1. Publication Nº	2. Version	3. Date	5. Distribution
INPE-2266-PRE/047		*Nov., 1981*	☐ Internal ☒ External
4. Origin *DSE/DIN*	Program *DIAGNOSTICO*		☐ Restricted

6. Key words - selected by the author(s)
FUZZY SETS
FUZZY INFERENCE
MULTI-VALUED LOGIC

7. U.D.C.: *681.3.019*

8. Title *INPE-2266-PRE/047*

CHANGING THE FUZZY RULE OF DETACHMENT

10. Nº of pages: *14*

11. Last page: *12*

12. Revised by

Paulo Ouvera Simoni

13. Authorized by

Nelson de Jesus Parada
Director

9. Authorship *Orion de Oliveira Silva*
Celso de Renna e Souza

Responsible author Orion de Oliveira Silva

14. Abstract/Notes

The "modus ponens" inference rule for fuzzy premises and conclusion is usually given by: "If x is P, then y is Q" (represented by "(x,y) is $(\overline{P}' \oplus \overline{Q})$*"); "x is R"; therefore, "y is* $R \circ (\overline{P}' \oplus \overline{Q})$*". It so happens that, for normal P's, if* $R = P$*, all values of the distribution* $P \circ (\overline{P}' \oplus \overline{Q})$ *are greater than or equal to the corresponding values from Q. From the logical essence of the rule of detachment, however, one would expect to have* $P \circ (\overline{P}' \oplus \overline{Q}) = Q$*, and not simply a distribution semantically implied by Q. The result* $P \circ (\overline{P}' \oplus \overline{Q}) \geqslant Q$ *(for normal P's) is partially due to a certain "information loss" brought about by the application of the bounded sum operator* $\oplus$*. In this paper, another operator is proposed, improving the result obtained when* $R = P$*. Moreover, the association of degrees of confidence, in the interval* $[0,1]$*, to premises and implication is allowed, thus generalizing the fuzzy "modus ponens" rule.*

15. Remarks *Submitted to the 12th. International Symposium on Multiple Valued Logic, Paris, France, May 1982.*

CHANGING THE FUZZY RULE OF DETACHMENT

by

Orion de Oliveira Silva

Celso de Renna e Souza

Instituto de Pesquisas Espaciais

Conselho Nacional de Desenvolvimento Científico e Tecnológico

Caixa Postal 515 - 12200 - São José dos Campos, SP

Tel. (0123) 22-9977 Extension 248

ABSTRACT

The "modus ponens" inference rule for fuzzy premises and conclusion is usually given by: "If x is P, then y is Q" (represented by "(x,y) is ($\overline{P}' \oplus \overline{Q}$)"); "x is R"; therefore, "y is R o ($\overline{P}' \oplus \overline{Q}$)". It so happens that, for normal P's, if R = P, all values of the distribution P o ($\overline{P}' \oplus \overline{Q}$) are greater than or equal to the corresponding values from Q. From the logical essence of the rule of detachment, however, one would expect to have P o ($\overline{P}' \oplus \overline{Q}$) = Q, and not simply a distribution semantically implied by Q. The result P o ($\overline{P}' \oplus \overline{Q}$) $\geqslant$ Q (for normal P's) is partially due to a certain "information loss" brought about by the application of the bounded sum operator $\oplus$. In this paper, another operator is proposed, improving the result obtained when R = P. Moreover, the association of degrees of confidence, in the interval $[0,1]$, to premises and implication is allowed, thus generalizing the fuzzy "modus ponens" rule.

INTRODUCTION

There is no need to stress the importance of the rule of detachment in classical two-valued logic. There are systems for the predicate calculus that use "modus ponens" as the only inference rule needed. In fuzzy logic, the adaptation of the rule to fuzzy premises and conclusion, with complete confidence in the implication, is usually given by (Zadeh, 1975a):

a) "If x is P then y is Q", represented by "(x,y) is $(\overline{P}' \oplus \overline{Q})$"

b) "x is R", therefore "y is $R \circ (\overline{P}' \oplus \overline{Q})$";

where P' is the complement of P; $\overline{P}'$ is the cylindrical extension of P' over the (non-fuzzy) basis set of Q; $\overline{Q}$ is the cylindrical extension of Q over the basis set of P; and $\oplus$ is the bounded-sum operator.

