The Third AFSS(1998), 719-724

Some new similarity based approaches in approximate reasoning and their applications to pattern recognition

Swapan Raha[1], Nikhil R. Pal[2] and Kumar Sankar Ray[3]
1. Department of Mathematics, Visva - Bharati, Santiniketan 731 235, WB, INDIA
2. Machine Intelligence Unit, Indian Statistical Institute, Calcutta - 35, INDIA
3. Electronics & Commn. Sc. Unit, Indian Statistical Institute, Calcutta - 35, INDIA
email: nikhil@isical.ernet.in, ksray@isical.ernet.in, bipasha@vbharat.ernet.in
2 Author for correspondence

Abstract

This paper presents a systematic development of a formal approach to inference in approximate reasoning. We introduce some measures of similarity and discuss their properties. Using the concept of similarity index we formulate two methods for inferring from vague knowledge. In order to illustrate the effectiveness of the proposed technique we use it to develop a vowel recognition system.

1 Introduction

In this paper we have considered approximate reasoning with only vague concepts that often appears in a subject-predicate formulation of natural language and for this we have considered a semantic representation for vague concepts, i.e., we have used a set to exemplify a concept and then use laws of the underlying set theory, already established, to manipulate them. Here we use fuzzy sets for representation and the theory of fuzzy sets for manipulation of vague predicates.

The proposed method of inference is based on a similarity measure. In practice, in a rule-based system, the rule (a condition) is first expressed as an implication (or a triangular norm) to be selected suitably and in a resolution - based system, the disjunction is expressed as a triangular co-norm to be selected suitably in order to generate a meaningful possibility distribution and in both cases we have interpreted them as a conditional possibility distribution. Then new facts are used to modify the above possibility distribution and the result is interpreted as the induced possibility distribution. Here in computing the induced possibility distribution we have used the concept of similarity measure between a pair of fuzzy sets.

Such similarity related work goes back to Turksen & Zhong [4]. Recently, Tsang et al. [3] also proposed a similar scheme for similarity based reasoning. From a given fact the conclusion is derived using a measure of similarity between the fact and the antecedent. They associated a threshold value τ to each rule. If the degree of similarity exceeds τ of the underlying rule, then only that rule is assumed to be fired. As an illustration, consider the statement

p : if X is A then Y is B, τ
and q : X is A'.

Here A and A' are fuzzy sets defined over the same universe of discourse U and B is defined over the universe V. Let $S(A, B)$ be a measure of similarity [7] between two fuzzy sets A and B. Now, if $S(A, A') > \tau$ then Turksen & Zhong [4] compute the conclusion $B' = \{\mu_{B'}/v_i, i = 1, \cdots, n\} = \{\min(1, \frac{\mu_B/v_i}{S(A,A')}), i = 1, \cdots, n\}$.

To simplify notation, we will write it as $B' = \min(1, B/S)$.

These methods [3, 4] use the similarity measure for a direct computation of inference without considering the induced relation, i.e., how the underlying relation (condition) is modified by the given fact. Consequently, these methods provide the same conclusion if we interchange A and A' in the propositions concerned. This is not appealing. Moreover, the conclusion becomes independent of the relational operator used. This motivates us to measure the absolute change in linguistic labels, represented as fuzzy sets, and systematically propagate the same to the conditional possibility distribution as induced by the facts in order to obtain a modified induced possibility distribution. From this, a possible conclusion can be drawn using the projection principle and it has been shown that nothing better than what the rule says can be concluded.

Our exposition to inferring in approximate reasoning paradigm begins in Section 2 where we develop the concept of similarity index for measuring the likeness of fuzzy sets over a given universe of discourse. We also discuss some basic properties and results in connection with similarity measures. In the next Section we present a new technique of inferring based on similarity measures together with the formulation of two deductive processes — generalized modus ponens and generalized disjunctive syllogism, the basic rules of inference based on the above technique of inferencing. In Section 4 we demonstrate an application to the vowel recognition problem and discuss the results. The paper is briefly concluded in Section 5.

