Geoscience Journal
Vol. 2, No. 4, p. 175–183, December 1998

Integration and fusion of geological exploration data: a theoretical review of fuzzy logic approach

Wooil M. Moon *Department of Earth System Science, Seoul National University, Seoul 151-742, Korea (e-mail:wmoon@eos1.snu.ac.kr)*

ABSTRACT: A geological survey project, including the exploration of nonrenewable resources, typically includes the following steps: planning of the survey, field mapping, acquisition of data, achieving of data, processing and fusion of all relevant information, which are closely followed by an interpretation of the final survey results. In this process, important spatial data layers are topographic map, often digital elevation model (DEM), geological map, several sets of geochemical and geophysical survey maps, airborne or space-borne remote sensing data and, sometimes, old archived data. Each of these multiple layers of geological exploration data can first be preprocessed, and input into a chosen geographic information systems (GISs). This will be followed by information representation and fusion steps for the final imaging of the processed and fused information. Most commercially available GISs however do not have information fusion and focusing capabilities. Users of the slected GIS must first be able to represent the preprocessed geological survey information with respect to the target geological hypothesis. In the case of mineral exploration, the target hypothesis can be a specific mineral deposit(s) being explored.

Compared to the model driven exploration approach, the data driven geological exploration utilizes multiple sets of exploration data and there are several mathematical tools available for information representation and fusion. Some of these include the traditional probability approach, evidential belief function methods, AI/Expert systems and the fuzzy logic approach.

In this paper, the fuzzy logic approach of quantifying the exploration information with respect to the target hypothesis, and several of the fuzzy operators which are frequently used for geological resource exploration are critically reviewed. Although our understanding of the fuzzy information representation and fuzzy operators is still incomplete, many case studies of applying fuzzy logic approaches to various exploration projects have concluded that the final fuzzy membership function maps or the final fuzzy theme maps are considerably more accurate than the results obtained using any conventional intuitive approach, including the brute stack of the exploration (spatial) data using a GIS. In general, the efficiency and accuracy of the final fused information with respect to the target hypothesis increases with the increasing number of geological exploration data layers.

If a specific Earth system model is available, or a mineral deposit model in the case of mineral exploration projects, is available or is known a priori, fuzzy logic approach can easily be combined with currently popular machine-reasoning processes such as the neural network approach. In this type of quantitative spatial reasoning and fuzzy logic information fusion, uncertainty and error analysis is also important. In most cases, propagation of errors and uncertainties can be handled and estimated in the same manner as the processing and fusion of main exploration data.

Key words: data fusion, fuzzy logic, GIS applications, spatial reasoning

1. INTRODUCTION

In resource exploration, it has been customary for a geologist to estimate and combine field survey information qualitatively, purely based on one's knowledge and experience. In digital information processing, however, there have been serious attempts to quantify the observed information and associated uncertainties using entropy as demonstrated by Shannon (1948). Many scientific disciplines including geophysical inversion theory, however, have not adopted Shannon's concept until recently. Instead, they tried to define the available information from various types of laboratory experiments and field observations and to draw qualitative conclusions. In geological resource exploration, many field surveys are limited to near-surface regions and produce data sets with incomplete spatial coverage even at the surface, due to logistical constraints. As a result, geologists routinely use intuitive expertise in geological reasoning and decision making processes. Many traditional data processing techniques adopted in geophysics have focused on inversion and statistical approaches (Mathai and Rathie, 1975), and they have very limited applications, often lacking spatial coverage and associated uncertainties.

Geologists have recognized and accepted situations involving incomplete data sets and imperfect knowledge since the early days of geological science. The problem of imperfect knowledge has been rigorously tackled by philosophers, logicians and mathematicians for more than two thousand years. Although increasing numbers of geological scientists now use mathematical analysis tools and techniques that are based on statistics and probability, most applications are based on classical set theory as established by Georg Cantor (1845–1918). Recently, however, following developments in computer science and increasing volume and types of exploration data, it has become necessary for geological scientists to accept the new concepts of reasoning with imperfect knowledge from incomplete data sets, and to develop more accurate and reliable decision-making process. There have been many approaches to the problem of imperfect geological knowledge: traditional statistical theory, evidential belief function theory, Boolean reasoning, rough set theory (Pawlak, 1996), and fuzzy set theory proposed by Zadeh (1968). Fuzzy logic is also utilized in the classification of spatial themes in satellite data and also in many geophysical data interpretation (Ishibuchi et al., 1996). In this paper, I will discuss the basic theory of the fuzzy logic approach in geological resource exploration applications.

2. CLASSICAL SET—A HISTORICAL BACKGROUND

Traditionally, geological scientists have relied on the classical set as developed by Georg Cantor. A set is a collection of objects or thoughts, within a certain realm, taken as a whole. Each object in the collection is called an element (or member) of the set. The notation

$$a \in A$$

denotes that a is an element of the set A. In this case, we understand that a is a member of A or a belongs to A. This method of describing an object is comparable to the way a geologist maps an outcrop in the field and determines to which rock types the outcrop belongs, i.e.,

$$\text{outcrop } "a" \in \text{rock type } "A".$$

In general, this is a definite statement that the outcrop "a" belongs to a rock type "A" as determined by the geologist.