The first part of the rule may be considered an adaptation, to fuzzy sets, of Lukasiewicz' material implication in multi-valued logic. In his logic, $v(r \rightarrow s) \triangleq \min(1, 1-v(r)+v(s))$, where $v(r)$ and $v(s)$ are truth values of r and s, both in the interval $[0,1]$. In fact,

$$\mu_{\overline{P}' \oplus \overline{Q}}(x,y) = \min(1, 1-\mu_P(x)+\mu_Q(y)) = (1-\mu_{\overline{P}}(x,y)) \oplus \mu_{\overline{Q}}(x,y)$$

$$= (1-\mu_P(x)) \oplus \mu_Q(y) \;;$$

(see also Zadeh, 1975b).

Consider now the fuzzy proposition:

"If x is old then x has high blood pressure",

with "x is old" and "x has high blood pressure" interpreted as fuzzy sets. The subsequent assertion of:

"Zadeh is old",

should bring, as an obvious conclusion, "Zadeh has high blood pressure". Unfortunately (or fortunately for Zadeh), the distribution of possibilities (fuzzy set) for Zadeh's blood pressure obtained using the rule above does not correspond exactly to "high blood pressure". In fact, with R = P, one can be sure only that, in this case, all values of $P \circ (\overline{P}' \oplus \overline{Q})$ are larger than or equal to those of Q.

Another unconfortable restriction of the usual definition is that complete confidence in the "if ... then" of item a) is assumed. The authors, studying the problem of fuzzy production-rule induction (Michalski, 1972, 1977), found it necessary to deal with degrees of confidence associated with the implication itself, as neatly allowed for in MYCIN-style production-rule systems (Shortliffe, 1976). Thus, as not all old x's have high blood pressure, one would like to write:

"If x is old then (0,4) x has high blood pressure",

meaning, roughly, that one has 40% of confidence in the conclusion, if the premise is true; (the value 0.4 was supplied by an M.D.).

The question is, then, knowing a distribution of possibilities for Zadeh's age, and the given fuzzy and uncertain implication above, to find a distribution of possibilities for Zadeh's blood pressure, say given as combination of systolic and diastolic pressures, in some convenient scale.

All the following discussion is valid for normal premises P, that is, for those P's in which at least one element has possibility 1; for these cases, it can be easily shown that $P \circ (\overline{P}' \oplus \overline{Q}) \geqslant Q$.

PROPOSED CHANGES

The undesirable fact that $P \circ (\overline{P}' \oplus \overline{Q}) \neq Q$ stems partially from a "loss of information" brought about by the use of the bounded sum $\oplus$. In order to introduce also a degree of confidence in the implication, $K \in [0,1]$, the natural way would be to change the operator $\oplus$ to a ternary operator $F(K,x_{ij},y_{ij})$ as follows:

a) "If x is P then (K) y is Q", or "(x,y) is $F(K,x_{ij},y_{ij})$", where $x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$ and $K \in [0,1]$;

b) "x is R", therefore "y is $R \circ F(K,x_{ij},y_{ij})$",

where it is assumed that $\overline{P}'$ and $\overline{Q}$ are given by two-dimensional arrays.

While the choice of F may be somewhat arbitrary (more comments on this shortly), the F chosen should satisfy some expected properties:

1) F should be a true extension of a two-valued logic operator. That is, $F(1,0,0) = 0$; $F(1,0,1) = 1$; $F(1,1,0) = 1$; $F(1,1,1) = 1$.

2) When $K = 1$, with maximum confidence in the implication, F should do better than the bounded sum $\oplus$, in terms of making the conclusion of the "modus ponens" closer to Q when $R = P$, for normal P's. Specifically, if $K = 1$, F should satisfy: $\forall\, x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$, $y_{ij} \leq F(1,x_{ij},y_{ij}) \leq x_{ij} \oplus y_{ij}$; thus, with $R = P$, it can be shown that $Q \leq P \circ F(1,x_{ij},y_{ij}) \leq P \circ (\overline{P}' \oplus \overline{Q})$.

3) When $K = 0$, with no confidence in the implication, nothing should be obtained about the possibilities in the conclusion; that is, all values of Q should be fully possible. This can be done with $\forall\, x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$, $F(0,x_{ij},y_{ij}) = 1$.

