PERGAMON

Expert Systems with Applications 20 (2001) 251–260

Expert Systems with Applications

www.elsevier.com/locate/eswa

An intelligent decision support system for management of petroleum-contaminated sites

Liqiang Geng[a], Zhi Chen[b], Christine W. Chan[a,*], Gordon H. Huang[b]

[a]Department of Computer Science, University of Regina, Regina, Sask, S4S 0A2, Canada
[b]Faculty of Engineering, University of Regina, Regina, Sask S4S 0A2, Canada

Abstract

Groundwater and soil contamination resulted from LNAPLs (light nonaqueous phase liquids) spills and leakage in petroleum industry is currently one of the major environmental concerns in North America. Numerous site remediation technologies have been developed and implemented in the last two decades. They are classified as ex-situ and in-situ remediation techniques. One of the problems associated with ex-situ remediation is the cost of operation. In recent years, in-situ techniques have acquired popularity. However, the selection of the optimal techniques is difficult and insufficient expertise in the process may result in large inflation of expenses. This study presents an expert system (ES) for the management of petroleum contaminated sites in which a variety of artificial intelligence (AI) techniques were used to construct a support tool for site remediation decision-making. This paper presents the knowledge engineering processes of knowledge acquisition, conceptual design, and system implementation. The results from some case studies indicate that the expert system can generate cost-effective remediation alternatives to assist decision-makers.

Keywords: Expert system; Selection of remediation; Petroleum contamination; Fuzzy set

1. Introduction

Automation of engineering selection is important for the petroleum industries in which decision for a desired remediation technology at a contaminated site is critical for ensuring safety of the environment and the public. A variety of remediation methods/technologies are available. However, different contaminated sites have different characteristics depending on pollutants' properties, hydrological conditions, and a variety of physical (e.g. mass transfer between different phases), chemical (e.g. oxidation and reduction), and biological processes (e.g. aerobic biodegradation). Thus, the methods selected for different sites vary significantly. The decision for a suitable method at a given site often requires expertise on both remediation technologies and site hydrological conditions (Sims, Suflita & Russell, 1992).

In general, soil and groundwater remediation techniques can be divided into two classes depending on whether the pollutant is directly removed/degraded in-place or not, i.e. in-situ or ex-situ. One of the main problems associated with ex-situ remediation is its high operating cost for activities like soil excavation and groundwater pumping. In recent years, in-situ techniques have become popular. However, with in-situ remediation methods, knowledge on processes and factors controlling the results is lacking, which translates to much inflated expenses. Several mathematical models have been proposed to furnish representations as close as possible to reality of the effects of widely known remediation techniques. Some quantitative models have also been proposed for coupling multiphase flow and transport in a porous medium, with consideration of various remediation strategies such as water pumping, vapor and air venting, and steam injection. All of these techniques rely on human intervention for removing the contaminant. These techniques are fast, but costly. Moreover, most of them are too complex and not easily comprehensible for managers and engineers in industries and government. Therefore, a new approach is needed for developing useful, cost-effective, and user-friendly systems which can be readily adopted by industry and/or government to support decision-making on site remediation techniques. This paper presents the development of a rule-based decision support system using Artificial Intelligence (AI) techniques for solving this problem of petroleum waste management.

This paper proceeds as follows. Section 2 reviews some background literature. Section 3 describes the problem domain. The next three sections discuss knowledge engineering for the expert system: Section 4 discusses the process

* Corresponding author. Tel.: +1-306-585-5225; fax: +1-306-585-4745.
E-mail address: chan@cs.uregina.ca (C.W. Chan).

0957-4174/01/$ - see front matter
PII: S0957-4174(00)00063-4

of knowledge acquisition, Section 5 describes the process of conceptual design and Section 6 presents details on system implementation. Section 7 discusses validation results using two case studies. Section 8 concludes the paper and discusses some directions for future work.

2. Background

Automation of the selection task has been studied in the process industries. Some examples are discussed as follows. Feng, Yang and Rao (1998) developed a fuzzy expert system for monitoring chemical processes, predicting incidents and providing operation support for process operators. The reasoning strategy of the system involves using a dynamic membership function of a fuzzy system. Lau and Wong (1998) presented a methodology for implementation of a fuzzy expert system, with a non-mathematical approach which is able to handle complex closed-loop control situations. The system also demonstrates its advantages over conventional Proportional-Integral-Derivative (PID) closed-loop control. Zhang and Zhao (1999) developed an expert system called Coal Mining Expert and Optimization Consultation System (CMEOC). The system determines the underground mining methods, open-pit mining transportation systems, and mining machinery for given conditions and integrates fuzzy sets and optimization methods. Other examples of automated systems developed for selection in process design include the selection of activation systems in oil production (Hoffmann & Valentin, 1987), parameter selection in metal cutting (Malakooti, 1989), and selection of solvent for removal of acidic gases (Chan & Tontiwachwuthikul, 1995; Chan & Lau, 1997).

