Approximate Reasoning by Linear Rule Interpolation and General Approximation

László T. Kóczy
Technical University of Budapest, Hungary
Kaoru Hirota
Hosei University, Tokyo, Japan

ABSTRACT

The problem of sparse fuzzy rule bases is introduced. Because of the high computational complexity of the original compositional rule of inference (CRI) method, it is strongly suggested that the number of rules in a final fuzzy knowledge base is drastically reduced.

Various methods of analogical reasoning available in the literature are reviewed. The mapping style interpretation of fuzzy rules leads to the idea of approximating the fuzzy mapping by using classical approximation techniques.

Graduality, measurability, and distance in the fuzzy sense are introduced. These notions enable the introduction of the concept of similarity between two fuzzy terms, by their closeness derived from their distance.

The fundamental equation of linear rule interpolation is given, its solution gives the final formulas used for interpolating pairs of rules by their α-cuts, using the resolution principle. The method is extended to multiple dimensional variable spaces, by the normalization of all dimensions.

Finally, some further methods are shown that generalize the previous idea, where various approximation techniques are used for the α-cuts and so, various approximations of the fuzzy mapping $\mathscr{R}$: $X \to Y$.

KEYWORDS: ***Approximate reasoning, fuzzy rule base, sparse rules, interpolation, resolution principle, fuzzy distance of fuzzy sets, approximation of fuzzy mapping***

Address correspondence to László T. Kóczy, Department of Telecommunication, Technical University of Budapest, Sztoczek u. 2, Budapest H-1111, Hungary.
Received May 1, 1991; accepted March 12, 1993.

International Journal of Approximate Reasoning 1993; 9:197–225

655 Avenue of the Americas, New York, NY 10010 0888-613X/93/$6.00

1. INTRODUCTION

Expert systems support decisions in many different fields: from medical diagnosis to industrial economical problems, from metallurgical quality control to robot technology, etc. In most of these cases, the problem requiring an expert's opinion is too complex just to be treated by some well-defined algorithm, however a human expert often solves the problem of deciding in very complex situations satisfactorily.

Similarly, many control problems cannot be treated effectively by traditional control algorithms, because of the extreme complexity of the system or the difficulty in modeling its behavior. Such problems, again, often can be solved acceptably by human operators. A simple example: driving a car can be managed by most grown-up people (at least after a period of appropriate training), but nobody can so far solve the fully automatic control of driving a car in a real traffic environment. The system consisting of car, road, weather, other vehicles, and persons taking part in the traffic, traffic signals, etc., seems to be too complicated to be modeled satisfactorily by any known mathematical method.

In recent years, a good many successful control and expert system applications have invaded the market that have the common feature of using the idea of linguistic/approximate reasoning formalized by fuzzy rules and inference.

The basic idea of fuzzy algorithm (rule-based fuzzy inference) was proposed originally by Zadeh [1, 2]. First applicational results were produced in a laboratory environment by Mamdani and colleagues (e.g., [3, 4]). In the last years, the center of gravity for applications has shifted unambiguously to Japan where hundreds of real industrial applications based on the research work done by Sugeno and Nishida [5], Hirota et al. [6], and others appeared in the middle of the 1980s.

In real applicational fuzzy inference algorithms, one of the crucial problems is the computational speed of the applied method. If the speed is not sufficient, real time control or practical use of an expert system is impossible. Computational speed is mathematically described by algebraic complexity; acceptable speed is achieved only if this complexity is at most polynomial. Our investigations have shown that methods with good sensitivity have also high complexity. For example, the compact rule method proposed by Kóczy and Hirota in a modified form [7] has very good features form the point of view of sensitivity in respect to the rules and the observation. In its original form, however, it has exponential complexity, that makes it untreatable in practice. The same algorithm combined with some boundedness-type restrictions leads to an acceptable complexity [8, 9] if the number of rules is not too high and especially if the support sizes in the rules and observations are small enough. An analysis of rule-based fuzzy control from the point of view of complexity was given in [10].

Such restrictions change the general image of "rule space." They may lead to low density of the rules both in the observation and conclusion spaces. The observations might not overlap with the condition parts of the rules. This situation raises a new problem in obtaining applicable control algorithms; the methods well-known from effective applications are founding namely on examining the "degree of overlapping" by taking the (min) intersection of rules and observation and weighting the conclusion parts of the rules by some typical parameter of this intersection.

Rule interpolation opens a new door in the treatment of such cases and gives us a new and universal algorithm. The topic of this paper is a family of such algorithms.

