Tatra Mt. Math. Publ. **21** (2001), 1–19

TATRA MOUNTAINS
Mathematical Publications

GENERALIZATION OF THE RULE INTERPOLATION METHOD RESULTING ALWAYS IN ACCEPTABLE CONCLUSION

Domonkos Tikk — Péter Baranyi —
Gedeon D. Tamás — Muresan Leila

ABSTRACT. This paper deals with the expansion of the MACI (Modified Alpha Cut based Interpolation) method. The original MACI method (see [P. Baranyi, D. Tikk, Y. Yam, L. T. Kóczy, and L. Nádai: *A new method for avoiding abnormal conclusion for α-cut based rule interpolation*, in Proc. of the 8th IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'97), volume 1, pages 383–388, Seoul, Rep. of Korea, 1999], [D. Tikk, P. Baranyi: *Comprehensive analysis of a new fuzzy rule interpolation method*, IEEE Trans. on Fuzzy Sets, **8** (2000), 281–296], [Y. Yam, P. Baranyi, D. Tikk, and L. T. Kóczy: *Eliminating the abnormality problem of α-cut based interpolation*, in Proc. of the 8th IFSA World Congress, volume 2, pages 762–766, Taipei, Taiwan, 1999]) was derived from the original alpha cut based interpolation method, usually termed as KH interpolation (see, e.g., [L. T. Kóczy, K. Hirota: *Approximate reasoning by linear rule interpolation and general approximation*, Internat. J. Approx. Reason. **9** (1993), 197–225], [L. T. Kóczy, K. Hirota: *Size reduction by interpolation in fuzzy rule bases*, IEEE Trans. on SMC, **27** (1997), 14–25]), to eliminate the abnormality of the KH interpolation method, however, maintaining its advantageous computational behaviour. The MACI (and the KH) method work over convex and normal fuzzy (CNF) sets. In certain systems, e.g., in hierarchical reasoning schemes, these properties cannot always be ensured. Hence, in this paper we propose the modification of the MACI algorithm to expand its applicability for such systems where nonconvex and subnormal fuzzy sets can occur. The skeleton of the algorithms can be used also for other fuzzy reasoning techniques.

1. Introduction

The classical inference methods in fuzzy control (Zadeh's CRI, Mamda-

2000 Mathematics Subject Classification: 94D05, 11Y16.
Key words: fuzzy rule interpolation, subnormality, nonconvexity.

This research was supported by the Australian Research Council, by the National Science Research Fund (OTKA) Grant No. D/034614, and the Hungarian Ministry of Education (MKM) FKFP 0422/1997

ni and Sugeno) deal with dense rule bases. If the universe of discourse is not completely covered by the rule antecedents, it can happen that for an observation no rule is specified. Fuzzy rule interpolation, proposed first by Kóczy and Hirota [5, 6], is an inference technique for fuzzy rule bases, whenever the antecedents do not cover the whole input universe, i.e., for so-called sparse rule bases.

The KH method has two main drawbacks: for certain locations of the input sets the calculated conclusion is abnormal (see Fig. 1), and the shape of the input sets is restricted to CNF sets. After the publication of the KH method, several authors proposed conceptually different fuzzy rule interpolation algorithms, but none of them preserved the advantageous computational behaviour of the original approach (see, e.g., [1, 7, 8]).

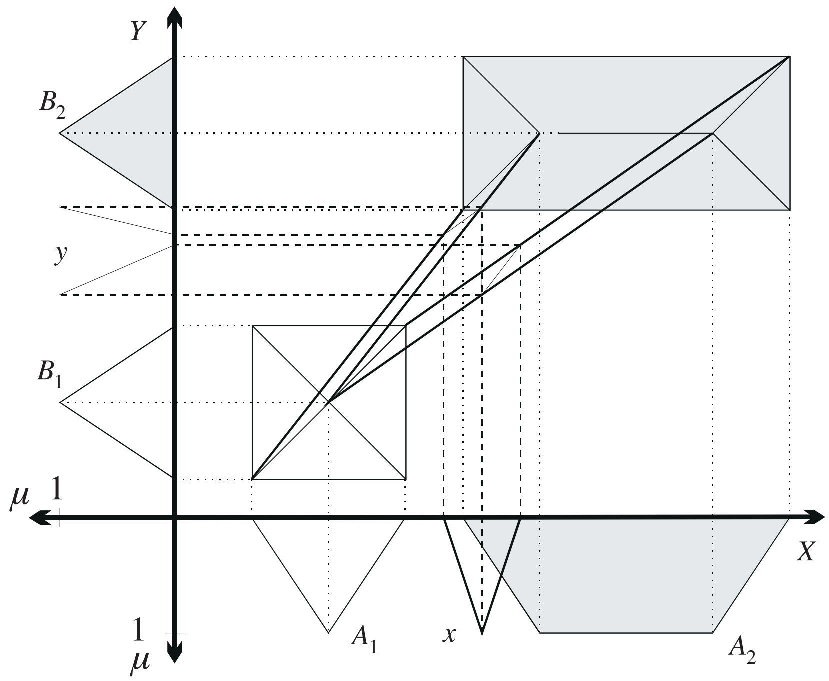

FIGURE 1. Abnormal conclusion generated by the KH method.

Recently, a new interpolation technique was proposed [2, 10, 13], which, on the one hand, eliminates the abnormality of the conclusion, while, on the other hand, maintains the favourable low complexity of the KH method. This approach works with vector description of the fuzzy sets involved [14], and termed the MACI (Modified Alpha Cut based Interpolation) method, which is a Vector Representation Based Interpolation (VRBI) approach. The main idea of this method is recalled in Section 2. Briefly, it is based on the transformation of the vectors representing the consequent fuzzy sets to another coordinate system, where the abnormality of the determined conclusion can be excluded. Finally, the "well-behaved" conclusion is described in the original coordinate system.

The aim of this paper is to increase further the applicability of the aforementioned MACI method. Therefore procedures are proposed to allow the new method to be able to handle subnormal and nonconvex fuzzy sets, as well.

The treatment of subnormal fuzzy sets is solved with a normalization procedure of the involved fuzzy sets before the calculation of the conclusion. Hence, the conclusion can be determined by means of the normalized fuzzy sets. Finally, a denormalization procedure is applied depending on the heights of the fuzzy set at hand. These results are presented in Section 3.

