ASSESSMENT OF GROUND WATER POTENTIAL FROM REMOTE SENSING DATA USING CASE BASED REASONING TECHNIQUE

J.L.Bind. Santanu Chaudhury, and V.K.Panchal
DTRL/DRDO, Delhi
IIT Delhi

KEY WORDS: Decision Support System, CBR, Ground Water Assessment, Reasoning Modality, Similarity Matching, Geographical Information System (GIS)

ABSTRACT

In this article we propose a Decision Support System using Case Based Reasoning (CBR) Technique for assessing ground water potential. Ground water is found in bedrock or in thick masses of overburden many feet below the surface. Theoretically ground water is function of various terrain parameters viz. Land Use/ Land Cover, Geology, Soil, Landform and Lineament. Using this theoretical framework a Decision Support System (DSS) is being proposed to assess the ground water availability. Human decision-making process makes use of multiple reasoning modalities. Among them the more significant ones are: Rule Based Reasoning, Case Base Reasoning and Model Based Reasoning. Case Based Reasoning (CBR) approach makes efficient use of past experiences for interpreting present situation. Past experiences are stored in the case base in the form of cases. The cases are prepared on the basis of real data collected from the fields, covering five types of terrain, i.e., plain, desert, marshy, costal and deltaic regions. Case structure uses six terrain parameters based on their roll for ground water availability. These include Land Use/ Land Cover, Geology, Soil, Slope, Landform and Lineament. Finally, the results are tested in the field and presented in the end of this article.

1. INTRODUCTION

Human decision-making process makes use of multiple reasoning modalities. Among them the more significant ones are: Rule Based Reasoning, Case Based Reasoning and Model Based Reasoning. Rule based paradigm is useful for utilizing generic knowledge for the interpretation of queries. Rules are used for encoding information that contains the management decisions under specific circumstances. Model based reasoning can also play an important role in the decision support system. In the absence of complete information or numerical simulation models for all possible scenarios, model based causal reasoning can be used to estimate or predict behavior or parameters of the physical system.

An important requirement for making decision by human being is the use of past experiences for interpreting present situations. Past experiences can be exploited efficiently by using Case Based Reasoning (CBR) methods [1]. Correlations between past and present scenarios can be established using cases depicting abstracted principles. Heuristic models can be used for assessing applicability of an example based solution strategy in a new scenario. These features provide the justification for structuring our tool as a Case Based Reasoner. CBR solves new problems by adapting previously successful solutions to similar problems. CBR does not require an explicit domain model, as it retrieves the most similar cases from case base then reuse the case(s) to attempt to solve the problem. Revise the case(s) using query case and retrieve case for proposed solution then retain the new solution as part of a new case. A summary of the above approaches may be found in [1].

This proposed method is suitable for complex scenario like terrain where no clear rules can be formulated, hence the main motivation to use CBR technique for predicting ground water availability. Ground water under the bedrock is function of various terrain parameters. Terrain parameters used for case formulation are cause of ground water or indicative of it. Some parameters have bee derived from remote sensing data, slope is derived from SRTM data and Geology is derived from Geological Survey of India archival. This proposed DSS is applied for assessing the ground water potential in Indian sub-continent. Finally result is tested and presented in this article. Results shows that the CBR techniques is an attractive and effective way to deals with a complex scenario like terrain where no clear cut rules can be formulated for a rule based Decision Support System. Finally the result is tested and found approximately 76% accurate.

In section 2 we describe the basics of CBR system. Complete system description is given in section 3. Section 4 concludes the paper.