2 Similarity index

In this section we introduce a notion of similarity between fuzzy sets defined over the same universe of discourse. For the sake of simplicity, we assume that the universe of discourse is a finite set, although many of our results are also true for infinite sets. Let $A = \sum_{u \in U} \mu_A(u)/u$ and $A' = \sum_{u \in U} \mu_{A'}(u)/u$ be two fuzzy sets defined over the universe of discourse U. A similarity index between the pair $\{A, A'\}$ denoted by $S(A, A')$ is a mapping $S : \mathcal{P}(U) \times \mathcal{P}(U) \longrightarrow [0, 1]$, where $\mathcal{P}(U)$ is the fuzzy power set of U. S should satisfy the following properties.

P1. $S(A, B) = S(B, A)$ for all fuzzy sets A and B.

P2. For all fuzzy sets A and B, $S(A, B) = S(A^c, B^c)$,A^c being the complement of A.

P3. For all fuzzy sets A and B, $0 \le S(A, B) \le 1$.

P4. Two fuzzy sets A and B are equal if and only if $S(A, B) = 1$.

P5. For all fuzzy sets A and B, if $S(A, B) = 0$ then either $A \cap B = \Phi$ or $A^c \cap B^c = \Phi$ or $B = 1 - A$.

There could be many functions satisfying properties **P1** through **P5**. One such measure of similarity satisfying properties **P1** through **P5** is given next.

Definition 1. Let $A = \sum_{u \in U} \mu_A(u)/u$ and $A' = \sum_{u \in U} \mu_{A'}(u)/u$ be two fuzzy sets defined over the universe of discourse U. The similarity index of the pair $\{A, A'\}$ is defined by $S(A, A') = \min\{\alpha(A, A'), \alpha(1 - A, 1 - A')\}$ where

$$\alpha(A, A') = \left\{ \frac{\sum_{u \in U}\{\mu_A(u) \times \mu_{A'}(u)\}}{\sum_{u \in U}\{\max(\mu_A(u), \mu_{A'}(u))\}^2} \right\}^{(1/2)} \quad (1)$$

and $\mu_{1-A}(u) = 1 - \mu_A(u)$. In case $\sum_{u \in U}\{\max(\mu_A(u), \mu_{A'}(u))\}^2 = 0$ we find that A and A' are null fuzzy sets and we set $\alpha(A, A') = 1 = S(A, A')$. It is easy to see that **P1** through **P5** are satisfied by Definition 1.

Although the last property **P5** is a plausible and an intuitively appealing one, we can possibly argue in favor of a stricter condition for which $S(A, B)$ should be zero. Two crisp sets A and B are completely dissimilar only when $A \cap B = \Phi$. If $A \cap B \neq \Phi$, then they have some similarity as A and B have some elements in common. The similarity between the two increases as the number of elements by which the two sets differ decreases. The similarity becomes maximum (the maximum value may be thought of as 1) when the two sets are identical, i.e., $| A \cap B | = | A | = | B |$. We now demand a direct extension of this concept to fuzzy sets. We know that for two non-crisp sets A and B always $A \cap B \neq \Phi$,i.e., two fuzzy (non-crisp) sets always have some degree of overlapping (fuzzy subsethood). Therefore, it is reasonable to assume, that two non-crisp sets always have some degree of similarity. The similarity should be zero if and only if $A \cap B = \Phi$, i.e., only when $A = B^c$ and A and B are crisp. We can now reformulate property **P5** as

P5′. For all fuzzy sets A and B, $S(A, B) = 0$ iff $A \cap B = \Phi$.

The need, thus, arises to find measures of similarity satisfying properties **P1** through **P4** and **P5′**. There could be several such measures, a family of such simple measures is given by the next definition.

Definition 2 : Let $A = \sum_{u \in U} \mu_A(u)/u$ and $B = \sum_{u \in U} \mu_B(u)/u$ be two fuzzy sets defined over the same universe of discourse U. The similarity index of the pair $\{A, B\}$ is defined by $S(A, B)$

$$= 1 - \left(\frac{\sum_u | \mu_A(u) - \mu_B(u) |^q}{n} \right)^{(1/q)}$$

where n is the cardinality of the universe of discourse and q is the family parameter. Next we discuss some properties of S without proofs which are reported in [8].

Theorem 1. If $S(A, B) = 1$ and $S(B, C) = 1$ then $S(A, C) = 1$.

Of course, in general $S(A, B)$ and $S(B, C)$ cannot determine $S(A, C)$.

Theorem 2. For all fuzzy sets A, B and C if either $A \subseteq B \subseteq C$ or $A \supseteq B \supseteq C$ then $S(A, C) \le \min\{S(A, B), S(B, C)\}$.

