PROCESSING FUZZY SET MEMBERSHIP FUNCTIONALS AS VECTORS

Rui J. P. de Figueiredo
Laboratory for Intelligent Signal Processing and Communications
California Institute for Telecommunications and Information Technology
University of California, Irvine
Irvine, CA 92697-2800
Email: rui@uci.edu
Tel: (949) 824-7043 Fax: (949)854-6528

Abstract: *A framework is presented for processing fuzzy sets for which the universe of discourse $X = \{x\}$ is a separable Hilbert Space, which, in particular, may be a Euclidian Space. In a given application, X would constitute a feature space .The membership functions of sets in such X are then "membership functionals", that is, mappings from a vector space to the real line. This paper considers the class Φ of fuzzy sets A, the membership functionals μ_A of which belong to a Reproducing Kernel Hilbert Space (RKHS) $F(X)$ of bounded analytic functionals on X, and satisfy the constraint $0 \le \mu(x) \le 1, x \in X$. These functionals can be expanded in abstract power series in x, commonly known as Volterra functional series in x. Because of the one-to-one relationship between the fuzzy sets A and their respective μ_A, one can process the sets A as objects using their μ_A as intermediaries. Thus the structure of the uncertainty present in the fuzzy sets can be processed in a vector space without descending to the level of processing of vectors in the feature space as usually done in the literature in the field. Also, the framework allows one to integrate human and machine judgments in the definition of fuzzy sets, and to use concepts similar to probabilistic concepts for the combination of fuzzy sets. Some analytical and interpretive consequences of this approach are presented and discussed. A result of particular interest is the best approximation $\hat{\mu}_A$ of a membership functional μ_A in $F(X)$ based on interpolation $\hat{\mu}_A(v^i) = u^i$ on a training set $\{(v^i, u^i), i = 1,..., q\}$ and under the positivity constraint. The optimal analytical solution comes out in the form of an Optimal Interpolative Neural Network (OINN) proposed by the author in 1990 for best approximation of pattern classification systems in a $F(X)$ space setting. An example is briefly described of an application of this approach to the diagnosis of Alzheimer's disease.*

Keywords: **Zadeh, Fuzzy, Sets, Membership, Functionals, Nonlinear, Analytic, Hilbert Space, Neural Networks, Alzheimer's**

1 INTRODUCTION

From the time they were invented by L. A. Zadeh four decades ago [1], fuzzy sets have played a prominent role in the modeling of uncertainty in the processing of data and information (See, for example, [2-5]).

To enhance their applicability, we present a new framework for processing fuzzy set membership functionals as vectors. In a moment we will explain the use of the more general term "membership functional" for the membership function of a fuzzy set.

The above framework allows one to include a completely additive class of events or attributes as well as human judgment in the definition of fuzzy sets. Also, under appropriate conditions, it permits to view a fuzzy set membership functional as a generalization of the concept of a conditional probability, and fuzzy set combinations as a generalization of analogous conditional-probability-based combinations of sets. Finally, it allows one to assign vector-inspired and appropriately-interpreted attributes to fuzzy sets, such as the *pseudo-norm* of a fuzzy set, the *pseudo-distance* between two fuzzy sets, and the *pseudo-scalar-product* of two fuzzy sets.

To achieve this goal in very explicit terms, we focus attention on the case in which the universe of discourse X is a vector space endowed with a scalar product. Then, under appropriate conditions, we create a Hilbert Space $F(X)$ to which the membership functions of the fuzzy sets under consideration may belong. In such a setting, membership functions become "membership functionals", i. e., mappings from a vector space to the real line. This, in turn, allows us to exploit the one-to-one relationship existing between fuzzy sets and their membership functionals, to obtain the results that we mentioned above.

Let Φ denote the class of fuzzy sets under consideration in this paper. Specifically, we define Φ as the class of fuzzy sets satisfying the following conditions:

(a) The universe of discourse X for the fuzzy sets is a separable Hilbert Space, which may, in particular, be a (possibly weighted) Euclidian Space. Note that then, in general, the membership function μ_A pertaining to any given fuzzy set A is a nonlinear (not necessarily linear) "functional" (rather than a mere function) from the vector space X to the real line R^1. In an application-specific context, X would be called a "feature space".

(b) For any given fuzzy set A, the membership functional μ_A is (or may be approximated by, a) bounded *analytic functional* on X belonging to a Reproducing Kernel Hilbert Space (RKHS) $F(X)$ of bounded analytic functionals on the Hilbert Space X. In addition, μ_A is required to satisfy the constraint $0 \le \mu(x) \le 1, x \in X$. The space $F(X)$ constitutes a generalization of the Symmetric Fock Space, the state space of non-self-interacting Bosons in quantum field theory. This generalization was introduced in 1980 by de Figueiredo and Dwyer [6] for nonlinear signal and system analysis (For some of the other related work of the author, see [7]-[15]). From now on, $F(X)$ will be denoted simply by F when its argument X is clear from the context.

As a fundamental result arising from this generic fuzzy set model, we show that, for any given fuzzy set $A \in \Phi$, the mapping $\mu_A : X \to [0.1]$ can be optimally represented and hence realized by an artificial neural network.