Section 4 investigates the case of nonconvex inputs. Nonconvex fuzzy sets can be considered as the union of convex fuzzy sets. This property is exploited in order to solve the problem of non-convexity. Each peak of a fuzzy set is treated as a local reference point and the connecting flanks can be split into a monotone decreasing and a monotone increasing part. The VRBI method is applied to the sequence of these increasing and decreasing parts. Sufficient conditions for the connectedness of the aggregated conclusion is also presented.

The question of the required computational effort is also addressed. The complexity of the algorithms for nonconvex sets is relatively higher than for CNF sets, but this is an acceptable price to pay for the expanded applicability.

2. The MACI method

The MACI method modifies the KH interpolation in such a manner that the produced conclusion is always a fuzzy set.

The vector description of α-cut based interpolation was used [14] to solve the problem of abnormal conclusions (see Fig. 1). Suppose that A is a piecewise linear CNF set described by characteristic points a_i $(i = -m, \dots, 0, \dots, n)$. (We remark that in [9] the vector description for arbitrary continuous CNF sets is presented, but here, for brevity, we restrict ourselves to the piecewise linear case.) Then fuzzy set A is represented by a vector $\underline{a} = [a_{-m}, \dots, a_0, \dots, a_n]$ where a_k $(k \in [-m, n])$ are the characteristic points of A and a_0 is the reference point of A with membership degree one. From this, $\underline{a}_L = [a_{-m}, \dots, a_0]$, and $\underline{a}_U = [a_0, \dots, a_n]$ represent the left flank and right flank of A, respectively. If A is a CNF then, e.g., for the right flank, $a_i \leq a_j$, $i < j \in [0, n]$ should hold with monotone decreasing α levels. When it is not ambiguous we omit the subscripts L and U.

Let us suppose further that two rules are given in form "If x is A_i then y is B_i" $(i = 1, 2)$, and that the location of the observation and the neighbouring fuzzy sets fulfils conditions:

$$A_1 \prec A^* \prec A_2 \qquad \text{and} \qquad B_1 \prec B_2 \,. \tag{1}$$

with some apparent ordering $\prec$ (see also Fig. 2).

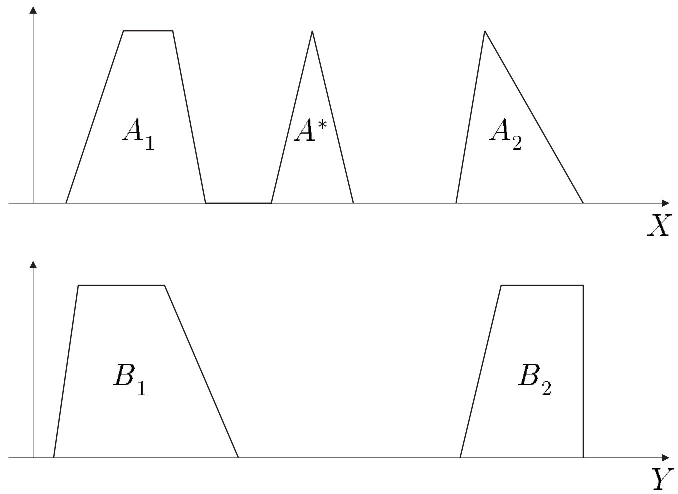

FIGURE 2. Location of involved fuzzy sets.

The conclusion B^* should fulfil the condition

$$b_i^* \le b_j^* \qquad \forall\, i \le j \in [-m, n] \tag{2}$$

for avoiding abnormality. The Condition 2 is ensured by a proper transformation of the output space. The MACI method transforms the points representing the consequents into a new coordinate system, in which (2) can be guaranteed easily, then the conclusion is calculated, and finally the conclusion is transformed back into the original coordinate system. The coordinates of the conclusion's right flank can be obtained as [9, 10]

$$b_k^* = {}^{\mathrm{KH}}b_k^* + \sum_{i=0}^{k-1} (\lambda_i - \lambda_{i+1})(b_{2i} - b_{1i})\,, \tag{3}$$

$k \in [0, n]$, where

$$\lambda_k = \frac{a_k^* - a_{1k}}{a_{2k} - a_{1k}} \tag{4}$$

and

$${}^{\mathrm{KH}}b_k^* = (1 - \lambda_k) b_{1k} + \lambda_k b_{2k} \tag{5}$$

is the value of the kth coordinate calculated according to the original KH approach. Analogously, for the left flank

$$b_k^* = {}^{\mathrm{KH}}b_k^* + \sum_{i=k+1}^{0} (\lambda_i - \lambda_{i-1})(b_{2i} - b_{1i}) \tag{6}$$

holds for $k \in [-m, 0]$.

The characteristic points of the conclusion generated by the described MACI method monotone increase (see, e.g., [9, 10]). This is illustrated in the example of Fig. 1, where the normal conclusion generated by the MACI method is depicted in Fig. 3.