Consider, in turn, the same geologist who takes a magnetometer to the same outcrop and takes a reading in the hope of determining relative content of magnetic minerals in the outcrop. The magnetometer reading, X, read by the geologist, is a number calibrated with respect to the magnetometer design parameters. If this number is properly calibrated, it can indeed provide the geologist with a definite percentage of the magnetic mineral. However, if the geologist wishes to connect the magnetic survey data (i.e., the magnetometer reading) to a specific mineral deposit type, the magnetic survey data X can not provide 100% evidence that a specific mineral deposit exists at the location. The magnetometer reading X can, at best, provide an indirect indication of the target mineral deposit. In this context, the exploration data (i.e., magnetometer reading X) represent partial information or "fuzzy" information about the exploration target and the mineral deposit hypothesis.

The classical set satisfies algebraic rules, such as the commutative, associative, distributive, and absorption laws. Operations which assign each element of a set A to each element of a set B are said a mapping or transformation from A to B. Union and intersection of sets A and B are also defined as in the classical set theory (Iyanaga and Kawada, 1980). The first publications in fuzzy set theory were by Zadeh (1965) and Goguen (1967, 1969) and included the generalization of classical definitions of a set and propositions to accommodate fuzziness of the newly defined sets.

3. THE BASIC CONCEPT OF A FUZZY SET

Let us consider characteristic features of a real exploration process; real exploration data are very often uncertain or vague in a number of ways. Geological mapping of rocks in the field can be viewed as a definitive and deterministic step, even when one can clearly see the transitional change from one rock type to another, because each rock type is defined by certain ranges of component minerals. In many real cases, however, problems become more difficult due to many factors which cannot be controlled by a geologist. In most exploration areas, bedrock is at least partially covered by thick overburden material and/or heavy vegetation. In such cases, even a sharp geological boundary has to be extrapolated under the overburden cover and the exact spatial extent of the geological unit becomes vague.

Most geophysical survey data pose a similar problem to exploration hypotheses. The design of most geophysical instruments is precise and calibration of these instruments is also exact with respect to chosen references. However, the information content of each data set is vague and non-deterministic with respect to exploration hypotheses (e.g., lithologies, mineral deposit models, complex geological factors associated with hydrocarbon traps, etc.). Zadeh (1987b) appropriately wrote: "As the complexity of a system increases, our ability to make a precise and yet significant statement about its behavior diminishes until a threshold is reached beyond which precision and significance become almost mutually exclusive characteristics".

Due to incomplete information, the "final outcome" of an exploration task can not be known exactly. This type of inherent uncertainty, in the classical set theoretical sense, has often been handled by probability and statistical methods. In this situation, the events (elements of sets) or the statements are well defined and this kind of uncertainty or vagueness is called "stochastic uncertainty". In contrast, vagueness concerning the description of the semantic meaning of the events, phenomena or statements themselves, is what we call "fuzziness". Fuzziness can be found in many aspects of geological exploration, such as porosity of formation rocks, electrical capacitance of disseminated mineral ores, reflectance of outcrop rocks, and other deposit model parameters in exploration problems.

Let X be a collection of objects denoted generically by

x then a fuzzy set $\tilde{A}$ in X is a collection of ordered pairs:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\},$$

where $\mu_{\tilde{A}}(x)$ is called the membership function or grade of membership (or degree of truth) of x in $\tilde{A}$ which maps X to the membership space M. (If M contains only two points and 1, $\tilde{A}$ is a non-fuzzy classical set). The range of the membership function is a subset of the nonnegative real numbers whose supremum is finite (Zadeh, 1965, 1987a, c; Zimmermann, 1984).

4. CLASSICAL MEASURES OF FUZZY EVENTS

Zadeh (1968) published on the probability of a fuzzy event and the axiomatic definition of fuzzy probability measures and he demonstrated that the fuzzy probability measures can be characterized uniquely either by a probability measure or a Markoff kernel. Zadeh (1968) defined the probability $\tilde{P}$ of a fuzzy event $\tilde{A}$ or a fuzzy set $\tilde{A}$ whose membership function μ_A is measurable by

$$\tilde{P}(\tilde{A}) = \int \mu_A(x)\, dp(x)$$

where p is some classical probability measure (Zadeh, 1968, 1987a, b, c; Klement, 1980). The probability $\tilde{P}$ has the following properties:

(i) $\tilde{P}(\tilde{0}) = 0$ and $\tilde{P}(\tilde{1}) = 1$
(ii) $\tilde{P}(\tilde{A} \cup \tilde{B}) + \tilde{P}(\tilde{A} \cap \tilde{B}) = \tilde{P}(\tilde{A}) + \tilde{P}(\tilde{B})$
(iii) if $(\tilde{A}_n)_{n \in IN}$ is an increasing sequence of fuzzy events then

$$\tilde{P}\left(\bigcup_{n \in IN} \tilde{A}_n\right) = \sup_{n \in IN} \tilde{P}(\tilde{A}).$$

The union and intersection of fuzzy sets are used in the usual sense here, i.e., union and intersection are expressed by the maximum and minimum of the membership functions respectively.