2. CBR SYSTEM

2.1 Definitions

- Expert or knowledge based systems (KBS) are one of the success stories of Artificial Intelligence research. However, despite the undoubted success of model-based KBS in many sectors, developers of these systems have met with several problems:
- · Knowledge elicitation is a difficult process often being referred to as the knowledge elicitation bottleneck.
- · Implementing KBS is a difficult process requiring special skills and often taking a lot of time.
- · Once implanted model-based KBS are often slow and are unable to access or manage large volumes of information.
- · Once implemented they are difficult to maintain.
- Solutions to these problems have been sought through better elicitation techniques and tools, better KBS shells and environments, improved development methodologies, knowledge modeling languages and ontologies, facilitating the co-operation between KBS and databases in expert databases and deductive databases and techniques and tools for maintaining systems [2, 3].
- However over the last few years an alternative reasoning paradigm and computational problem solving method

has increasingly attracted more and more attention. Case based reasoning (CBR) solves new problems by adapting previously successful solutions to similar problems. It has following advantages:

- CBR does not require an explicit domain model and so elicitation becomes a task of gathering case theories.
- Implementation is reduced to identifying significant features that describe a case, an easier task than creating an explicit model.
- By applying the database techniques large volumes of information can be managed.
- CBR systems can learn by acquiring new knowledge as cases, thus making maintenance easy.
- A Case Based Reasoner solves new problems by adapting solutions used to solve similar old ones. Hence, it can be stated that CBR system is a problem solving paradigm which involves matching the current problem against problems that are solved successfully in the past [1]. The processes involved in a typical CBR can be represented by a schematic cycle shown in Fig. 1.
- CBR is a cyclic process comprising of the following steps:
- RETRIEVE the most similar case(s)
- REUSE the case(s) to attempt to solve the problem
- REVISE the proposed solution if necessary
- RETAIN the new solution as a part of a new case
- A new problem is matched against cases in the case base and one or more similar cases are retrieved. A solution suggested by the matching cases is then reused and tested for success. Unless the retrieved case is a close match, the solution requires to be revised, usually with some form of human interpretation producing a new case that can be retained. Case revision is also called adaptation.

Figure 1: A typical CBR Cycle

In the Fig. 1 above, an initial description of the problem defines a new case. The new case is used to RETRIEVE a case from the collection of the previous cases. The retrieved case is combined with the new case - through REUSE, into a solved case (that is, the proposed solution to the initial problem). Through the REVISE process these solutions is tested for success (say, by applying it to some unseen/unknown examples) and are repaired if they fail. During RETAIN useful experience is retained for future reuse and the case base is updated by a new learned case, or by modification of some existing cases.

2.1. Cases and Their Representation

A case is a contextualized piece of knowledge representing an experience. It contains the past lesson that is the content of the case and the context in which the lesson can be used. Typically a case comprises of:

Problem: This describes the state of the world when the case occurred,

Solution: This states the derived solution to the problem, and/or

Outcome: This describes the state of the world after the case has occurred.

Cases which comprise of problems and their solutions can be used to derive solutions to new problems whereas cases comprising of problems and the outcomes can be used to evaluate new situations. If, in addition, such cases contain solutions, they can be used to evaluate the outcome of the proposed solutions and prevent potential problems. Cases can be represented in a variety of forms using the full range of AI representational formalisms including frames, objects, predicates, semantic nets and rules [1].

2.2. Case Retrieval

Given a description of a problem, a retrieval algorithm using the indexes in the case-memory, should retrieve the case most similar to the current problem or situation. Similarity can be modeled by various approaches like attribute based, structure based similarity. The retrieval algorithm relies on the indices and the organization of the memory to direct the search to potentially useful cases. Among well known methods for case retrieval is: nearest neighbor, induction, knowledge guided induction and template retrieval. These methods can be used alone or combined into hybrid retrieval strategies. The present article uses nearest neighbor method and is discussed below in brief.

Nearest neighbor: Similarity between stored cases and the new input case, based on matching a weighted sum of features. The similarity (that is, the proximity) of the target case to a source case for each attribute is determined. This measure may be multiplied by a weighing factor. Then the sum of similarity of all attributes is calculated. This can be represented by the equation

(1)

Where T is the target case, S is the source case, n is the number of attributes in each case, i is an individual attribute from 1 to n, f is a similarity function for attribute i in cases T and S, w is the weighting of the attribute i. Algorithms similar to this are used by most CBR tools to perform nearest neighbor retrieval. Similarities are usually normalized to fall within a range of 0 and 1 (where 0 is totally dissimilar and 1 is an exact match) or depicted as a percentage similarity [1, 4].