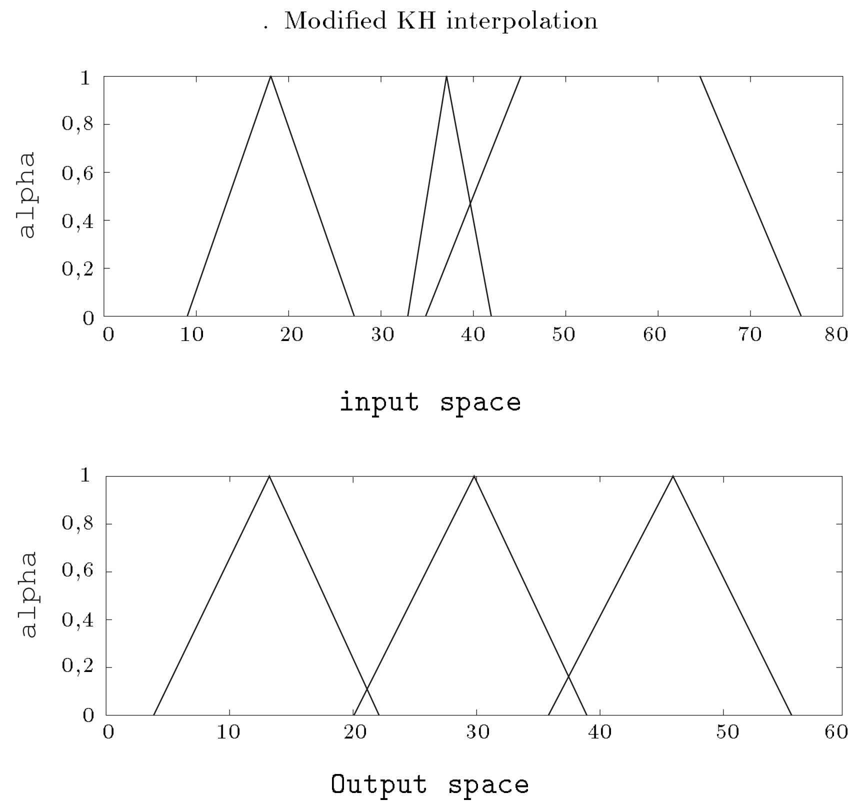

FIGURE 3. The conclusion obtained by the MACI method for the situation of Fig. 1. (This figure is generated by MATLAB).

The conclusion calculated by the MACI algorithm can be described based on the original KH method as follows. For the reference point the two algorithms give the same output. The first coordinates, in each direction, are determined by the distances between corresponding coordinates and reference points of the flanking consequents. The following coordinates can be determined analogously using the distance between the previous and the current coordinates of the consequents. For further details see Fig. 4.

One possible benchmark of a rule interpolation algorithm is whether it satisfies the Modus Ponens, a tautology from symbolic logic. This simple logical scheme works when an implication and its antecedent part are known to be

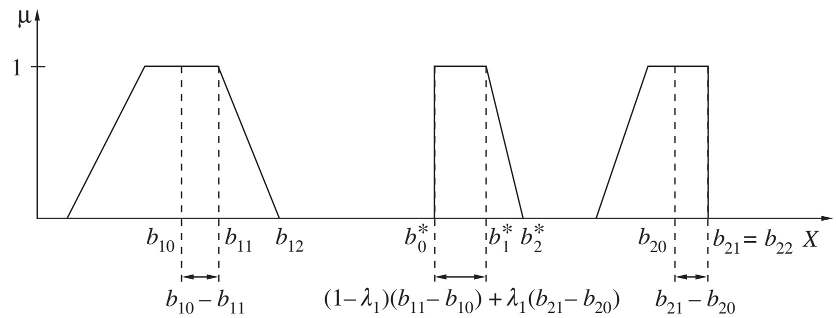

FIGURE 4. The construction of the conclusion by the MACI algorithm.

true; then, its consequent part is concluded to be also true:

$$\frac{\begin{array}{l} A \to B \\ A \end{array}}{B} .$$

In fuzzy context this tautology is usually modified as

$$\frac{\begin{array}{l} A \to B \\ A' \end{array}}{B'} ,$$

where the pairs A, A' and B, B' are similar with respect to an appropriate similarity function. This generalization is in accordance with the gradual semantic interpretation of fuzzy rules [4]. The idea of gradual rules supposes that the IF ... THEN rules implicitly contain the semantics that the terms in the rules might be satisfied in different degrees, and the more the antecedent A is true in the case of the observation, the more should also the corresponding consequent B be satisfied in the calculated conclusion. (For further references see, e.g., [3, 4, 11].) As an extreme situation (which coincides with the classical Modus Ponens of predicate calculus), if the observation is identical with a rule antecedent, this concept requires the conclusion to be identical with the corresponding rule consequent.

The MACI method satisfies the Modus Ponens. To see this, we first remark that in this case each λ_k $\big(k \in [-m, n]\big)$ in (4) is 0 or 1 depending on which antecedent the observation coincides with. Without loss of generality, let us suppose that $A^* = A_1$, hence, $\lambda_k = 0$ $(k \in [-m, n])$. Thus the sums in expressions (3) and (6) vanish and $b_k^* = {}^{\mathrm{KH}}b_k^* = b_{1k}$ for all $k \in [-m, n]$ by means of equation (5), i.e., the conclusion is identical with consequent B_1. It is also worth noting how one should assign the characteristic values to a certain set of fuzzy sets, whose breakpoint sets are different. Since in (3) the coefficients λ_k correspond to the input variable, namely to the kth coordinate of the antecedents and the observation, but the calculated value corresponds to the kth coordinate of the conclusion, in order to avoid confusion a common breakpoint level set should be determined for both spaces, which is the union of (perhaps different) breakpoint level sets for each variable. Such a situation is depicted in Fig. 5.

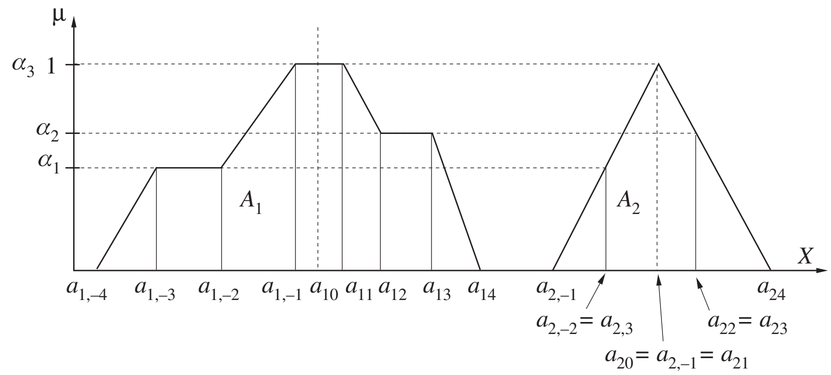

FIGURE 5. The determination of the characteristic points under different break point sets.

Finally, we remark that multivariable antecedents can be handled analogously as the transformation described in this section affects only the consequent part. Common combined antecedent sets (and the observation) can be calculated from the corresponding antecedents (observation) of each variable using Minkowski-type distance, where the weights are identically one ($w = 1$), in order to preserve the linearity of the method. Hence, for example, the coordinates of the common combined observation can be calculated as

$$a_k^* = \frac{1}{r} \sum_{i=1}^{r} a_{ik}^* , \tag{7}$$

where r is the number of variables, and a_{ij}^* is the jth coordinate of the observation of the ith variable.

In the next two sections we generalize the MACI method for arbitrary shaped fuzzy sets.

