REFERENTIAL MODES OF REASONING

Kaoru Hirota* Witold Pedrycz**

*Dept. of Instrument and Control Engineering, College of Engineering
Hosei University, Koganei-city, Tokyo 184, Japan
**Dept. of Electrical and Computer Engineering
University of Manitoba, Winnipeg, Canada R3T 2N2
pedrycz@eeserv.ee.umanitoba.ca

Abstract. We will propose and study models of referential structures and referential modes of reasoning for fuzzy data. The style of information processing considered there allows for reasoning about some global properties of the spaces in which the fuzzy data are placed (like similarity, dominance, inclusion, etc.). The distributed models given in terms of logic-based neural networks realize mappings of these properties between the spaces. The scheme of this type embraces reasoning about similarity, difference, dominance, inclusion of the conclusions that is based on the corresponding relationships and their strength discerned for the antecedents. For instance, the conclusions are of the form: **b** and **b'** are *similar*, **b** and **b'** are *different*, etc. It will be also clarified how, considering the available degrees of satisfaction of these properties, the corresponding fuzzy sets of conclusion can be determined.

Keywords: referential reasoning, fuzzy reasoning, global properties of spaces, matching, inclusion, dominance

1. Introductory remarks and problem statement

Let us consider fuzzy sets defined in two discrete input and output universes of discourse

$$(\mathbf{a}_1, \mathbf{b}_1)\ (\mathbf{a}_2, \mathbf{b}_2)\ ,\dots,(\mathbf{a}_N, \mathbf{b}_N)$$

where $\mathbf{a}_k \in [0,1]^n$ and $\mathbf{b}_k \in [0,1]^m$, $k=1,2,\dots,N$. The frequently accepted approach being used in describing these fuzzy data leads to identification of functional or relational relationships between the variables, say $\mathbf{b}=f(\mathbf{a})$ or $\mathbf{b}=\mathbf{a}\circ\mathbf{r}$, where $\mathbf{r}$ denotes relation existing between the data while "$\circ$" is used to denote a set-to-relation composition operator. Another alternative look at the fuzzy data can be derived by determining transformation of certain global properties defined in the input and output spaces. Formally speaking we will treat this as a mapping,

$$g: \mathcal{F}(\mathbf{a}', \mathbf{a}'') \rightarrow \mathcal{F}(\mathbf{b}', \mathbf{b}'')$$

where $\mathcal{F}(.,.)$ is used to underline a binary operation applied to any two elements in $[0,1]^n$ (and $[0,1]^m$, respectively) and describing a certain property of the space like nearness (similarity), difference, inclusion, dominance, etc. These properties constitute global rather than local (pointwise) characteristics of the space. They do not depend upon the single points of the space but take on a homogenous form valid across the entire universe. Depending on the nature of the mapping $\mathcal{F}$ one can envision several specific interpretations. Take, for instance, the property of similarity. This leads us directly to the scheme of reasoning by analogy [4]. In this form of reasoning one looks at the situation for which a single pair $(\mathbf{a}', \mathbf{b}')$ is completely specified while the second one, say $(\mathbf{a}, \mathbf{b})$ is given partially, namely $\mathbf{b}$ has to determined while $\mathbf{a}$ is provided. In other words, $\mathbf{b}$ should adhere to the property of similarity (analogy), particularly the relationship $\mathcal{F}(\mathbf{b}, \mathbf{b}')$ should be fulfilled to a certain degree depending on the level of similarity achieved for the pair $\mathbf{a}$ and $\mathbf{a}'$, $\mathcal{F}(\mathbf{a}, \mathbf{a}')$. Once the degree of satisfaction of $\mathcal{F}(\mathbf{b}, \mathbf{b}')$ has been computed then the fuzzy set $\mathbf{b}$ can be reconstructed as a solution to a so-called inverse problem.
The paper is structured into sections. The sections will be dealing with the general architecture of the referential mode of reasoning involving their logic-based neural networks, knowledge representation aspects, learning issues and detailed schemes of reasoning.

