatureProxy(row);
         // row.getOID();
         // System.out.println(row.getOID());
            feature = new IFeatureProxy(row);
      //Create new fields with the outliers detected in the new shape file
         fields = feature.getFields();
         count = fields.findField("RESULT_VAL");
         //System.out.println(feature.getValue(count));
         if( checkData((Double)(feature.getValue(count))))
         {
            count = fields.findField("LONGITUDE_");
            Longitude = (Double)(feature.getValue(count));
            LatLong.print(Longitude+" ");

            count = fields.findField("__LATITUDE");
            Latitude = (Double)(feature.getValue(count));
            LatLong.print(Latitude+" ");

            count = fields.findField("RESULT_VAL");

            LatLong.println((Double)(feature.getValue(count)));

         }
         row = cur[0].nextRow();
```

```
            }

            }



            }
            catch(AutomationException as)
            {
                System.out.println("Error-->"+as.getDescription()+as.getSource());
                System.out.println("Run Button Pressed");


            }
            catch(IOException ai)
            {


            }
            //calling function to ceate shapfile from text which contains the data
//related to the outliers
            shapeFromText = new CreateShapeFromText();
            fClass = shapeFromText.createShapefile();
    if (fClass != null) {
      shapeFromText.AddPoints(fClass, false);
    }
    else {
      System.out.println("Unable to create feature class");
    }
        try{
                ShapefileWorkspaceFactory wsf = new ShapefileWorkspaceFactory();
                Workspace work = new Workspace(wsf.openFromFile("C:\\shapefile\\",0));
                IFeatureClass featureClass = work.openFeatureClass("ARSENIC.shp");



            FeatureLayer featureLayer = new FeatureLayer();
featureLayer.setFeatureClassByRef(featureClass);
map.addLayer(featureLayer);
tocb.repaint();
        }
        catch(IOException FLA)
        {
```

```
        }
          }
  });


          }
        catch(IOException s)
        {
            System.out.println("ERROR"+s);
        }
      this.setVisible(true);
    //}
    //if(actionName.equals("DQ"))
    //{
    // this.setVisible(true);
      }

    else
      {

      }
    }


  }

public static void main(String[] args) throws IOException{

EngineInitializer.initializeVisualBeans();
        final AoInitialize aoInit = new AoInitialize();
        aoInit.initialize(esriLicenseProductCode.esriLicenseProductCodeEngine);
JFrame.setDefaultLookAndFeelDecorated(true);
  JDialog.setDefaultLookAndFeelDecorated(true);
  IFrame FC = new IFrame();
  FC.setSize(800,600);
  FC.setVisible(true);



  }
}
```

References

1) Geographical Information Systems - www.gis.com

2) *Spatial Databases A Tour* - Shashi Shekhar and Sanjay Chawla

3) ArcGIS Desktop- http://www.esri.com/software/arcgis/about/desktop.html

4) ArcObjects - http://www.esri.com/software/arcgis/about/ArcObjects.html

5) ArcObjects- http://edndoc.esri.com/ArcObjects/9.1/

6) ArcGIS Engine - http://www.esri.com/software/arcgis/arcgisengine/

7) *ArcGIS Server Administrator and Developer Guide* – ESRI Press

8) *Principles of Artificial Intelligence and Expert Systems Development* – David W.Rolston

9) *Artificial Intelligence and Expert Systems for Engineers* – C.S.Krishnamoorthy,S. Rajeev

10) Basin water quality terminology

   http://bcn.boulder.co.us/basin/natural/wqterms.html#do

11) *Expert Systems ,Introduction to the technology and applications*– Dieter Nebendahl

12) Rebecca Bari Buchheit, *Vaccum – Automated Procedures for assessing and cleansing Civil Infrastructure data,* Carnagie Mellon Univeristy, 1998

13) *Data Mining Concepts and Techniques* – Jiawei Han , Micheline Kamber

14) Suresh C Jayanty and Dr. Uta Ziegler, *A Rule Based Expert System Framework for Small Water Systems*, Western Kentucky University, 2005

15) Yichun Xie and George D. Graettinger, *An Integrated ARCVIEW Expert System for Analyzing Contaminated Sediments in the Great Lakes Basin*, 1996 ESRI User Conference

16) ARCVIEW, http://www.esri.com/software/arcview/

17) Aditya Tadakaluru, Karla Andrew, Mostafa Mostafa, Andrew Ernest, *"GeoExpert A Framework for Data Quality in Spatial Databases," cimca*, pp. 557-561, International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce Vol-2 (CIMCA-IAWTIC'06), 2005.

18) The Model View Controller architecture-

   http://rd13doc.cern.ch/Notes/004/Note004-7.html

19) JSR-94, http://www.jcp.org/aboutJava/communityprocess/review/jsr094/

20) Java Server Faces- By Hans Bergsten

21) http://java.sun.com/javaee/Java Serverfaces

22) http://Tomcat.apache.org/

23) http://herzberg.ca.sandia.gov/Jess/

24) http://java.sun.com/developer/technicalArticles/J2SE/JavaRule.html

25) http://nhd.usgs.gov/

26) *Expert systems for the evaluation of data quality for establishing the Recommended Dietary Allowances.* JNutr.1996 Sep;126(9 Suppl):2329S-2336S. PMID: 8811794 [PubMed - indexed for MEDLINE]

27) *Data quality fuzzy expert system* - F. Moura-Pires, R. A. Ribeiro, A. Pereira, F. J. Varas, G. Mantovani, A. Donati - Proceedings of the 10th Mediterranean Conference on Control and Automation - MED2002 Lisbon, Portugal, July 9-12,

2002.

28) *GIS and Expert System Integration for Land Development Village, Thailand*

- Kanjana Chuenpichai, Sutisak Chuenpichai and Ard Somrang

29) *A Software Engineering View of Data Quality* -Mónica Bobrowski, Martina

Marré, Daniel Yankelevich ,Departamento de Computación, FCEyN, Universidad

de Buenos Aires, Argentina

30) *Data Quality Assessment*- Leo L. Pipino, Yang W. Lee, and Richard Y. Wang

31) Software Quality Metrics for Object-Oriented Environments - Dr. Linda H.

Rosenberg, Lawrence E. Hyatt