3. Algorithms for handling subnormal sets

The problem of subnormality can be treated in the vector representation based fuzzy interpolation algorithm at different levels. These levels depend on the number of the subnormal set to cope with in the current situation, e.g., whether just the observation or also the components of the rules (antecedents, consequents) are subnormal. We will distinguish the following cases

1. The only subnormal set is the observation.
2. Arbitrary fuzzy sets can be subnormal, i.e., the observation, as well as the antecedents and the consequents. We divide this case into three subcases:
   (a) Only the observation's subnormality is not taken into consideration, the subnormality of the rules is disregarded;
   (b) The consequents' and observation's subnormality is taken into account, the subnormality of the antecedents is disregarded;
   (c) All subnormality is taken into consideration.

The reason to distinguish these cases comes from practice. In the case of hierarchical reasoning algorithms the calculated conclusion of the first reasoning level is used as the observation of the next level. Most of the classical fuzzy control algorithms can result in subnormal fuzzy sets (e.g., Zadeh's CRI, Mamdani, Larsen method). So then, assuming the rule base to be normal, only the observation is subnormal. This situation corresponds to Case 1.

In the case where rules are generated from an input-output data sample set, the normality of the terms cannot always be guaranteed. Then, Case 2 applies.

3.1 Case of subnormal observation.

The skeleton of the algorithm, being suitable for Case 1 and with apparent modification to Case 2, is the following:

Algorithm 1 (**for Case 1, subnormal observation**)

STEP 1 Normalize A^* with a normalization function N with normalization factor β, in the simplest case just multiply A^* by β as (see Fig. 6):

$$(A^*)' = N(A^*) = \beta \cdot A^*, \qquad \text{where} \quad \beta = \frac{1}{h(A^*)}. \tag{8}$$

$h(A)$ denoting the height of the fuzzy set A: $h(A) = \sup\{A(x) \colon x \in X\}$.

STEP 2 Determine the reference and characteristic points of the input and output fuzzy sets.

STEP 3 Calculate the normalized conclusion $B^{*\prime}$ according to the original MACI algorithm.

STEP 4 Denormalize the conclusion $B^{*\prime}$ by the inverse normalization operator N^{-1}, in the simplest case by the reciprocal of the normalization factor β as:

$$B^* = N^{-1}(B^{*\prime}) = \frac{1}{\beta} \cdot B^{*\prime} = h(A^*) \cdot B^{*\prime}. \tag{9}$$

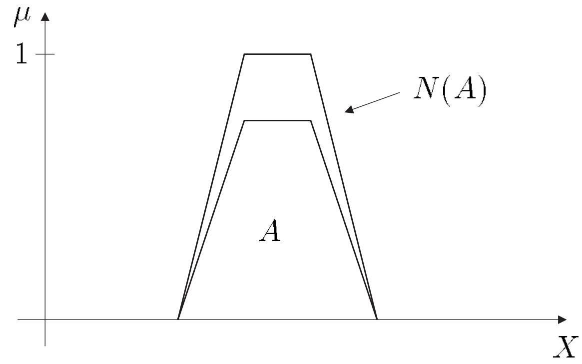

FIGURE 6. Normalization of a subnormal fuzzy set according to (8).

Let us make a short detour to analyse normalization Step 1 of the algorithm. It is clear that the algebraic product operation in the normalization and denormalization steps (8) and (9) can be substitutet by any reasonable invertable operation $N : \mathcal{F}(X) \to \mathcal{F}(X)$ producing normal fuzzy set.

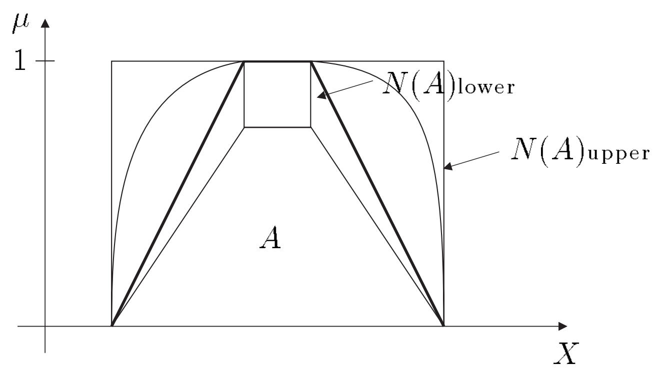

FIGURE 7. Lower and upper bounds for the normalized fuzzy set. The thickened lines shows the algebraic normalization, the curve depicts an arbitrary normalization between the upper and the lower bounds.

The possible lower and upper bounds of the normalized fuzzy set $N(A)$ of the subnormal fuzzy set A are depicted if Fig. 7. We require the normalization operator $N(A)$ to fulfil the following boundary conditions:

$$N\big(A(x)\big) = 1\,, \quad \text{if} \quad A(x) = h(A)\,, \tag{10}$$

$$N\big(A(x)\big) = 0\,, \quad \text{if} \quad A(x) = 0\,. \tag{11}$$

When the membership value of A is within the interval $\big(0, h(A)\big)$ we can choose an arbitrary function ensuring the monotonicity condition:

$$N\big(A(x_1)\big) \le N\big(A(x_2)\big) \quad \text{if} \quad A(x_1) \le A(x_2) \in \big(0, h(A)\big)\,. \tag{12}$$

(Observe that these conditions do not require $h(A)$ to be less than 1, i.e., in general the function N satisfying conditions (10)–(12) performs a fuzzy-to-fuzzy transformation by "augmenting" or "thickening" the membership function.)

Hence, we can have for example the following classes of normalization functions:

1. Convex combination:

$$N_\lambda\big(A(x)\big) = \begin{cases} 1 & \text{if } A(x) = h(A)\,, \\ A(x)(1-\lambda)+\lambda\,,\ \lambda \in [0,1] & \text{if } A(x) \in \big(1, h(A)\big)\,, \\ 0 & \text{if } A(x) = 0\,. \end{cases}$$

This class includes the two extreme cases (see Fig. 7) when $\lambda = 1$ and $\lambda = 0$.

2. Truncated factorization:

$$N_\kappa\big(A(x)\big) = \begin{cases} 1 & \text{if } A(x) = h(A)\,, \\ \min\big(A(x)\big(\frac{1}{h(A)}\big)^\kappa, 1\big),\ \kappa \in [0,\infty] & \text{if } A(x) \in \big(1, h(A)\big)\,, \\ 0 & \text{if } A(x) = 0\,. \end{cases}$$

This class produces $N(A)_{\text{lower}}$ and $N(A)_{\text{upper}}$ for $\kappa = 0$ and $\kappa = \infty$, respectively. When $\kappa = 1$, $N_\kappa(A)$ yields the algebraic product normalization.