3. GROUND WATER PREDICTION SYSTEM

- Due to complex nature of terrain it is very difficult to model the behavior of it, hence, in such scenario Case Based Reasoning (CBR) methodology is most appropriate approach to solve a problem related to terrain. For assessment of ground water potential it is not possible to predict exactly the ground water table, hence we are proposing a decision support system, which will predict the ground water availability in the form of low, moderate and high. Here an assumption is made in the form of low, moderate, and high possibilities of ground water in respect to ground water table. Ground water low possibility indicates 20 mtr below the ground level, medium possibility indicates between 10 mtr to 20 mtr below the ground level and high possibility indicates 0 mtr to 10 mtr below the ground level. Number of cases is generated using field data and the same is stored in case base file. The typical structure of a case is shown below.
- LAND_FORM : FLOOD PLAIN
- GEOLOGY : YOUNGER ALLUVIUM
- SOIL : ALLUVIAL CLAY
- LAND_USE : AGRICULTURAL
- SLOPE : GENTLE
- LINEAMENT : ABSENT
- PROBLEM : GROUND WATER POSSIBILITY
- SOLUTION: HIGH

- Cases may comprise of problem space and solution space. Problem space may contain those terrain parameters, which are responsible for affecting the ground water table or those terrain parameters, which may indicate the status of ground water. Solution space of case may contain the status of ground water. Thematic maps of terrain parameters are derived from remotely-sensed images and stored in a Geographical Information System (GIS) framework. Image data used to derive the thematic maps is IRS LISS III [5]. In CBR system, the query case may comprise only problem space and it retrieve the similar case from the case base in response to a query case using similarity matching algorithm which is given below. Retrieved case is revised using query case to find the solution.
- System Architecture and Design of proposed work is given below which comprises mainly three parts.
- (1) Case Based Application Developer
- (2) Case Based Problem Solver
- (3) GIS Interface
- Architecture diagram of the developed system is given below.
- Case based reasoning (CBR) is a technology that allows to finds the analogies between a current working case and past experiences. It makes direct use of past experiences to solve a new problem by recognizing its similarity with specific known problems and by, at least partially, applying the known solution for the new problem. The general task that case based reasoning methods have to deal with to identify the current problem situation and find a past case similar to the current problem, evaluate the proposed solution and update the system by learning from this experience.
-
- 3.1. Case Based Application Developer
- The Case Base Application Developer module provide a shell with facilities for encoding of past cases, rules and modules and an in-built inference scheme which will be incorporated into each application being developed. The nature of the problem and the knowledge base itself will change but the change will not affect the shell which will be developed. The knowledge of past cases is stored in a case base. How to create a new case base, update an existing case base, mechanisms for encoding rules and models. Models are created for handling different applications. Architecture diagram of Application developer is given below.
-
- 3.2. Case Based Problem Solver
- Once the knowledge base has been set up, at run time the user is to be permitted to pose a query to the application. The Problem Solver needs to retrieve all cases in the case base similar to the query entered by the user. Then on the basis of the constraints imposed by the user, either a direct solution is provided from one of the cases, or the solutions are modified or the cases are combined to provide a solution. Generation of a solution can involve use of physical models. A user interface is required to provide.
-
- 3.3. Geographical Information System (GIS) Interface
- In this proposed work the thematic overlays of land use, land form, Soil, Geology, Slope and Lineament is stored in the GIS. A case generation module is provided to generate the quarry case which comprises a Data Access Object (DAO) driver to interface with GIS. This quarry generation module retrieves the terrain information from thematic maps stored in GIS corresponding to a location where the ground water is to be assessed [6]. In this proposed work, Cases stored in the case base are comprises of five types of terrain i.e. Plain, Desert, Marshy, Deltaic and Coastal.