In section 2, we present our formulation for the important case in which X is a (possibly weighted) Euclidian N-space E^N. The case in which X is the Hilbert Space L^2 of square-integrable functions, such as waveforms or video images, is discussed briefly in section 3 on applications. In this section we also briefly describe the results of a computational-intelligence-based study on the diagnosis of Alzheimer's disease in the setting of the approach presented here. In section 4, we summarize our conclusions. Finally, we provide, in an Appendix, a brief overview of some of the fundamentals on the space F underlying our presentation

2. FUZZY SETS IN A EUCLIDIAN SPACE

Let the universal set X be a N-dimensional Euclidian space E^N over the reals, with the scalar product of any two elements $x = (x_1, ..., x_{N)})^T$ and $y = \left(y_{1,...,}y_N\right)^T$ denoted and defined by

$$< x, y >= x^T y = \sum_{i=1}^{N} x_i y_i \, , \tag{1}$$

where the superscript T denotes the transpose. If E^N is a weighted Euclidian Space with a positive definite weight matrix R, the scalar product is given by

$$< x, y >= x^T R^{-1} y \tag{2}$$

Typically, such X would constitute the space of finite-dimensional feature vectors x associated with the objects alluded in a given discourse, as in the following example.

Example 1: **The Space X of Feature Vectors in a Patient Database** The database of a health care clinic for the diagnosis of the conditions of patients w. r. t. various illnesses could constitute a universal set X. In this space, each patient p, would be represented by a feature vector $x = x(p) \in E^N$, consisting of N observations made on p These observations could consist, for example, of lab test results and clinical test scores pertaining to p. Then, a subset A in the feature space X with a membership functional μ_A could characterize the feature vectors of patients possibly affected by an illness Γ_A. For a given patient p, a physician or a computer program would assign a membership value $\mu_A(x(p)) = \mu_A(x)$ to x indicating the extent of sickness, w. r. t. the illness Γ_A, of the patient p (e.g., gradual scoring from very well to fairly well to, okay to slightly sick to very sick). ▌[1]

2.1 Membership Functionals as Vectors in the RKH Space *F(E^N)*

2.1.1 Abstract Formulation

Returning to our formulation, in the present case, F is the space of bounded analytic functionals on a bounded set $\Omega \subset E^N$ defined by

$$\Omega = \left\{ x \in E^N : \|x\| \le \gamma \right\}, \tag{3}$$

where γ is a positive constant selected according to a specific application. Typically, γ may be viewed as the radius of the uncertainty ball in the feature space. The conditions under which F is defined and the mathematical properties of F are given in the Appendix.

Under the assumption made in this paper, namely that the fuzzy sets A under consideration belong to Φ, the membership functionals μ_A belong to F .This is not a serious restriction because, in most instances, μ_A can be approximated by members in F (in a way analogous to that in which a square pulse can be approximated to any desired degree of accuracy by an analytic function in a given region of interest). For simplicity in notation, we will denote μ_A simply by μ when its relationship to a set A is clear from the context

Remark 1. It is important to note that not all the members of F are membership functionals. Only those objects f in F that satisfy the condition

$$f(x) \in [0,1] \ \forall\, x \in \Omega \subset E^N \tag{4}$$

qualify for membership functionals. As explained in Remark 2, this condition is taken into account in the RLS algorithm of section 2.3, which estimates the membership functional of a fuzzy set in Φ from the training data.▌

. In view of the analyticity and other conditions satisfied by members of F, stated in the Appendix, any μ pertaining to a fuzzy set belonging to Φ can be represented in the form of an

---

[1] Ends of formal statements will be signaled by ▌

N -variable power series, known as an abstract Volterra series, in these N variables, absolutely convergent at every $x \in \Omega$, expressible by:

$$\mu(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \mu_n(x), \tag{5}$$

where μ_n are homogeneous Hilbert-Schmidt(H-S) polynomial of degree n in the components of x, a detailed representation of which is given in the Appendix.

The scalar product of any μ_A and $\mu_B \in F$, corresponding to fuzzy sets A and B is defined by

$$\langle \mu_A, \mu_B \rangle_F = \sum_{n=0}^{\infty} \left(\frac{1}{\lambda_n} \frac{1}{n!} \sum_{|k|=n} \frac{|k|!}{k!} c_k d_k \right) \tag{6}$$

where. λ_n . $n = 0, 1, 2, \ldots$, is a sequence of positive weights expressing the prior uncertainty respectively in the terms μ_n, n = 0, 1, 2,…, satisfying (A11) in the Appendix, and c_k and d_k are defined for μ_A and μ_B in the same way as c_k is defined for μ in.(A4) and (A8) in the Appendix.

The reproducing kernel $K(x,z)$, in F is

$$K(x,z) \overset{def}{=} \varphi\left(\langle x,z \rangle\right) = \sum_{n=0}^{\infty} \frac{\lambda_n}{n!} \left(\langle x,z \rangle\right)^n, \tag{7}$$

where $\langle x,z \rangle$ denotes the scalar product of x and z in E^N .

In the special case in which $\lambda_n = \lambda_0{}^n$, φ is an exponential function, and thus $K(x,z)$ becomes

$$K(x,z) = \exp\left(\lambda_0 \langle x,z \rangle\right). \tag{8}$$

We note that $K(x,z)$ is called a Reproducing Kernel (RK) because: (a) $K(x,.)$ is a member of F , and (b) it has the “reproducing property” expressed by the equation (A16) of the Appendix, which we quote below

$$< K(x,.), \mu(.) >_F = \mu(x)$$

Example 2: **Fuzzy Set of “Good Weather” Days.** Let days d be represented by their respective feature vectors $x = x(d)$, the universe of discourse being $X = \{x = x(d)\}$. The components $x_i, i = 1..., N,$ of x would be observation variables, such as temperature, pressure, humidity, wind velocity, etc., associated with the day d that x represents.