4) By the same token, as K increases, the possibilities should decrease, approaching Q, with R = P, when K approaches 1. That is, if $K_m \geqslant K_n$, then $\forall\, x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$, $R \circ F(K_n, x_{ij}, y_{ij}) \geqslant$ $\geqslant R \circ F(K_m, x_{ij}, y_{ij})$, and, with R = P, $P \circ F(K_n, x_{ij}, y_{ij}) \geqslant$ $\geqslant P \circ F(K_m, x_{ij}, y_{ij})$. By property 2), $P \circ F(K_n, x_{ij}, y_{ij}) \geqslant$ $\geqslant P \circ F(K_m, x_{ij}, y_{ij}) \geqslant P \circ F(1, x_{ij}, y_{ij}) \geqslant Q$.

The properties 1) through 4), though clearly desirable, do not define uniquely the form of F. There are infinitely many functions, in fact, satisfying the properties. From the point of view of performance in the "modus ponens", it can be said that "F is better than F'" if, for R = P, $\forall\, x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$, $Q \leqslant P \circ F(1, x_{ij}, y_{ij}) \leqslant P \circ F'(1, x_{ij}, y_{ij})$ and, for some x_{ij}, y_{ij}, $P \circ F(1, x_{ij}, y_{ij}) < P \circ F'(1, x_{ij}, y_{ij})$. The authors have not found yet a mathematically elegant fifth property that would yield a single "best" F.

Theorem 1: The function

$$F(K,x,y) = \begin{cases} y^K & \text{if} \quad x \leqslant y \\ x^{(1+x-y)K} & \text{if} \quad x > y \\ 1 & \text{if} \quad K = x = y = 0 \end{cases}$$

is continuous in the intervals desired and satisfies properties 1) through 4).

Proof: Continuity along the diagonal x = y is easily checked through limits; elsewhere, the exponentials are clearly continuous. The properties desired can be checked directly, using proof by cases.

1) For K = 1, $F(1,0,0) = 0^1 = 0$; $F(1,0,1) = 1^1 = 1$; $f(1,1,0) = 1^2 = 1$ and $F(1,1,1) = 1^1 = 1$.

2) Both cases must be verified:

i) if $x_{ij} \leq y_{ij}$ then $y_{ij} \leq F(1,x_{ij},y_{ij}) = y_{ij} \leq x_{ij} \oplus y_{ij}$

ii) if $x_{ij} > y_{ij}$, $1+x_{ij}-y_{ij} > 1$ and $x_{ij}^{1+x_{ij}-y_{ij}} \leq x_{ij} \leq x_{ij} \oplus y_{ij}$ and hence $F(1,x_{ij},y_{ij}) \leq x_{ij} \oplus y_{ij}$; on the other hand, it can be seen that $\frac{\partial}{\partial x} F(1,x_{ij},y_{ij}) > 0$ and, therefore, F grows with x_{ij}; since, in the lower limit, for $x_{ij} = y_{ij}$, F is equal to y_{ij}, then $\forall\, x_{ij} > y_{ij}$, $F(1,x_{ij},y_{ij}) \geq y_{ij}$.

Therefore, $\forall\, x_{ij} \in \overline{P}'$, $y_{ij} \in \overline{Q}$, $y_{ij} \leq F(1,x_{ij},y_{ij}) \leq x_{ij} \oplus y_{ij}$.

3) For K = 0, if $x_{ij} = y_{ij} = 0$ then $F(0,x_{ij},y_{ij}) = 1$ by definition. If not, then $y_{ij}^0 = 1$ or $x_{ij}^0 = 1$ and $\forall\, x_{ij},y_{ij}$, $F(0,x_{ij},y_{ij}) = 1$.

4) It is enough to check that F decreases as K increases. In fact, the form of F is Z^K, where $Z \in [0,1]$, a non-decreasing function of K. Therefore, for $K_m \geq K_n$, $F(K_n,x_{ij},y_{ij}) \geq F(K_m,x_{ij},y_{ij})$ and, for P normal, it can be easily shown that $P \circ F(K_n,x_{ij},y_{ij}) \geq$ $\geq P \circ F(K_m,x_{ij},y_{ij})$ ∎

CHANGING CONFIDENCE LEVELS OF PREMISE AND CONCLUSION

The association of a degree of confidence to a fuzzy "if... then" rule seems to be a natural step. The effect of K, through a suitable operator F, is to alter the membership function of the fuzzy set that represents the conclusion. This can be made to be, as it will be seen shortly, tantamount to associating a degree of confidence to the conclusion fuzzy set, assuming that the degree of confidence in the premise is 1.