3 Proposed Method

Now let us see how conclusions can be obtained from given premises with the help of such a similarity measure. Let us consider two linguistic variables X, Y and let $\mathcal{U}$, $\mathcal{V}$ respectively denote the universes of discourse. Consider two typical propositions p : if X is A then Y is B and q : X is A'. From p

and q we would like to derive a conclusion r : Y is B'. Let $\mathcal{U} = \{u_1, u_2, \cdots, u_l\}$, $\mathcal{V} = \{v_1, v_2, \cdots, v_m\}$. Translating the inexact concepts in the propositions p and q into appropriate possibility distributions, we have $A = \sum_{i=1}^{l} \mu_A(u_i)/u_i$; $A' = \sum_{i=1}^{l} \mu_{A'}(u_i)/u_i$; $B = \sum_{i=1}^{m} \mu_B(v_i)/v_i$. From p and q the conclusion r can be computed using the following steps :

Step 1. Translate premise p and compute $R(A, B)$ using any suitable translating rule possibly, a T-norm operator.

Step 2. Compute $S(A, A')$ according to either Definition 1 or Definition 2 or some other Definition.

Step 3. Modify $R(A, B)$ with $S(A, A')$ to obtain the conditioned relation $R(A' \mid A, B)$ according to some scheme.

Step 4. Project $R(A' \mid A, B)$ on V to obtain $\mu_{B'}(v) = \sup_u \mu_{R(A'|A,B)}(u, v)$.

Given the fact q : X *is* A', we propose two schemes $C1$ and $C2$ for computation of the conditioned relation $R(A' \mid A, B)$ in Step 3.

Scheme C1

The first scheme C1 is based on a concept similar (but NOT identical) to the method of Turksen & Zhong [4]. Recall that Turksen & Zhong computed the conclusion $B' = \min(1, B/S)$, where S is the measure of similarity between fuzzy sets A and A'. Here we propose

$$R(A' \mid A, B) = [r'_{u,v}]_{l \times m} = \begin{bmatrix} r'_{u,v} & = & \min(1, r_{u,v}/S) \text{ if } S > 0 \\ & = & 1 \text{ otherwise.} \end{bmatrix} \quad (2)$$

Note the differences of our scheme from that of Turksen & Zhong. They did not compute the conditioned relation. Moreover, in their scheme when S exceeds a threshold then only the conclusion is computed. It is clear that our scheme, unlike the schemes in [3, 4], does not necessarily produce the same conclusion when A and A' are interchanged.

Scheme C2

We believe that a scheme for computation of the conditioned relation based on similarity should satisfy the following basic properties :

A1 : If $S(A, A') = 1$, i.e., if $A = A'$, then the strength of association $\mu_{R(A'|A,B)}(u, v) = \mu_{R(A,B)}(u, v) \quad \forall (u, v) \in U \times V$.

A2 : If $S(A, A') = 0$, i.e., when $A' = A^c$ and A is crisp then $\mu_{R(A'|A,B)} = 1 \quad \forall (u, v) \in U \times V$.

A3 : As $S(A, A')$ increases from 0 to 1, $\mu_{R(A'/A,B)}(u, v)$ should decrease uniformly from 1 to $\mu_{R(A'|A,B)}(u, v)$; $\quad \forall (u, v) \in U \times V$.

A1 needs no explanation, A2 asserts that when A' is completely dissimilar to A, nothing can be concluded. A2 says that as the fact changes from the most dissimilar case to the most similar one, the inferred conclusion changes from the UNKNOWN case to B and also says that whatever be A', $R(A' \mid A, B) \geq R(A, B)$, i.e., the conditioned relation cannot be more specific than what was given.

For notational simplicity, let us denote $S(A, B) = s$ and $R_{A'|A,B} = r'$. Now A3 *uniquely* suggests a function of the form $\frac{dr'}{ds} = k \Rightarrow r' = k\ s + c$, c is a constant. When, $s = 1, r' = r$ (for A1) and when $s = 0, r' = 1$ (for A2). This gives $c = 1$ and $r' = 1 - (1 - r)s$. Therefore, A1 through A3 *uniquely* suggest the scheme C2 as

$$\mu_{R(A'|A,B)} = 1 - (1 - \mu_{R(A,B)})S(A, A'). \quad (3)$$

From the above when $S(A, A') = 0$ we find that $B' = V$, in other words, it is impossible to conclude anything when $\{A, A'\}$ are dissimilar. It is also easy to see when $S(A, A')$ is close to unity, then $R(A' \mid A, B)$ is close to $R(A, B)$ and hence $S(B, B')$ would also be close to unity. A3 also suggests that a small change in the input produces a small change in the output and hence in this sense the above mechanism of inference is stable.