A fuzzy probability measure is a function

$$m : A \rightarrow [0, 1],$$

which fulfills the above properties (i), (ii), and (iii) among many others which are satisfied by many classical probability spaces (Klement, 1980).

5. FUZZY OPERATOR

The membership function is the fundamental component of a fuzzy set and operations with fuzzy sets are defined via their membership functions (Zadeh, 1965).

The membership function $\mu_{MIN}(x)$ of intersection of a fuzzy set $\tilde{A}$ is defined by

$$\mu_{MIN}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}, \quad x \in X.$$

This intersection operator is also called the "fuzzy AND" operator or "MIN" operator in some publications (An et al., 1991; Bonham-Carter, 1994). The membership function $\mu_{MAX}(x)$ of the union $D = A \cup B$ is point-wise defined by

$$\mu_{MAX}(x) = \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}, \quad x \in X.$$

As above, this operation is called the "fuzzy OR" or "MAX" operator. This operator behaves like the Boolean OR and the output membership values are controlled by the maximum values of the input information. The membership function of the complement of a fuzzy set A, $\mu_c(x)$ is defined by

$$\mu_c(x) = 1 - \mu_A(x), \quad x \in X.$$

In addition to the above basic operators, there is a large number of operators, many of which are specially designed for specific applications in complicated engineering problems. Some of these operators are also used effectively in the spatial data fusion applications of geological exploration data (An, 1992; Bonham-Carter, 1994). A combined membership function $\mu_{comb}(x)$ of a fuzzy set A is defined as

$$\mu_{comb}(x) = \prod_{i=1}^{n} \mu_i$$

where $\mu_i(x)$ is the fuzzy membership function for the ith subset of A. (In a geological exploration application, the subscript i may represent individual map layers to be combined). In this case, the combined fuzzy membership values tend to be very small due to the effect of multiplying several numbers smaller than 1.0. This fuzzy operator is also called fuzzy algebraic product operator. The next useful fuzzy operator is the algebraic sum operator, which is defined by

$$\mu_{A\text{-}SUM}(x) = 1 - \prod_{i=1}^{n} (1 - \mu_i).$$

The algebraic sum operator is complementary to the algebraic product operator and the result is always greater than or equal to the largest contributing fuzzy membership value. This operator can be used when there is an evidence that supports the chosen exploration hypothesis and the combined evidence is more supportive than individual pieces of evidence. Another useful fuzzy operator is the γ (Gamma) operator which was proposed by Zimmermann and Zysno (1980). The membership function $\mu_\gamma(x)$ of the combined fuzzy algebraic product and fuzzy algebraic sum operation is defined by

$$\mu_\gamma(x) = (\textit{Fuzzy algebraic sum})^{\gamma} * (\textit{Fuzzy algebraic product})^{(1-\gamma)},$$

where γ is a parameter chosen in the range (0, 1). When γ is 1, the combination is the same as the fuzzy algebraic sum and when γ is 0, the combination is same as the fuzzy algebraic product result (An, 1992). Therefore, careful choice of γ can produce output membership values which can ensure a flexible compromise.

In geological application of digital data fusion, the available data or spatial data have, in most cases, varying degree of information content with respect to the chosen exploration hypothesis. In such cases, it becomes necessary to use several different fuzzy operators separately or a combination of selected operators depending on the characteristics of each data layer (Moon and Jiang, 1995; Moon et al., 1998) The situation is similar for oil and gas exploration application, utilizing a fuzzy logic approach. If one attempts an integrated imaging of well-log data, surface seismic and VSP data, and associated reservoir characteristics, the information content of each data set will be considerably different with respect to the hydrocarbon deposit model of the study site. In such a case, the relationship between each data set with respect to the specific hydrocarbon deposit model will favor a certain fuzzy operator or a combination of several fuzzy operators.

6. FUZZY INFORMATION REPRESENTATION

Most geological exploration tasks include multiple sets of spatial data layers. They include geological maps, geochemical data, geophysical data and other auxiliary information. To carry out a systematic integration (or fusion) of these types of spatial data towards a chosen exploration hypothesis, there are several steps to follow: preprocessing of individual data layers, information representation, integration (digital fusion), visualization, and decision making (Moon, 1995). Assuming that all exploration data are properly digitized, the preprocessing and geocoding steps with the exception of certain specialized processing tasks can be carried out utilizing any of a number of available commercial and public domain geographic information system (GIS).

6.1. Exploration Data—Spatial Data

Definition of spatial data may vary slightly depending on applications. In resource exploration, the data gathering step usually comes first along with field work. Information layers obtained from field observation or field measurements using specific instrumentation are called "spatial data" which are often symbolic models or a collection of symbols. As an example, a geological map is a symbolic model that includes all the geological symbols mapped in the field. Most survey data for non-renewable resource exploration are thus spatial data, whereas many of the data sets used in renewable resource studies also include (multi-) temporal information. In hydrocarbon and mineral exploration, the most important and basic information layer is the geological map. The next most important spatial data is the topographic map, including both paper and digital maps, necessary as a base map.