However, in order to preserve the low computational cost of the original MACI algorithm, it is reasonable to use the most simple algebraic product.

3.2 Case of subnormal rules.

To cope with Case 2, as a first possibility, one can ignore the rules' subnormality and just follow the scheme of Algorithm 1. This solution may be justified by arguments on computational effectiveness. However, if the heights of the involved fuzzy sets are different, then the calculation time lengthens, as the various heights determine different α-level, and hence, the number of characteristic points increases. This implies the restricted applicability of this simple solution, which is confined to the case where the height of the involved sets is identical.

Further, this method does not correspond to the gradual semantic interpretation of fuzzy rules [4] and Modus Ponens (see Section 2). In our case, if the observation is identical with one of the (possible subnormal) antecedents and the consequent of this rule has a different height than the antecedent, the calculated conclusion does not coincide with the consequent, simply because their heights differ.

This observation leads to the second solution (Case 2b) when subnormal consequents are taken into account in the calculation of the conclusion. We will return to this approach after having discussed Case 2c, the general solution.

In Case 2c, we use all subnormal information expressed by the rules. The algorithm of this method is the following

Algorithm 2 (for Case 2c, subnormal rules and observation)

STEP 1 Determine the normalization factors for each involved fuzzy set. We let β_i $(i = 1, \ldots, 5)$ be these factors, where, for convention, subscripts 1 and 2 belong to the antecedents, subscripts 3 and 4 belong to the consequents, and 5 to the observation. So we have:

$$\beta_i = \frac{1}{h(A_j)}\ (i,j = 1,2); \quad \beta_i = \frac{1}{h(B_j)}\ (i = 3,4;\ j = 1,2); \quad \beta_5 = \frac{1}{h(A^*)}\,.$$

STEP 2 Calculate $\overline{\beta}$, the aggregated normalization factor as:

$$\overline{\beta} = g(\beta_1, \ldots, \beta_5)\,, \tag{13}$$

where function g denotes an arbitrary aggregation operator.

STEP 3 Normalize fuzzy sets where $\beta_i > 1$ with a common normalization function N.

STEP 4 Determine the reference and characteristic points of the normalized input and output fuzzy sets.

STEP 5 Calculate the normalized conclusion $B^{*\prime}$ according to the original MACI algorithm.

STEP 6 Denormalize the conclusion $B^{*\prime}$ by the inverse of the common normalization operator N^{-1} with aggregated normalization factor $\overline{\beta}$, in the simplest case by the reciprocal of the aggregated normalization factor $\overline{\beta}$ as:

$$B^* = N^{-1}(B^{*\prime}) = \frac{1}{\overline{\beta}} \cdot B^{*\prime}. \tag{14}$$

This algorithm has two degrees of freedom: the selection of the normalization function and the aggregation function. We propose to use here the simplest normalization function, the multiplication with the height's reciprocal, being also the most computational effective possibility. But the choice of the aggregation function is not so obvious.

As a simple solution, we can take the arithmetic or geometric mean of values β_i by letting g be or

$$g_{\text{geom}}(\beta_1, \ldots, \beta_5) = \sqrt[5]{\prod_{i=1}^{5} \beta_i}\,.$$

However, these aggregation operators are not suitable to incorporate any information regarding the location of the observation in terms of the distance from the antecedents. Nevertheless, this might be a useful information which helps to fulfil the Modus Ponens. An appropriate solution of this problem is the choice of the standard weighted aggregation operator. This operator assigns normalized weights $0 \le w_j \le 1$ to the elements to be aggregated, and calculates the value

$$A_w = (x_1, \dots, x_n) = \frac{\sum_{j=1}^{n} w_j x_j}{\sum_{j=1}^{n} w_j}. \tag{15}$$

An even more sophisticated aggregation, Yager's OWA (ordered weighted averaging) aggregation operator [12], can also be applied, which is also linear and generalizes the arithmetic mean. In this paper, for simplicity, we deal only with (15).

Applying operator (15) as

$$\overline{\beta} = \frac{\sum_{i=1}^{5} w_i \beta_i}{\sum_{i=1}^{5} w_i} \tag{16}$$

we can ensure that the closer the observation is located to one antecedent, the more similar the conclusion is to the corresponding consequent. To do this, we have to define the weights assigned to the membership functions appropriately based on their mutual locations. We let d_1 and d_2 be the distance between the reference points of A_1 and A^*, and A_2 and A^*, respectively (see Fig. 8). Then let $w_1 = w_3 = 1 - \frac{d_1}{d_1+d_2}$, $w_2 = w_4 = 1 - \frac{d_2}{d_1+d_2}$, and finally $w_5 = \sqrt{w_1 \cdot w_2}$. Having these weights, if the observation is located right in between the two antecedents $d_1 = d_2$, then all weights are the same ($= 0.5$) and the aggregated value is equal to the arithmetic mean of β_i. When, e.g., $d_1 = 0$ (the observation's and one antecedent's reference points are identical), then $w_1 = w_3 = 1$, $w_2 = w_4 = w_5 = 0$. These weights ensure that the height of the conclusion is determined only by the rule which is fired in that situation.

However, the aggregation function (16) fails to solve the problem of exact fitting described earlier, if the heights of the rule's antecedents and consequents differ, because (16) does not take into consideration the particular height values. If, e.g., $\beta_1 = \beta_5 \neq \beta_3$, then $\overline{\beta} = \frac{\beta_1+\beta_3}{2} \neq \beta_3$, further, if $\beta_1 = \beta_3 \neq \beta_5$, then $\overline{\beta} = \beta_1 = \beta_3$, which implies that the conclusion can be identical with B_1 (if the shape of the involved functions is appropriate), though $A^* \neq A_1$.

