DESIGN OF INTERFACES FOR FUZZY INFORMATION PROCESSING

Witold Pedrycz

Dept. of Electrical & Computer Engineering
University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2
pedrycz@eeserv.ee.umanitoba.ca

Abstract The paper develops the concepts of the input and output interfaces for fuzzy information processing. These interfaces play an essential role in coupling fuzzy processing (that is carried out at a global and abstract level of linguistic labels) with the pointwise numerical processing being characteristic for the external environment. The design criteria for the input interface pertain to its robustness and abilities to handle uncertainty residing within input data. The sampling mechanism in the input interface is also studied along with an analysis of reconstruction properties conveyed by the family of the sampling objects (signals) used in the sampling procedure. The design of the output interface is focussed primarily on achieving its zero-error mapping capabilities. Some well known transformation mechanisms (including a Centre of Gravity method) are reviewed and assessed in the light of the proposed mapping criterion.

Keywords fuzzy information processing, fuzzy sampling, reconstruction, uncertainty representation, possibility, necessity, robustness

1. Introduction - a paradigm of fuzzy information processing

The general scheme of fuzzy information processing as carried out at a conceptual and highly abstract level of linguistic quantities embodies the three essential functional blocks:

<u>input interface</u>: The role of this block is to transform (internalize) the input information into an appropriate format to be handled properly by the processing block. The input interface realizes a process of knowledge representation by transforming all available data into coherent pieces of information. The standard way of this internalization is to define a collection of fuzzy sets (linguistic labels) and determine levels of matching occurring between each of them and the input datum. It is worth underlying that the data might be heterogeneous; along with the numerical pieces of knowledge they could include fuzzy sets as well as interval-valued quantities.

<u>processing block</u>: At this level all the computations are carried out for the linguistic labels defined for the input interface. The main objective at this stage is to reveal logical relationships between fuzzy sets of the fuzzy partitions defined for the individual variables expressed by the interfaces.

<u>output interface</u>: The aim of this processing is to re-transform the internal quantities produced at the level of the processing block into some other formats including the numerical (pointwise) one as required by the environment of the processing block.

While analyzing this general architecture of information processing, it becomes evident that many of the algorithms existing in fuzzy sets technology falls under this category. Fir instance, in fuzzy controllers one can easily identify fuzzification and defuzzification procedures forming the interfaces while the processing level is limited to a simple recall mechanism implemented via the standard max-min composition.

It is also remarkable that most of the optimization effort has been aimed at mastering the processing procedures situated at the level of the second block while relatively much less attention has been paid to studying and optimizing the interfaces themselves.

The main objective of this paper is to study the input and output interfaces, analyze their features, provide a way of their characterization along with a discussion of the optimization tasks emerging there. In Section 2, we will depart from the analysis of the input interface that covers several information processing aspects such as sampling properties and representing uncertainty of the input data. Subsequently, the design of the output interface is discussed in the context of its error-free linguistic-numerical transformation (Section 3).

2. Input interface: fuzzy set-based information representation

As mentioned in Section 1, the aim of the input interface is to collect input data and express them consistently in terms of a predefined collection of fuzzy sets (a so-called linguistic labels). These quantities are viewed as basic information granules [14] being recognized as conceptual entitites playing an essential role in problem description. They are usually referred to as a cognitive perspective (frame of cognition), cf. [4] [5]. The frame of cognition is constructed as a finite family of unimodal and overlapped fuzzy sets, namely $\mathcal{A}=\{A_1, A_2, ..., A_c\}$. Quite often one additionally assumes that $\mathcal{A}$ is a fuzzy partition which means that its elements satisfy the property of ortogonality, viz. the condition $\sum_{k=1}^{c} A_k(x_j)=1$ that holds for each x_j of the universe of discourse. This condition is automatically fulfilled when the elements of $\mathcal{A}$ are generated via fuzzy clustering, cf. [1].

In general, the available pieces of input information could be highly heterogeneous and may include both numerical as well as non-numerical (interval-valued or fuzzy sets) entities. The common transformation exploited in designing the input interface is the one based on possibility and necessity measures, cf [15] [3]. Nevertheless the use of some other mappings can be anticipated as well.