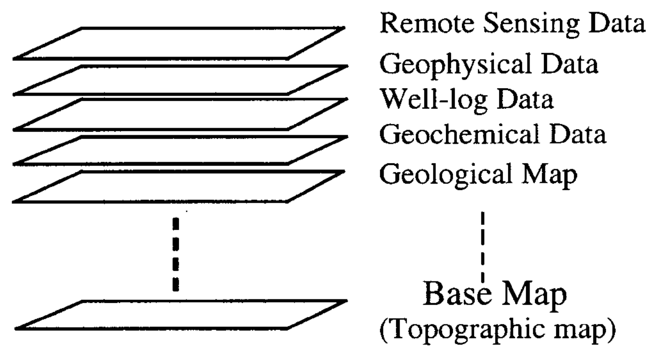

Fig. 1. Typical spatial data layers in resource exploration.

Other types of exploration data include geological structure maps, geochemical survey data, well-log data and geophysical survey data, including various types of seismic survey sections (Fig. 1). These real world observations and measurements are characteristic of the instrument adopted and the personnel involved in the survey, and are usually represented by field variables and spatial objects. Definitions and detailed description of spatial objects are given in many textbooks (e.g., Bonham-Carter, 1994; Burrough, 1986) and will not be repeated here.

Spatial objects are usually described in terms of dimensions (0-D, 1-D, 2-D, and 3-D) and in terms of mode of occurrence at the time of observation. Observed gravity values at each gravity station and geochemical analysis results of samples at each sampling site represent point data and they are typical 0-D data. Lithological boundaries of geological formations are 1-D spatial objects and anomalous zones on a geochemical map represent 2-D spatial objects. Similarly, a geometric description of a subsurface ore body constitutes a 3-D spatial object. Salt domes and underground cavities are also 3-D spatial objects. In applications where the geometric dimension, n, is more than 3, such as multi-spectral remote sensing data, a feature defined by each spectral window datum can be described as a hyperspectral or n-D spatial object (Benediktsson et al., 1997).

6.2. Spatial Reasoning in Non-renewable Resource Exploration

An exploration project involves several steps: planning, field survey, data processing, digital data integration or data fusion, visualization of the fused information with respect to the target proposition, interpretation and decision making. Among these steps, the information representation, digital data fusion, visualization, interpretation and decision making steps are often collectively referred to as a spatial reasoning process, a term probably derived from the

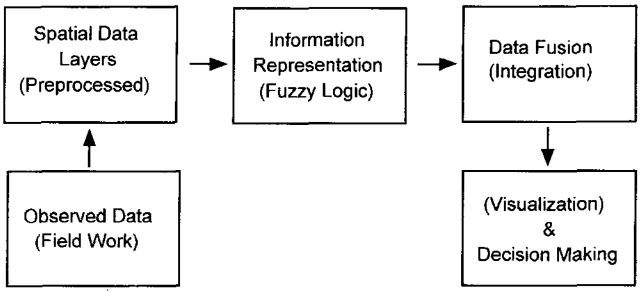

Fig. 2. Block diagram of spatial data processing steps.

discipline of artificial intelligence (AI) and expert system study in computer engineering (Fig. 2).

In a mathematical form, the spatial data fusion steps may also be explained as a mapping or a transformation between the raw field data and the final, digitally fused, information with respect to a chosen exploration hypothesis. Suppose that n maps represent various exploration data such as geology, specific geochemical survey data, etc. in a prospecting area A. Evidence contained in the ith layer of the n map database can be denoted by E_k (k= 1, 2, 3, ⋯, n). The complete set of the n layer database in A may be written as

$$E=\{E_1, E_2,\cdots, E_n\}.$$

For each layer of exploration data, information representation may be scaled with respect to the target hypothesis in such a way that

$$g_k : E_k \Rightarrow [0, 1],$$

where $\gamma_k(e)$ represents the actual information content with respect to the exploration target model. Here $g_k(e)$ also defines a mapping $G(E)$ from the observed data space to the exploration information space.

Suppose that $d_k(e)$ is defined as the "probability" that e of E_k is related to having an exploration target at a point p in A where the observation e is made. Here the word "probability" does not necessarily have the same meaning as in "probability and statistics". The mapping d_k for each layer of exploration evidence, E_k represents the degree of "compatibility" of the observation e at p in A (Moon, 1993). Let us now define a proposition ET (exploration target) that a point p in A belongs to a deposit of the target type. A membership function

$$U_k(ET\,|\,e)=d_k(e)$$

then represents the degree of certainty that p is a member of the set of points which belong to a deposit, given e in A.

Now consider a set of n observed values $\{e_1, e_2, \cdots, e_n\}$ of pieces of evidence $\{E_1, E_2, \cdots, E_n\}$ at a point p in A. Supposeare $\{d_1, d_2, \cdots, d_n\}$ are defined on $\{E_1, E_2, \cdots, E_n\}$. We then have representations $\{d_1(e_1), d_2(e_2), \cdots, d_n(e_n)\}$ at the point p where the observations are made for ET. We now wish to integrate these n representations into one single function. The integration rules depend on the interpretation of the mapping used (Moon, 1993). In general, a traditional probabilistic approach has been used most frequently, however an evidential belief function approach has also been used (Moon, 1990; An, 1992; Bonham-Carter, 1994). Although the basic concept is very similar, the mathematical details and interpretation of the results are quite different.