Finally Quarry Generation module provides the interface between the developed system and GIS which is used to generate the quarry case. GIS comprises the thematic information of terrain parameters. This work is developed in VC++ 6.0 and Geo media GIS is used to store and manage the thematic data.

- 3.4. Results

Quarry case has been prepared of Punjab plain area and output from this developed system was 0.7 similarity matching. Final result was tested from the field data and it was found approx. 70 % accurate result. Developed system was also tested on standard data sets [7]. It gives approx. 70 to 80 % accuracy. Accuracy of results is fully dependent on cases generated from collected field data. Although sufficient case gathering representing all scenario of terrain may give good result and over the period of time learning from system itself, a robust system can be evolved [8].

4. CONCLUSIONS

This paper presents a DSS based on CBR approach. Cases have been prepared from real field data and quarry case is generated from remote sensing data, hence error due to generation of quarry case will lead to final results. Since output is obtained from past experience, hence very much suitable for terrain type application. The algorithm is suitable for terrain analysis because of its complex nature of data. The complexity is due to the fact that terrain data is large both in dimension and size, where the requirement is to achieve approximate, but effective solution fast.

5. REFERENCES

J.L. Kolodner, Case-Based Reasoing. San Mateo, Calif.: Morgan Kaufmann, 1993.

[2] S.K. Pal, T.S. Dillon, and D.S. Yeung, Soft Computing in Case-Based Reasoning. London: Springer Verlag, 2000.

[3] J.R.Quinlan, C4.5:Program for Machine Learning, San Mateo, CA:Morgan Kaufmann, 1993.

[4] R.K. De and S.K. Pal, “A Connectionist Model for Selection of Cases,” Information Sciences, vol. 132, pp. 179-194, 2001.

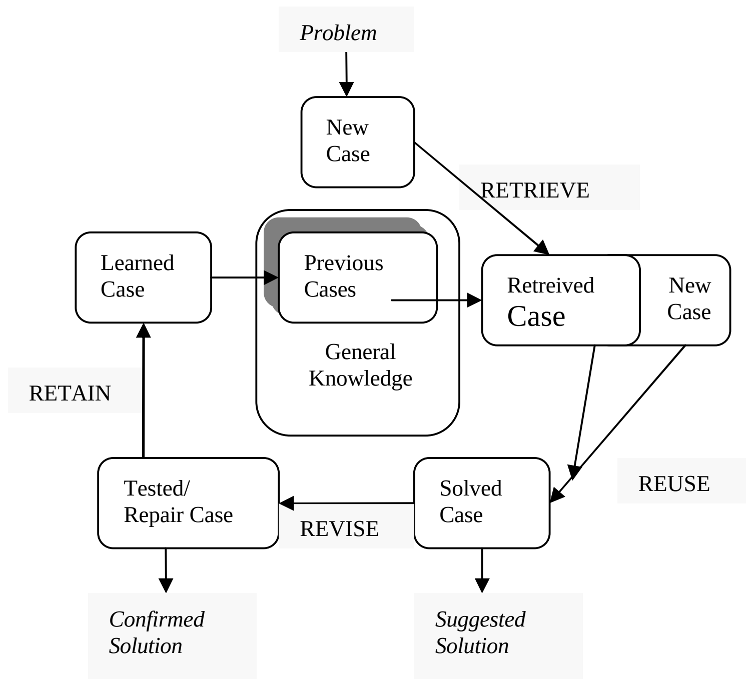
[5] James R. Irons, Richard A. Weismiller, Gary W. Petersen, “Theory and Applications of Optical Remote Sensing”, edited by Ghassem Asrar: John Wiley & Sons.

[6] Peter M. Atkinson, “Extracting Information from Remotely Sensed and GIS data,” Edited by Peter M. Atkinson: John Wiley & Sons, pp.263-268.