Next as Theorem 3 we present a very basic and desirable property of the inferred proposition that nothing better than what the rule says can be concluded.

Theorem 3. If X is A then Y is B
X is A'

Then the conclusion B' obtained by the proposed scheme always satisfies $B' \supseteq B$.

3.1 Different Models

In this section we discuss the application of the proposed inference technique and formulate deductive processes using the theory of possibility and the concept of similarity index. We shall consider two types of models : rule-based and resolution based.

Rule-based model

Let us consider the following model :

if X_1 is A_{11} & X_2 is A_{12} & $\cdots$ & X_n is A_{1n}
then Y is B_1 else
if X_1 is A_{21} & X_2 is A_{22} & $\cdots$ & X_n is A_{2n}
then Y is B_2 else
⋮ ⋮
if X_1 is A_{m1} & X_2 is A_{m2} & $\cdots$ & X_n is A_{mn}
then Y is B_m
X_1 is A_1 & X_2 is A_2 & $\cdots$ & X_n is A_n
Conclusion : Y is B

Under the conventional fuzzy reasoning paradigm, for each rule the consequent fuzzy set is modulated by the firing strength of each rule (computed usually by some T-norm) and then the union (usually computed by some S-norm) of all modulated fuzzy sets is taken as the conclusion which is then crispified by some defuzzification scheme.

In the present case of similarity based reasoning we cannot do this as the possibility distribution computed from the conditioned relation becomes less and less specific as the similarity between the facts and antecedent of a rule decreases. In conventional paradigm also, the possibility distribution over various alternatives becomes ambiguous (more alternatives with similar possibility values) with the reduction of the firing strength, but the possibility values at which the ambiguity occurs becomes less. For example, in case of Mamdani-type of reasoning, if the firing strength of a rule is, say 0.3, then all alternatives which have possibility values greater than equal to 0.3 take possibility values of 0.3. On the other hand, in the present case, if the similarity value is 0.3, then the computed possibility values for all alternatives will be greater than or equal to those in the possibility distribution on the consequent variable of the concerned rule. This means that with decrease in similarity the computed possibility distribution will be more close to the least specific possibility distribution (with possibility values of 1 for all alternatives). Here based on the principle of least specificity, we propose the following scheme for computing the final conclusion. Our method is based on relative specificity of the conditioned possibility distributions. By relative specificity $RS(A, B)$ of a possibility distribution A with respect to another possibility distribution B, where B is less specific than A (i.e., $A \subseteq B$) we mean a concept which increases as B approaches A and it attains the minimum value when B is the least specific fuzzy set with possibility values of 1 for all alternatives.

Let for the ith rule, $P_i = \{p_{1i}, \cdots, p_{ki}\}$ be the possibility distribution on the consequent and $\pi_i = \{\pi_{1i}, \cdots, \pi_{ki}\}$ be the conditioned possibility distribution, then a measure of relative specificity can be defined as

$$RS_i = S(P_i, \pi_i).$$

With this concept of relative specificity we propose the following scheme for computation of the conclusion.

Step 1 : For each rule R_i, we compute the conditioned possibility distribution (π_i) as described earlier.

Step 2 : Compute the relative specificity, RS_i, $\forall i$, i.e., for all rules.

Step 3 : Find the final possibility distribution π as $\pi = \pi_m = \{\pi_{1m}, \cdots, \pi_{km}\}$ where $m = Argmax_i\{RS_i\}$.

Step 4 (optional) : The crisp output, if needed, can be computed as u_l such that $\pi_l = \max_j\{\pi_{jm}\}$.

In Step 3, we actually find the rule whose antecedent best matches (in terms of the given similarity measure) the given fact and accept the conclusion suggested by that rule. Ties, both in Step 3 and Step 4, can be broken arbitrarily.