2. Principal elements of the architecture

The way of handling the global property of the spaces can be concisely summarized according to the format,

$$\mathcal{F} \rightarrow g \rightarrow \mathcal{F}^{-1}$$

Each processing stage distinguished in this scheme takes on a specific interpretation:
(i) $\mathcal{F}$ operates on any two elements of $[0,1]^n$,say a and a' and returns a degree of satisfaction of the property described by $\mathcal{F}$,
(ii) the transformation "g" maps the level of satisfaction of $\mathcal{F}$ achieved in the input space into the level of satisfaction of $\mathcal{F}$ in the output space,
(iii) finallly, $\mathcal{F}^{-1}$ stands for the inverse operation to that carried out for the global property.The inverse problem arising now pertains to the determination of a specific object ,say b',based on the level of satisfaction of the global property elicited in the output space.
Being acquainted how the scheme functions, we will now look at its functional blocks in more detail.

2.1.Representing global properties of similarity,difference,inclusion and dominance.

The global properties are of a referential nature,namely they are determined with respect to any two points of the universe of discourse.The relationship $\mathcal{F}(a',a)$ can be also expressed coordinatewise,

$$\mathcal{F}(a',a)= [\mathcal{F}(a_1',a_1)\ \mathcal{F}(a_2',a_2)\ ...\mathcal{F}(a_n',a_n)]$$

We will distinguish several basic global properties:
-similarity (equality).The degree of its satisfaction by x and a, $x,a \in [0,1]$,EQ(x,a),is embodied by considering the equality index defined as [6],

$$EQ(x,a)=a \equiv x = \frac{1}{2}\left[(a\ \varphi\ x) \wedge (x\ \varphi\ a) + (\bar{a}\ \varphi\ \bar{x}) \wedge (\bar{x}\ \varphi\ \bar{a})\right]$$

where $\wedge$ stands for minimum, overbar denotes complement and φ is used to denote pseudocomplement (implication) ,cf.[6], $a\varphi x=\sup\{c \in [0,1] \mid a t c \le x\}$.
-inclusion.The degree of inclusion of x in a, INCL(x,a), is expressed as

$$INCL(x,a)= x\varphi a$$

difference.The difference is considered to form a dual feature to the property of similarity,namely

$$DIFF(x,a)=1-EQ(x,a)$$

-dominance.The degree of dominance,DOM(x,a) ,is defined as

$$DOM(x,a)= a\varphi x$$

It represents a degree to which x dominates a.
The plots of the two first properties,EQ and INCL ,are shown in Fig.1(both of them are implemented with the use of the Lukasiewicz implication,viz. the φ-operator is given by

$$a\varphi x=\begin{cases} 1-a+x \text{ ,if } a \ge x \\ 1\text{,otherwise} \end{cases}$$

This form of the implication operator implies a piecewise linear form of the derived indices.

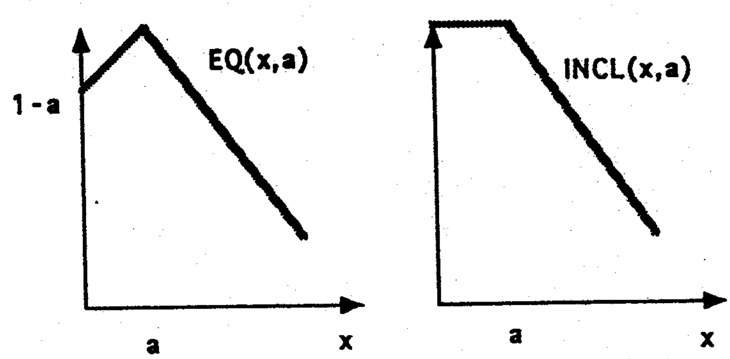


Fig.1.Plots of EQ(x,a) and INCL(x,a).

2.2.Mapping the global properties between the spaces

The global properties will be mapped from the input to the output space by considering logic-based neural networks.Denote by **x** and **y** the results of satisfaction of the global properties in the respective spaces,namely $\mathbf{x}=\mathcal{F}(\mathbf{a},\mathbf{a}')$,$\mathbf{y}=\mathcal{F}(\mathbf{b},\mathbf{b}')$.We will start from the two types of neurons that will be directly used in designing the network.