In comparison to these studies that focus on automation of the selection task in the process industries, however, research that focus on applying a knowledge-based approach to automating the selection task in the problem domain of petroleum waste management is scarce. In this paper, we describe our efforts at developing a knowledge-based decision support system for selection of remediation technologies for petroleum contaminated sites.

3. Problem domain: selection of remediation methods

Nonaqueous phase liquids (NAPLs) are hydrocarbons that exist as a separate, immiscible phase in contact with water and/or air. Differences in the physical and chemical properties of water and NAPLs result in the formation of a physical interface between the liquids which prevents the two fluids from mixing. Nonaqueous phase liquids are typically classified as either light nonaqueous phase liquids (LNAPLs) which have densities lower than that of water, or dense nonaqueous phase liquids (DNAPLs) which have densities greater than that of water.

Accidental emission of LNAPLs has affected groundwater quality at many sites across North America. The most common LNAPLs-related groundwater contamination problems result from the release of petroleum products. These products are typically multi-component organic mixtures composed of chemicals with varying degrees of water solubility. The LNAPLs inside subsurface represent potential long-term sources for continued groundwater contamination in many areas.

The problem domain in this study involves a vast amount of knowledge and decision tools related to site remediation practices. Concisely, the site contaminated by petroleum products is the target for remediation. Information on the site hydrogeology, subsurface contamination, and contaminant transport and conversion are factors relevant for determining the remediation technology. These factors were all considered in the development of the expert system.

Knowledge engineering for constructing the decision support system on remediation technique selection involves three stages: knowledge acquisition (KA), conceptual design, and system implementation. In the knowledge acquisition phase, the objects and decision processes were clarified and determined. In the conceptual design stage, the knowledge was formalized and represented with various representation methods, such as decision trees and fuzzy membership functions. Then the formalized knowledge was represented in production rules in the knowledge base of the system. The three phases are presented in detail as follows.

4. Knowledge acquisition

Knowledge acquisition took the form of interviews conducted over four weekly sessions, with each session lasting 2 h. The first author acted as the knowledge engineer and the second author acted as the domain expert. The first interview was expert-driven in that the expert prepared the material before the meeting and introduced and explained the concepts and tasks of the problem domain to the knowledge engineer. In the later interviews, when the knowledge engineer had gained some familiarity with the domain, he analyzed the materials obtained in the previous knowledge acquisition sessions and prepared the questions for the subsequent meeting. This process continued until the knowledge engineer was satisfied that the material was sufficient for conceptual modeling.

In addition to the human expert, a secondary knowledge source was found in the remediation database, which is a commercial database of several hundreds of remediation methods. The database includes descriptions of the remediation methods and the conditions in which they are suitable. This system contains much useful information but has poor retrieval mechanism. Hence, we used the database as reference material. The names of the remediation methods as stated in the records were put in the proper position in the main decision tree of the system according to the expert's opinion.

A number of heuristics were clarified through knowledge accquisition. The domain expert decided that there are a number of factors crucial for the selection of remediation technologies. They are discussed as follows.

4.1. Contaminated site

There are three types of contaminated sites: soil, groundwater, or both soil and groundwater. Therefore, there are three possible polluted situations: soil is contaminated, groundwater is contaminated, and both soil and groundwater are contaminated.

4.2. Site hydraulic condition

The hydraulic properties of a site include the following considerations: (1) soil permeability determines if it is easy to transport solute or fluid in the soil, (2) whether the site is homogeneous, i.e. soil property is the same at different locations, or heterogeneous, i.e. soil property varies at different locations, and (3) if the site is isotropic, i.e. soil property is the same in different directions, or anisotropy, i.e. soil property varies in different directions.

We can classify a contaminated site according to these properties into two classes: simple media and complex media. For example, if the site media have low soil permeability and are homogeneous and isotropic, the site has simple media; but if the media have high soil permeability and are heterogeneous and anisotropy, the site has complex media.

4.3. Estimated volume of contaminated soil and groundwater

The remediation technique is also determined by the range of the contaminated site. If the area of the site is less than 1600 m^2 and the volume is less than 25,000 m^3, we define it as a small site, and an ex-situ remediation technique can be used. If the contaminated area and volume exceed 1600 m^2 and 2500 m^3, respectively, then we define it as a large site and in-situ remediation technique is preferred.

4.4. Density of the immisible petroleum contaminant

If the density of the immisible petroleum contaminant is lower than water, then relatively simple methods can be applied. If the density of the immisible petroleum contaminant is higher than water, then some advanced remediation methods have to be considered.

4.5. The immisible contaminants are present as free phase or residual phase

If the immisible contaminants are present as free phase, then oil recovery need to be considered. If the immisible contaminants are present as residual phase, then more remediation actions like integrated technology are needed.

4.6. Concentration range of chemicals in soil and groundwater

We classified the concentration range of chemicals in soil and groundwater into three classes: low (0–5 times of standard), high (5–50 times of standard), extremely high (greater than 50 times of the standard). These ranges are used by the system for generating an overall contamination level using a fuzzy membership function and statistical methods. Different contamination levels require different remediation methods. The standard is assumed to be the Saskatchewan standard for the maximum acceptable level of the contaminant.