Resolution based model

Let us now consider resolution based models for reasoning[1]. Consider two typical propositions p and q as

$$p : X \text{ is } A \text{ or } Y \text{ is } B \ ;$$
$$q : X \text{ is } A' \ .$$

From p and q we would like to derive a conclusion

$$r : Y \text{ is } B' \ .$$

according to the following steps. First we compute $S(A, A')$ and set

$$\beta = 1 - S(A, A') \neq 0.$$

In this case the inferential procedure also remains the same. In other words, under scheme C1 we compute $R(A' \mid A, B)$ substituting β for S in (2) and under scheme C2 substituting β for S in (3). The conclusion B' is finally obtained projecting $R(A' \mid A, B)$ on V.

When $\beta = 0$ nothing in particular can be concluded and we set $B' = V$, if we use C1. When we use C2, automatically B' is derived as V. If $\beta = 1$, i.e., $S(A, A') = 0$, i.e., A and A' are complementary fuzzy sets then we find $B' = B$ for both the schemes which is in accordance with the law of generalized disjunctive syllogism[1].

4 An Application in Vowel Recognition

We present an illustrative application of the proposed similarity based inferencing scheme to recognition of Telugu vowels [2]. Let $X = \{x_1, ..., x_n\} \in R^p$ be the training data and let there be c classes. We represent p real-valued features by $\{F_1, F_2, ..., F_p\}$. The problem is to design a rule-based system using the similarity based reasoning scheme so that unknown points can be classified. The data set consists of 800 samples of discrete phonetically balanced speech samples for the Telugu vowels in consonant-vowel nucleus-consonant (CNC) form.

The system consists of rules of the form *If* X_1 *is LOW .... X_p is MEDIUM then* π_i. Here π_i is a possibility distribution on the set of classes. In order to design such a rule-based system first of all we partition each feature space to define fuzzy linguistic values like LOW, MEDIUM etc. for each linguistic variable (feature). An initial set of such rules may be either obtained by exploratory data analysis or from experts and then may further be tuned to refine the performance of the system.

The initial rule set is so designed that it covers the entire input space. The possibility distributions can also be assigned by experts or may be learnt. Since our intention here is not to address all these issues but to show just an application of the proposed inference scheme we use a rule set similar to the one used in [2] and apply the similarity based reasoning scheme.

Following authors in [2] we are considering here only the first and second formant frequencies. Thus we a have two-dimensional data set with 800 points. For linguistic variables F_1 and F_2 we consider 5 and 7 linguistic values respectively and generated thirty five rules as given in Tables 1 and 2. In order to classify an unknown pattern, we first fuzzify the given feature values by triangular fuzzy sets (one can safely use other methods like fuzzy singleton as well). Then apply the similarity based approximate reasoning so as to obtain a possibility distribution on different classes. At the time of making a non-fuzzy decision we can select the class with maximum possibility value. Ties, if arises, may be broken arbitrarily.

∧	ZE : Zero	BL : Below Low
BZ	.5/o+1/u	.1/e+.5/o+1/u
ZE	.5/e+1/i	.5/e+1/i
BL	1/e+.1/o+.1/u+.5/∂	1/e+.5/i+.5/∂
LO	.5/a+.1/e+1/o +.1/u+.1/∂	.5/a+.5/e+1/o +.1/u+1/∂
ME	1/e+.5/i	.1/a+.5/o
HI	1/e+1/∂	1/e+.1/∂
AH	1/a+.1/o+1/∂	1/a+1/∂

Table 1. Rule-base (contd. to Table 2)

In Tables 1 and 2 BZ = Below Zero, ZE = Zero, BL = Below Low, LO = Low, ME = Medium, HI = High, and AH = Above High

∧	LO : Low	ME : Medium	HI : High
BZ	.5/e+.5/o+1/u	.5/e+.1/o+.1/u	.5/e+.5/i
ZE	1/o+1/u +1/u+.1/∂	.1/e+1/o +.5/u+.5/∂	1/e+1/o
BL	1/e+1/i+.1/∂	1/e+1/i	.1/a+1/o+.1/u
LO	.5/a+1/e +.1/o+1/∂	1/e+.1/i+1/∂	1/e+.5/i+.1/∂
ME	1/a+1/e+.1/∂ +.5/o+1/∂	1/a+.1/e +.1/o+1/∂	1/a+0.5/e
HI	.5/e	.1/a+.1/o	1/a+.1/o+.1/∂
AH	.5/e+.5/∂	.5/e+.1/∂	.1/e

Table 2: Rule-base (continued from Table 1)