OR neuron.The OR neuron is described as

$$y = \mathrm{OR}\,(\mathbf{x}; \mathbf{w})$$

or coordinatewise

$$y = \mathrm{OR}[x_1 \text{ AND } w_1, x_2 \text{ AND } w_2, \ldots, x_n \text{ AND } w_n]$$

where $\mathbf{w} = [w_1, w_2, \ldots, w_n] \in [0,1]^n$ forms a vector of connections (weights) of the neuron. The standard implementation of the fuzzy set connectives standing in the formula of the neuron involves the usage of triangular norms,

$$y = \underset{i=1}{\overset{n}{S}} [x_i \, t \, w_i]$$

AND neuron.In comparison to the OR neuron ,the AND neuron uses t-norm to aggregate partial results,

$$y = \mathrm{AND}(\mathbf{x}; \mathbf{w})$$

Again in the notation of the triangular norms this translates into

$$y = \underset{i=1}{\overset{n}{T}} [x_i \, s \, w_i]$$

The aim of the logic-based neural network consisting of these two type of the neurons is to approximate logical relationships between **x** and **y**,**y**=g(**x**).The basic architecture encompasses a single hidden layer composed of the AND nodes (neurons) followed by the OR nodes constituting an output layer of the network.Each AND neuron functions as a generalized minterm.The number of the neurons determines the size of the hidden layer and affects the representation capabilities of the network. The OR neurons placed in the output layer are used to build the generalized union of the minterms.Recall that any Boolean function can be represented as a sum of minterms;here the generalization includes standard forms of minterms and naxterms implemented via the AND and OR neurons.Thus the resulting neural network represents fuzzy function and is used to approximate the available fuzzy data.

2.3.Inverse problem

The global property conveyed by $\mathcal{F}$ implies an associated inverse problem that is formulated as follows:Let us assume that the property $\mathcal{F}$ operating on **b** and **b'** satisfies a certain condition Γ.One of its arguments ,say **b**, is known.Determine all **b'** s satisfying the condition

$$\mathcal{F}(\mathbf{b}',\mathbf{b}) \in \Gamma$$

In particular, depending on the character of this property,we will be interested in more specific problems where

$$\mathrm{EQ}(\mathbf{b}',\mathbf{b}) \geq \gamma \quad \mathrm{INCL}(\mathbf{b}',\mathbf{b}) \geq \gamma \quad \mathrm{DOM}(\mathbf{b}',\mathbf{b}) \geq \gamma \quad \mathrm{DIFF}(\mathbf{b}',\mathbf{b}) \leq \gamma$$

The conditions given above reduce to the families of scalar problems with inequality constraints

$$\mathcal{F}(b'_j,b_j) \leq \gamma_j \qquad \text{or} \qquad \mathcal{F}(b'_j,b_j) \geq \gamma_j$$

j=1,2,...,m.

Depending on the character of $\mathcal{F}$ the following observations hold:

(i) similarity property: the solution set to $b'_j \equiv b_j \geq \gamma$ reduces to a single element $b'_j=b_j$ iff $\gamma_j=1$ .Otherwise the solution to this inverse problem forms a subinterval of [0,1] ,cf.[6]

(ii) the results for the difference property are dual to these summarized in (i),namely $b'_j=b_j$ iff $\gamma_j=0$

In all the remaining situations higher values of γ induce the membership intervals originating around b_j .

(iii) inclusion.This property induces the inverse problem in which $\gamma_j=1$ produces the interval of membership values equal to [0,a].Lower values of γ generate broader ranges of admisible values of membership.

(iv) dominance.For the dominance property the condition $\gamma_j=1$ produces subinterval [a,1]. Again,the lower the value of γ ,the broader the range of the membership values.

It is worth mentioning that except $\gamma=0$ (difference) and $\gamma=1$ (all the remaining problems), the solution to the inverse problem happens to constitute a certain interval of [0,1].This means that even though **b** is a fuzzy set, the result becomes an interval-valued fuzzy set [9].