2. THE IMPORTANCE OF RULE INTERPOLATION

A classical problem for illustrating fuzzy inference is the generalized modus ponens in the "tomato classification" example using the colors of tomatoes to classify them according to the degree of ripeness. (This problem was proposed by Zimmermann and Mizumoto [11, 12].) Let us compare three basic types of reasoning:

1. Simple *modus ponens.*

   If a tomato is red then the tomato is ripe.
   This tomato is red.

   This tomato is ripe.

In this reasoning pattern there is no philosophical difficulty at all, as the observation (red) is identical with the condition part of the rule. Conclusion should be identical with the consequence part.

2. Generalized *modus ponens.*

   If a tomato is red then the tomato is ripe.
   This tomato is very red.

   This tomato is very ripe.

The linguistic term "very" modifies the meaning of "red," thus the conclusion is obtained by an identical modification of the consequence part "ripe." It is essential, that the condition part ("red") and the observation part "very red" contain a clear semantical overlapping. It must be mentioned that the "generalized *modus ponens*" is not accepted by many fuzzy scientists as a correct reasoning pattern. It can be considered as an open problem for further investigations in fuzzy linguistic computing.

3. Open problem—no conclusion.
   If a tomato is red then the tomato is ripe.
   If a tomato is green then the tomato is unripe.
   This tomato is yellow.

   This tomato is ???

Our main purpose is to present a method that can treat reasoning type 3. Let us illustrate this reasoning type by a simple figure using triangular membership functions (Figure 1). Observation space X contains colors from a deep green to a deep red, and linguistic fuzzy terms of colors can be introduced over this space: very green, green, greenish yellow, yellow, orange, red, very red, etc. Conclusion space Y contains degrees of ripeness over which such terms can be introduced: unripe, almost unripe, little ripe, halfripe, quite ripe, almost ripe, ripe, etc. Rules R_1 and R_2 are represented by membership function pairs in X and Y, resp., observation O is a membership function in X. There is no overlapping between "yellow" and "green", nor between "yellow" and "red."

In most of the industrial fuzzy systems a way of reasoning is applied that builds on evaluating the "degree of overlapping" of observation and condition parts of the rules. Such a degree might be

$$ol(O, I_i) = \max_x\{\min\{O(x), I_i(x)\}\}$$

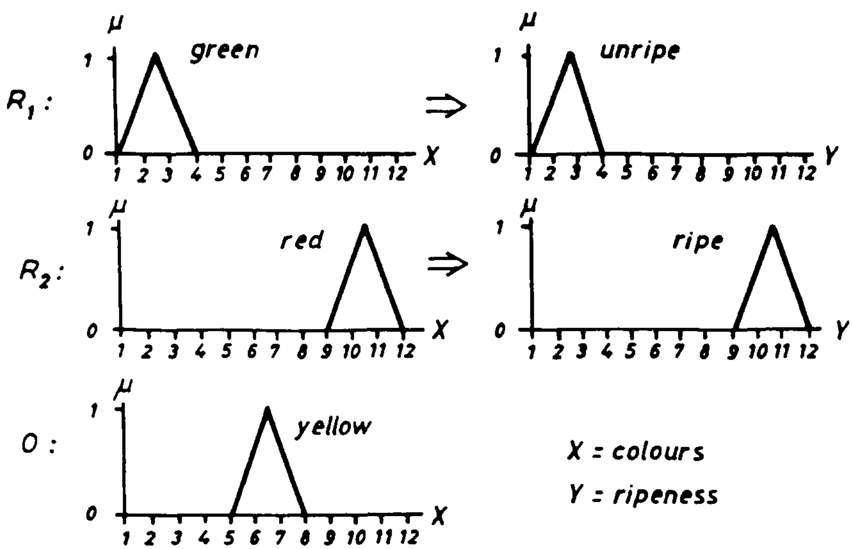

Figure 1.

where O is the observation in $X = \{x\}$ and I_i is the condition part ("if-part") of rule i (in X). The degree expressing the overlapping (ol) serves as a weight for the consequence part in the same rule (T_i). Final conclusion can be obtained by combining these weighted consequences.

Using this last reasoning algorithm for the "yellow tomato case," because of no overlapping the conclusion is a membership function obtained by combining the consequence membership functions weighted by 0, so the conclusion is also identically 0, i.e., no conclusion whatsoever can be calculated. On the other hand we feel intuitively that a conclusion like "*This tomato is halfripe*," would be reasonable.

A solution of this contradiction can be achieved by rule interpolation. The idea is that if observation is in some sense between the two "if-parts" of rules R_1 and R_2 so the conclusion is expected to be similarly between the "then-parts." We attempt to give an exact formulation of this rather intuitive statement in the following section.