Assume that x is a random vector with a co-variance Σ , and denote by $v^i, i = 1,..., N$, the eigenvectors of Σ listed in the order of decreasing eigenvalues. Assume also that most of the relevant information lies in the subspace of X spanned by the first q , say three, eigenvectors of Σ .

Then a fuzzy set A of days that are considered to be “Good” may be represented by the membership functional

$$\mu_A(x) = w_1\varphi\left(\langle v^1, x \rangle\right) + w_1\varphi\left(\langle v^2, x \rangle\right) ... + w_q\varphi\left(\langle v^q, x \rangle\right) \tag{9}$$

where φ is as defined in (7), and the w_i, as in (25), are obtained by choosing a set of prototype days $d_1,...,d_q$,corresponding to different types of weather, and interpolating the membership values $u^1,...,u^q$, assigned by a weather expert to those days, at the feature vectors $x(d_1),...,x(d_q)$ corresponding to those days.∎

2.1.2 The Case in which Fuzzy Sets are Intervals on Real Line

To illustrate the above formulation, we now consider the case in which the universe of discourse X is the real one-dimensional Euclidian space E^1 . Then if x and y are two real numbers constituting vectors in E^1 , their scalar product and the norm of x are defined by

$$< x, y >= xy \qquad (10)$$

$$\|x\| = |x| \qquad (11)$$

The uncertainty region $\Omega \subset E^1$ may be represented by an appropriate interval $|x| \leq \gamma$ of the real line.

Then $\mu(x)$ is represented by (5), where $\mu_n(x)$ simplifies to the expression

$$\mu_n(x) = c_n x^n \qquad (12)$$

The condition (A12) takes the form

$$\sum_{n=0}^{\infty} \frac{1}{n!\lambda_n} {c_n}^2 < \infty \qquad (13)$$

If, accordingly, the membership functionals for two intervals A and B are represented by

$$\mu_A(x) = \sum_{n=0}^{\infty} c_n x^n \qquad (14)$$

$$\mu_B(x) = \sum_{n=0}^{\infty} d_n x^n \ , \qquad (15$$

and the scalar product between μ_A and μ_B is

$$< \mu_A, \mu_B >= \sum_{n=0}^{\infty} \frac{1}{n!\lambda_n} c_n d_n \qquad (16)$$

For example, suppose that A and B are the interval $[0,1]$ with the membership functions

$\mu_A(x) = \exp(-ax)$ and $\mu_B(x) = \exp(-bx)$, and $\lambda_n = \lambda_0^{-n}$, where a, b and λ_0 are positive constants. Then, since using the formula for the power series expansion of an exponential function, we obtain the following expression for the scalar product (16) for this example:

$$< \mu_A, \mu_B >= \sum_{n=0}^{\infty} \frac{1}{n!}(ab/\lambda_0)^n = \exp(ab/\lambda_o) \quad (17)$$

Going further with this simple illustration, one may obtain, in general, a best approximation for a fuzzy interval along the lines described in section 2.3 , as a linear combination of exponentials of the form

$$\mu(x) = \sum_{i=1}^{q} w_i \exp(v^i x), \quad (18)$$

where given a training set of pairs of real numbers $\{(v^i, u^i) : i = 1,...,q\}$, the constants w_i are obtained by requiring that (18) interpolate μ with the values u^i at v^i.

2.2 Pseudo-Norm, Pseudo-Distance, and Pseudo-Scalar-Product of Fuzzy Sets in Φ

The following important questions arise regarding the formulation of membership functionals as vectors:

(a) What is the meaning of vector addition in F of two membership functionals μ_A and μ_B belonging to fuzzy sets A and B ?

(b) Is it possible to assign a meaning to their point-wise multiplication $\mu_A(x)\mu_B(x)$?

(c) What is the meaning of the multiplication by scalar of a membership functional μ_A in F ?

For this purpose, we resort to a probability-like interpretation to the membership function by defining a fuzzy set, in our specific way, as follows:

Definition 1. Let there be given a universe of discourse $X = \{x\}$, a measure space ***M*** = $(\Psi = \{\xi\}, C = \{\tilde{A}\}, m : C \to [0,1])$, where Ψ denotes a set of elements ξ; C a completely additive class of subsets $\tilde{A}$ of Ψ, each $\tilde{A}$ representing an attribute or event; and m a measure on C satisfying the usual axioms of a probability measure. Let there also be a set $\Upsilon = \{J\}$ of human- or/and machine-based judgment criteria J. Then, given an attribute or event $\tilde{A}$ and a judgment J, on the basis of which set membership decision is made, a fuzzy set A of objects x of X is defined to be one for which there is a membership functional $\mu_A : C \times \Upsilon \times X \to [0,1]$, such that $\mu_A(\tilde{A}, J; x)$ expresses the extent, on a scale from 0 to 1, to which x belongs to A. From now on, we will denote $\mu_A(\tilde{A}, J; x)$ simply by $\mu_A(x)$ when the longer notation is clear from the context. ▌

Note that in a probabilistic setting, x would represent values of measurable functions (random vectors) from Ψ to X

The above definition is broad and yet precise enough to incorporate other formulations of human-based uncertainty, such as belief theory, in the structure of fuzzy sets.