5. Conceptual design

5.1. Task decomposition

In the knowledge engineer's analysis of the elicitation results, we used both a top-down analysis approach and an object oriented technique of knowledge analysis known as Inferential Modeling Technique (Chan, 1992, 1995). By adopting the top-down approach, the knowledge engineer began by identifying the main task of the system and then subdividing the main task into several subtasks. This process of decomposition continued until every subtask can be easily implemented. For the system, the main task is to determine the appropriate remediation method for a given contaminated site with specific characteristics. Remediation method selection depends on parameters about the media condition and pollution situation of the contaminated site, such as the media type, the contamination level and the size of the media. The decision processes involving these parameters constitute the subtasks. The task structure of the system is shown in Fig. 1.

Both the main task and the subtasks can be represented as decision trees. The main decision tree corresponds to the main task of the system which is the decision processes for determining the remediation method for a given situation; the tree is shown in Fig. 2. The nodes of the tree represent either actions to be performed by the system or parameters whose values are to be provided by system users as inputs. Since determination of the remediation method depends on values of media parameters, the system performs inferencing by gathering input data and then making a recommendation. In other words, forward chaining was used as the inference mechanism for the system.

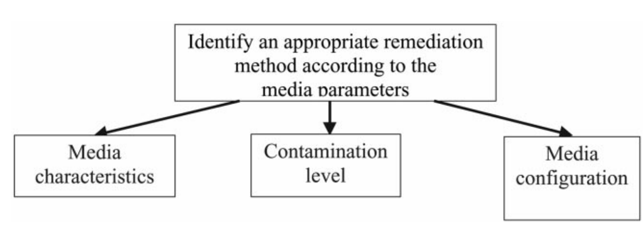

Fig. 1. Structure of tasks.

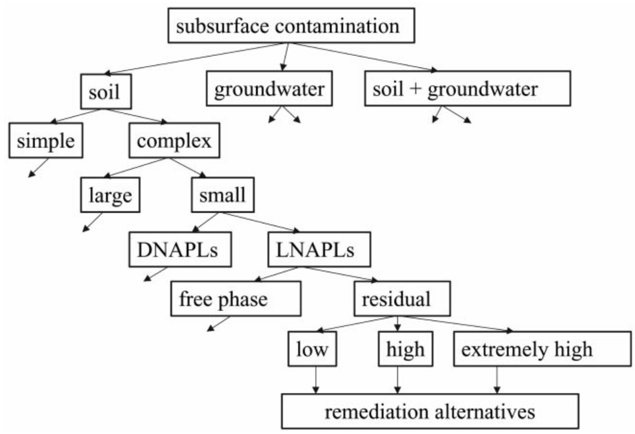

Fig. 2. Main decision tree.

Since the determination of remediation methods depends on media parameters, a number of auxiliary decision trees were formulated which correspond to the subtasks of the system. For example, one subtask is to decide the type of the media, whether it is complex or simple. The decision tree representing this subtask is shown in Fig. 3.

5.2. Object-oriented modeling

We also adopted an object-oriented approach to knowledge modeling during the conceptual design phase because it is an intuitive way of conceptualizing the domain and also because the implementation tool, G2 (trademark of Gensym Corp., USA), is an object-oriented real-time expert system shell. The Inferential Modeling Technique (Chan, 1992, 1995) is the object-oriented approach for analyzing domain knowledge adopted in the project. Some knowledge elements specified using the technique include:

- Objects including media, petroleum waste and remediation methods
- Relationships between the objects of media and remediation methods

Since determining the level of contamination that result from the four contaminants is highly subjective and based on heuristics, fuzzy logic is useful for representing this imprecise knowledge. Fuzzy expert systems show exceptional performance for working with processes which are adequately defined in qualitative terms and for which no precise mathematical model of the process exists (Zimmerman, 1985). This characterization exactly describes the process of determining the pollution level that results from the four most common and important contaminants: benzene, toluene, ethyl-benzene, and xylenes. The four pollutants contribute differently to the final level of contamination, which is defined by the weighted sum of the fuzzy value of each pollutant. The values of the weights are given by the domain expert. The fuzzy functions in the system used for determining the level of contamination is discussed as follows.