Suppose we now have the fuzzy membership function, $U_k(ET|e)=d_k(e)$ for all k = 1, 2, 3, ⋯, n for each layer of evidence $\{e_1, e_2, \cdots, e_n\}$ at a point p in A. Now, our task is to define a membership function from membership functions $U_k(ET|e_k)$. This can be accomplished using many different types of operators available in fuzzy set theory as discussed above. Choice of operator or operators depends on the exploration hypotheses (model) and the data sets available. Among the fuzzy operators listed above, the fused membership function using the algebraic sum operator takes the form of

$$U(ET\,|\,e_1, e_2, \cdots, e_n) = \sum_{k=1}^{n} U_k(ET\,|\,e_k) + \sum_{k=1}^{n}\sum_{j=k+1}^{n} U_k(ET\,|\,e_k)\,U_j(ET\,|\,e_j) + \cdots + (-1)^n U_1(ET\,|\,e_1)\cdots U_n(ET\,|\,e_n)$$

and similarly the fused membership function using the γ-operator is

$$U(ET\,|\,e_1, e_2,\cdots, e_n) = \left[\prod_{k=1}^{n} U_k\,(ET\,|\,e_k)\right]^{(1-\gamma)} - \left[1-\prod_{k=1}^{n}(1-U_k(ET\,|\,e_k))\right]^{\gamma}$$

where $0 \le \gamma \le 1$.

Spatial information representation and fusion procedures using classical probability and evidential probability are similar in their basic approach. The data layers collected from an exploration project are represented with conditional probabilities $Prob_k(e|ET)$ that the observation e of E_k is made at a point p with a condition that the location p contains target deposit. One then can combine the individual target probability contributions using Bayes theorem (Moon, 1993). Similarly, if an evidential belief function is used for representing multiple layers of exploration data, Dempster's rule can be employed for fusion of the multiple layers of exploration data (Moon, 1990; 1993). If there are unsurveyed or missing sections of data, an evidential belief function approach is particularly effective because of further quantization of spatial information (or unknowns) in terms of plausibility and ignorance functions. These two additional mappings of plausibility and ignorance along with the weighted evidential belief function representation make the final decision making process considerably more accurate and precise. This unique feature may be crucially important with exploration projects with sparse survey coverage.

7. APPLICATION FORMULATION

Digital data fusion of multiple layers of exploration data is, in fact, a mapping process, which transforms multiple layers of unweighted survey data into a single fused layer of information tuned towards the exploration target. Let $[E, U(ET|E)]$ be the multiple sensor input and output vector pairs where $E=\{E_1, E_2, \cdots, E_n\}$ is the input exploration data and $U(ET|e_1, e_2, \cdots, e_n)$ is the final fused information. The mapping $\hat{G}$ can then be defined such that

$$U(ET\,|\,e_1, e_2, \cdots, e_n)=\hat{G}\,(w, E)+\varepsilon,$$

where the final fused information layer $U(ET|e_1, e_2, \cdots, e_n)$ is the final m dimensional output depending on the problem formulation, w represents the weighting for each exploration data layer with respect to the chosen exploration model (or target hypothesis), and $\varepsilon=\{\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_m\}$ is an error vector associated with $U(ET|e_1, e_2, \cdots e_n)$. Even in a single exploration area with a given set of survey data, w can be quite different depending on the target deposition model, therefore, the mapping $\hat{G}$ includes two stage mapping: information representation and digital fusion (Moon, 1993).

8. FUZZY-NEURAL NETWORK

Fuzzy set theory has relatively well understood for approximate reasoning and fusion of geological exploration data (An et al., 1991; An, 1992; Moon, 1993; Bonham-Carter, 1994). Although fuzzy set theory has also been applied successfully in various fields of science and engineering, there still exists two difficulties: lack of guidelines to decide or adjust the membership functions of fuzzy variables, and short of algorithms for automatic rule generation. Therefore, many investigators have recently tried to combine the concepts of neural network with fuzzy set theory to overcome problems (Takagi and Hayashi, 1991; Yager, 1992).

Generally neural networks are good at approximating nonlinear systems as the fuzzy set theory is used for simplifying nonlinear information structure. Learning ability, parallel processing and distributed knowledge representation are the major features of neural networks. On the other hand, the neural network suffers from unstructured knowledge representation, opposite to the structured knowledge representation in fuzzy systems. Therefore, the combination of the two theories would be able to provide a new paradigm to model realistic, often