Let A denotes one of the elements of $\mathcal{A}$ (the index of this label, as not relevant in further investigations, is ignored). The input datum X is then "translated" into an internal logical format as follows, cf [15]

$$\text{Poss}(X|A) = \sup_{x\in\mathbf{X}} [\min(X(x), A(x))] \qquad (1)$$

$$\text{Nec}(X|A) = \inf_{x\in\mathbf{X}} [\max(1-X(x), A(x))] \qquad (2)$$

The basic properties of these measures have been thoroughly studied in the literature; for more details the reader can refer e.g., to [3]. Let us remind that the possibility measure evaluates a degree of overlap of X and A while the necessity measure is involved in expressing a degree of inclusion of X in A. In particular, Poss (X|A) ≥ Nec (X|A). These two generic definitions can be immediately generalized by replacing the lattice (max and min) operators used in the above definitions by the triangular norms. This approach is useful in capturing a global aspect of evaluation of the above matching properties. Observe that due to the sup and inf operations the above expressions are noininteractive and the final numerical results produced there depend solely upon a single element of the universe of discourse- thus the aggregation operations are rather limited with this respect. The sound generalization would be of the type

$$\text{Poss}(X|A) = \underset{x\in\mathbf{X}}{S} (X(x)\ t\ A(x))$$

$$\text{Nec}(X|A) = \underset{x\in\mathbf{X}}{T} ((1-X(x))\ s\ A(x))$$

that involves the s-t and t-s composition, respectively.

The primordial question one can pose about the development of the suitable input interface utilizing these measures can be conveniently handled by splitting it into several fundamental issues such as (i) analyzing sampling properties residing within the given fuzzy partition (and the relevant translation mechanisms); (ii) expressing uncertainty and incompleteness of the input data being processed by the interface, and (iii) determining robustness of the frame of cognition.

2.1. Sampling problem

The input (sampling) data $X_1, X_2, ..., X_N$ expressed via the possibility and necessity measures being taken with respect to a single element of the frame of cognition $\mathcal{A}$ can be conveniently looked at as a stream of sampling signals applied to A. The question formulated within this context pertains to how to reconstruct A based on X_k's and the associated possibility and necessity values λ_k= Poss(X_k|A) and μ_k= Nec(X_k|A).Furthermore, the sampling objects themselves could be distributed in many different ways with respect to A as well as posses different level of granularity. These two factors affect the performance of reconstruction to a great extent.

The reconstruction problem of A can be placed in the framework of the general task of membership function estimation in which A has to be determined based on the results of sampling collected through a series of experiments. In this way X_k can be viewed as fuzzy "probes" that are utilized to probe the environment manifested by A. As it will become clear later on, the selection of the sampling data and their granularity (specificity) depends to a significant extent on the granularity of A. The specificity of fuzzy sets can be computed as e.g., proposed in [11].

In a formal setting one can reformulate the sampling (reconstruction) problem accordingly

- for X_k and the given outcomes of sampling (λ_k, μ_k), $k = 1,2,\ldots,N$ where $\lambda_k \geq \mu_k$, determine A for which the following series of conditions holds

$$\lambda_k = (\mathrm{Poss}\,(X_k|A) \quad \mu_k = \mathrm{Nec}\,(X_k|A) \qquad (3)$$

$k = 1,2,\ldots,N$

Firstly, it should be emphasized that the solution to the problem is not unique. This means that usually A cannot be reconstructed in a unique manner based on the outcomes of the sampling experiment. To be more systematic, one can look at the system of conditions (3) as a family of fuzzy relational equations to be solved with respect to A. Assuming the consistency of these conditions (that is equivalent to the assumption of solvability of the system of equations (3)), the extremal solutions to it can be derived immediately, cf. [2] [6]