Elsewhere [16], In the context of the above definition, we have interpreted the membership functional $\mu_A(x)$ as a generalization of the posterior probability $P(\tilde{A}/x)$, i.e.,

$$\mu_A(\tilde{A}, J; x) \Leftrightarrow P(\tilde{A}/x) \tag{19}$$

We also proposed that, in such a probabilistic setting, for given fuzzy sets A and B , fuzzy intersections, fuzzy unions, and fuzzy complements be assigned membership values as

$$\mu_{A\cap B}(x) = \mu_A(x)\mu_B(x) \tag{20}$$

$$\mu_{A\cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x) \tag{21}$$

$$\mu_{A^C}(x) = 1 - \mu_A(x) \tag{22}$$

Note that the fuzzy sets resulting from these definitions are the same as the ones resulting from the conventional definitions of fuzzy intersections, fuzzy unions, and fuzzy complements.

Returning to the three questions posed earlier, the answer to question (b) is that point-wise multiplication of membership functionals is useful for the interpretation (20); the answer to question (a) is that vector addition of membership functions is useful for the interpretation (21). Finally, the answer to question (c) is that a scalar α can be used as a weight in the mathematical processing of membership functions. Also, in some applications, it may be interpreted as the probability $P(x)$ in the computation of the probability of the joint event $P(\tilde{A}, x)$. i.e.,

$$P(\tilde{A}, x) = P(x)P(\tilde{A}/x) \Leftrightarrow \alpha\mu_A(x) \quad ▌ \tag{23}$$

Finally, the entities and their interpretation in the following definition may be useful in applications.

Definition 2. Let A and B denote fuzzy sets in Φ . Then the "pseudo-norm", "pseudo - distance", and "pseudo-scalar-product" in Φ may be defined by the rules:

$$\|A\| = \|\mu_A\|_F \tag{24}$$

$$\|A - B\| = \|\mu_A - \mu_B\|_F \quad (25\}$$

$$< A, B > = \langle \mu_A, \mu_B \rangle_F \, , \tag{26}$$

Interpretations:

$\|A\|$ = "Membership Load" of fuzzy set A

$\|A - B\|$ = "Load of Membership Difference" between fuzzy sets A and B

$< A, B >$ = "Correlation in Membership" between fuzzy sets A and B . ▌

.

2.3 Best Approximation of a Membership Functional in *F*

Now we address the important issue of the recovery of a membership functional μ , associated with a fuzzy set $A \in \Phi$ from a set of training pairs (v^i, u^i) , $i = 1, 2, ..., q$, i.e.,

$$\left\{ (v^i, u^i) : v^i \in A \subset E^N, u^i = \mu(v^i) \in R^1, i = 1, \cdots, q \right\}, \tag{27}$$

and under the positivity constraint (4) which we re-write here as

$$0 \leq \mu(x) \leq 1, \quad x \in A \tag{28}$$

The problem of best approximation $\hat{\mu}$ of μ can be posed as the solution of the optimization problem in F of the problem

$$\inf_{\mu \in F} \sup_{\tilde{\mu} \in F} \| \mu - \tilde{\mu} \|_F \quad \text{subject to (18) and (19)} \tag{29}$$

This is a quadratic programming problem in F that can be solved by standard algorithms available in the literature. However, a procedure that usually works well (see [10]) is the one which solves (29) recursively using a RLS algorithm under (27) alone. In such a learning process setting, the pairs in (18) are used sequentially and a new pair is added whenever (19) is violated. This procedure is continued until (27) is satisfied. Such a procedure leads to the following closed form for $\hat{\mu}$, where the pairs (v^i, u^i) are all the pairs used until the end of the procedure,

$$\hat{\mu}(x) = u^T G^{-1} \tilde{K}(x) \tag{30}$$

where

$$\tilde{K}(x) = \begin{pmatrix} \varphi\left(\left\langle v^1, x \right\rangle\right) \\ \varphi\left(\left\langle v^2, x \right\rangle\right) \\ \vdots \\ \varphi\left(\left\langle v^q, x \right\rangle\right) \end{pmatrix} = (s_1(x), s_2(x), ..., s_q(x))^T \tag{31}$$

$$u = \begin{pmatrix} u^1 \\ \vdots \\ u^q \end{pmatrix} \tag{32}$$

and G is a $q \times q$ matrix with elements $G_{ij},\ i, j = 1, \cdots, q$, defined by

$$G_{ij} = \varphi\left(\left\langle v_i, v_j \right\rangle\right) \tag{33}$$

In terms of the above, another convenient way of expressing (33) is

$$\begin{aligned} \hat{\mu}(x) = {} & w_1 \varphi\left(\left\langle v^1, x \right\rangle\right) \\ & + w_1 \varphi\left(\left\langle v^2, x \right\rangle\right) \\ & \cdots \\ & + w_q \varphi\left(\left\langle v^q, x \right\rangle\right) \\ = {} & w_1 s_1(x) + w_2 s_2(x) + ... w_q s_q(x) \end{aligned} \tag{34}$$

where w_i are the components of the vector w obtained by

$$w = G^{-1} u \tag{35}$$

Remark 2: We repeat that It is important to note that the functional μ expressed by (30) or, equivalently (34), need not satisfy the condition (28) required for it to be a membership functional. If this condition is violated at any point, say $x_0 \in X$ then x_0 and its observed membership value at x_0 are inserted as an additional training pair $(v^{q+1,}, u^{q+1})$ in the training set (27) and the procedure of computing the parameter vector w is repeated. This process is continued using the RLS algorithm in until it converges to a membership functional that satisfies (28). For details on a general recursive learning algorithm that implements this process see [10]. ▌

….Remark 3: In most applications, a fuzzy set in the feature space is a union of fuzzy sets and its membership function is multiple-valued conditional membership function, i. e., a n-tuple of membership funcdtionals. In a probabilistic interpretation, each value corresponds respectively to the conditional probability that a given feature vector x belong to a respective fuzzy set. This interpretation provides further motivation for the setting for the following realization of the membership functional as a neural network, the latter being obtained strictly on mathematical basis.