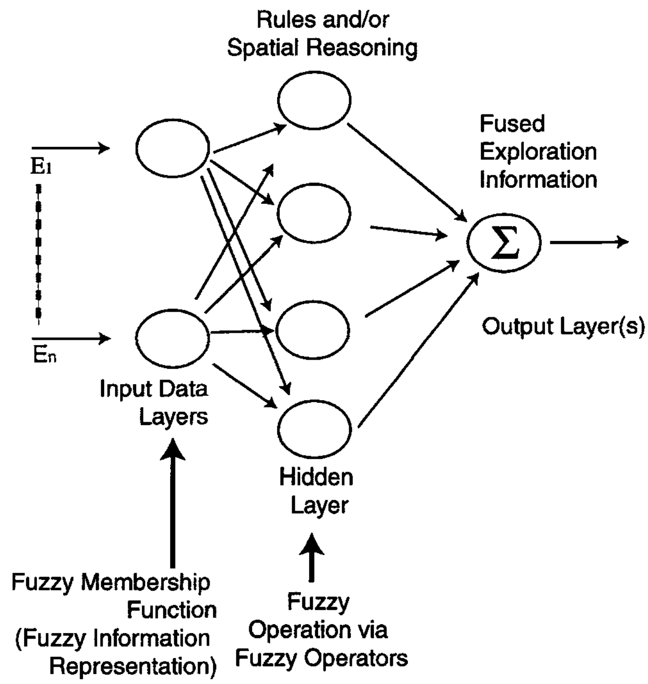

Fig. 3. Structure of three layer fuzzy neural network for exploration data processing and fusion.

nonlinear exploration models with high degrees of uncertainty. Fuzzy neural networks are increasingly used in many other disciplines, particularly in engineering control system applications (Yager, 1992; Hauptmann and Heesche, 1995; Yao and Kuo, 1996).

A schematic diagram of a three-layer fuzzy neural network structure is shown in Figure 3, representing a multiple data set exploration project. In many engineering problems, the hidden layer can be adequately trained with the known input and an expected or desired output, however, many geological exploration tasks depend on either a selected deposit model or models, in which case the hidden layer will closely represents the deposit model and the exploration reasoning process will depend largely on the model parameters. If one chooses to adopt a data-driven exploration approach, the hidden layer can pose problems because the final output is completely unknown in many resource exploration problems and the hidden layer cannot be trained adequately.

With the evolution of second generation fuzzy systems, a considerable amount of research has focused on the integration of fuzzy logic and artificial neural networks, giving birth to neuro-fuzzy systems or an Integrated Neuro-Fuzzy System (INFS). The term neuro-fuzzy generally refers to the fusion of fuzzy systems and neural networks with the aim of combining the advantages of both paradigms and at the same time compensating for the inadequacies with respect to the target hypothesis. In many exploration problems, the deposit models are neither robust nor universal, and an application of a neural network approach of data fusion does not appear to be suitable. But some of the data layers such as a set of geophysical data can be modeled using a specific mathematical model, in which case an optimized inversion of limited amount of observation may be used for training of the hidden layer. In the data-driven exploration model, we usually do not have any background information with respect to the exploration hypothesis and we have to guess the information content of the observed input data. Subsequently, the choice of fuzzy operators that will replace some of the hidden elements of the neural network to be adopted can be somewhat arbitrary.

9. ERROR AND UNCERTAINTY ANALYSIS

There are, in general, two types of uncertainties and errors in integration and fusion of exploration data sets. The first type of error and uncertainty is introduced to the data from the beginning when experiments are carried out or when the field surveys are performed. Analysis of this type of error and uncertainty is relatively well established and there are several methods of estimating and evaluating them. The second type of error and uncertainty is the one introduced during the information representation and digital fusion of the preprocessed information layers because of vagueness of the original information contained in the data sets with respect to the exploration target. The impreciseness of mathematical operators employed during the digital fusion stage introduces additional uncertainty (An et al., 1994). This has been an inherent problem with data fusion using both fuzzy logic and evidential belief function approaches. In a neural network approach, the design of the network and training of the neurons using both the input and output information can minimize this type of error or uncertainty.

The fusion of exploration data discussed above includes independent interpretation of each data layer E_k ($k=1, 2, 3, \cdots, n$) and mapping of the data into a fuzzy space, assigning membership functions to each data object. These errors and uncertainties associated in the fuzzy membership functions subsequently propagate through the entire reasoning process towards the final membership function (An, 1992). In the above, the algebraic product of n membership functions is given as

$$\mu_{comb}(x)=\prod_{i=1}^{n}\mu_i \,.$$

Let μ_{-i} be the algebraic product with the ith term deleted i.e.,

$$\mu_{-i}=\mu_1,\cdots,\ \mu_{i-1},\ \mu_{i+1},\cdots,\ \mu_n \,.$$

Then, we have

$$\frac{\partial\mu_p}{\partial\mu_i}=\mu_{-i} \,.$$

Let ε_i be the error term associated with the membership function from the ith E_i. The error, ε_p, associated with algebraic product can be estimated using Taylor expansion with the second and higher terms ignored and is given as

$$\varepsilon_{comb}=\sum_{i=1}^{n}\mu_{-i}\,\varepsilon_i =\mu_p\sum_{i=1}^{n}\frac{\varepsilon_i}{\mu_i} \,.$$

The ε_p can then be further normalized using the resulting membership function μ. The normalized error is, in fact, the relative error. Similarly, if we denote the algebraic sum discussed above as

$$\mu_{A-SUM}=1-\prod_{i=1}^{n}(1-\mu_i) =1-\nu$$

where $\nu=\prod_{l=1}^{n}(1-\mu_i)$, the relative error associated with the algebraic sum operation becomes

$$\varepsilon_{A-SUM}=\frac{\nu}{1-\nu}\sum_{l=1}^{n}\frac{\varepsilon_i}{1-\mu_i} \,.$$

In the case of the γ-operator, the relative error in the final fused information becomes

$$\varepsilon_{\gamma}=(1-\gamma)\,\varepsilon_{comb}+\varepsilon_{A-SUM}$$

(An, 1992; Moon, 1993). Numerical computation of the above errors which propagate through the fusion (or aggregation) processes is a relatively easy task, whereas quantitative estimation of the errors and uncertainties introduced through information representation and data fusion operators is considerably difficult.