Firstly, the system calculates the fuzzy value of each pollutant based on the concentration of the pollutants which the user inputs into the system. The fuzzy function for calculating the contamination level of benzene is shown in Fig. 4. The *X*-axis denotes the number of times the concentration of benzene exceeds the standard acceptable level of the compound. The maximum acceptable level, or

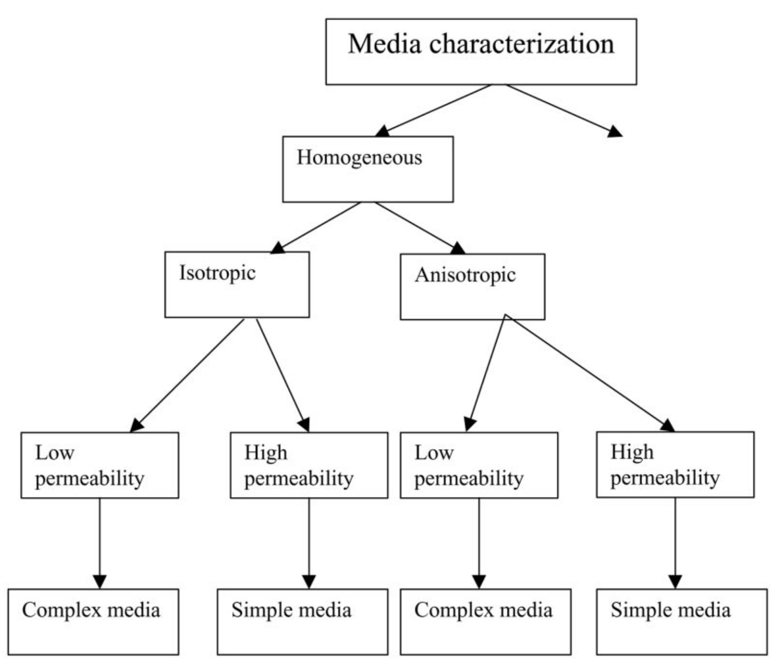

Fig. 3. Decision tree for subtask of media type determination.

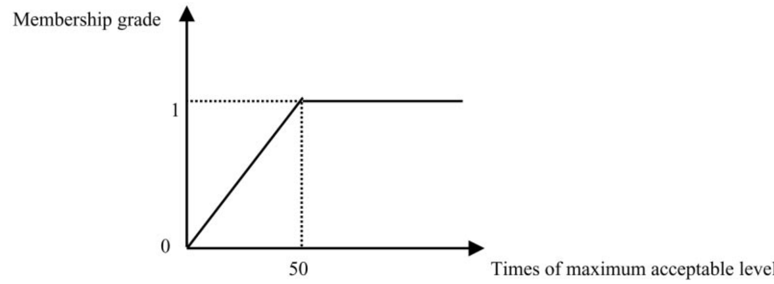

Fig. 4. Fuzzy membership function.

the standard of benzene concentration is 50 mg/l (Chen Huang & Chakma, 1998; Huang, Chen, Tontiwachwuthikul & Chakma, 1999) which is a fixed value. If the concentration the user inputs is 100 mg/l, the concentration will be two times that of the acceptable level ($100/50 = 2$). So X equals 2. The Y-axis denotes the fuzzy membership grade of the contamination level of benzene called 'fb' in the system.

To determine the membership grade for contamination level F that results from all four pollutants, the following equation is used:

$$F = \text{weight\_}b^{*}\text{fb} + \text{weight\_}t^{*}\text{ft} + \text{weight\_}e^{*}\text{fe} + \text{weight\_}x^{*}\text{fx}$$

where weight_b, weight_t, weight_e and weight_x are the weights of the four pollutants contributing to the integrated contamination level. They are determined by the expert to be 0.5, 0.2, 0.15, and 0.15, respectively. Fb, ft, fe, and fx are the membership grade of the contamination level of each pollutant.

The result of F in the equation determines the general contamination level. The expert also decided the thresholds to categorize the contamination grades. If $F < 0.1$, then the contamination grade is 'low'; if $F > 0.8$, the contamination grade is 'extremely high'; otherwise, it is considered as 'high'.

6. System implementation

System implementation involves mapping the result of the design process into the knowledge base of the system developed with G2. First the G2 shell is briefly introduced.

G2 (trademark of Gensym Corp., USA) is an object-oriented real time expert system shell and a development environment for creating and deploying intelligent real time applications. It offers components such as class hierarchy, objects, rules, procedures and interface items and can be used for building real time or non-real-time systems for monitoring, scheduling, diagnosis and decision support. System implementation using this tool includes the following steps.

6.1. Define classes of objects

According to the conceptual design, there are three classes of objects in this system: (1) media, (2) petroleum waste and (3) remediation. Each class consists of attributes or properties which are presented as follows; the letter following each attribute indicates its system status ('U' for user input value, 'I' for intermediate result calculated or derived by the system, and 'E' for input value provided by the environmental engineer):