- the upper bound of the reconstruction, say $\hat{A}$ results from the intersection of the successive α-compositions of X_k and λ_k

$$\hat{A} = \bigcap_{k=1}^{N} (X_k \,\alpha\, \lambda_k)$$

namely,

$$\hat{A}(x) = \min_{k=1,2,\ldots,N} [X_k(x) \,\alpha\, \lambda]$$

- the lower bound of reconstruction is taken as a union of the minimal solutions to the individual condition

$$\tilde{A} = \bigcup_{k=1}^{N} (\overline{X}_k \,\varepsilon\, \mu_k)$$

$$\tilde{A}(x) = \max_{k=1,2,\ldots,N} [(1 - X_k(x)) \,\varepsilon\, \mu_k]$$

where the α– and ε -operations are defined pointwise as

$$a \,\alpha\, \lambda = \begin{cases} 1, & a \leq \lambda \\ \lambda, & a > \lambda \end{cases}$$

and

$$a \,\varepsilon\, \mu = \begin{cases} a, & a > \mu \\ \lambda, & a \leq \mu \end{cases}$$

$a, \lambda, \mu \in [0,1]$

The reconstruction results could be directly extended to the sup-t and inf-s composition as used in (3). The boundary solutions are determined now as,

$$\hat{A} = \bigcap_{k=1}^{N} (X_k \,\varphi\, \lambda_k)$$

and

$$\tilde{A} = \bigcup_{k=1}^{N} (\overline{X}_k \,\beta\, \mu_k)$$

where $a \,\varphi\, \lambda = \sup \{ q \in [0,1] | \, atq \leq \lambda \}$ and $a \,\beta\, \mu = \inf \{ q \in [0,1] | \, asq \geq \mu \}$.

The analytical solutions are not available neither for the general s-t composition nor its dual t-s form.

The question of reconstruction of A and its quality depends on the character and distribution of X_k's. The character of the reconstruction can be conveniently illuminated by the following example. We will study several families of sets distributed over **X** ; in case of **X** being a real line these sets are just intervals. In this somewhat specific environment of the sampling signals (objects) it is clearly visible that each X_k considered individually is capable of "reconstructing" A within its region of operation (more precisely, its support). Note that this reconstruction (for A assumed to be a continuous membership function) produces

$$\hat{A}_k(x) = \begin{cases} \max\limits_{x \in X_k} A(x) \\ 1, \quad \text{otherwise} \end{cases}$$

$$\tilde{A}_k(x) = \begin{cases} \min\limits_{x \in X_k} A(x) \\ 0, \quad \text{otherwise} \end{cases}$$

so in fact, the derived bounds depend directly upon the cardinality of X_k as well as the variability of A that is reported over its support . Outside the support the produced estimates (bounds) are fairly meaningless as they lead to the unit interval formed by these bounds. The quality of reconstruction relies heavily on a relative information granularity of X_k's considered with respect to A, see Fig.1. The regions of X_k that are too "broad" with respect to A, do not contribute to its meaningful reconstruction. When the sampling objects (signals) do not "cover" the entire universe, this sampling scenario generates some reconstruction gaps, as shown in Fig.1.