3. GRADUAL RULES AND ANALOGICAL REASONING

In order to study the problem further, let us refer to the works of Dubois and Prade [13, 14] that offer some very interesting thoughts concerning the semantics of *If ... then* rules. These rules refer to some *gradual* property, and in this sense

If X is A then Y is B

can be read in the way

The more X is A the more Y is B.

This paper catches the essential property enabling the solution of the "yellow tomato problem." Graduality in the rules hides graduality of the properties, a structure of the variables, and the variable spaces themselves. We shall discuss this aspect in the next sections.

Another approach by Türkşen and Zhong [15] uses the term *similarity* in order to express the semantics of inference. According to this, the rules can be interpreted as

The more similar is X to A the more similar is Y to B.

Here, similarity of two fuzzy sets of the same universe is expressed by the following:

$$\textit{similarity measure} = (1 + \textit{distance measure})^{-1}$$

The advantage of this approach is that it includes a quantitative measure to express the degree of analogy. This measure assumes values in [0, 1] if the distance is normed, as well. Its crucial point is the distance applied, If a distance based on the disconsistency measure is applied, this analogous reasoning technique delivers the maximum-overlapping technique used in numerous industrial applications of fuzzy control as a simplified approximation of the compositional rule of inference method (CRI), as a special case. However, a disadvantage is that the distance used is a crisp value, and so, the original fuzzy nature of the linguistic terms is hidden. Also, in the case of the referred examples, similarity becomes 0 if the supports of the sets are disjoint.

Despite this fact the basic idea to express the semantical overlapping of two linguistic values is very important in the further study of interpolative reasoning. Instead of the crisp distance, however, a fuzzy notion should be introduced that is suitable to form the base of a fuzzy approach to establishing some kind of similarity concept between fuzzy linguistic terms. This aspect will be discussed in the next sections.

There is one more important approach that shows related features: the work done by Shen et al. [16, 17]. This method uses the analogy of a single rule by constructing the semantic curve, a function between the input and the output variables. If an observation is given, it is possible to draw the corresponding conclusion by using the same semantic curve. Clearly, this technique is limited by the number of variables (it is impossible to draw the conclusion in multiple dimensions) and also by the fact that directly no two rules can be considered simultaneously. The method has interesting perspectives nevertheless, as it can be obviously extended to a more general approach constructing the semantic curve associated with the whole rule base by averaging various semantic curves generated by the individual rules.

Reasoning in sparse rule bases where the observation is often disjoint with all antecedents, is an important problem from the point of view of control and expert system applications (cf. [18]). As an extension of the idea of analogical reasoning according to [15], a method must be found where the degree of analogy is always in $(0, 1)$ (except maybe some extremely different sets), even if there is no overlapping between the two sets. In such a sense, only those fuzzy sets have a degree of analogy equal to 0, where the distance is somehow maximal.

4. GRADUAL AND METRIC VARIABLES, FUZZY DISTANCE OF FUZZY TERMS

After this brief overview of reasoning techniques based on some kind of analogy with the yet known rules, let us return now to the problem of the yellow tomato.

Because tomatoes ripen gradually, the colors indicating the degree of ripeness are *gradual properties*. It is the graduality of the properties represented by the variables that enables the application of the theory of gradual interpretation of *If... then* rules. Gradual reasoning based on gradual rules can be applied whenever the input and output variables are gradual.

In most industrial applications variables like spatial position, velocity, acceleration, pressure, temperature, etc. are used; in all of them, a natural full ordering exists. Moreover, in practical applications, domain and range of the variables are finite, so $\max\{X\}$ and $\min\{X\}$, $\max\{Y\}$ and $\min\{Y\}$ exist. If X is compound (i.e., a cross-product of several variables), every dimension in itself can be represented as a bounded set with a full ordering. So in the total cross-product X, there is a partial ordering in the sense

$$x_1 < x_2 \text{ iff } \forall i\colon x_{1,i} < x_{2,i}$$

The overall minimum and maximum are:

$$\min\{X\} = \left(\min\{X_1\}, \ldots, \min\{X_{k_1}\}\right)$$

and

$$\max\{X\} = \left(\max\{X_i\}, \ldots, \max\{X_{k_1}\}\right)$$

It is possible to determine the least upper bound and greatest lower bound for any pair x_1 and x_2. This leads to the following statement:

STATEMENT 1: *The space generated by the cross-product of an arbitrary number of gradual variables is a lattice.*

The behavior of the lattice of gradual variable spaces will be treated in detail in [19].

