01 Jan 1995

Using Fuzzy Set Theory: Exploring How Generalties Can Pinpoint Where to Look for the Answers

Cosmin Radu

Ralph W. Wilkerson
Missouri University of Science and Technology, ralphw@mst.edu

Using fuzzy set theory

Exploring how generalties can pinpoint where to look for the answers

Cosmin Radu and
Ralph Wilkerson

©Guido Rossi/The Image Bank

Very often, precise, quantitative analysis proves not to have much relevance when solving people's real world problems. In these instances, a fuzzy approach attempts to address this aspect of human thinking typically neglected. It is based on the premise that humans don't see classes of objects as totally disjointed but rather as sets where transitions from membership to non-membership is gradual. This and the observation that humans do not use the traditional two-valued logic—1 or 0—have helped produce a new mathematical domain: fuzzy mathematics.

Two main trends should be mentioned here: the fuzzy set theory and the fuzzy logic. Both build upon set theory and logic, respectively. Three features distinguish the approaches:

1. the use of so called linguistic variables, instead of or together with numeric variables;
2. the use of fuzzy conditional statements to represent simple relations between variables;
3. the characterization of complex relations by fuzzy algorithms.

Fuzzy linguistic variables and fuzzy algorithms offer an effective, more flexible way to describe a system's behavior too complex for a classical mathematical model. They are very successful in economics, management science, artificial intelligence, information retrieval systems, pattern recognition, image processing, psychology, biology, and other fields rendered inherently fuzzy do to the unpredictable behavior of their components.

Expert systems

Expert systems are computer programs that emulate the reasoning process of a human expert or perform in an expert manner in a domain for which no human expert exists. Typically, they reason with uncertain and imprecise information. The knowledge is often inexact in the same way that a human's knowledge is imperfect. The facts—or user supplied information—are also uncertain.

A close examination of trends in expert systems development shows that expert systems have evolved mainly into fuzzy expert systems and hybrid systems. As opposed to the probabilistic approaches, fuzzy mathematics offer a more intuitive mapping to real world problems.

The most successful examples of dealing with imprecision come from the field of fuzzy below control. The diagram illustrates the core of a fuzzy expert system:

Inputs ⇒ Fuzzifier → Interface Engine → Defuzzifier ⇒ Outputs
↖ ⇑ ↗
Knowledge Base

The system's inputs go through a fuzzifier. The inference engine works with attribute values, which have fuzzy memberships attached. They may be created from real-valued attributes which have been partitioned into individual fuzzy sets. The inference engine provides a fuzzy output which may need to be defuzzified. Certainly, in the control domain the output must be defuzzified so that there is a single, well defined control action taken.

Case study: oil exploration

For this oil exploration expert system, two major requirements need to be fulfilled: a) the multidisciplinary data must be available (meaning knowledge of possibly hundreds of experts) plus a way to integrate it; and b) a way to handle the non-precise, subjective nature of the rules.

The first problem can be addressed by using evidential reasoning (Dempster—Shafer Theory—DST), while the second suggests the use of fuzzy mathematics. Fuzzy expert systems eliminate the need to introduce hundreds of rules to represent a simple concept; using fuzzy rules inference becomes a "process of propagation of elastic constraints," according to Dr. Lofti Zadeh, father of fuzzy sets.

Evidential reasoning can be considered a special case of fuzzy logic; its foundations have been put by Dempster and Shafer. It represents an effective way of representing "ignorance," incomplete information or inexact rules; it also handles conflicting data and rules (conflict management).

The main difference between DST and fuzzy set theory is that, while in the latter framework $\mu(\bar{A}) = 1-\mu(A)$, i.e. the measure of an object membership to

0278-6648/95/$4.00 © 1995 IEEE

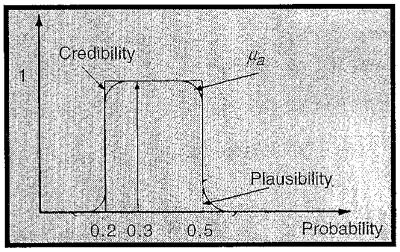

Fig. 1 Comparison of Probability function in Classical Theory(P=0.3), DST(P: (0.2, 0.5)) and Fuzzy Logic(μ_a= Membership Function)

$\overline{A}$ is determined, in DST knowledge of probability of *A*(*Belief*(*A*)) doesn't provide knowledge of *Belief* ($\overline{A}$), i.e. *Belief*(*A*) + *Belief* ($\overline{A}$) + *Ignorance*(*A*) = 1. Generally, we denote *Belief*(*A*) + *Ignorance*(*A*) as *Plausibility*(*A*).