This problem can be solved by modifying the function (16) as

$$\overline{\beta} = \frac{\sum_{i=3}^{5} w_i \beta_i}{\sum_{i=3}^{5} w_i} + \delta_w \cdot \sum_{i=1}^{2} w_i (\beta_i - \beta_5)(\beta_{i+2} - 1), \tag{17}$$

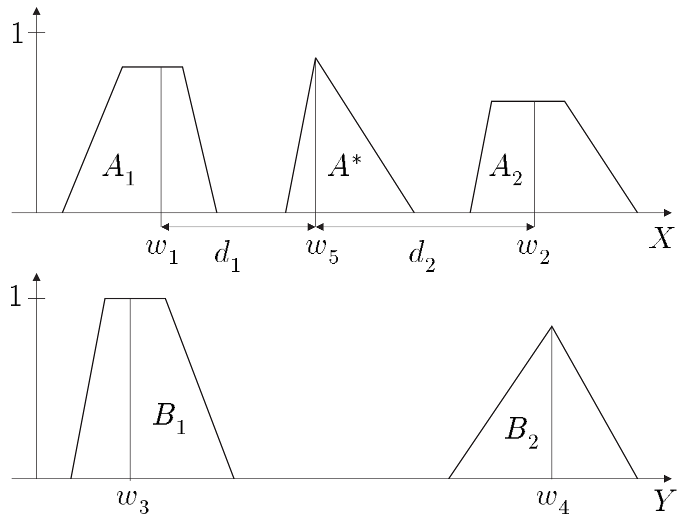

FIGURE 8. Weights are determined by the relative position of the observation with respect to the flanking antecedents.

where $\delta_w = 1$ if and only if $w_1 = 1$ or $w_2 = 1$. In non extreme cases this function determines the denormalization factor based only on the relative position of the observation, because by δ_w, the second member of (17) vanishes. When the reference points of the observation and one of the antecedents coincide (extreme case) the second tag of (17) plays the role of adjustment. If also the heights of the observation and, e.g., antecedent A_1 are identical then this tag vanishes and $\overline{\beta} = \beta_3$. If $\beta_1 \neq \beta_5$ then $\overline{\beta}$ differs from β_3 as it increases (the conclusion's height decreases) when $\beta_5 < \beta_1$, and it decreases (the conclusion's height decreases) when $\beta_5 > \beta_1$.

It is worth noting that despite this decrease $\overline{\beta} \geq 1$ always holds. Let $w_1 = w_3 = 1$ and $w_2 = w_4 = w_5 = 0$. In this case $\delta_w = 1$ and $\overline{\beta} - 1 = \beta_3 + (\beta_1 - \beta_5)(\beta_3 - 1) - 1$ must be nonnegative, where $\beta_3 \geq 1$ and $\gamma = \beta_1 - \beta_5 \in [-1, 1]$. This expression is identical with $(\beta_3 - 1)(1 + \gamma)$ which is nonnegative due to the conditions on β_3 and γ.

Observe that in the extreme cases the conclusion is identical with the corresponding consequents if and only if the observation and the antecedents are identical. The common reference point and height alone do not imply the identity.

Finally, we remark that Case 2b can be treated analogously to Case 2c, but the antecedents' subnormal heights are not taken into consideration in the appropriate steps of Algorithm 2.

4. Algorithms for handling nonconvex sets

In this section, we present the necessary modifications of the MACI algorithm to be able to cope with nonconvex input, as well. The basic algorithm to be proposed below deals with nonconvex observations. We outline the appropriate steps to generalize this result for the general nonconvex situation, where rule antecedents and consequents are also allowed to be nonconvex. Nevertheless, we note that nonconvex rules are seldom used in practice, and the required computational effort is considerably high compared to the effectiveness of their use. However, nonconvex observations often occur, i.e. when fuzzy systems are used in a hierarchical structure as described previously. This fact fortifies our investigation in the direction of this topic.

4.1 The case of nonconvex observations.

Nonconvex fuzzy sets can be considered as a union of convex fuzzy sets. This property is exploited in order to solve the problem of non-convexity. Each peak of a fuzzy set is treated as a local reference point and the connecting flanks can be split into a monotone decreasing and a monotone increasing part. A pair of these parts forms a so-called subobservation. The subobservations are convex, thus the original MACI method can be applied to them.

The basic algorithm is the following:

Algorithm 3 (for nonconvex observation)

STEP 1 Divide the nonconvex observation A^* into convex parts according to Fig. 9. Denote the convex subobservations by $A_1^*, \ldots, A_k^*$.

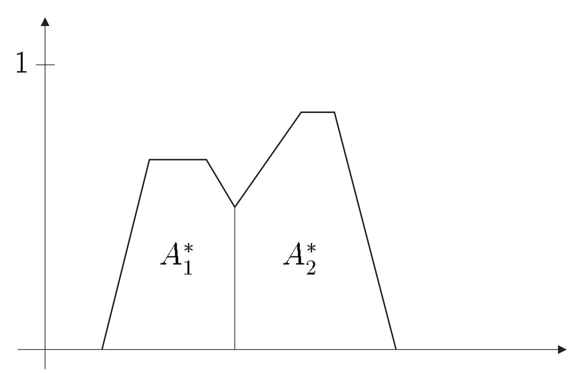

FIGURE 9. Division of a nonconvex set into convex parts.

STEP 2 For every A_i^* calculate the subconclusion B_i^* $(i = 1, \ldots, k)$ by using the algorithm for convex sets. When the subobservation A_i^* is subnormal, use the modified algorithm described in the previous Section.

STEP 3 Join the subconclusions B_i^* $(i = 1, \dots, k)$ to create the final conclusion by any reasonable t-conorm operator as:

$$B^* = \bigotimes_{i=1}^{k} B_i^*, \tag{18}$$

($\otimes$ denotes an arbitrary t-conorm).

The simplest choice for $\otimes$ is the maximum t-conorm. It is interesting to consider whether the final conclusion is connected, or which conditions ensure its connectivity.

PROPOSITION 1. *If the supports of the antecedents are identical then the final conclusion is always connected.*

The proof of Proposition 1 can be found in the Appendix.

We can generalize Algorithm 3 for nonconvex rules, but it is more reasonable to work with the convex hull of the rules than with nonconvex rules. In the latter case one has to calculate the subconclusions to be aggregated for all possible choice of nonconvex parts of the rules. If we have membership functions with only two peaks, the number of calculated subconclusions is $2^5 = 32$. This example demonstrates the reduced effectiveness of involving nonconvex rules into the calculation with the MACI algorithm.