It must be mentioned however, that not all the real life variables fulfill the requirements summarized in the previous sentences. We mention two counterexamples: The colors green, yellow, and red play an important role in traffic lights. There is however, no internal structure in the set {*red*, *yellow*, *green*} in the sense that yellow is *not* between green and red in the sense of an ordering (except the standard position of the yellow lamp in the usual traffic light complex...) as each of the lights require another, independent behavior from the side of the drivers, etc. Another example could be a pendulum system with full freedom in turning around an axis. The angular position of this pendulum is a cyclic variable, where two positions cannot be compared by any ordering (maybe locally, in small areas, yes), and there are no maximum and minimum positions. The considerations in this paper will not conform to the behavior of systems including cyclic variables and variables with values without an internal, ordered (or partially ordered) structure.

In order to extend analogical reasoning to arbitrary terms (at least, in the world of gradual variables), it is necessary to introduce the fuzzy notion of distance. We can observe that variables in control applications are usually measurable. In some other examples, like the tomatoes, because of the full ordering (ripe > unripe, red > yellow > green, etc.), it is possible to map the range of the variable to an interval, e.g. [0, 1]. Then, measurability can be introduced with help of an isomorphism. An example is

$$r\colon [\textit{completely unripe}, \textit{completely ripe}] \to [0,1]$$

where e.g.,

$$\textit{distance}(\textit{half ripe}, \textit{completely unripe}) = 0.5$$

It has to be stressed that orderedness and measurability are not tied together. Cyclic variables are usually measurable, or can be mapped into a measurable scale.

Variables with measurability in this extended sense will be called *metric*. It is advisable to norm the range always, so for example when X_j is the speed of a particular type of car, we can apply

$$f\colon [0km/s, 200km/s] \to [0,1]$$

Normality is especially important if the variable space is compound. Distance in the traditional sense has no meaning if the various axes in the vector space have various dimensions. Also, if dimensionality is omitted, but the numerical values map to a different scale, distance of two points (vectors) can be hardly interpreted in the original context, as clearly the dimensions with large absolute values will be dominant, even, if the distance is small in its importance.

In the context of metric variables, it is possible to define the distance of two fuzzy sets. Distance of two values in X_j can be described by the following axiomatic properties:

$$d\colon X_i^2 \to [0,1]$$

$$d(x,x) = 0 \ (x \in X_i)$$

$$d(x_1,x_2) < d(x_3,x_4) \text{ if } x_3 < x_1 < x_2 < x_4 \ (x_j \in X_i)$$

In the case of discrete variables (with a finite number of possible values), it might be reasonable to define the distance as

$$d(x_i,x_j) = |i-j|/(n-1)$$

where $|X| = n$.

For compound variables, the family of Minkowski distances can be applied:

$$D(x_1, x_2) = \left(\sum_{i=1}^{k} |x_{1,i} - x_{2,i}|^w \right)^{1/w}$$

Because of normalization in every component, for D

$$D\colon [0,1]^k \to [0, k^{1/w}]$$

When $w = 2$, we obtain the Euclidean distance d as the most obvious choice for practical applications.

If both ordering and distance in X_i exist, a partial ordering among the fuzzy sets (the linguistic terms) of the universe X_i, even $X = \prod_{i=1}^{k} X_i$ can be defined. Let us denote the set of all normal and convex fuzzy sets of the universe X_i by $\tilde{\mathscr{P}}(X_i)$. Then for $A, B \in \tilde{\mathscr{P}}(X_i)$, $A \prec B$ iff

$$\forall \alpha \in (0,1]\colon \mathit{inf}\{A_\alpha\} < \mathit{inf}\{B_\alpha\}$$

and

$$\mathit{sup}\{A_\alpha\} < \mathit{sup}\{B_\alpha\}$$

It is always possible to find both upper and lower bounds to any two elements, e.g.

$$C = \min\{A, B\} \text{ if } \forall \alpha \in (0,1]\colon$$
$$C_\alpha = [\min\{\inf\{A_\alpha\}, \inf\{B_\alpha\}\}, \min\{\sup\{A_\alpha\}, \sup\{B_\alpha\}\}]$$

Moreover, if $\mathit{first}\{X\}$ is defined as the element of $\tilde{\mathscr{P}}(X)$ for which

$$\forall A \in \tilde{\mathscr{P}}(X)\colon \mathit{first}\{X\} \prec A$$

and similarly for $\mathit{last}\{X\}$

$$\forall A \in \tilde{\mathscr{P}}(X)\colon A \prec \mathit{last}\{X\}$$

and X is gradual, both $\mathit{first}\{X\}$ and $\mathit{last}\{X\}$ exist and are unique. The membership function of $\mathit{first}\{X\}$ is, e.g.:

$$\mu_{first\{X\}} = \begin{cases} 1 & \text{if } x = (\min\{X_1\}, \ldots, \min\{X_k\}) \\ 0 & \text{otherwise} \end{cases}$$

These two are lower and upper bounds to all elements in $\tilde{\mathscr{P}}\{X_i)$, respectively.