Another further complication is the fact that, as discussed above in Section 7, the multiple layers of data sets which we collect in the field for a specific exploration project are often inter-linked through a very complex deposit model. Some of these data sets are even dependent on each other, making the whole data fusion process even more complicated (Jiang et al., 1997).

As we discussed at the beginning of this paper, the membership functions are relative. A fuzzy membership function close to 1.0 does not necessarily mean certainty with respect to the target hypothesis. If, however, the final fused membership value is higher at one location than at another, it does mean that the possibility for the target proposition is higher at the former location than at the latter. The errors or uncertainties propagated through the fusion of various exploration data layers do not necessarily have the same precise meaning. In this sense, the interpretation of the errors and uncertainties in the final results is more vague than the error analyses of many physical experiments.

10. CONCLUSION

Exploration strategies for non-renewable resources have been changing rapidly along with the accelerating innovations in computer hardware and information processing technology. Although the basic philosophy of geological exploration remains the same, the methodology has been changing to the extent that we now utilize multiple sets of exploration data simultaneously, and visualize the final target information in an optimum way with minimized uncertainties. There are still a number of areas for further investigation and many of the mathematical tools we employ require further research. The main problem we can identify at this point is the exact quantization of the field survey data. There are several ambiguities. First, most exploration data we collect in the field are 1-D, 2-D, and rarely 3-D, whereas almost all exploration targets are 3-D. This requires interpolation of the available information by a geologist utilizing, in the past, his or her own expert intuition, but recently more utilizing accurate and more sophisticated mathematical approach. Second, most exploration equipment we routinely use is designed based on simple principles with a limited number of parameters representing the overall physical characteristics. Therefore, the survey data we collect with these instruments can provide us with only partial information.

To overcome the inherent ambiguities in modern exploration techniques (see above), exploration geologists have used several mathematical tools such as statistical and probability theory, fuzzy logic approach of quantizing vagueness, and evidential belief function method in addition to the conventional data processing techniques.

The fuzzy set approach of converting exploration field data into exploration information is not exact in mathematical sense but provide us with an opportunity to carry out spatial reasoning process with less subjective human intervention. It also improves the consistency of exploration data processing and considerably reduces mistakes. However, one of the problems of using fuzzy logic approach to processing exploration data lies in the availability of optimum operators with respect to the target proposition. Many of the simple fuzzy operators being used by geologists now are designed by engineers for different applications and there is no one fuzzy operator which can adequately be utilized for all deposit models (Jiang et al., 1997). It may in fact be impossible to expect one fuzzy operator to integrate and fuse multiple sets of exploration data. Even for one type of base metal exploration, several fuzzy operators are used in parallel and some in combination (Moon and Jiang, 1995; Jiang et al., 1997). Application of fuzzy operators both in parallel and/or in combination can adequately fuse partial information contained in each data layer with respect to the target hypothesis. The problem of conditional dependency of certain data layers can also be resolved in this approach. A marriage of fuzzy information representation and fuzzy reasoning process with neural networks can provide more effective and accurate solution, however, development of a neuro-fuzzy system for resource exploration problems requires further research and investigation.

Error and uncertainty analysis is an important step in any integrated exploration data processing. There are two types of errors and uncertainties in fuzzy data fusion and information imaging of exploration data. The errors and uncertainties originating from direct experiments and field survey can be adequately estimated following the traditional approach. The errors and uncertainties introduced from the fuzzy membership assignment and during the application of fuzzy operator(s) are, however, considerably more complicated and again require further investigation and research.

ACKNOWLEDGMENTS: This research is supported by an NSERC of Canada operating grant #7400 to W.M. Moon. Most of the material in this paper was prepared during my three short visits to the Seoul National University from Canada during the 1997–1998 winter. The author would like to thank my colleague and friend, Professor Chang-Eob Baag, who has been most kind throughout and made my visit pleasant.

REFERENCES

An, P., 1992, Spatial Reasoning and Integration Techniques for Geophysical and Geological Exploration Data. Ph.D. thesis, University of Manitoba, Winnipeg, 280 p.

An, P., Moon, W.M. and Rencz, A., 1991, Integration of geological, geophysical and remote sensing data using fuzzy logic. Journal of Canadian Society of Exploration Geophysicists, 27, 1–11.

An, P., Moon, W.M. and Bonham-Carter, G.F., 1994, Uncertainty management of exploration data using the Belief Function. Nonrenewable Resources, 3, 60–71.