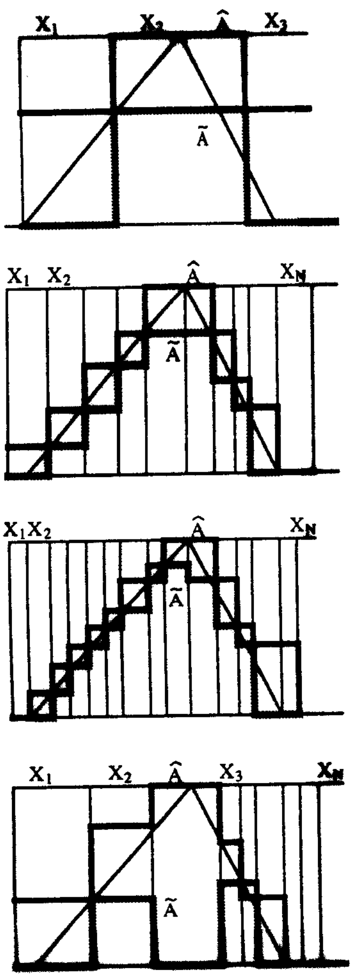

Fig.1. Reconstruction results for different sampling scenarios

When the intervals X_k become narrow (this number increases) while they still cover the entire universe of discourse, the reconstruction produces the bounds of A that are more precise. The bounds, as they rely on variability observed within the membership function, tend to become more specific (narrow). A careful inspection of the quality of reconstruction related with the nature of the sampling data sheds some light on general recommendation on how to carry out sampling. The granularity of the sampling signals should be then selected with respect to the granularity of the concept. While the increasing granularity help improve the quality of the reconstruction of the concept, this is valid only for the regions of activation generated by X_k. When the number of experiments is kept constant, the increasing granularity give rise to some reconstruction gaps resulting again in a quite poor overall performance of the reconstruction. On the other hand, too low granularity of X_k's reduces the number of the experiments but simultaneously results in a low quality of reconstruction. These comments clearly highlight a need for a rational trade-off between granularity of the sampling data; these should be selected with respect to the character of the concept under analysis. The importance of this compromise has become obvious in fuzzy modelling incorporating a hierarchy of fuzzy models, cf. [8] [12].

In many practical situations the assumptions of the sampling consistency (or equivalently, the solvability of fuzzy relational equations) are too strong and cannot be satisfied. In this case the membership function of A has to be determined numerically. The approximation of A emerges out of the optimization task in which the bounds A_+ and A_- are derived by solving the following tasks that minimize the MSE performance criterion formulated separately for the possibility and necessity conditions,

$$\min_{A_+} \sum_{k=1}^{N} [\lambda_k - \mathrm{Poss}(X_k \mid A)]^2$$

$$\min_{A_-} \sum_{k=1}^{N} [\mu_k - \mathrm{Nec}(X_k \mid A)]^2$$

The problem should also incorporate an additional requirement that calls for a preservation of inclusion between A_+ and A_-, namely, $A_+ \supseteq A_-$.

2.2. Representing and handling uncertainty through the input interface

The possibility and necessity measures processed together can be useful in handling uncertainty, in particular the aspects of ignorance and conflict manifested in the available input information X. Again, these two notions are context-dependent and as such should be analyzed with respect to the given fuzzy set A. The context-dependency implies also that the numerical qualifications of these phenomena depend upon the environment (the frame of cognition) within which they are embedded. Let us define two indices

$$\lambda = Poss(X|A)$$

$$\xi = 1 - Nec(X|A)$$

as expressing relationships occurring between X and A. For a pointwise character of X, λ and ξ are linked together via a straightforward relationship

$$\lambda + \xi = 1$$

(for any numerical information X both the measures coincide).

In general, when the datum is of a general (viz. nonpointwise) character then we end up having one of these inequalities

$$\lambda + \xi < 1, \quad \lambda + \xi > 1$$

These cases are worth studying since they tackle the situations including information ignorance and conflict:

Let $\lambda + \xi > 1$ that can be expressed as $\lambda + \xi = 1 + \gamma$ where $\gamma \in [0,1]$. The higher the value of γ, the higher the level of conflict emerging out of X placed in the context of A; γ denotes this level .

The case in which $\lambda + \xi < 1$, with $\lambda + \xi = 1 + \gamma$, $\gamma \in [0,1]$, articulates a situation of ignorance arising from expressing X via A. More precisely, γ denotes this level of ignorance.

3. Design of the output interface

The optimization of the output interface is carried out with respect to the transformation of the linguistic representatives (treated as the levels of activation of the linguistic labels of the fuzzy partition) into a numerical format. From a formal point of view one can look at the function of the output interface as the one carrying out a mapping from a unit hypercube into a line of reals. Denote the mapping by $\mathcal{L}$, $\mathcal{L}: [0,1]^m \rightarrow \mathbf{R}$. The fundamental requirement we will be concerned with can be formulated as a zero error of reproducibility of the original numerical quantity. To elaborate on this formulation , let us consider a single numerical quantity $y \in \mathbf{R}$. The transformation of y through the output interface consisting of "m" fuzzy sets yields a vector $\mathbf{y} = [y_1, y_2, \ldots, y_m] \in [0,1]^m$. Subsequently, the numerical representative $\hat{y}$ is computed as $\hat{y} = \mathcal{L}(\mathbf{y})$. The error-free output interface should yield the relationship

$$\mathcal{L}(\mathbf{y}) = y$$

to be satisfied for all $y \in \mathbf{R}$.