2.4 Neural Network Realization of a Membership Functional in the Space *F*

The structure represented by (30) or, equivalently (34), is shown in the block diagram of Fig, 1. It is clear from this figure that this structure corresponds to a two-hidden layer artificial neural network, with the synapses and activation functions labeled according to the symbols appearing in (30) and (34). Such an artificial neural network, as an optimal realization of an input output map in F based on a training set as given by (27), was first introduced by de Figueiredo in 1990 [8-9] and called by him an Optimal Interpolative Neural Network (OINN). It now turns out, as developed above, that such a network is also an optimal realization of the membership functional of a fuzzy set in Φ .

3 APPLICATIONS

Two important cases of applications in which the universal set X is infinite-dimensional are those in which the objects in X are waveforms $x = \{x(t) : a \le t \le b\}$ or images $x = \{x(u,v) : a \le u \le b, c \le d\}$. By simply replacing the formulas for the scalar product in X given by (1) or (2) for the Euclidian case to the present case, all the remaining developments follow in the same way as before, with the correct interpretation of the inner products and with the understanding that summations now become integrations.

For example, the scalar product, analogous to (1) and (2), for the case of two waveforms x and y would be

$$\int_a^b x(t)y(t)dt \tag{36}$$

and

$$\int_a^b x(t)R^{-1}(t,s)y(s)dtds \tag{37}$$

Then the synaptic weight summations in Fig. 1 would be converted into integrals representing matched filters matched to the prototype waveforms in the respective fuzzy set. The author called such networks dynamic functional artificial neural networks (D-FANNs) and provided a functional analytic method for their analysis in [12] and application in [14]..

Due to limitations in space, it is not possible to dwell on other applications, except to briefly mention the following example of an analysis of a database of Brain Spectrogram image feature vectors x belonging to two mutually exclusive fuzzy sets A and B of images of patients with possible Alzheimer's and vascular dementia (see [13]).

Fig. 2 shows the images of 12 slices extracted from a patient brain by single photon emission with computed tomography (SPECT) hexamethylphenylethyleneneaminieoxime technetium-99, abbreviated as HMPAO-99Tc. The components of the feature vector x are the average intensities of the image slices in the regions of interest in the templates displayed in the lower part of Fig. 2. This feature vector was applied as input to the OINN shown in Fig. 3, which is self explanatory. The results of the study, including a comparison of the performance of our machine-based algorithms and a human expert in achieving the required objective, are displayed in Table 1. It turned out that, in this experiment, the machine performed better than the human expert. The performance of the expert, reported in the Table, represents an improvement of his diagnostic ability achieved by using this machine learning algorithm as a tool.

4 CONCLUSION

Assuming that the universe of discourse X is a vector space endowed with a scalar product, the membership functions μ_A of fuzzy sets A are nonlinear membership functionals from X to the real line satisfying the usual positivity constraint. Under appropriate conditions, we have allowed such μ_A to belong to a Reproducing Kernel Hilbert Space $F(X)$, and process the uncertainty present in the fuzzy sets A by nonlinear functional analytic processing of their μ_A in the space $F(X)$. We have shown how such a processing of μ_A affects the processing of the fuzzy sets A themselves. We have described some of the benefits of this approach toward providing a common rigorous platform for human and machine intelligence, and toward generalizing some of the concepts of probability theory. In this way, our approach is different from the alternative approaches in the literature (see, e.g., [5]), whereby the processing of uncertainty takes place in the feature space rather than in the membership functional space. A more detailed discussion of our approach will appear in future publications such as [16].

REFERENCES

[1] Zadeh, L.A., "Fuzzy sets", *Information and Control,* **8**, pp.338-353, 1965

[2] Klir, G. J. and Yuan, B., *Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems – Selected Papers by Lotfi A. Zadeh,* World Scientific Publishing Co., 1996.

[3] Klir, G.J. and T .A. Folger, *Fuzzy Sets, Uncertainty, and Information*, Prentice Hall, 1988.

[4] Ruan, Da (Ed.) "*Fuzzy Set Theory and Advanced Mathematical Applications*", Kluwer Academic, 1995

[5] Bezdek, J. C. and Pal, S. K. (Eds.), *"Fuzzy Models for Pattern Recognition",* IEEE Press, 1992

[6] R.J.P. de Figueiredo and T.A.W. Dwyer "A best approximation framework and implementation for simulation of large-scale nonlinear systems", *IEEE Trans. on Circuits and Systems*, vol. CAS-27, no. 11, pp. 1005-1014, November 1980.

[7] R.J.P. de Figueiredo "A generalized Fock space framework for nonlinear system and signal analysis," *IEEE Trans. on Circuits and Systems*, vol. CAS-30, no. 9, pp. 637-647, September 1983 (Special invited issue on "Nonlinear Circuits and Systems")

[8] "R.J.P. de Figueiredo A New Nonlinear Functional Analytic Framework for Modeling Artificial Neural Networks" (invited paper), *Proceedings of the 1990 IEEE International Symposium on Circuits and Systems*, New Orleans, LA, May 1-3, 1990, pp. 723-726.

[9]. R.J.P. de Figueiredo "An Optimal Matching Score Net for Pattern Classification", *Proceedings of the 1990 International Joint Conference on Neural Networks (IJCNN-90)*, San Diego, CA, June 17-21, 1990, Vol. 2, pp. 909-916.