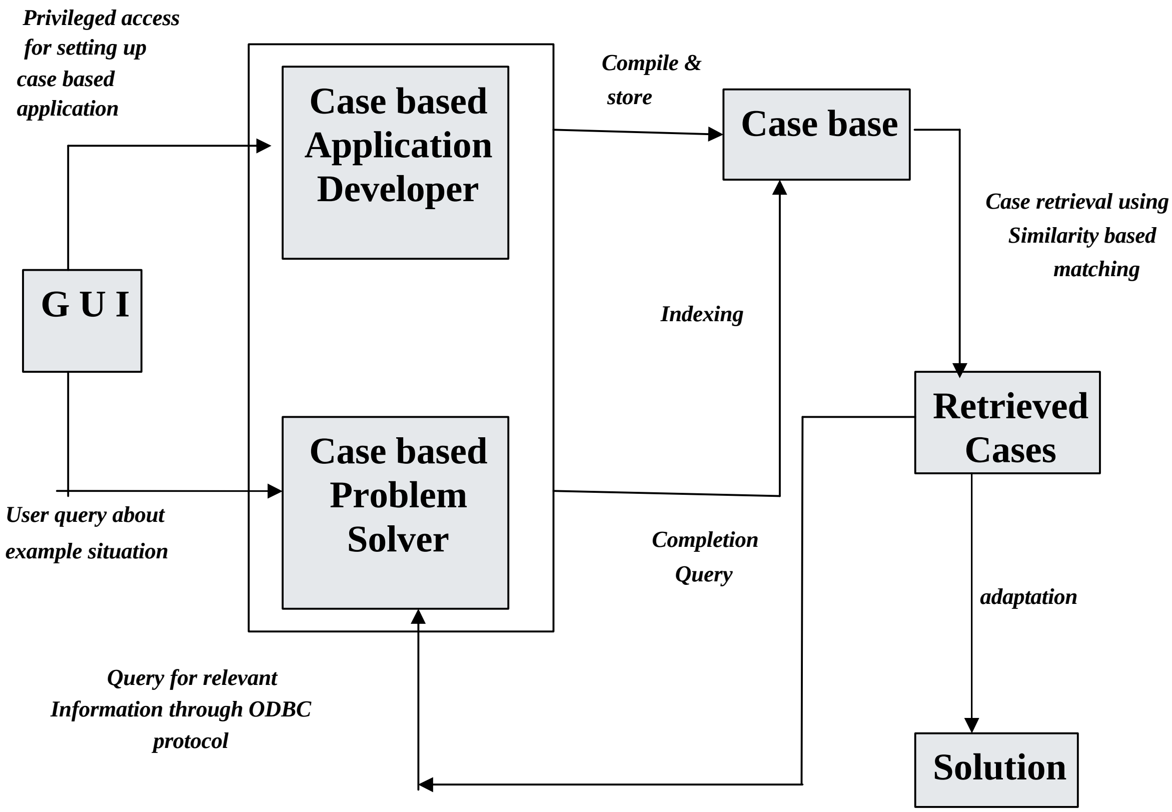
[7] C.L. Blake and C.J. Merz, UCI Repository of Machine Learning Databases, Univ. of California, Irvine, Dept.