Table 3 shows the performance of the proposed classifier driven by similarity based reasoning. While generating Table 3, as mentioned earlier, ties were broken arbitrarily. From Table 3 we find that the average recognition score is about 65%. The performance of the classifier is quite satisfactory because in the 2-space there are significant overlaps between different pairs of classes [2]. Some improvement in performance may be realized through tuning of the membership functions. But for any classifier which uses only two features some misclassifications are bound to result. However, use of the third formant frequency may help further improve the performance.

	a	e	i	o	u	∂	Total	Score(%)
a	41	0	0	7	0	35	83	49.40
e	1	131	20	10	0	38	200	65.50
i	0	33	100	0	0	0	133	75.19
o	3	7	0	71	17	18	116	61.21
u	0	8	0	23	79	2	112	70.54
∂	12	7	0	12	0	35	66	53.03
Total	57	186	120	123	96	128		64.37

Table 3 : Recognition score for Telugu vowels

	a	e	i	o	u	∂	Total	Score(%)
a	72	0	0	8	0	3	83	86.75
e	0	184	0	5	0	11	200	92.00
i	0	8	125	0	0	0	133	93.98
o	1	0	0	112	0	3	116	96.55
u	0	8	0	11	92	1	112	82.14
∂	0	2	0	0	0	64	66	96.97
Total	73	202	125	136	92	82		91.41

Table 4 -Recognition score of Telugu vowels assuming correct classification when the suggested choices include the actual class

If the proposed scheme is a consistent one, then the classifier is likely to suggest more than one choice with the highest possibility for points lying in the overlapped regions. In order to establish that it is indeed the case, we now assume that if the alternatives suggested by our rule-based system include the correct class, then the system output is correct. Table 4 is generated keeping this in mind. To make it more clear, we computed the recognition scores in Table 4 as follows. If the rule-base suggests only one class and it is the correct class then the recognition score is increased. If the system suggests more than one class containing the correct class then also the recognition score is increased. The dramatic improvement in the recognition score suggests that, when the proposed scheme generates more than one choice, it usually includes the correct choice and possibly they correspond the overlapped region.

5 Conclusions

In this paper we have proposed two new schemes for approximate reasoning with vague knowledge which are based on a measure of similarity. These schemes are applicable to both rule-based and resolution-based systems. In this regard we have defined axiomatically several measures of similarity between fuzzy sets and investigated their properties. One of the proposed schemes can be viewed as a modification/ extension of a similarity based reasoning system suggested by Turksen and Zhong; while the other scheme axiomatically suggested a unique method for propagating a measure of similarity between the antecedent and the facts into the inferred conclusion. Several interesting properties of the proposed schemes have been studied. Finally, as an illustration of their effectiveness, we have shown an application of the proposed schemes in designing a vowel recognition system.

References

[1] S.Raha and K.S.Ray, ' Approximate reasoning based on generalized disjunctive syllogism ', Fuzzy Sets and Systems, 1994, vol.61, no.2, pp. 143-151.

[2] K.S.Ray and J.Ghoshal, 'Neuro fuzzy approach to pattern recognition, Neural Networks, 1997, vol 10, no.1, pp. 161-182.

[3] F.C.C. Tsang, J.W.T.Lee and D.S.Yeung, 'Similarity based fuzzy reasoning methods for fuzzy production rules', Proc. IFSA, world Congress, Sao Paolo, Brazil, 1995, pp. 157-160.

[4] I.B.Turksen and Z.Zhong, 'An approximate analogical reasoning approach based on similarity measures', IEEE Trans. on SMC, vol.18(6), 1988, pp. 1049-1056.

[5] L.A.Zadeh, ' Fuzzy logic and approximate reasoning ', Synthese, 1975, vol. 30, pp. 407-428.

[6] L.A.Zadeh, ' A theory of approximate reasoning ', in : J.E.Hayes, D.Mitchie and L.I.Mikulich,Eds., Machine Intelligence Vol. 9, 1979, (Wiley,New York) pp. 149-194.

[7] R.Zwick, E.Carlstein and D.V.Budescu, 'Measures of similarity among fuzzy concepts : A comparative analysis', Inter. Jour. of App. Reasoning, vol 1, 1987, pp. 221-242.

[8] S. Raha, N. R. Pal and K. S. Ray, 'Similarity based approaches to inference in approximate reasoning', IEEE Transactions on Fuzzy Systems (Communicated).