[10] S.K. Sin and R.J.P. de Figueiredo, "An evolution-oriented learning algorithm for the optimal Interpolative neural net", *IEEE Trans. Neural Networks*, Vol. 3, No. 2, March 1992, pp. 315-323

[11] R. J.P. de Figueiredo and Eltoft, T., "Pattern classification of non-sparse data using optimal interpolative nets", *Neurocomputing*, vol.10, no.4, pp. 385-403, April 1996

[12] R. J.P. de Figueiredo, "Optimal interpolating and smoothing functional artificial neural networks (FANNs) based on a generalized Fock space framework," *Circuits, Systems, and Signal Processing*, vol.17, (no.2), pp. 271-87, Birkhauser Boston, 1998.

[13] R. J.P. de Figueiredo, W.R. Shankle, A. Maccato, M.B. Dick, P.Y. Mundkur, I. Mena, C.W. Cotman, "Neural-network-based classification of cognitively normal, demented, Alzheimer's disease and vascular dementia from brain SPECT image data", *Proceedings of the National Academy of Sciences USA,* vol 92, pp. 5530-5534, June 1995.

[14] T. Eltoft and R.J.P de Figueiredo, "Nonlinear adaptive time series prediction with a Dynamical- Functional Artificial Neural Network", *IEEE Transactions on Circuist and Systems*, Part II, Vol.47 no.10, Oct.2000, pp.1131-1134

[15] R.J.P. de Figueiredo "Beyond Volterra and Wiener: Optimal Modeling of Nonlinear Dynamical Systems in a Neural Space for Applications in Computational Intelligence", in *Computational Intelligence: The Experts Speak*, edited by Charles Robinson and David Fogel, volume commemorative of the 2002 World Congress on Computational Intelligence, published by IEEE and John Wiley & Sons, 2003.*emy of Sciences USA*, vol 92, pp. 5530-5534, June 1995.

[16] R. J. P. de Figueiredo, "A Nonlinear Functional Analytic Framework for Modeling and Processing Fuzzy Sets", in the book "Forging the New Frontiers: Fuzzy Pioneers I", *Studies in Fuzziness and Soft Computing*, edited by M. Nikravesh, J. Kacprzyk, and L. A. Zadeh, Springer Verlag, 2006.

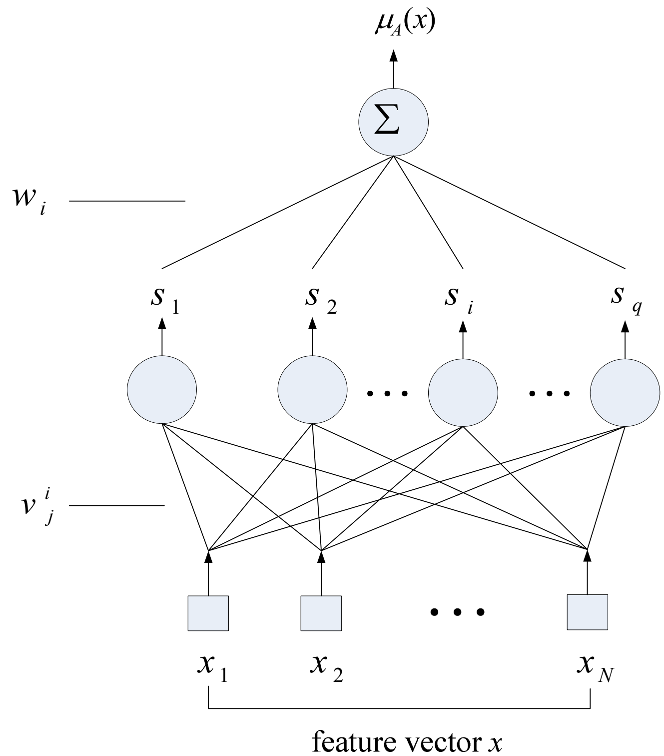

Fig. 1 Optimal realization of the membership functional μ_A of a fuzzy set $A \in \Phi$.