Because of the above properties of $\prec$, it defines a lattice over the convex and normal fuzzy sets of X.

With help of $\prec$, it is possible to define the relation "comparable pairs" ($\tilde{\mathscr{R}}_{\prec}$) which is a subset of $\tilde{\mathscr{P}}^2(X)$ and contains all such pairs of fuzzy sets over X that are comparable in the sense $A \prec B$:

$$\tilde{\mathscr{R}}_{\prec} = \{(A, B) | A, B \in \tilde{\mathscr{P}}(X), A \prec B\}$$

$\tilde{\mathscr{R}}_{\prec}$ is a crisp relation.

For pairs of fuzzy sets in $\tilde{\mathscr{R}}_{\prec}(A, B)$, the *lower and upper fuzzy distances* of A and B can be defined with the help of the *Resolution Principle*:

$$\tilde{d}_L(A, B), \tilde{d}_U(A, B): \tilde{\mathscr{R}}_{\prec} \to \tilde{\mathscr{P}}([0,1]),$$

$$\mu_{\tilde{d}_L(A,B)}(\delta) = \sum_{\alpha \in [0,1]} \alpha / D(\inf\{A_\alpha\}, \inf\{B_\alpha\}), \delta \in [0, k^{1/w}]$$

Similarly,

$$\mu_{\tilde{d}_U(A,B)}(\delta) = \sum_{\alpha \in [0,1]} \alpha / D(\sup\{A_\alpha\}, \sup\{B_\alpha\}), \delta \in [0, k^{1/w}]$$

D denotes the Minkowski distance.

The fuzzy lower and upper distances defined in this way can be considered an extension of the notion of distance introduced for values of a gradual variable space. In the following, a few important properties of $\tilde{d}$ are summarized:

1. *The lower and upper distance of any fuzzy set from itself is the crisp set first*$\{[0, k_1^{1/w}]\}$ *i.e., a set with membership function* 1 *over* 0 *and* 0 *over any other possible distance.*

This property is the extension of $D(0,0) = 0$.

2. *The lower and upper distance of first*$\{X\}$ *and last*$\{X\}$ *is* $k^{1/w}$. *So for any distance:*

$$\tilde{d}_L(A, B) \leq \tilde{d}_L(\mathit{first}\{X\}, \mathit{last}\{X\}) = d_{max}$$

and

$$\tilde{d}_U(A, B) \leq \tilde{d}_U(\mathit{first}\{X\}, \mathit{last}\{X\}) = d_{max}$$

3. *For arbitrary* $A, B, C, D \in \tilde{\mathscr{P}}(X)$

$$\text{If } A \prec B \prec C \prec D \text{ then } \tilde{d}_{L/U}(B,C) \prec \tilde{d}_{L/U}(A, D)$$

(Here subscript L/U means "L or U").

The fuzzy distance introduced here is essentially different from the crisp distance as in [15], because it is a fuzzy set of distances, with different values having different membership levels α. It is necessary to introduce a pair of distances (L and U), as α-cuts of convex fuzzy sets are intervals.

5. REPRESENTATION OF RULES IN $X \times Y$ AND REASONING BY "CLOSENESS"

The resolution principle plays a very important role in this approach. Taking an arbitrary rule of the form

R: *If X is A then Y is B*

where A (over X) and B (over Y) are fuzzy sets, R can be viewed as a family of "α-rules":

R_α: *If* X *is* A_α *then* Y *is* B_α

where $\alpha \in (0, 1]$. The α-cuts A_α and B_α are represented k- and k'-dimensional hyperintervals in $X \times Y$. Every hyperinterval has its infimum and supremum, so if α is set fix, every rule can be unambiguously described by a *pair of points in* $X \times Y$ i.e., one for the infima ("lower point") and one for the suprema ("upper point"). As in the general case α can take any value in $(0, 1]$, theoretically, every rule is represented by an infinite number of point pairs in $X \times Y$. However, if the level sets of both A and B have a finite cardinality then it is sufficient to represent every rule by $2|\Lambda_A \cup \Lambda_B|$ points altogether. For example, if the fuzzy sets in the rules are piecewise linear, as the most frequently used shapes, triangular or trapezoidal, $|\Lambda_A|$ and $|\Lambda_B|$ are finite, these values are 3 and 4 in the latter cases.