In particular, we will be concerned with the averaging transformation $\mathcal{L}$ described as

$$\hat{y} = \frac{y_1 r_1 + y_2 r_2 + \ldots + u_m r_m}{y_1 + y_2 + \ldots + y_m} \qquad (4)$$

where $r_1, r_2, \ldots, r_m$ are some prototypes (usually taken as modal values of the membership functions forming the interface). This transformation is well-known in fuzzy controllers as a Centre of Gravity Method and as such has been investigated vigorously in many applications, cf. [13] [16].

Usually the requirement of the strict equality formulated for the output interface does not always hold. The differences encountered there give rise to the transformation error. The values of the error as well as the distribution of it depend heavily on the fuzzy partition of the universe utilized within this transformation as well as the transformation itself.

It also becomes obvious that the values of the error depend on the specificity of the fuzzy sets (and subsequently the level of overlap between them). For the Gaussian-like membership functions, the transformation error depends upon the distribution of these fuzzy sets: both too low or too high overlap may have a very destructive impact on the quality of this transformation, see Fig.2.

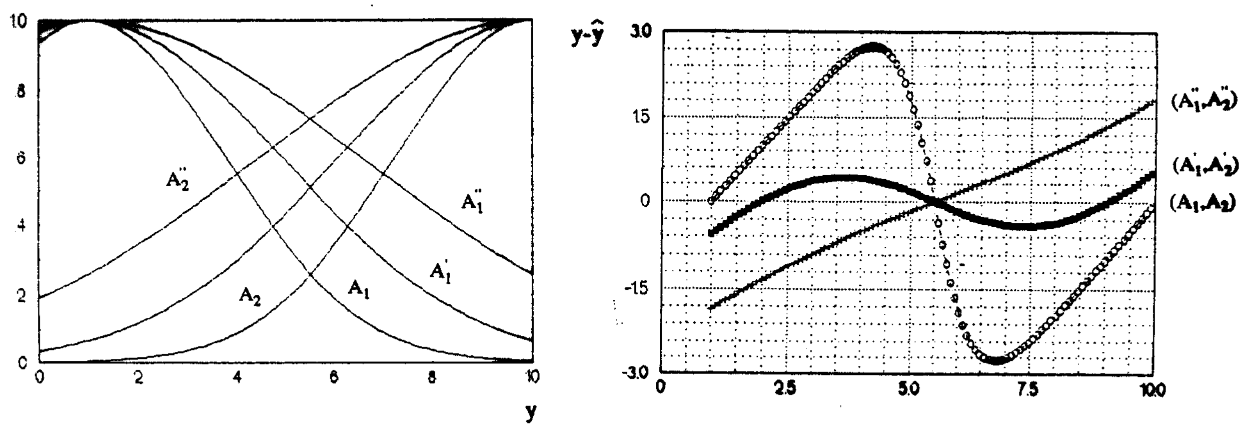

Fig.2. Error distribution $y-\hat{y}$ for different membership functions

The output interface formed by a collection of piecewise linear membership functions leads to several patterns of the representation error of the output interface, see Fig.3

The particular distribution of these membership functions produces a zero value of this error. This could simplify the design process of the interface as it calls only for the shift of the position of the modal values of the fuzzy sets. Note also that in this optimal situation the successive membership functions must intersect at the membership level of 0.5. Thus the arbitrary selection of the class of linear membership functions with this particular level of overlap, as found quite frequently in many applications, gains now an additional strong and well-sound justification established on the ground of the nonzero representation capabilities of the output interface. Unfortunately, this form of the fuzzy partition could be quite rigid as it does not provide any freedom when the modifications of these fuzzy sets should be made with regard to any additional criteria. Some other interfaces found in fuzzy controllers such as e.g., Mean of Maxima (MOM) are less suitable when evaluated in the light of the reproducibility criterion.