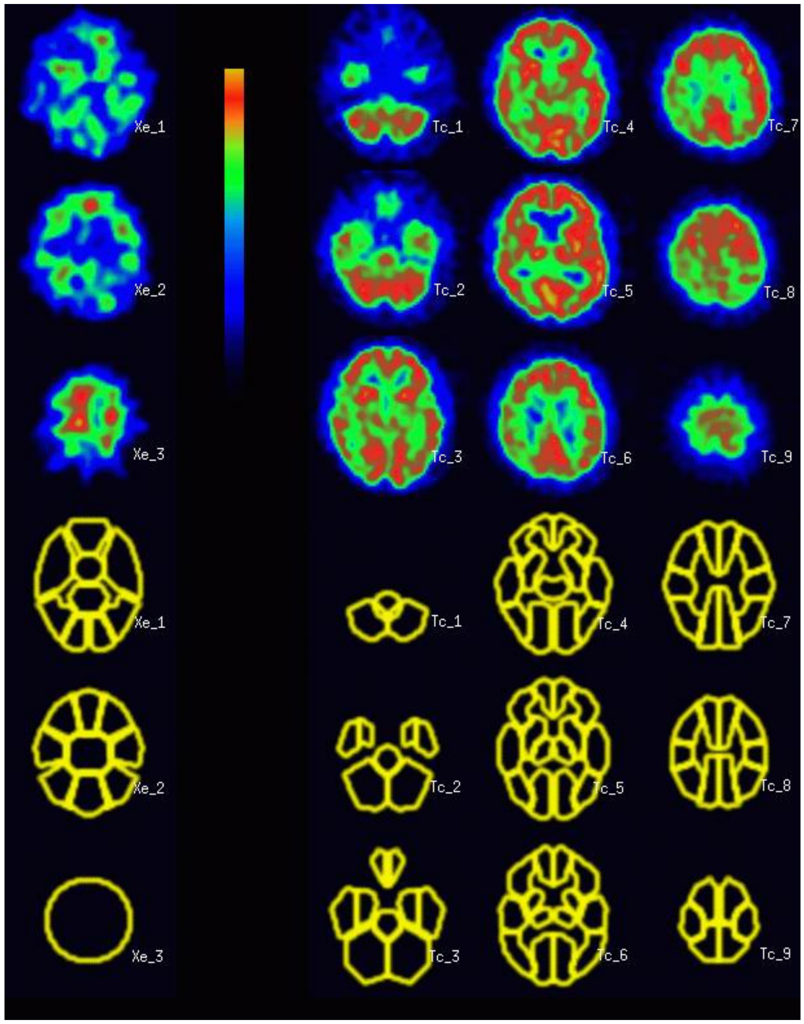

Fig. 2. Images of slices of Brain Spectrogram images of a prototype patient. The template frames are used to extract the components of the feature vector x characterizing the patient.

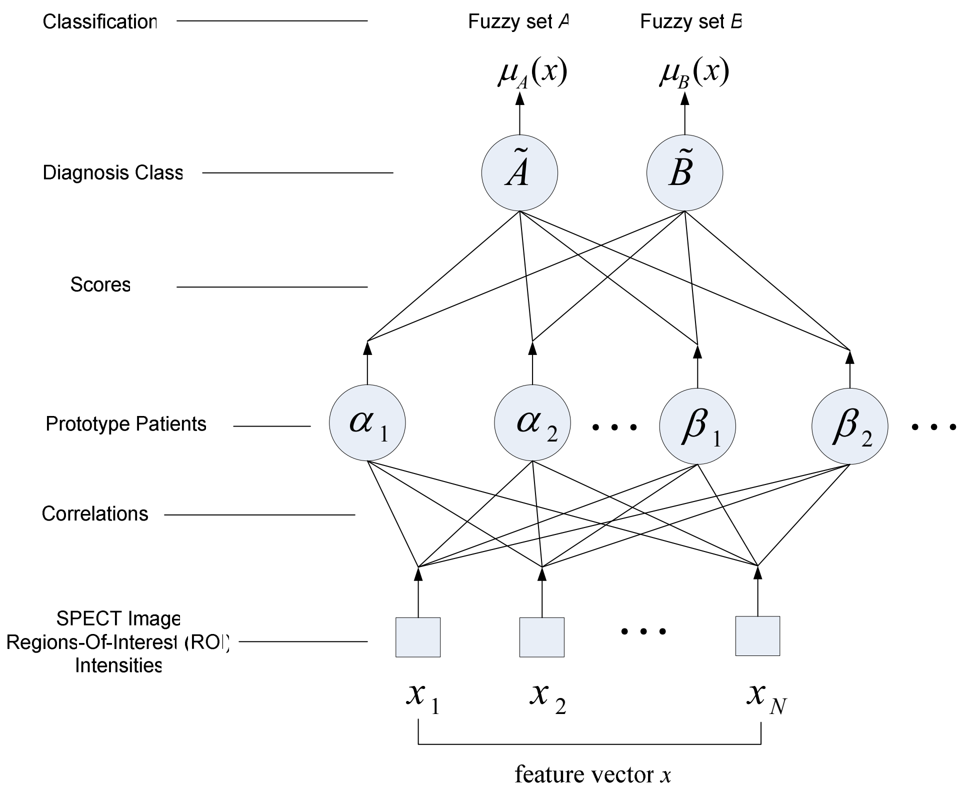

Fig. 3. An OINN (Optimal Interpolative Neural Network) used to realize the membership functionals of two fuzzy sets of patients with two different types of dementia

Table: SPECT Data Analysis

Source of Classification	Rate of Correct Classification*
Fuzzy-Set-Based Classification	81%
Expert Radiological Diagnostician	77%

*41 subjects: 15 Probable AD, 12 Probable VD, 10 Possible VD, 4 Normal.

Table 1 Results of the study of dementia based on the fuzzy set vector functional membership analysis presented in the paper

APPENDIX

BRIEF OVERVIEW OF THE SPACE F

In this Appendix we briefly review some definitions and results on the space $F(X)$ invoked in the body of the paper.

For the sake of brevity, we focus on the case in which X is the N-dimensional Euclidian space E^N. The developments also apply, of course, when X is any separable Hilbert space, such as l_2, the apace of square-summable strings of real numbers of infinite length (e.g., discrete time signals), or L^2, the space of square-integrable waveforms or images.

Thus let F consist of bounded analytic functionals μ on E^N [We denote members of F by μ because, in this paper, they are candidates for membership functionals of fuzzy sets in Φ].Then such functionals μ can be represented by abstract power (Vollerra functional) series on E^N satisfying the following conditions.