1. (1) Media class
   - Media component (soil, groundwater, or both) (U)
   - Media property I (heterogeneous or homogeneous) (U)
   - Media property II (isotropic or anisotropic) (U)
   - Media permeability (high or low) (U)
   - Media type (complex or simple) (I)
   - Media volume (a numeric value) (U)
   - Media area (a numeric value) (U)
   - Zone type (small or large) (I)
2. (2) Petroleum waste class
   - Oil type (dnapl or lnapl) (U)
   - Present form (residual or free phase) (U)
   - Concentration (low, high, or extremely high) (I)
   - Threshold between low concentration and high concentration (a numeric value) (E)
   - Threshold between high concentration and extremely high concentration (a numeric value) (E)
   - Benzene concentration (a numeric value) (U)
   - Benzene concentration standard (a numeric value) (E)
   - Weight of Benzene to comprehensive contaminated level (a numeric value) (E)
   - Ethyl-benzene concentration (a numeric value) (U)
   - Ethyl-benzene concentration standard (a numeric value) (E)
   - Weight of Ethyl-benzene to comprehensive contaminated level (a numeric value) (E)
   - Toluene concentration (a numeric value) (U)
   - Toluene concentration standard (a numeric value) (E)
   - Weight of Toluene to comprehensive contaminated level (a numeric value) (E)
   - Xylenes concentration (a numeric value) (U)
   - Xylenes concentration standard (a numeric value) (E)

- Weight of Xylenes to comprehensive contaminated level (a numeric value) (E)

3. (3) Remediation class
   - Technology name (text) (E)
   - Technology description (text) (E)

This list consists of attributes deemed essential by the domain expert for selecting remediation technologies. Conceptually the attributes can be divided into those parameters that pertain to the contaminated site and pollutant and those that are thresholds set by the environmental engineer. This distinction is reflected in the implementation, and the properties in the above list are classified into the three categories of (1) user input properties, (2) intermediate results derived from calculation based on the user input values and (3) properties whose values are set by the environmental engineer.

1. *User input properties*: These properties are indicated by 'U' in the list. It is important to obtain a complete list of these properties during knowledge acquisition, because changing these properties changes the basic structure of the system, and hence can lead to much work in system maintenance.
2. *Intermediate results*: These properties are computed with rules and procedures based on the user input properties. They include both system output values and intermediate calculated results. They are indicated in the list as 'I'.
3. *Engineer input properties*: These properties refer to the threshold values and coefficients of the fuzzy functions, which embody the experts' heuristics and can be adjusted during testing and validating. They are indicated in the list as 'E'.

The different kinds of properties requires different handling in both system design and implementation. For example, the engineer input properties are initialized before running the program and can be adjusted when running, the user input properties are connected to the interface parameters and the intermediate results are calculated or derived using rules and procedures.

6.2. Create objects

Object is an instantiation of a class. This system consists of one media object, one waste object and 29 remedy objects. That is, at any time during runtime, there is only one contaminated site, one compound of petroleum waste, and 29 remedy methods under consideration.

6.3. Define rules

Rules are the important components in the knowledge base. The rules are derived from the decision trees, a sample of which is shown in Fig. 2. Each path in the decision tree determines a rule. Each non-leaf node denotes a relevant property of the media and petroleum waste class objects, which when combined with other nodes on the branch determines the comprehensive condition for a remediation action. The following are two sample rules. Example 1 was drawn from the highlighted branch of the main decision tree shown in Fig. 2, and example 2 was drawn from the supporting decision tree shown in Fig. 3. Each example is presented first in English and then in the G2 implementation language.

1. Example 1
   - In English: If the component of media is soil and the type of media is complex and the zone of media is small and the type of petroleum waste is LNAPL and the present form of petroleum waste is residual phase, then the remediation method to use is bio-treatment
   - In G2 implementation: If the media_component of media1 is soil and the media_type of media1 is complex and the zone_type of media1 is small and the oil_type is waste1 is lnapl and the present_form of waste1 is residual then inform this workspace that 'The remedy is [the remedy_name of bio-treament]'
2. Example 2
   - In English: If the media is homogeneous and permeability of media is high then the media type is simple.
   - In G2 implementation: If the h_c1 of media1 is homogeneous and the permeability of media1 is high then conclude that the media_type of media1 is simple.

The rules derived from the decision trees can be combined according to propositional logic truth table values. Combining rules can make the knowledge base more compact and efficient. However, if the decision trees are modified and increase in size later, they are more difficult to maintain. This is a tradeoff to be considered in combining rules. Rules are combined under two conditions: (1) the new rule does not destroy or violate the semantics of either of the original rules from which it is derived, and (2) the combination does not involve too many logical manipulations. The following is an example for combining two rules in the supporting decision tree.

- Rule 1: if the h_c1 of media1 is homogeneous and the h_c2 of media1 is isotropic and the permeability of media1 is high then conclude that the media_type of media1 is simple.
- Rule 2: if the h_c1 of media1 is homogeneous and the h_c2 of media1 is anisotropic and the permeability of media1 is high then conclude that the media_type of media1 is simple.
- Combined Rule: if the h_c1 of media1 is homogeneous and the permeability of media1 is high then conclude that the media_type of media1 is simple.