Motivated by work on fuzzy production rule induction, the authors found necessary to deal with modifications in the possibility values of a distribution, caused by the existence of a degree of certainty associated to it. For instance, given a distribution of possibilities over an age scale describing the linguistic assertion "x is old", it might be of interest to alter it to translate "the degree of certainty in the old age of x is K", which is quite different than, say, "x is middle-aged".

The modification considered had to be consistent with that caused by the operator F in the uncertain fuzzy "if ... then" rule, and to act independently upon each possibility value of a given distribution P. The consistency desired can be expressed by considering a modifying operator $g: [0,1] \times Q \to [0,1]$ (where Q stands for its set of possibility values) such that, for any $K \in [0,1]$ and any two distributions P and Q, $g(K,F(1,x_{ij},y_{ij})) = F(K,x_{ij},y_{ij})$.

Besides the consistency property, g should, as F, satisfy the following properties:

1) With K = 1, g should not alter the distribution. That is, $\forall\, y_i \in Q$, $g(1,y_i) = y_i$.
2) With no confidence in the distribution, all values of the modified distribution should be fully possible. That is, $\forall\, y_i \in Q$, $g(0,y_i) = 1$.
3) As the value of K increases, the modified possibilities should decrease, approaching continuously the original values as K approaches 1. That is, if $K_m \geq K_n$, $\forall\, y_i \in Q$, $g(K_m,y_i) \leq g(K_n,y_i)$.

The consistency with F points to the adoption of an exponential form for g, even though there may be infinitely many functions satisfying the properties 1) through 3) above.

<u>Theorem 2</u>: The function

$$g(K,y_i) = y_i^K \quad , \quad y_i, K \in [0,1]$$

is consistent with the function F(K,x,y) of Theorem 1, is continuous and satisfies properties 1) through 3) above.

Proof: Firstly, it must be seen whether $F(K,x,y) = g(K,F(1,x,y))$. This is clearly the case, since

$$F(1,x,y) = \begin{cases} y & \text{if} \quad x \leq y \\ x^{(1+x-y)} & \text{if} \quad x > y \end{cases}$$

and

$$g(K,F(1,x,y)) = \begin{cases} y^K & \text{if} \quad x \leq y \\ x^{(1+x-y)K} & \text{if} \quad x > y \\ 1 & \text{if} \quad K = 0 \; , \end{cases}$$

as desired.

1) With $K = 1$, $g(K,y_i) = y_i$, and the distribution is not altered.

2) With $K = 0$, $f(0,y_i) = y_i^0 = 1$ (if $y = 0$, postulate $0^0 = 1$).

3) If $K_m \geq K_n$, $y_i^{K_m} \leq y_i^{K_n}$, since $y_i \in [0,1]$ and $g(K_m,y_i) \leq g(K_n,y_i)$ ∎.

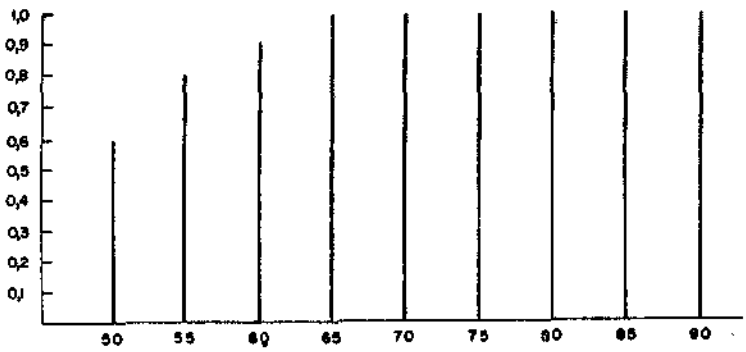
As an example, let P represent the assertion "x is old". The corresponding fuzzy set can be given by: $P = 0.6/50 + 0.8/55 + 0.9/60 + 1/65 + 1/70 + 1/75 + 1/80 + 1/85$. The conclusion "x has high blood pressure" can be given by: $Q = 0.2/(14/7) + 0.3/(15/9) + 0.7/(15/10) + 0.8/(16/10) + 1/(20/11) + 1/(20/12) + 1/(22/13) + 1/(22/14)$ (supplied by an M.D.). The fuzzy and uncertain implication:

"If x is old then (0.4) x has high blood pressure",

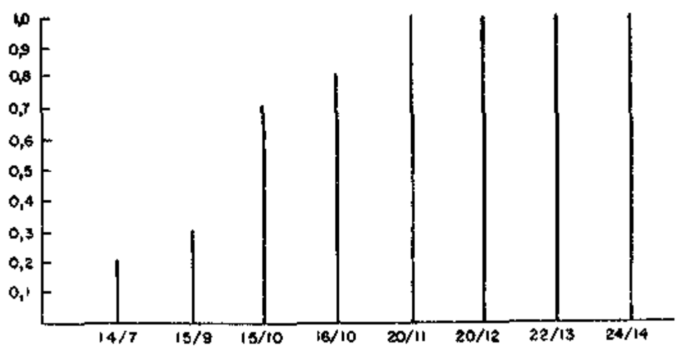
because of the consistency between F and g, can be replaced by:

"If x is old then x's blood pressure is $Q^{0.4}$"

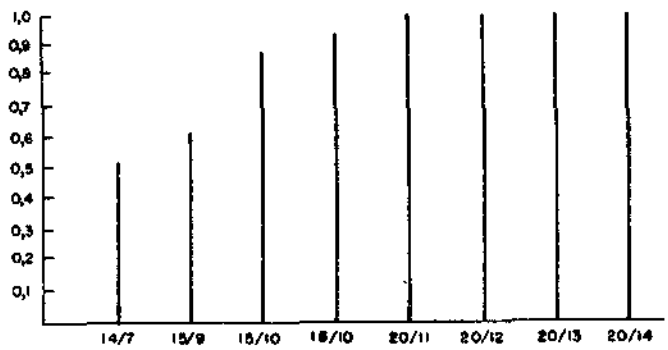
with full confidence in the implication, for use in the "modus ponens", where $Q^{0.4} = g(0.4, Q) \cong 0.53/(14/7) + 0.62/(15/9) + 0.87/(15/10) + 0.91/(16/10) + 1/(20/11) + 1/(20/12) + 1/(22/13) + 1/(22/14)$, where the values were rounded off. Figure 1 shows the three distributions plotted.



P, "x is old"



Q, "x has high blood pressure"



$Q^{0.4}$

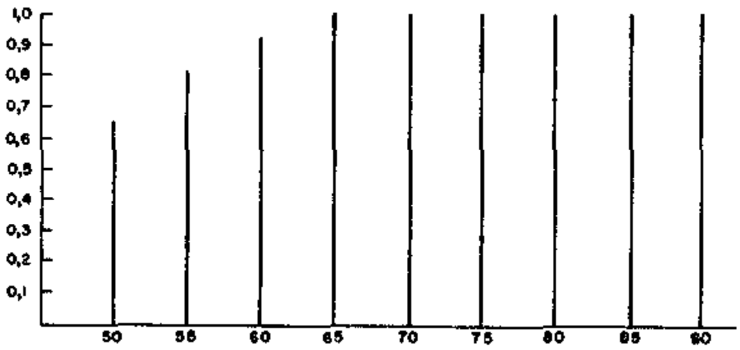
Fig. 1 - Distributions for age and blood pressure.

Associating a degree of certainty to the premises brings in a new set of problems. Systems of the MYCIN type usually multiply the certainties of premise and implication to arrive at a value associated with the conclusion. Though practical, the authors feel that the approach is too naive to be applied to fuzzy rules.

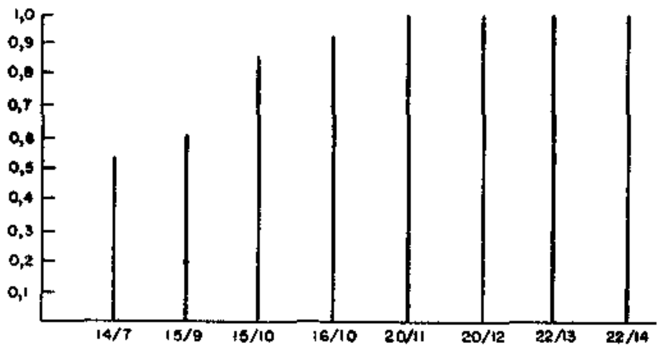
If the base rule of a desired deduction has a premise with certainty 1, the modifying function g can be used in the application of the "modus ponens" as follows: let us suppose that the certainty in the assertion "Zadeh is old" is 0.85 — this number was arrived at by consensus and does not necessarily reflect his actual age — and g is applied to the distribution for "x is old".