We will conclude this section with the detailed architecture of the referential structure,see Fig.2.

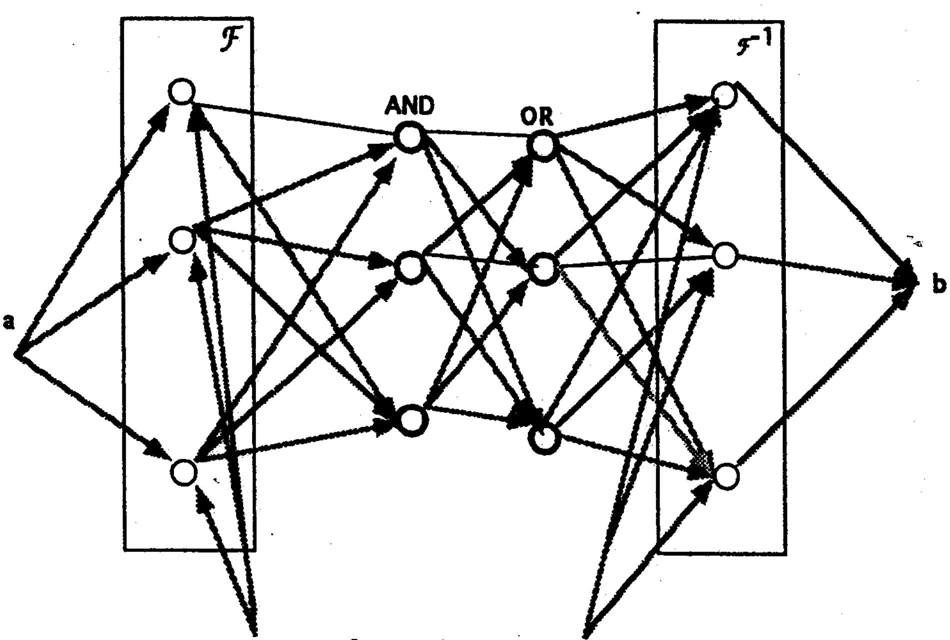


Fig.2.Detailed architecture of the referential structure

3.Learning in the referential structure

The learning of the neural network is based on available cases and is carried out in a supervised mode. The original learning set consists of N cases (pairs) $r_k=(a_k,b_k)$.The process of learning proceeds in a referential format.This involves all different pairs of the cases. These cases used for training can be conveniently arranged in a tabular form

$$\begin{bmatrix} (r_1, r_1) & (r_1, r_2) & (r_1, r_N) \\ (r_2, r_1) & (r_2, r_2) & (r_2, r_N) \\ (r_N, r_1) & (r_N, r_2) & (r_N, r_N) \end{bmatrix}$$

Each pair (r_i, r_j) represents an individual training situation in which one of the cases, say r_i, is viewed as a point of reference.The matrix given above is symmetrical therefore only a single element out of each pair (r_i, r_j) and (r_j, r_i) will be utilized for training.This in total results in $N(N+1)/2$ situations available in the training phase.

The learning set collected in that way is used to adjust the connections of the logic neurons of the network.Assuming a certain performance index Q we will be modifying the connections to achieve minimum of Q,

$$\min_{\text{connections}} Q$$

The standard gradient-like method yields the updates of the connections that are proportional to

$\frac{\partial Q}{\partial \text{connections}}$. The resulting learning scheme is standard to a high extent; the reader is referred to [7] with respect to more specialized computational details as well as some improvements which could be particularly essential for some types of triangular norms being used in the network.