Benediktsson, J.A., Ingimundarson, J.I. and Sveinsson, J.R., 1997, Classification and feature extraction of hyperdimensional data using LOOC covariance estimation. IGARSS'97 (Proceedings), Singapore, August, p. B10.03.

Bonham-Carter, G.F., 1994, Geographic Information Systems for Geoscientists. Pergamon Press, Kidlington, 398 p.

Burrough, P.A., 1986, Principles of Geographic Information Systems for Land Resources Assessment. Clarendon Press, Oxford, 194 p.

Goguen, J.A., 1967, L-fuzzy sets. Journal of Mathematical Association of America, 18, 145–174.

Goguen, J.A., 1969, The logic of inexact concepts. Syntheses, 19, 325–373.

Hauptmann, W. and Heesche, K., 1995, A neural net topology for bidirectional fuzzy-neuro transformation. IEEE International Conference on Fuzzy Systems (Proceedings, Vol. 3), Yokohama, March, 1995, p. 1511–1518.

Ishibuchi, H., Morisawa, T. and Nakashima, T., 1996, Voting schemes for fuzzy-rule-based classification systems. IEEE International Conference on Fuzzy System (Proceedings, Vol. 1), New Orleans, September, p. 614–620.

Iyanaga, S. and Kawada, Y., eds., 1980, Encyclopedic Dictionary of Mathematics. MIT Press, Cambridge, p. 1167–1170.

Jiang, W.W., Moon, W.M. and Feng, L.N., 1997, Data fusion Research for Mineral Exploration Data from Hwanggang-ri Region, Korea. Technical Report #17, University of Manitoba, Winnipeg, 62 p.

Klement, E.P., 1980, Fuzzy measures by classical measures. In: Wang, P.P. and Chang, S.K. (eds.), Fuzzy Sets. Plenum Press, New York, p. 25–33.

Mathai, A.M. and Rathie, P.N., 1975, Information Theory and Statistics. John Wiley and Sons, New York, 137 p.

Moon, W.M., 1990, Integration of multiple sets of geophysical information using evidential belief function. IEEE Transaction of Geoscience and Remote Sensing, 28, 272–278.

Moon, W.M., 1993, Mathematical basis for geophysical information representation and integration. Canadian Journal of Remote Sensing, 19, 63–67.

Moon, W.M., 1995, Information representation and integration of multiple sets of spatial geoscience data. IGARSS'95 (Proceedings), Florence, August, 1995, p. 1021–1022.

Moon, W.M. and Jiang, W.W., 1995, Integrated Exploration of Base Metal Deposits in East Java, Indonesia. Technical Report #14, University of Manitoba, Winnipeg, 34 p.

Moon, W.M., Jiang, W.W., Yun, S.G. and So, C.S., 1998, Data fusion modeling of the Hwanggang-ri mineralization zone. Canadian Journal of Remote Sensing. Remote Sensing. (submitted)

Pawlak, Z., 1996, Why rough sets? IEEE International Conference on Fuzzy System (Proceedings, Vol. 2), New Orleans, September, p. 738~743.

Shannon, C.E., 1948, A mathematical theory of communication. Bell System Technology Journal, 27, 379–423, 623–656.

Takagi, H. and Hayashi, I., 1991, NN-driven fuzzy reasoning. Journal of Approximate Reasoning, 5, 191–212.

Wang, P.P. and Chang, S.K., eds., 1980, Fuzzy Sets. Plenum Press, New York, 413 p.

Yager, R.R., 1992, Implementing fuzzy logic controllers using a neural network framework. Fuzzy Sets and Systems, 48, 53–64.

Yao, C.C. and Kuo, Y.H., 1996, A fuzzy neural network model with three-layered structure. IEEE International Conference on Fuzzy System (Proceedings, Vol. 3), New Orleans, September, p. 1503–1510.

Zadeh, L.A., 1965, Fuzzy sets. Information and Control, 8, 338–353.

Zadeh, L.A., 1968, Probability measures of fuzzy events. Journal of Mathematical Analysis and Application, 23, 421–427.

Zadeh, L.A., 1987, Fuzzy sets. In: Yager, R.R., Ovichinikov, S., Tong, R.M. and Nguyon, H.T. (eds.), Fuzzy Sets and Applications. John-Wiley, New York, p. 29–44.

Zadeh, L.A., 1987, Fuzzy sets as a basis for a theory of possibility. In: Yager, R.R., Ovichinikov, S., Tong, R.M. and Nguyon, H.T. (eds.), Fuzzy Sets and Applications. John-Wiley, New York, p. 193–218.

Zadeh, L.A., 1987, A theory of approximate reasoning. In: Yager, R.R., Ovichinikov, S., Tong, R.M. and Nguyon, H.T. (eds.), Fuzzy Sets and Applications. John-Wiley, New York, p. 368~412.

Zimmermann, H.J., 1984, Fuzzy Set Theory and Its Applications. Kluwer-Nijhoff Publishing, Boston, 357 p.

Zimmermann, H.J. and Zysno, P., 1980, Latent connectives in human decision making. Fuzzy Sets and Systems, 4, 37–51.

Manuscript received August 20, 1998
Manuscript accepted September 17, 1998