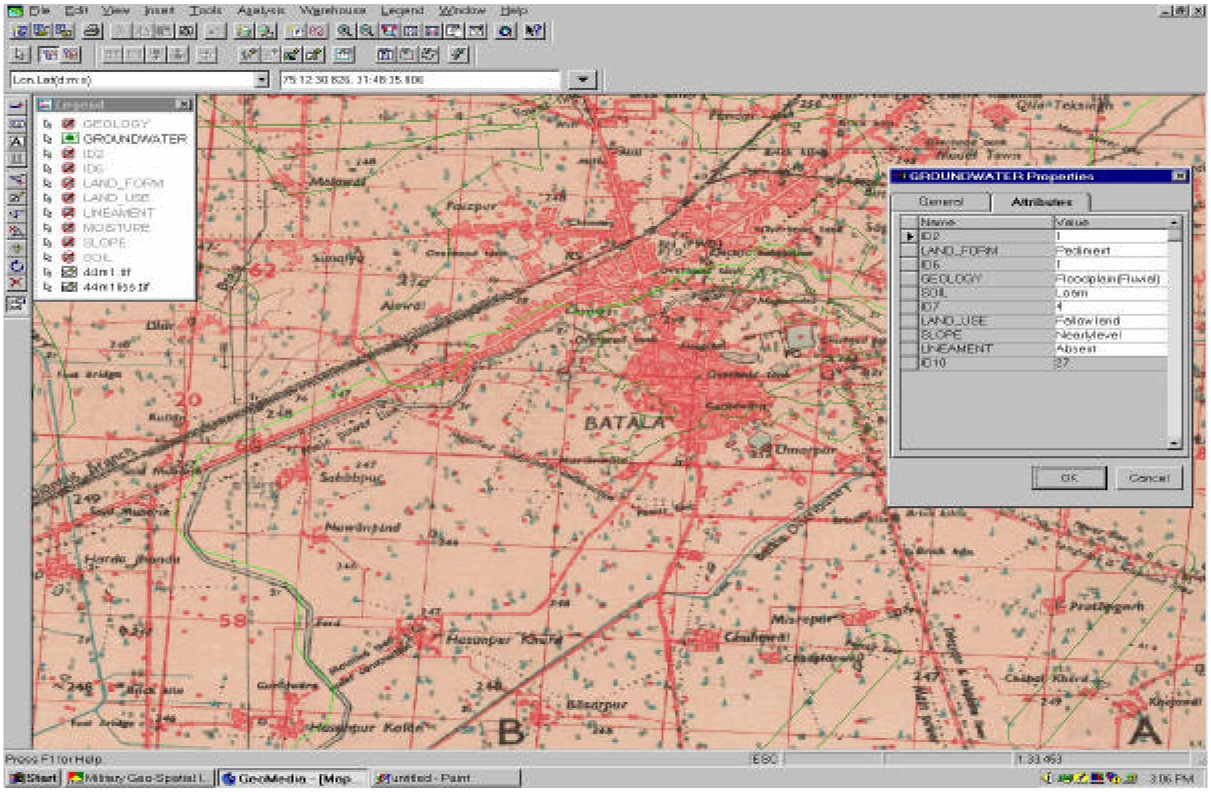
of Information and Computer Sciences, http://www.ics.uci.edu/~mlearn/ MLRepository.html, 1998.