4.Reasoning in the referential structure

Using the structure given in Fig.2 we will discuss in detail how the reasoning process is realized.The basic idea is to cycle through all the cases r_k and for each of them summarize outcomes about the level of the global property satisfied in the output space.We will focuss our analysis on a single element of the output space.Let us start from the first case $\mathbf{a}_1$ (input) and b_1(output).The fuzzy set for which the reasoning will be realized is given as **a**.The corresponding output of the network obtained in this situation is γ_1.Cycling through all the available cases r_k,k=1,2,...,N, we obtain the pairs,

$$b_1, \gamma_1$$
$$b_2, \gamma_2$$
$$\vdots$$
$$b_N, \gamma_N$$

For all the discussed properties but difference we will select a maximal value in the above collection of γ_i's as the one that produces the most specific reasoning results (viz. narrow intervals of membership grades constituting solutions to the inverse problem).Let us first assume that there exists only a single maximal element in the entire collection of γ_i's,

$$\underset{i_0}{\exists!}\ \gamma_{i_0} = \max_{i=1,2,\dots,N} \gamma_i$$

Since γ_{i_0} and b_{i_0} are provided, the inverse problem can be solved.

For the difference property one is looking for the minimal value in the collection of γ_i's.Assuming its uniqueness,namely

$$\underset{i_0}{\exists!}\ \gamma_{i_0} = \min_{i=1,2,\dots,N} \gamma_i$$

the emerging inverse problem takes on the same format as in the previous situation.

The algorithm requires a slight modification if there are several cases for which this maximum (or minimum,respectively) is observed.

Denote by I_0 the elements such that

$$I_0 = \{l=1,2,\dots,p / \gamma_l = \max_{i=1,2,\dots,N} \gamma_i\}$$

or,for the difference property

$$I_0 = \{l=1,2,\dots,p / \gamma_l = \min_{i=1,2,\dots,N} \gamma_i\}$$

The inverse problem has to be tackled with respect to the set of the corresponding membership values.Let those associated values be equal to $b_{i_1}, b_{i_2}, \dots, b_{i_p}$, $i_1, i_2, \dots, i_p \in I_0$.The inverse problem solved separately for each of them gives rise to the interval of the membership values,

$$[b^-_{i_j}, b^+_{i_j}],\ j=1,2,\dots,p$$

The values are aggregated in a conservative way by forming the broadest possible interval out all of them,

$$[\ \min b^-_{i_j},\ \max b^+_{i_j}\ ]$$

j=1,2,...,p

5.Illustrative numerical studies

This section will be devoted to the reasoning about similarities in the input and output spaces.This pertains to reasoning by analogy carried out for fuzzy data,cf.[1][2][3][5][8].We will be using a simple numerical data set discussed in a two-dimensional space as shown in Fig.3.Here n=2 and m=1.The property of similarity has been characterized through the use of the equality index implemented in terms of the Lukasiewicz implication,see Fig.1.The learning has been perfomed in the neural network with the two nodes placed in the hidden layer (the triangular norms used there

were set as: t-norm:product,s-norm:probabilistic sum).The optimized performance index was taken as a sum of squared errors.The obtained connections are equal to:
output-hidden layer:

$$w = [0.022 \quad 0.982]$$

hidden-input layer:

$$v = \begin{bmatrix} 0.553 & 0.000 \\ 0.000 & 0.545 \end{bmatrix}$$

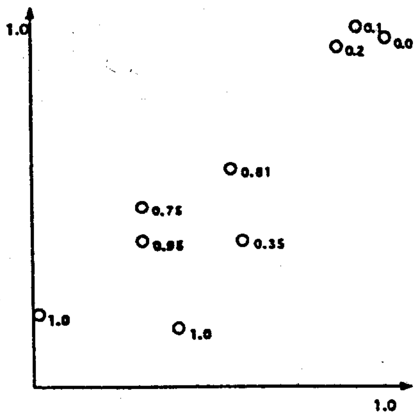
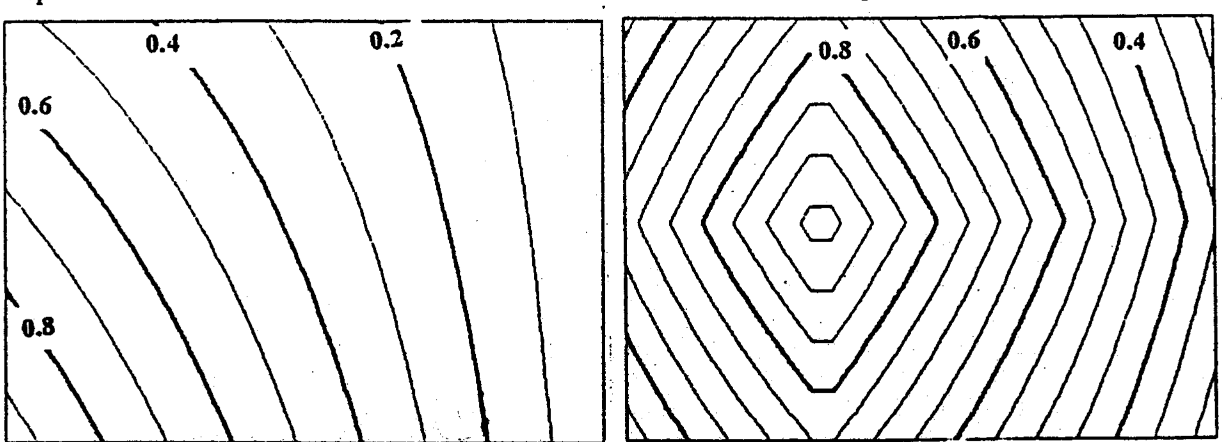