An additional consideration in constructing the knowledge base was to group the rules from each decision tree into a separate workspace. Since each decision tree

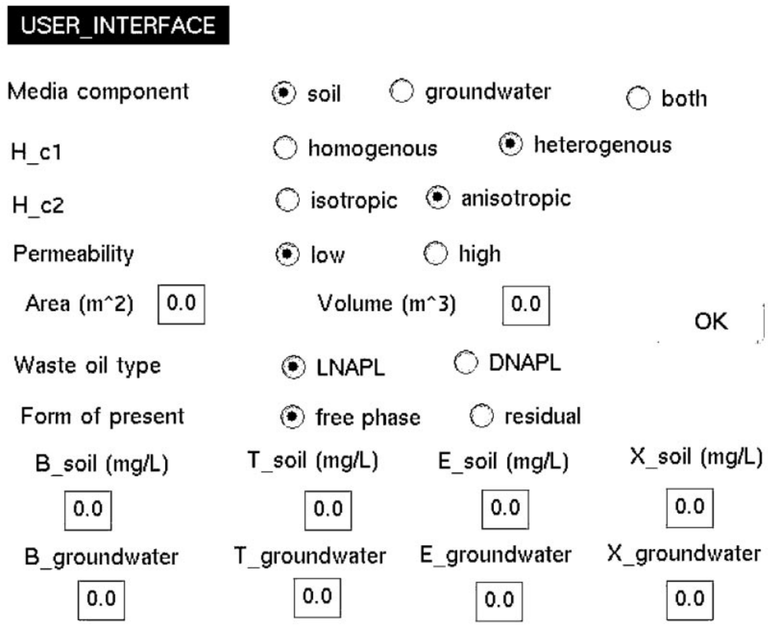

Fig. 5. User interface.

illustrates the problem solving process for accomplishing one task objective, the rules are grouped according to objectives such as determining the type of media or determining the final selection of remediation technique. This facilitates invoking rule groups and enables easier maintenance.

6.4. Define interface

In order to support the two different types of input and increase flexibility of the system, we constructed two separate system interfaces: one for the user and one for the engineer. The user interface enables the user to enter values describing the contaminated site and contaminants. The engineer interface enables the environmental engineer to set the values in the fuzzy functions inside the system that are used to decide the contamination level.

The user interface queries the user for details about the contaminated site and the contaminants. Through this interface, the user provides information about the site such as the media component (whether it is soil, groundwater, or both), properties about the media such as whether it is homogenous or heterogenous, isotropic or anisotropic, whether permeability of the media is low or high, and the area and volume of the media. The user also answers queries about the contamination such as the type of petroleum waste, whether the waste is in free phase or residual phase, and the amount of the four contaminants of benzene, toluene, ethyl-benzene, and xylenes in the soil and groundwater. The system's user interface is shown in Fig. 5.

The environmental engineer can provide three different kinds of information through the engineer interface. First, the two threshold values of 'low' and 'high' discretise the membership grade of the contamination level to 'low', 'high', and 'extremely high'. The second type of input provides information on the maximum acceptable levels of contamination. The third type of input measures relative contribution of each contaminant to the final contamination level. These are labelled 'B-weight' for the weight of benzene, 'T-weight' for the weight of toluene, 'E-weight' for the weight of ethyl-benzene, and 'X-weight' for the weight of xylenes. All three types of inputs can be altered during runtime. Any change in the input to this engineer interface can result in recommendation of a different remediation method by the system. The engineer interface is shown in Fig. 6.

6.5. Define procedures to connect user interface with G2 objects

There are three procedures in the remediation advisor. The main procedure is used to transfer values from the interface items to the objects defined in the system, thereby triggering the forward chaining process. The user interface accepts input values which trigger the inference mechanism of forward chaining. However, the G2 environment requires that interface items connect to parameters and variables but not to system defined objects. Hence an explicit link has to be set up between the parameters or variables and the objects. For example, the following command assigns the values from interface item media_component to the property media_component of object media1:

Conclude that the media_component of media1 = media_component

Procedures were also used to implement the fuzzy

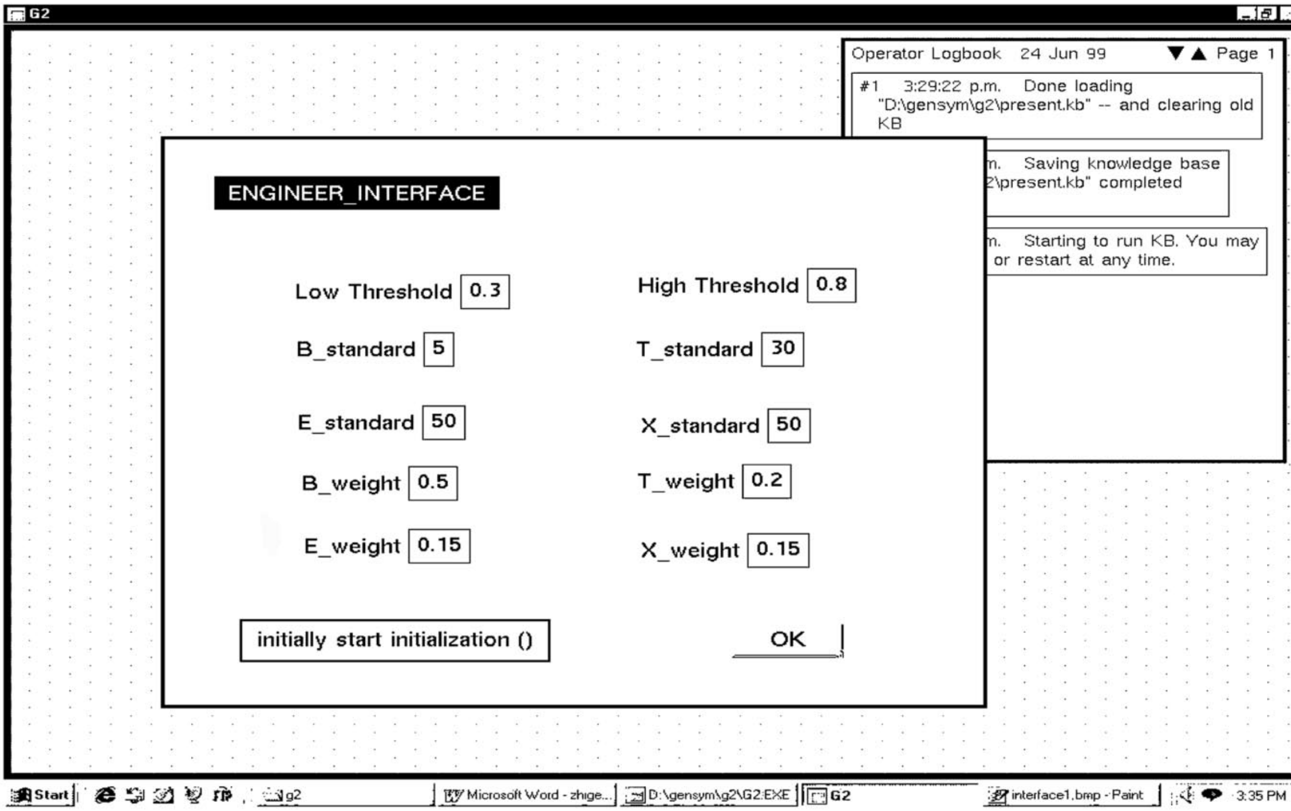

Fig. 6. Engineer interface.

functions. The following is the procedure to determine the contamination level of benzene. It is triggered when the user inputs a value for benzene concentration in the soil denoted in the system as 'b_soil' of petroleum waste. The procedure then generates the fuzzy value of benzene contamination denoted as 'B_soil_fuzzy'.

```
Decide_b_concentration()
B_times: quantity;
B_soil_fuzzy: quantity;
Begin
B_times = (the b_soil of waste1)/b_standard;
B_soil_fuzzy = 1/50* b_times;
If B_soil_fuzzy <0 then
   B_soil_fuzzy  = 0;
If B_soil_fuzzy >1 then
   B_soil_fuzzy = 1;
End.
```

7. Demonstration and validation of the system

In this section, two cases are presented to illustrate system feedback to given sets of user inputs. In both cases, the resulting recommendations are appropriate, demonstrating applicability of the developed system for decision support.

7.1. Case 1

7.1.1. Input data set

Media	soil
HC1	homogenous
HC2	isotropic
Permeability	low
Area	1000 m^2
Volume	8000 m^3
Waste oil type	LNAPL
Form of present	free phase

7.1.2. System output

The recommended alternative is static pile bioremediation technology.

7.1.3. Discussion of output

The expert system recommends static pile bioremediation technology for this contaminated site. This recommendation is valid because free phase oil exists in the soil and needs to be cleaned. Also, the contaminated area and volume are not very large, which indicates that ex-situ technology can be used. Free phase oil is semi-volatile in soil. The static pile bioremediation technology is designed to decompose the oil through microorganism activities. It is the most widely used technology for the treatment of petroleum-contaminated soil. During the treatment process, the contaminated soil is placed into piles within which a perforated pipe can be installed to supply oxygen to soil microbes. Moisture and nutrients are added by drip irrigation. The soil piles are constructed on a pad or liner to prevent groundwater contamination. The pile may also be covered to reduce volatilization. Compared to landfarming which is a widely used method for treatment of petroleum contaminated soils, this method requires less space, and is more cost-effective.

7.2. Case 2

7.2.1. Input data set

Media	soil
HC1	heterogeneous
HC2	anisotropic
Permeability	high
Area	500 m^2
Volume	25000 m^3
Waste oil type	DNAPL
Form of present	residual
Observed concentration of pollutants	
benzene	8.9 mg/l
toluene	10.1 mg/l
E-benzene	4.9 mg/l
xylenes	12.4 mg/l

7.2.2. System output

The recommended alternative is Integrated Shallow Soil Mixing, Thermally Enhanced Vapor Extraction, and Soil Vapor Extraction.