5. Conclusion

Although these procedures increase the computation time of the MACI algorithm (particularly the second one), the resulting method is applicable to all possible control situation. Further, in the case of hierarchical fuzzy systems, it is advantageous to have this property, as, for instance, the inputs of a control level (which uses fuzzy interpolation) might be the outputs of Mamdani-like control strategies, and those outputs are often nonconvex and subnormal.

We remark that the skeletons of the algorithms presented in the paper can be used in other fuzzy reasoning methods to expand their applicability for subnormal and nonconvex sets. A trivial example for such a method is the KH interpolation.

Appendix

Proof of Proposition. Without loss of generality, we assume that the observation has only two peaks, i.e., the observation can be divided into two convex fuzzy sets, A_1^* and A_2^*. These sets fulfil the condition $a^*_{1,n_1} = a^*_{2,-m_2}$, where subscripst n_1 and $-m_2$ refer to the rightmost and to the leftmost coordinates of A_1^* and A_2^*, respectively (see also Fig. 10). Let us denote the rightmost and the leftmost coordinates of the two adjacent subconclusions by b^*_{1,n_1} and $b^*_{2,-m_2}$, defined as:

$$b^*_{n_1} = \left(1-\lambda^{(1)}_{n_1}\right)b_{1,n_1} + \lambda^{(1)}_{n_1} b_{2,n_1} + \sum_{i=0}^{n_1-1}\left(\lambda^{(1)}_i - \lambda^{(1)}_{i+1}\right)\left(b_{2i}-b_{1i}\right);$$

$$b^*_{-m_2} = \left(1-\lambda^{(2)}_{-m_2}\right)b_{1,-m_2} + \lambda^{(2)}_{-m_2} b_{2,-m_2} + \sum_{i=-m_2+1}^{0}\left(\lambda^{(2)}_i - \lambda^{(2)}_{i-1}\right)\left(b_{2,i}-b_{1,i}\right),$$

where

$$\lambda^{(j)}_k = \frac{a^*_{j,k} - a_{1,k}}{a_{2,k} - a_{1,k}} \qquad \forall k \in [-m_j, n_j],\ j \in \{1,2\}.$$

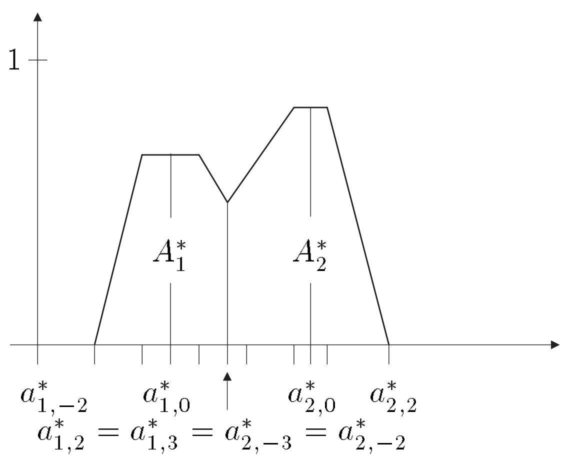

FIGURE 10. Coordinates of the two subobservations. Here $m_1 = n_2 = 2$ and $m = n = n_1 = m_2 = 3$.

We refer to the maxima and to the minima of the possible different characteristic point systems generated by the two subobservation, A_1^* and A_2^* by $n_1; n_2$ and $-m_1; -m_2$, respectively. However, as we investigate just the right slope of A_1^* and its corresponding subconclusion B_1^* and, analogously, the left slope of A_2^* and its corresponding subconclusion B_2^*, notation n_2 and $-m_1$ are unused. This allows us to simplify the notation, so henceforth we use subscript n associated with A_1^* and B_1^* and subscript m associated with A_2^* and B_2^*.

Let us suppose that $\mathrm{supp}(A_1) = \mathrm{supp}(A_2)$, i.e., $a_{1,n} - a_{1,-m} = a_{2,n} - a_{2,-m}$ (the values of the extremes are independent of the characteristic point system generated by the actual (sub)observation). We may ensure the connectivity of the final conclusion by proving that the difference $b^*_{1,n} - b^*_{2,-m}$ is positive.

$$b^*_{1,n} - b^*_{2,-m} = \left(1-\lambda^{(1)}_n\right)b_{1,n} + \lambda^{(1)}_n b_{2,n} - \left(1-\lambda^{(2)}_{-m}\right)b_{1,-m} - \lambda^{(2)}_{-m} b_{2,-m}$$
$$+\sum_{i=0}^{n-1}\left(\lambda^{(1)}_i - \lambda^{(1)}_{i+1}\right)(b_{2,i}-b_{1,i}) - \sum_{i=-m+1}^{0}\left(\lambda^{(2)}_i - \lambda^{(2)}_{i-1}\right)(b_{2,i}-b_{1,i})\,. \tag{19}$$

Let us estimate in both sums of expression (19) the difference $(b_{2i} - b_{1i})$ by $d_{\min} = b_{2,-m} - b_{1,n}$, the distance between the rightmost point of the left consequent and the leftmost point of the right consequent. Further, we let d have the value $a_{2,-m} - a_{1,-m} = a_{2,n} - a_{1,n}$, because of the identical supports (see Fig. 11).

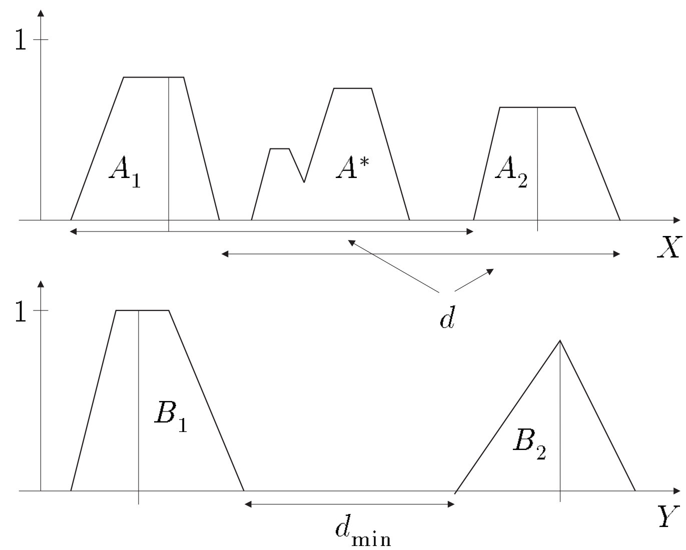

FIGURE 11. Meaning of the notations used in the proof $(\mathrm{supp}(A_1) = \mathrm{supp}(A_2))$.