Every rule base consisting of r rules and having l elements in the level sets, can be represented in $X \times Y$ *by maximally 2rl points.*

The introduction of fuzzy distance makes it possible to translate the semantics of an *If ... then* rule into the form:

The closer A^* *to* A_i *the closer is* B^* *to* B_i

This interpretation is a quantitative extension of the idea of gradual rules and also an extension of the analogical reasoning by similarity. From here, instead of similarity, the reciprocal concept of distance will be used.

On the basis of the above interpretation, it is possible to introduce a large variety of function approximation techniques for estimating $\mathscr{R}(x)$, and for using it to the construction of $B^* = \mathscr{R}(A^*)$ by applying the resolution principle. Such techniques include interpolation of two or more rules, extrapolation and mixed inter/and extrapolation, further use of regression or other techniques for estimating the "average tendency" of rule bases where evidence in the rules is (partially) conflicting—even techniques where different tendencies in the same rule base are detected simultaneously, and so alternative conclusions are constructed in parallel. In the following, the basic idea of linear interpolation will be discussed.

6. LINEAR INTERPOLATION OF TWO RULES

Suppose that we have two rules, which are disjoint in X and we have an observation between these two (in the sense of the partial ordering $\prec$ discussed previously). The interesting case is when also the observation is disjoint with both of the condition parts, i.e., there is no overlapping between observation and rule.

Let us denote the condition parts ("if-parts") of the rules by I_1 and I_2, the consequence parts ("then-parts") by T_1 and T_2, respectively. Linear interpolation of the two rules can be intuitively defined following the idea of closeness in the fuzzy sense:

$$\tilde{d}(O, I_1) : \tilde{d}(O, I_2) = \tilde{d}(C, T_1) : \tilde{d}(C, T_2)$$

where $\tilde{d}$ stands for the fuzzy distance of the two membership functions indicated, or the supposed function C of the conclusion. This is the fundamental idea of linear rule interpolation. (From here, instead of $\tilde{d}$ we simply write d.)

Interpolation can be done if the observation is flanked by two rules (if-parts). Then, it is also expected that the conclusion will be also flanked by the two rules (then-parts). In order to interpolate in this sense, it is necessary that both the if-parts and the then-parts should be comparable in the sense of the partial ordering in the respective space. Suppose that our two rules R_1 and R_2 are such that

$$\min\{\text{supp}(I_1)\} \prec \min\{\text{supp}(O)\} \prec \min\{\text{supp}(I_2)\} \quad \text{and}$$

$$\max\{\text{supp}(I_1)\} \prec \max\{\text{supp}(O)\} \prec \max\{\text{supp}(I_2)\}$$

further on that

$$\min\{\text{supp}(T_1)\} @ \min\{\text{supp}(T_2)\} \quad \text{and}$$

$$\max\{\text{supp}(T_1)\} @ \max\{\text{supp}(T_2)\}$$

where @ means either $\prec$ or $\succ$.

It is also reasonable to restrict investigations to the case where all membership functions are convex and normal. In that case, the fuzzy distance between any pair will be also a normal fuzzy set. It is a more complicated problem, how the various lower and upper distances can be treated in a compact way. Both type distances are defined by their α-cuts, if all are known, the entire distance function can be reconstructed by the resolution principle.

Let us illustrate now the yellow tomato problem with help of the notions of fuzzy lower and upper distance. Let X and Y be the finite intervals $[x_1, x_{12}]$ and $[y_1, y_{12}]$, i.e.,

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}\}$$

and

$$Y = \{y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8, y_9, y_{10}, y_{11}, y_{12}\}$$

We suppose that in X, $\prec$ stands for "more green" or "less red" i.e., x_1 is a bluish green and x_{12} is a violet red color. Between them, we have various shades of green, yellow, orange, and red. As a matter of course, yellow and yellowish colors will be in the middle of the ordered set. In Y, $\prec$ denotes the degree of ripeness so that y_1 stands for completely not ripe and y_{12} for completely ripe. If because of some reason, this discretization of the color or ripeness scale is not sufficient, for simplicity, the inclusion of elements like $x_{4.5}$ etc. is also possible, so that the distances $d(y_i, y_j)$ and $d(x_i, x_j)$ will be still defined as $|i - j|$. (By this, we only formally violate the originally proposed concept of distance by subscripts as by reindexing the elements we obtain an isomorphic structure that completely conforms with the idea.)