7.2.3. Discussion of output

The remediation alternative recommended by the expert system is an integrated approach which consists of shallow soil mixing, thermally enhanced vapor extraction, and soil vapor extraction, which is an in-situ enhanced soil extraction technique. The site in this case is not only contaminated by waste heavy oil (DNAPLs) but also some oil constituents like benzene, toluene, E-benzene, and xylenes, all of which have concentrations higher than the acceptable levels. In addition, the soil condition is more complicated than that in Case 1, and the contaminated volume is very large. Hence, normal ex-situ techniques are not applicable. In general, the in-situ enhanced soil extraction technology can remove more than 70% of the pollutants in the soil. The clayey soils are treated by injecting hot air through a large mixing auger which penetrates the soils without excavation. This technology is especially suitable for this kind of contaminated sites with large volumes and high depths. Volatile organic compounds (VOCs) are removed by vapor extraction through wells as well as a special movable shroud which covers the work area during mixing. In general, this method is designed to enhance removal of organic compounds in clayey soils using a combination of soil mixing, soil vapor extraction, and hot air injection.

8. Conclusions

A rule-based expert system for the management of petroleum contaminated sites has been developed. In the development process, a variety of methodologies and tools are employed and integrated, e.g. IMT for knowledge analysis, decision trees and object-oriented methodology for knowledge representation, and fuzzy set theory for uncertainty modelling. To enhance flexibility of the system, a user interface and an engineer interface have been constructed for capturing different types of input. Finally the system was implemented with the real time expert system shell G2.

An important characteristic of this system is its capability for dealing with imprecision in the inputs on site conditions and contamination levels. Presently, the designs of site remediation technology are mainly based on the judgement of engineers confronted with the problems. Often the engineer tends to simplify the uncertainties in the situation and recommends a technique s/he is familiar with, which may not necessarily be the best one from both the environmental and economic points of view. The developed system contributes to the selection task by allowing more comprehensive and realistic consideration of the many factors related to the remediation selection decision. The results from the case studies indicate that the expert system can help the environmental engineer to identify the optimal alternatives for site remediation and thereby help the industries to reduce costs for site remediation practices. The system's inputs and outputs are also directly useful for further modeling studies on remediation processes and contaminant transport in soil and groundwater.

The system in its current version has been developed to deal with real-world site remediation problems. Some directions for future work include employing machine learning for automated maintenance, i.e. using other experts' feedback to train the system rather than to modify the conceptual design manually. We are also presently in the process of developing more user-friendly and powerful interfaces for the system so that more flexible functions can be incorporated to let the user define parameters according to his/her own experiences. This would further customise the system for individual user preference.

Acknowledgements

The generous support of a Strategic Grant from Natural Sciences and Engineering Research Council is gratefully acknowledged. We are also grateful for the support of the Faculty of Graduate Studies and Research, University of Regina to the first author.

References

Chan, C. W. (1992). Knowledge acquisition by conceptual modeling. *Applied Mathematics Letters Journal*, *3*, 7–12.

Chan, C. W. (1995). Development and application of a knowledge modeling technique. *Journal of Experimental and Theoretical Artificial Intelligence*, *7* (2), 217–236.

Chan, C. W., & Lau, P. (1997). An expert system architectural framework for engineering selection. *Engineering Applications of Artificial Intelligence*, *10* (4), 357–367.

Chan, C. W., & Tontiwachwuthikul, P. (1995). Expert system for solvent selection of CO_2 separation. *Expert System with Application*, *8*, 33–46.

Chen, Z., Huang, G. H., & Chakma, A. (1998). Integrated environmental risk assessment for petroleum-contaminated sites—a North American case study. *Water Science and Technology.*, *38*, 131–138.

Feng, E., Yang, H., & Rao, M. (1998). Fuzzy expert system for real-time process condition monitoring and incident prevention. *Expert Systems with Applications*, *15*, 383–390.

Hoffmann, F. C., & Valentin, E. P. (1987). Opus: An integrated assistance system for oil production. *Expert System*, *4*, 242–249.

Huang, G. H., Chen, Z., Tontiwachwuthikul, P., & Chakma, A. (1999). Environmental risk assessment for under ground storage tanks through an interval parameter fuzzy relation analysis approach. *Energy Resources.*, *21*, 75–96.

Lau, H., & Wong, T. (1998). A fuzzy expert system for complex closed-loop control: a non-mathematical approach. *Expert System*, *15* (2), 98–109.

Malakooti, B. (1989). An interactive multiple criteria approach for parameter selection in metal cutting. *Operations Research*, *37* (5), 805.

Sims, J. L., Suflita, J. M. & Russell, H. H. (1992). In-situ bioremediation of contaminated ground water, Ground Water Issue, EPA/540/S-92/003, U.S.EPA, R.S. Kerr Environ. Res. Lab., Ada, OK, 11 pp.

Zhang, H., & Zhao, G. (1999). CMEOC—An expert system in the coal mining industry. *Expert System with Applications*, *16*, 73–77.

Zimmerman, H. J. (1985). *Fuzzy set theory and its applications*, Kluwer-Nijhoff Publishing.