Letting $R = P^{0.85} \cong 0.65/50 + 0.83/55 + 0.91/60 + 1/65 + 1/70 + 1/75 + 1/80 + 1/85$, and using either form of the base rule, the question posed in the introduction may be answered, with the use of the proposed operator F:

$$\text{Distribution of Zadeh's blood pressure}_1 \cong 0.53/(14/7) + 0.62/(15/9) + 0.87/(15/10) + 0.91/(16/10) + 1.0/(20/11) + 1.0/(20/12) + 1.0/(22/13) + 1.0/(22/14).$$



"Zadeh's age"



"Zadeh's blood pressure"

Fig. 2 - 1st. Possibility distributions for Zadeh's age and blood pressure.

It is important to note that the distribution obtained in the above manner does not necessarily correspond to a direct modification of Q by the use of g; in other words, it does not correspond to the association of a degree of certainty to the conclusion. A first, simple idea, would be to find the largest c(Q) such that $g(c(Q),Q) = Q^{c(Q)}$ would still cover the distribution given by the "modus ponens". In the above case, such c(Q) is found to be nearly equal to .39 and a second distribution for Zadeh's blood pressure may be found (rounded off):

$$DZBP_2 \cong 0.53/(14/7) + 0.63/(15/9) + 0.87/(15/10) + 0.92/(16/10) + \\ + 1/(20/11) + 1/(20/12) + 1/(22/13) + 1/(22/14)$$

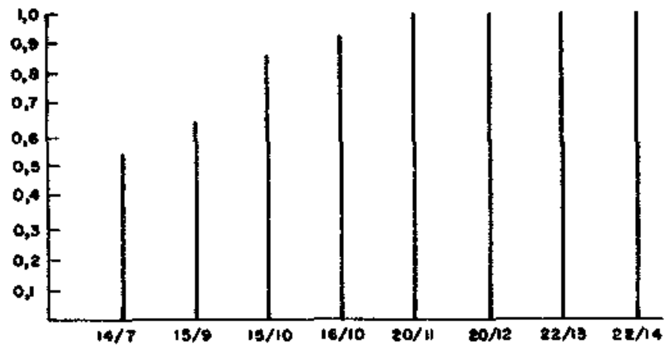


Fig. 3 - 2nd. Possibility distribution for Zadeh's blood pressure.

At present the authors are still studying a third approach to the problem, that would yield a degree of certainty for Q, using the operators g and F directly. The idea is to make the operators "commute", in the sense that applying g first to the premise and then the "modus ponens" would give the same result as applying the "modus ponens" first and then g to the conclusion.

CONCLUSIONS AND COMMENTS

The fuzzy "modus ponens" rule, as proposed in the literature, has the counterintuitive property of not reproducing the conclusion when the premise is asserted. In this paper, a ternary operator has been proposed to replace the bounded-sum operator $\oplus$, also allowing for a degree of confidence associated with the implication. Another consistent binary operator has been proposed, permitting the association of degrees of confidence to premises and conclusion.

The form of the operators proposed is not unique. The authors feel that an yet more elegant form for the operator F may be found.

The problem of transmission of degrees of confidence from premises to conclusion cannot be considered entirely solved, it being desirable that the application of the operators and of the "modus ponens" rule be commutative.

Several other examples of application of the proposed operators have been treated by the authors, with satisfactory results. It remains still to be seen whether Zadeh's real blood pressure (before reading this paper) would fall among the calculated possible values.

REFERENCES

MICHALSKI, R.S. *A variable-valued logic systems as applied to picture description and recognition.* Urbana, University of Illinois, 1972.

———. Variable-valued logic and its applications to pattern recognition and machines learning. In: RINE, D.C., ed. *Computer science and multiple valued logic.* Amsterdam, North-Holland, c1977. Cap. 18, pp. 507-534.

SHORTLIFFE, E.H. *Computer-based medical consultations: MYCIN.* New York, American Elsevier, 1976.

ZADEH, L.A. The concept of a linguistic variable and its application to approximate reasoning III. *Information Science,* 9(1): 43-80, 1975c.

ZADEH, L.A.; KING, S.; TANAKA, S. *Fuzzy sets and their applications to cognitive and decision processes.* New York, Academic, 1975b.