Fig.3. Set of two-dimensional cases

The next plot,Fig.4, summarizes the degrees of satisfaction of the similarity property in the output space for different entries of the neural network and selected reference points.



reference point= (0.0,0.0) **reference point= (0.35,0.50)**

Fig.4 . Values of γ for different input values and selected reference points

Fig.5 illustrates the lower and the upper bounds of the membership values derived from the values of the satisfaction of the similarity property produced by the network.

6.Referential schemes of reasoning

The general types of referential reasoning can be envisioned in the following formats,
(i)

$\mathcal{F}$(a' ,a)
$R=\{r_1,r_2,\ldots,r_N\}$

$\mathcal{F}$(b' ,b)= ? b'=?

where the case (**a**,**b**) is fully specified,.i.e,both **a** and **b** are available while R is used to express symbolically the cases encapsulated in the form of the neural network.Both the degree of the satisfaction of the property $\mathcal{F}$ and the values of the membership function of **b'** will be inferred.

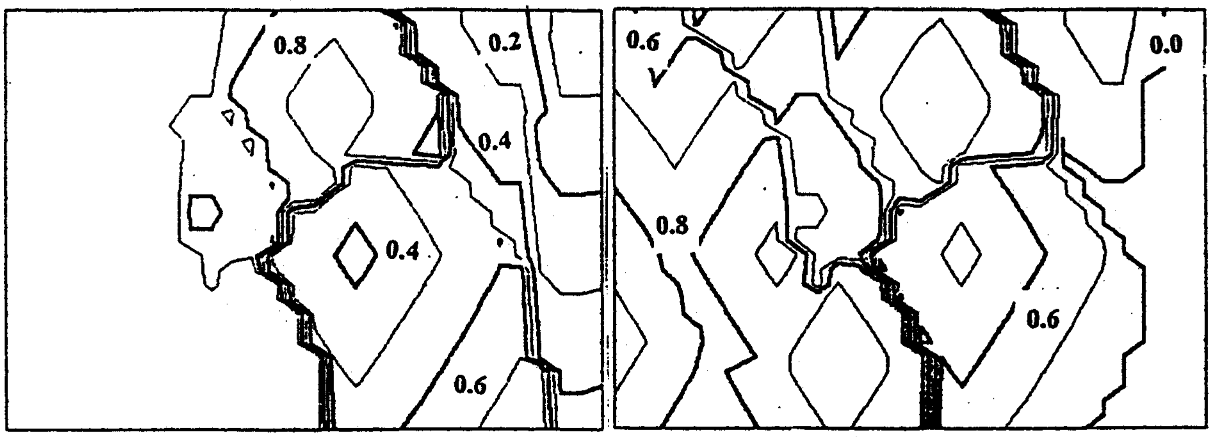


Fig.5.Upper and lower bound of the membership values for different values of the inputs

(ii). **a'** and **a** are given (viz.the case (**a**,**b**) is partially available).The inference is now focussed on inferring the degree of satisfaction of $\mathcal{F}$ in the output space,$\mathcal{F}$(**b'**,**b**),

$\mathcal{F}$(**a'**,**a**)
$R=\{r_1,r_2,\dots,r_N\}$

$\mathcal{F}$(**b'**,**b**)= ?