To understand the way evidential reasoning handles conflicting data let us perform the integration of information provided by a hypothetical geological and geochemical analysis. Assume that geological data indicates a 40 percent chance of "gas or oil" saturation, and the geochemical data a 50 percent chance of "water" saturation and 30 percent chance of "oil" saturation. There are 60 percent and 20 percent ranges of ignorance for the sets of data respectively. Table 1 shows how the knowledge from two different frames is combined. Notes:

1. the credibility of the result is the product of the credibility factors of the components;
2. combining a row and a column having common elements results in information about the particular common elements;
3. when the unidentified field is combined to anything else, the resulting credibility refers to that other element;
4. combining "gas or oil" with "water" results in a conflict, with a 0.2 factor.

The next step is to normalize the resulting credibility factors by eliminating the conflict. This is done by multiplying everything with a factor $\frac{1}{1 - Credconfl} = 1.25$.

The final results are:

gas or oil	0.1
water	0.375
oil	0.375
unidentified	0.15

Note that the ignorance range is smaller than both initial ignorance ranges. The final credibility and plausibility factors are: gas (0, 0.25), oil (0.375, 0.625), water (0.375, 0.525). This process can be repeated for any number of data sources. Furthermore, Aminzadeh proved that the result of integration is the same irrespective of the order in which the knowledge sources are combined.

Table 1 Integrating data using ER methods

	Water 0.5	Oil 0.3	Unidentified 0.2
Gas or Oil 0.4	Conflict 0.2	Oil 0.12	Gas or Oil 0.08
Unidentified 0.6	Water 0.3	Oil 0.18	Unidentified 0.12

Fuzzy neural computing

Current studies in neural networks can be classified into three main directions of study:

1. Modeling a single neuron (either as a static one, a dynamic one or a domain specific one).
2. Architectural issues.
3. The operational aspect of neural computation, in which fuzzy mathematics have been integrated to enhance the uncertain information process capabilities of neural networks.

This last direction was the starting point of a new class of neural networks: the fuzzy neural networks (FNN). But first, how can different models of neurons be integrated in a single unified framework. Two main neuronal morphologies are widely used: *Product-Summation-Monotonicity*(PSM) and *Difference-Summation-Radial*(DSR). Aside from the particular way of realizing the computations, all neuronal processing yields a degree of similarity between the input signals it receives and the neuron's weight vector. This similarity is obvious once we consider the neuron as a semilinear application from a vector space (determined by the input signals) to the real numbers vector space (its output).

The general operations executed at the neuron level are:

Synaptic operation: mapping the input signal, to a real value, using the weight corresponding to the specific synapse.

Somatic operations: aggregation (combining the results of all synaptic operations into a real value), *thresholding* and *nonlinear activation* (maps the previous result, to a real value, through a nonlinear real function).

The semantics of each operation is: the synaptic operation and the aggregation provide a *degree of mutual relationship* between the input vector and the strength of the synaptic weighting vector (which represents the accumulated past experiences). If this value is higher than a threshold value, the degree of similarity will be extracted and a graded output will be yielded through a nonlinear activation function.

We can now detail the two kinds of morphologies we specified earlier, in terms of this general framework:

PSM Neuronal Form: the degree of *similarity* it computes represents the inner product of the input vector and weight vector (the *projection* of the input vector on the direction of the weight vector), the activation function being a monotonical nonlinear function.

DSR Neuronal Form: a degree of *dissimilarity* is computed as the *Euclidean distance* between the two vectors in the n-dimensional Euclidean space, the activation function being a radial nonlinear function.

If we can define a vector space with fuzzy characteristics, we can generalize the above neuron structures to a fuzzy neuron (processing fuzzy inputs, through fuzzy weight vectors). Fortunately such a mathematical framework exists; it is the fuzzy space and fuzzy vectors theory. In this theory, concepts such as fuzzy vectors and fuzzy matrices have been defined as well as operations such as fuzzy vector projection, addition, subtraction, distance between two fuzzy vectors, inner product, vector-matrix multiplication, correlation product of two vectors.

In such a fuzzy vector space, a fuzzy set can be regarded as a vector in a unipolar hypercube. Each coordinate representing the value of the membership function for each fuzzy variable defined in any of the given universes of discourse. Indeed using these operations, fuzzy variants have been defined.

Based on the actual processing, several classes of FNN may be discerned:

Non-Fuzzy Neuronal Models are usually feedforward neural networks in which an n-dimensional input vector representing real-valued membership grades is mapped to an m-dimensional output vector.

The weights in this network can be real numbers, and the learning algo-

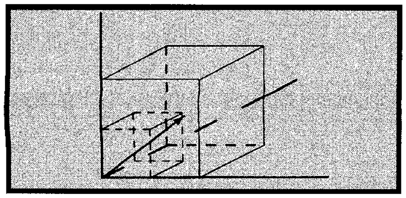

Fig. 2 A fuzzy vector in a 3-dimensional fuzzy hypercube; each dimension corresponds to a fuzzy variable defined in the given universes of discourse.

rithm any appropriate learning algorithm used in standard neural networks.