In the general setting, the output interface can be optimized parametrically through an adjustment of the membership functions of the fuzzy partition. This way has been pursued in [9] [10]. The optimization task studied there is formulated by considering an available finite collection of the numerical data $y_1, y_2, \ldots, y_N$ and specifying the following performance index

$$Q=\sum_{k=1}^{N} [y_k - \mathcal{L}(y_k)]^2$$

More generally, this criterion can be augmented by a probability density function (p.d.f.) p(y) of the numerical data so that it reads now as

$$\int [y - \mathcal{L}(y)]^2 p(y)dy$$

For the uniform p.d.f. this reduces into the form

$$\int [y - \mathcal{L}(y)]^2 dy$$

Another optimization option that is worth exercising lies in improving the form of the transformation $\mathcal{L}$ itself. While (6) constitutes a relatively simple averaging transformation not equipped with any parameters that could be modified in the

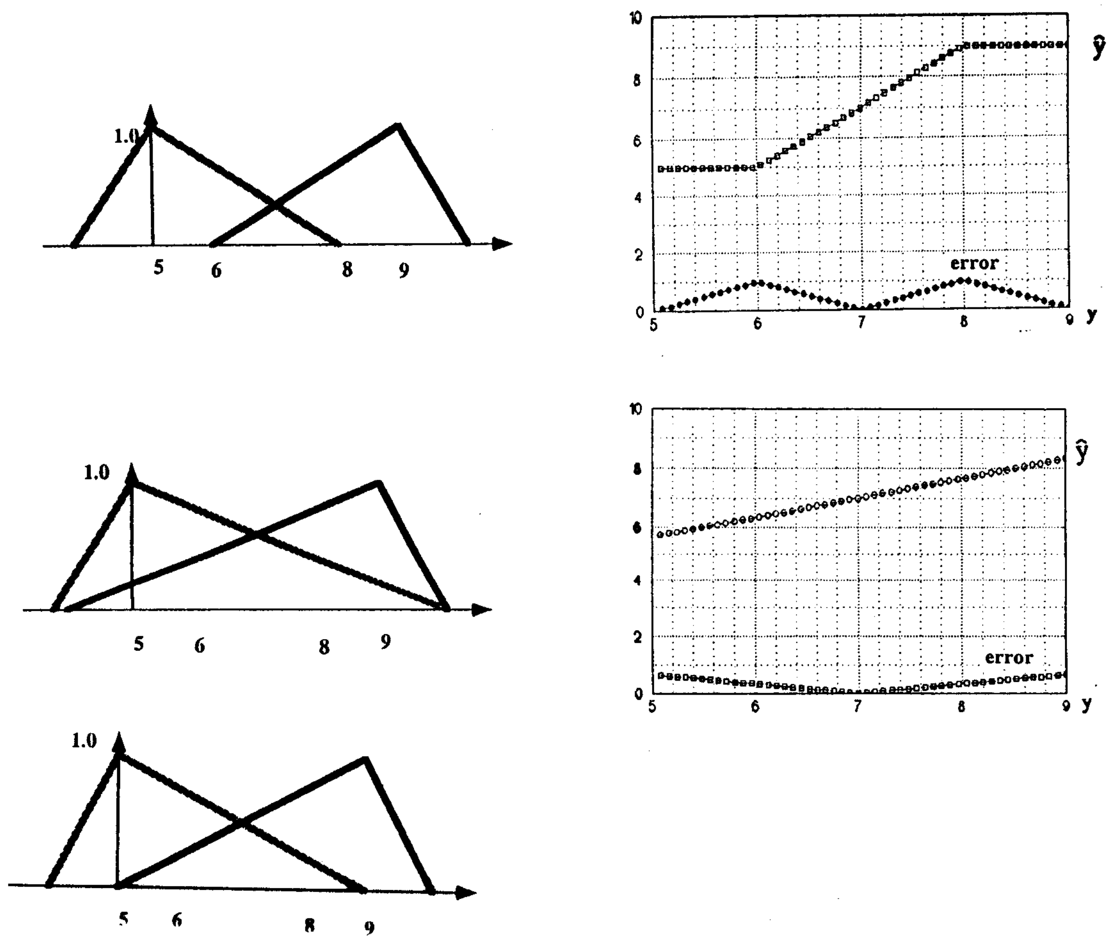

Fig.3. Mapping property realized by different piecewise membership functions

course of optimization, some other types of $\mathcal{L}$ might be more successful in bringing the transformation error to zero. One can think of a specialized neural network with y treated as its input and a single output producing y. The learning of the network is guided by the same performance index as used in the evaluation of the interface. Returning to the interface formed with the Gaussian membership functions as illustrated in Fig.2, a simple neural network with a single hidden layer consisting of the two neurons and equipped with a standard sigmoid nonlinearity was capable of producing the zero value of the discussed performance index.