(a) μ is a real analytic functional on a bounded set $\Omega \subset E^N$ defined by

$$\Omega = \left\{ x \in E^N : \|x\| \leq \gamma \right\} \tag{A1}$$

where γ is a positive constant. This implies that there is an N-variable power series known as a Volterra series in these N variables, absolutely convergent at every $x \in \Omega$, expressible by:

$$\mu(x) = \sum_{n=0}^{\infty} \frac{1}{n!} \mu_n(x) \tag{A 2}$$

where μ_n are homogeneous Hilbert-Schmidt(H-S) polynomials of degree n in the components of x, given by

$$\mu_n(x) = \sum_{\substack{k_1=0 \\ |k|=k_1+k_2+\cdots k_N=n}}^{n} \cdots \sum_{k_N=0}^{n} c_{k_1\cdots k_N} \cdot \frac{|k|!}{k_1!\cdots k_N!} x_N^{k_1} \cdots x_N^{k_N} \tag{A3}$$

$$= \sum_{|k|=n} c_k \frac{|k|!}{k!} x^k \tag{A4}$$

where, in the last equality, we have used the notation

$$k = (k_1, \cdots, k_N) \tag{A5}$$

$$|k| = \sum_{i=1}^{N} k_i \tag{A6}$$

$$k! = k_1!, \cdots, k_N! \tag{A7}$$

$$c_k = c_{k_1 \cdots k_N} \tag{A8}$$

$$x^k = x_1^{k_1} \cdots x_1^{k_N} \tag{A9}$$

(b) Let there be given a sequence of positive numbers

$$\lambda = \left\{ \lambda_0, \lambda_1 \cdots \right\}, \tag{A10}$$

where the weights λ_n express prior uncertainty in the terms μ_n , *n* = 0,1, 2,…, and satisfy

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n} \frac{\gamma^{2n}}{n!} < \infty \tag{A11}$$

Actually, some elements of λ, namely λ_k, $k \in S$, where S is a subset of non-negative integers, may be allowed to be zero, if we assume that f belongs to a subspace of F consisting of powers series in F with the terms of degree $k \in S$ deleted.

(c) Finally, the coefficients c_k in the terms μ_n of the abstract power series expansion of the membership function satisfy the restriction

$$\sum_{n=0}^{\infty} \frac{1}{n!\lambda_n} \sum_{|k|=n} \frac{|k|!}{k!} |c_k|^2 < \infty \tag{A12}$$

The above allows us to state the following theorem. For a proof, see [6].

Theorem 1 (de Figueiredo / Dwyer) [6].

Under (A1), (A2),(A11), and (A12), the completion of the set of nonlinear functionals μ in (A2) constitutes a Reproducing Kernel Hilbert (RKHS) $F(E^N) = F$ on Ω, with:

(i) The scalar product between any μ_A and $\mu_B \in F$, corresponding to fuzzy sets A and B is defined by

$$\langle \mu_A, \mu_B \rangle_F = \sum_{n=0}^{\infty} (\frac{1}{\lambda_n} \frac{1}{n!} \sum_{|k|=n} \frac{|k|!}{k!} c_k d_k) \tag{A13}$$

where c_k and d_k are defined for μ_A and μ_B in the same way as c_k is defined for defined for μ in.(A4) and (A8).

(ii) The reproducing kernel, $K(x,z)$, in F is

$$K(x,z) \overset{def}{=} \varphi\left(\langle x,z \rangle\right) = \sum_{n=0}^{\infty} \frac{\lambda_n}{n!} \left(\langle x,z \rangle\right)^n \tag{A14}$$

where $\langle x,z \rangle$ denotes the inner product in E^N .In the special case that $\lambda_n = \lambda_0{}^n$ φ is an exponential function, and thus $K(x,z)$ becomes

$$K(x,z) = \exp\left(\lambda_0 \langle x,z \rangle\right) \quad \blacksquare \tag{A15}$$

It is easily verified that the $K(x,z)$ defined as the reproducing property

$$< K(x,.), \mu(.) >_F = \mu(x) \tag{A16}$$

Methods for best approximation of a functional in F, on which the development s in section 2.3 are based, are described in [6-10] and [15].

SHORT BIO: Dr. Rui J.P. de Figueiredo, B.S. and M.S. (M.I.T.), and Ph.D.(Harvard) is Research Professor (Above Scale) of Electrical Engineering and Computer Science and of Mathematics, and Director of the Laboratory for Intelligent Signal Processing and

Communications, at the California Institute for Telecommunications and Information Technology, at the University of California, Irvine (UCI). Prior to joining UCI in 1990, Dr. de Figueiredo served as Professor of Electrical Engineering and Mathematical Sciences at Rice University, Houston, Texas (1965-90). Professor de Figueiredo has won numerous honors. These include: ***election to*** the UN-sponsored ***International Informatization Academy* (2003)**, the **1999 *IEEE Circuits and Systems (CAS) Society Golden Jubilee Medal***, the **2000 *IEEE Tri-Millennium Medal***, the **2003 *Gh. Asachi Medal*** from the Technical University of Iasi (TUI), Romania, from which he also received the title of ***Honorary Professor* (2003)**, the ***IEEE Fellow Award* (1976)**, the ***1994 IEEE CAS Technical Achievement Award***, the **2000 *IEEE Neural Networks Transactions Best Paper Award***, the ***2003 IEEE Circuits and Systems Transactions Guillemin-Cauer Best Paper Award***, the ***2002 IEEE CAS Society M. E. Van Valkenburg Society Award***, the ***1988 NCR Educator-of-the-Year Award***, his election to ***President of IEEE CAS Society* in 1998**, and his selection by IEEE to be one of its fifty leaders to present the IEEE vision of the new century in the book ***ENGINEERING TOMORROW:Today's Technology Experts Envision the Next Century, Janie Fouke, Editor, IEEE Press, 2000.***.