Then (19) can be estimated as

$$\left(1 - \frac{a^*_{1,n} - a_{1,n}}{d}\right)b_{1,n} + \frac{a^*_{1,n} - a_{1,n}}{d}b_{2,n} + \left(\lambda^1_0 - \lambda^1_n\right)d_{\min}$$
$$-\left(1 - \frac{a^*_{1,n} - a_{1,-m}}{d}\right)b_{1,-m} + \frac{a^*_{1,n} - a_{1,-m}}{d}b_{2,-m} - \left(\lambda^2_0 - \lambda^2_{-m}\right)d_{\min}$$

using $a^*_{1,n} = a^*_{2,-m}$. With apparent transformation we obtain that

$$\left(1 - \frac{a^*_{1,n}}{d}\right)(b_{1,n} - b_{1,-m}) + \frac{a^*_{1,n}}{d}(b_{2,n} - b_{2,-m})$$
$$+\ \frac{1}{d}\left(a_{1,n}(b_{1,n} - b_{2,n}) + a_{1,-m}(b_{2,-m} - b_{1,-m})\right)$$
$$+\ \left[\frac{a^*_{1,0} - a^*_{2,0}}{a_{2,0} - a_{1,0}} + \frac{a_{1,n} - a_{1,-m}}{d}\right]d_{\min}\,.$$

The elements with factor $\frac{a_{1,n}}{d}$ sum up as

$$\frac{a_{1,n}}{d}\big(b_{1,n} - b_{2,m} + b_{2,-m} - b_{1,n}\big) = \frac{a_{1,n}}{d}(b_{2,-m} - b_{2,n}) = \frac{a_{1,n}}{d}\big(-\operatorname{supp}(B_2)\big)\,.$$

Similarly,

$$\frac{a_{1,-m}}{d}\big(b_{2,-m} - b_{1,-m} - b_{2,-m} + b_{1,n}\big) = \frac{a_{1,-m}}{d}\big(\operatorname{supp}(B_1)\big)\,.$$

Finally we get

$$\Big(1 - \frac{a^*_{1,n} - a_{1,-m}}{d}\Big)\operatorname{supp}(B_1) + \frac{a^*_{1,n} - a_{1,n}}{d}\operatorname{supp}(B_2) \geq 0$$

where the supports are nonnegative and the fractions are from interval $[0,1]$ due to the location of A_1, A^* and A_2. This completes the proof. □

REFERENCES

[1] BARANYI, P.—KÓCZY, L. T.: *A general and specialized solid cutting method for fuzzy rule interpolation*, BUSEFAL **67** (1996), 13–22.

[2] BARANYI, P.—TIKK, D.—YAM, Y.—KÓCZY, L. T.—NÁDAI, L.: *A new method for avoiding abnormal conclusion for α-cut based rule interpolation*, in: Proc. of the 8th IEEE Int. Conf. on Fuzzy Systems, (FUZZ-IEEE '97), Vol. 1, Seoul, Rep. of Korea, 1999, pp. 383–388.

[3] DING, L.—SHEN, L.—MUKAIDONO, M.: *Revision principle for approximate reasoning, based on linear revising method*, in: Proc. of the 2nd Int. Conf. on Fuzzy Logic and Neural Networks, (IIZUKA '92), Iizuka, 1992, pp. 305–308.

[4] DUBOIS, D.—PRADE, H.: *Gradual rules in approximate reasoning*, Inform. Sci. **61** (1992), 103–122.

[5] KÓCZY, L. T.—HIROTA, K.: *Approximate reasoning by linear rule interpolation and general approximation*, Internat. J. Approx. Reason. **9** (1993), 197–225.

[6] KÓCZY, L. T.—HIROTA, K.: *Size reduction by interpolation in fuzzy rule bases*, IEEE Trans. on SMC **27** (1997), 14–25.

[7] KÓCZY, L. T.—KOVÁCS, Sz.: *Shape of the fuzzy conclusion generated by linear interpolation in trapezoidal fuzzy rule bases*, in: Proc. of the 2nd European Congress on Intelligent Techniques and Soft Computing, Aachen, 1994, pp. 1666–1670.

[8] SHI, Y.—MIZUMOTO, M.: *Some considerations on Kóczy's interpolative reasoning method*, in: Proc. of the 4th IEEE Int. Conf. on Fuzzy Systems, (FUZZ-IEEE/IFES '95), Yokohama, 1995, pp. 2117–2122.

[9] TIKK, D.: *Investigation of fuzzy rule interpolation techniques and the universal approximation property of fuzzy controller*, PhD thesis, Budapesti Műszaki Egyetem, Budapest, 1999.

[10] TIKK, D.—BARANYI, P.: *Comprehensive analysis of a new fuzzy rule interpolation method*, IEEE Trans. on Fuzzy Sets **8** (1000), 281–296.

[11] TÜRKŞEN, I. B.—ZHONG, Z.: *An approximate analogical reasoning approach of functions*, in: Proc. of the 2nd Int. Conf. on Fuzzy Logic and Neural Networks, (IIZUKA '92), Iizuka, Japan, 1992, pp. 629–632.

[12] YAGER, R. R.: *On ordered weighted averaging aggregation operators in multicriteria decision making*, IEEE Trans. on SMC **18** (1988), 183–190.

[13] YAM, Y.—BARANYI, P.—TIKK, D.—KÓCZY, L. T.: *Eliminating the abnormality problem of α-cut based interpolation*, in: Proc. of the 8th IFSA World Congress, Vol. 2, Taipei, Taiwan, 1999, pp. 762–766.

[14] YAM, Y.—KÓCZY, L. T.: *Representing membership functions as points in high dimensional spaces for fuzzy interpolation and extrapolation*, Technical Report CUHK-MAE-97--03, Dept. of Mechanical and Automation Eng., The Chinese Univ. of Hongkong, 1997.

Received April 28, 2000

Department of Telecommunication and Telematics
Budapest University
of Technology and Economics
H–1117 Budapest
Pázmány sétány 1/d
HUNGARY
E-mail: dtikk@central.murdoch.edu.au