Depending on the form of $\mathcal{F}$ one can study more specific examples of this class of inference schemes.For instance:

a' and **a** are *similar*
$R=\{r_1,r_2,\dots,r_N\}$

---

b' and **b** are *similar* = ?, **b'**=?

a' is *included* in **a**
$R=\{r_1,r_2,\dots,r_N\}$

---

b' is *included* in **b** = ?

a' is *different* from **a**
$R=\{r_1,r_2,\dots,r_N\}$

---

b' is *different* from **b**= ?, **b'**=?

In general,the available cases can be used in constructing several specialized inference schemes that handle separately various global properties.The schemes are also constructed on a basis of different mutually exclusive families of cases utilized for training purposes.For instance,the inference to be realized for **a'** can invoke a series of schemes of reasoning

$\mathcal{F}_1$(**a'**,**a**)
R_1

$\mathcal{F}_1$(**b'**,**b**)= ? **b'**=?

$\mathcal{F}_2$(**a'**,**a**)
R_2

$\mathcal{F}_2$(**b'**,**b**)= ? **b'**=?

...

$\mathcal{F}_c(a',a)$
R_c

$\overline{\mathcal{F}_c(b',b)= ?\ \ b'=?}$

The selection of the scheme that is the most suitable for a'can be accomplished by considering the maximum of the satisfaction levels obtained for the schemes.Denote the results produced by them by $\gamma_l=\mathcal{F}_l(b',b), l=1,2,\ldots,c$. The sum of the coordinates $\sum_{j=1}^{m}\gamma_{lj}$ computed for each of the schemes introduces a linear order among them. The reasoning scheme with the highest value of this sum becomes selected.

7.Conclusions

We have developed a general scheme of referential reasoning.It enables us to reason about some general properties satisfied for the conclusion as well as determine this conclusion in the form of interval-valued set.The distributed model of mapping the global properties has to be constructed with the aid of logic-based neural networks.The referential scheme of reasoning realized within the proposed framework embraces several types of global properties one can reason about.The relevancy of the results of reasoning are quantitatively expressed by interval-valued fuzzy sets.

Acknowledgements

Support from the Natural Sciences and Engineering Research Council of Canada and MICRONET is greatly appreciated.

8.References

1.Diamond,J.,McLeod,R.D.,Pedrycz,W.,A fuzzy cognitive structure: Foundations, applications and VLSI implementation, *Fuzzy Sets and Systems*,47,49-64, 1992.
2.Di Nola,A.,Pedrycz,W.,Sessa,S.,Knowledge representation and processing in frame-based structures, In: *Fuzzy Engineering Toward Human Friendly Systems*, vol.1,Proc.Int.Fuzzy Engineering Symp.'91, Yokohama,pp. 461-470,1991.
3. Hirota K., Pedrycz,W, "Concepts formation: representation and processing issues",*Int. J. of Intelligent Systems*, 7, 3-13,1992.
4.Kling, R.E.,A paradigm for reasoning by analogy, *Artificial Intelligence*,2, 147-178,1971.
5.Pedrycz,W.,A fuzzy cognitive structure for pattern recognition,*Pattern Recognition Letters*,9,305-313,1989.
6.Pedrycz,W.,Direct and inverse problem in comparison of fuzzy data,*Fuzzy Sets and Systems*, 34, 223-235, 1990.
7.Pedrycz,W.,Neurocomputations in relational systems,*IEEE Trans.on Pattern Analysis and Machine Intelligence*,13,289-296,1991.
8.Pedrycz,W.,Bortolan,G.,Degani,R.,Classification of electrocardiographic signals:a fuzzy pattern matching approach,*Artificial Intelligence in Medicine*,3,331-46,1991.
9.Sambuc,R.,Functions ϕ-flous:Application de l'Aide a Diagnostique en Pathologie Thyroidienne, These Univ.de Marseille,Marseille,1975.