4. Conclusions

We have introduced the notion of the input and output interfaces as well as studied their role in fuzzy information processing. Considering the need for developing fundamental links between numerical computational framework of the physical environment and the modelling environment constituted by fuzzy sets capturing the reality at the conceptual level, the optimization of these interfaces becomes indispensable. The paper looks at such essential and central design aspects of the development of the input interface such as robustness (fault tolerance), sampling properties, and representation of information uncertainty and incompleteness. The output interface is analyzed with respect to its mapping capabilities. It has been shown that the specific output interface formed with the aid of triangular membership functions intersected at the $\frac{1}{2}$ membership level was capable of generating the zero mapping error. Nevertheless, the minimization of the mapping error for the general form of the interface calls for the learning of the linguistic labels.

While our main purpose was to concentrate on the essential aspects of fuzzy information processing worked out at the level of the interfaces, more detailed numerical investigations are worth pursuing. The resulting optimization tasks that have pertained here exclusively to the interfaces should be also put into a broader optimization perspective of fuzzy modelling and embrace the optimization of the processing block along with the interfaces, as considered e.g. in [6] [7].

Acknowledgment

Support from the Natural Sciences and Engineering Research Council of Canada and MICRONET is gratefully acknowledged.

References

1. J.C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

2. A. Di Nola, S. Sessa, W. Pedrycz, E. Sanchez, Fuzzy Relational Equations and Their Applications in Knowledge Engineering, Kluwer Academic Press, Dordrecht, 1989.

3. D. Dubois, H. Prade, Possibility Theory - An Approach to Computerized Processing of Uncertainty, Plenum Press, New York, 1988.

4. W. Pedrycz, Fuzzy set framework for development of a perception perspective, Fuzzy Sets and Systems, 37, 1990, pp. 123-137.

5. W. Pedrycz, Selected issues of frame of knowledge representation realized by means of linguistic labels, Int. J.of Intelligent Systems, 7, 1992, 155-170.

6. W.Pedrycz, Fuzzy Control and Fuzzy Systems, 2nd edition, Research Studies Press/J.Wiley, Taunton/New York, 1993.

7. W. Pedrycz, Fuzzy neural networks and neurocomputations , Fuzzy Sets and Systems, 55, 1993, to appear.

8. W.Pedrycz, Hierarchical fuzzy modelling for heterogeneous information processing, In: R.R.Yager, L.A. Zadeh (eds), Fuzzy Sets, Neural Networks and Soft Computing, J.Wiley, submitted.

9. J. Valente de Oliveira, Neuron inspired learning rules for fuzzy relational structures, Fuzzy Sets and Systems, to appear

10. J. Valente de Oliveira, On optimal fuzzy systems I/O interfaces, Proc. IEEE Int. Conference on Fuzzy Systems, San Francisco, 1993

11. R.R. Yager, Entropy and specificity in a mathematical theory of evidence, Int. J. General Systems, 9, 1983, 249-260.

12. R.R.Yager, On a hierarchical structure for fuzzy modeling and control, Tech.Rep. IonaCollege, New Rochelle, MII-1213-E, NY1992.

13. R.R.Yager, D.P. Filev, Adaptive defuzzification for fuzzy systems modeling, Proc. NAFIPS'92 Puerto Vallarta, Mexico, Dec 15-17, 1992,vol.I, pp.135-142

14. L.A.Zadeh, Fuzzy sets and information granularity, in: M.M. Gupta, R.K. Ragade, R.R. Yager, eds., Advances in Fuzzy Set Theory and Applications, North Holland, Amsterdam, 3-18, 1979.

15. L.A. Zadeh, Fuzzy sets as a basis for a theory of possibility, Fuzzy Sets and Systems, 1, 1978, 3-28.