Fields for this kind of network include: the fuzzy decision making, diagnostic domain, fuzzy system modeling, and neuro-fuzzy controllers design.

Pattern recognition

Pattern recognition is part of most of today's hot fields of study:

Man-machine communication: automatic speech recognition, optical character recognition systems, speech understanding, picture understanding;

Crime and criminal detection: fingerprint, handwriting;

Biomedical application: ECG, EEG, EMG Analysis, X-ray analysis, diagnosis;

Military applications: detection of nuclear explosions, missile guidance and detection, radar and sonar detection, target identification, naval submarine detection;

Industrial applications: computer aided design and manufacturing, nondestructive testing;

Robotics and AI: intelligent sensor technology, natural language processing.

Computer pattern recognition can be viewed as a task consisting of a) learning the invariant and common properties of a set of samples characterizing a class, and b) deciding whether a new sample is a member of the class or not, by selecting and extracting its properties. Fuzzy set theory provides suitable tools and techniques for analyzing complex systems and decision processes where patterns are indeterminant due to inherent variability and/or vagueness rather than randomness.

Two general methods are used to approach this task:

Decision Theoretic approach is a series of mappings (transformations) conserving the class discrimination. The first mapping is from the measurement space to a feature space (usually a finite, lower dimensional space, containing sufficient information to successfully perform the classification problem), and then from this feature space to a decision space. This last mapping is done on the basis of a characterizing function, called a discriminant function in the case of deterministic classification technique, a probability density function in statistical decision theory and a membership function in the context of fuzzy set theory.

The *Syntactic approach* is used in problems where structural information plays an important role in describing the patterns. Typical examples are picture recognition, fingerprint recognition, chromosome analysis, character recognition, scene analysis. In such cases, where the patterns are complex and the number of possible descriptions very large, it is impractical to regard each description as defining a class. Rather, a description in terms of small sets of simple subpatterns (primitives) and grammatical rules derived from formal language theory become necessary.

The problem of extracting and selecting features in the decision theoretic approach is similar in nature to that of primitives in syntactic approach. However, the primitives reflect more local information, while the features may represent, in general, any set of numerical measurements taken from the pattern.

In practical situations most patterns are noisy or distorted. This means that the string corresponding to a noisy pattern may not be recognized by any of the pattern grammars. This problem can be dealt with by:

1. Using approximation in the early stages of processing (preprocessing and primitive extraction).
2. Using transformational grammars—defining a relation between noisy patterns and their noise-free correspondents; if this succeeds, the problem reduces to classify noise-free patterns.
3. Using stochastic grammars—this means to assign a probability to a string being a proposition of the generated language; the decision is made based on maximum probability.
4. Using fuzzy grammars—each proposition of the generated language is assigned a membership value to the set; as above, decision is made based on maximum value of membership function. Besides that, the pattern primitives may be fuzzy sets (e.g. "almost circular arcs, gentle slope").

An important point is the fact that pattern classification is intrinsically unsuited for precise mathematical approaches. Because of this, the conceptual structure of fuzzy sets theory may provide a more natural setting.

Conclusion

Today, the world of science is still debating what fuzzy set theory's place is in the big picture. Fields such as artificial intelligence, exclusively committed to symbolic manipulation, refuse any numerical methods, including fuzzy and neural methods. Despite this, fuzzy theory has proved effective in major control applications as well as in most intelligent systems.

The future might provide the cross-fertilization of fuzzy set theory and other areas. This trend is already visible in the newly conceived fuzzy neural architectures, the commercially available fuzzy controllers and the fuzzy expert systems that successfully perform in industry.

Read more about it

• Fred Aminzadeh and Mohammad Jamshidi, editors. *Soft Computing—Fuzzy Logic, Neural Networks and Distributed Artificial Intelligence.* Environmental and Intelligent Manufacturing Systems Series. Prentice-Hall, '94.

• R. L. Chang and T. Pavlidis. Fuzzy decision-tree algorithms. *IEEE: Syst., Man and Cyberns.,* SMC-4(1):28-35, '77.

• A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals Mathematical Statistics,* 38(2):325-329, 1967.

• M. Gupta and H. Ding. *Foundations of Fuzzy Neural Computation,* chapter 8. In Aminzadeh and Jamshidi, '94.

• Sankar K. Pal and D. K. Dutta Majumder. *Fuzzy Mathematical Approach to Pattern Recognition.* Halsted Prexpert systems, A Division of John Wiley & Sons, Inc., 1986.

• G. Shafer. *A mathematical theory of evidence.* Princeton University Prexpert systems, 1976.

• Lofti A. Zadeh. Knowledge representation in fuzzy logic. *IEEE: Knowledge and Data Engineering,* 1(1):89-100, 1989.

About the authors

Cosmin Radu is a graduate student in Computer Science and Engineering Management. Dr. Ralph Wilkerson is a Professor in the Computer Science Department. Both are at the University of Missouri-Rolla.