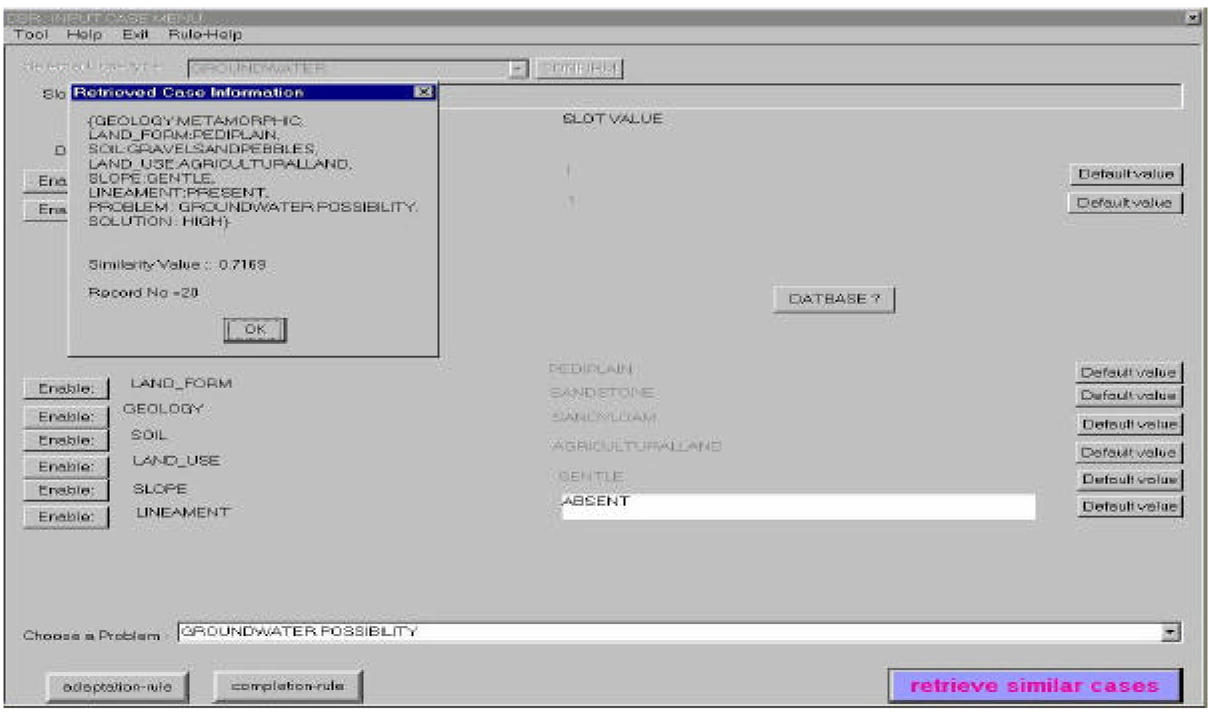
[8] W. Mark, E. Simoudis, and D. Hinkle, “Case-Based Reasoning: Expectations and Results,” AAAI Press, pp. 269-294, 1996.



Problem
New Case
RETRIEVE
Learned Case
Previous Cases
General Knowledge
Retreived Case
New Case
RETAIN
REUSE
Tested/ Repair Case
REVISE
Solved Case
Confirmed Solution
Suggested Solution


Privileged access
for setting up
case based
application
Case based
Application
Developer
Compile &
store
Case base
Case retrieval using
Similarity based
matching
G U I
Indexing
Retrieved
Cases
Case based
Problem
Solver
User query about
example situation
Completion
Query
adaptation
Query for relevant
Information through ODBC
protocol
Solution


File Edit View Insert Tools Analysis Warehouse Legend Window Help
GEOLOGY
GROUNDWATER
ID2
ID6
LAND_FORM
LAND_USE
LINEAMENT
MOISTURE
SLOPE
SOIL
GROUNDWATER Properties
General
Attributes
Name
Value
ID2
1
LAND_FORM
Pediment
ID6
1
GEOLOGY
Floodplain(Fluvial)
SOIL
Loam
ID7
4
LAND_USE
Fallow land
SLOPE
Near level
LINEAMENT
Absent
ID10
27
OK
Cancel
BATALA
Press F1 for Help
Start
3:06 PM


Tool Help Exit Rule-Help
GROUNDWATER
Retrieved Case Information
{GEOLOGY:METAMORPHIC,
LAND_FORM:PEDIPLAIN,
SOIL:GRAVELSANDPEBBLES,
LAND_USE:AGRICULTURALLAND,
SLOPE:GENTLE,
LINEAMENT:PRESENT,
PROBLEM: GROUNDWATER POSSIBILITY,
SOLUTION : HIGH}
Similarity Value :: 0.7169
Record No =20
OK
SLOT VALUE
Default value
DATBASE ?
Enable:
LAND_FORM
GEOLOGY
SOIL
LAND_USE
SLOPE
LINEAMENT
PEDIPLAIN
SANDSTONE
SANDYLOAM
AGRICULTURALLAND
GENTLE
ABSENT
Choose a Problem :
GROUNDWATER POSSIBILITY
adaptation-rule
completion-rule
retrieve similar cases