Let us define the linguistic term "green" by a simple triangular membership function (being on rather the left side of X, let in the sense of $\prec$) as

$$\{0/x_1, 0.67/x_2, 0.67/x_3, 0/x_4, 0/x_5, 0/x_6, 0/x_7, 0/x_8, 0/x_9,$$
$$0/x_{10}, 0/x_{11}, 0/x_{12}\}$$

and "yellow" as

$$\{0/x_1, 0/x_2, 0/x_3, 0/x_4, 0/x_5, 0.67/x_6, 0.67/x_7, 0/x_8, 0/x_9,$$
$$0/x_{10}, 0/x_{11}, 0/x_{12}\}$$

Then

$$\text{green}_{0.67} = \{x_2, x_3\} \quad \text{and} \quad \text{yellow}_{0.67} = \{x_6, x_7\}.$$

The lower and upper distances are

$$d_L(\text{green}_{0.67}, \text{yellow}_{0.67}) = d(\min\{x_2, x_3\}, \min\{x_6, x_7\}) = 4$$
$$d_U(\text{green}_{0.67}, \text{yellow}_{0.67}) = d(\max\{x_2, x_3\}, \max\{x_6, x_7\}) = 4$$

The 0.67-cuts of both the lower and upper distances between green and yellow are equal to 4.

After this example, let us continue with the idea of the proportional distances between observation and if-parts and conclusion and then-parts, formulated earlier. The fundamental formula delivering conclusion C is obtained by the solution of the equalities

$$d_L(O_\alpha, I_{1\alpha}): d_L(O_\alpha, I_{2\alpha}) = d_L(C_\alpha, T_{1\alpha}): d_L(C_\alpha, T_{2\alpha}) \quad \text{and}$$

$$d_U(O_\alpha, I_{1\alpha}): d_U(O_\alpha, I_{2\alpha}) = d_U(C_\alpha, T_{1\alpha}): d_U(C_\alpha, T_{2\alpha}) \qquad \text{for}$$

$$\forall \alpha \in [0,1]$$

Where the two rules are $R_1(I_1 \to T_1)$ and $R_2(I_2 \to T_2)$, the observation is O, and the conclusion is C.

If R_1, R_2 and O are ordered according to the inequalities stated earlier, these equalities can be solved for unknown C_α for every α.

Statement 2:

The solutions for $\min@\{C_\alpha\}$ *and* $\max@\{C_\alpha\}$ *are*

$$\min_{@}\{C_\alpha\} = \frac{w'^{\alpha}_{1L}.\min_@\{T_{1\alpha}\} + w'^{\alpha}_{2L}.\min_@\{T_{2\alpha}\}}{w'^{\alpha}_1 + w'^{\alpha}_2} \quad \textit{and}$$

$$\max_{@}\{C_\alpha\} = \frac{w'^{\alpha}_{1U}.\max_@\{T_{1\alpha}\} + w'^{\alpha}_{2U}.\max_@\{T_{2\alpha}\}}{w'^{\alpha}_1 + w'^{\alpha}_2}, \quad \textit{where}$$

$$w'^{\alpha}_{1L} = d(\min\{O_\alpha\}, \min\{I_{2\alpha}\}) \quad \textit{and} \quad w'^{\alpha}_{2L} = d(\min\{O_\alpha\}, \min\{I_{1\alpha}\});$$

$$w'^{\alpha}_{1U} = d(\max\{O_\alpha\}, \max\{I_{2\alpha}\}) \quad \textit{and} \quad w'^{\alpha}_{2U} = d(\max\{O_\alpha\}, \max\{I_{1\alpha}\}).$$

Proof is simple by rearrangement.

Weighting factors w'^{α}_{1L}, w'^{α}_{2L}, w'^{α}_{1U} and w'^{α}_{2U} can be substituted by $w^{\alpha}_{2L} = 1/w'^{\alpha}_{2L}$, $w^{\alpha}_{2L} = 1/w'^{\alpha}_{1L}$, $w^{\alpha}_{1U} = 1/w'^{\alpha}_{2U}$ and $w^{\alpha}_{2U} = 1/w'^{\alpha}_{1U}$, resp. without effecting the results. (Use of the reciprocal distance weights is advisable as so the weight is a function of the distance between the observation and the *own* rule.)

The α-cut of C is given then by

$$C_\alpha = [\min\{C_\alpha\}, \max\{C_\alpha\}]$$

i.e., either $[\min_@\{C_\alpha\}, \max_@\{C_\alpha\}]$ or $[\max_@\{C_\alpha\}, \min_@\{C_\alpha\}]$, depending on whether $@ = \prec$ or $\succ$.

We shall complete our tomatoes example, so let us now calculate the final step in the interpolation for the conclusion of reasoning type 3. For

α, we choose the set {0, 0.67} (0 is meant in the sense of strict 0-cut). Then

$$C_0 = [(1/4.y_1 + 1/4.y_9).2, (1/4.y_4 + 1/4.y_{12}).2] = [y_5, y_8]$$
$$C_{0.67} = [(1/4.y_2 + 1/4.y_{10}).2, (1/4.y_3 + 1/4.y_{11}).2] = [y_6, y_7]$$

(If we wanted to calculate C_1 as well, it would require a denser scale in X and Y, otherwise we would find that C_1 is the empty set.)

The reconstruction of C shows the membership function in Figure 2, which is obviously identical with the possible definition of "halfripe." Thus, the open conclusion in Reasoning 3 is as follows:

This tomato is halfripe.

Another example can be seen in Figure 3, where although only triangular membership functions are applied, they have varying widths. (For simplicity only the subscripts are marked.) From these two examples, we have the impression that it is enough to calculate the support and the maximum as linear interpolation of triangular membership functions always leads to triangular results. This can be stated also in general:

Statement 3:

If rules $R_1(I_1 \to T_1)$ and $R_2(I_1 \to T_2)$, further on the observation are defined by normal triangular membership functions, the interpolated conclusion will be also normal and triangular. A conclusion of this statement is that it is enough to calculate only two different α-cuts in order to reconstruct the full conclusion.

Ideal of the proof:

If $C_{\alpha 1}$ and $C_{\alpha 2}$ are calculated for $\alpha_1 \neq \alpha_2$, an arbitrary α_3 can be unambiguously decomposed as:

$$\alpha_3 = q_1\alpha_1 + q_2\alpha_2, \text{where } q_1 + q_2 = 1.$$

Then the weights belonging to α_3 will be also linearly decomposable by coefficients q_1 and q_2, and so one side of the triangular membership function defined by two points obtained from $C_{\alpha 1}$ and $C_{\alpha 2}$ (min or max) will contain the point defined by $C_{\alpha 3}$ (min or max), as well.

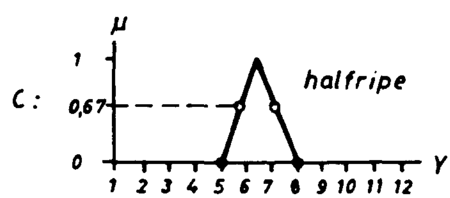

Figure 2.

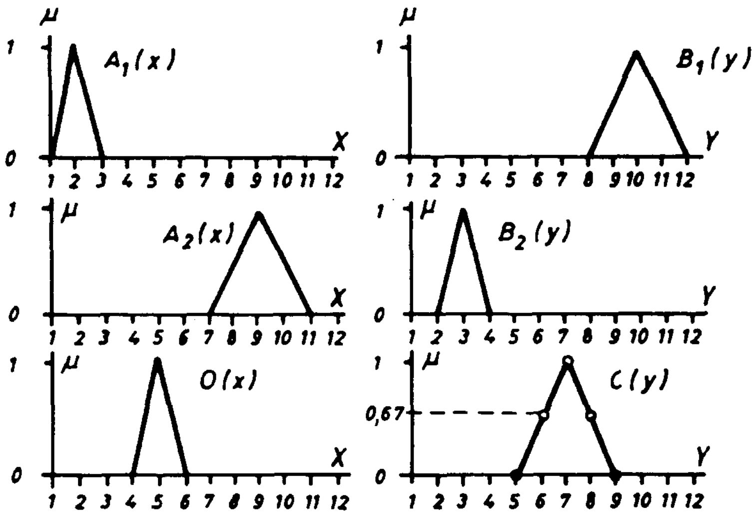

Figure 3.

In the example in the Figure, the two rules can be described briefly as

$$R_1\colon 1\Delta 3 \Rightarrow 8\Delta 12$$

$$R_2\colon 7\Delta 11 \Rightarrow 2\Delta 4$$

(where $i\Delta j$ stands for a symmetrical triangular membership function with support $[i, j]$). The observation is

$$O\colon 4\Delta 6.$$

Calculating the weights, we have:

$$w_{1L}^0 = 1/3, \qquad w_{2L}^0 = 1/3, \qquad w_{1U}^0 = 1/3, \qquad w_{2U}^0 = 1/5;$$

$$w_1^1 = 1/3, \qquad w_2^1 = 1/4,$$

so the conclusion is

$$C\colon 5\Delta 9.$$

An illustrative program has been prepared for interpolation of symmetrical triangular terms based rules. Figures 4a–h show the steps of defining two rules (a: if-part of R_1, b: then-part of R_1, c: if-part of R_2, d: then-part of R_2, e: the two rules shown together, f: if-parts and observation together, g: then-parts and calculated conclusion together, h: f and g in one screen picture). Figure 5 presents a few different cases. In Figure 4, we had rules where $@ = \prec$. In Figure 5a $@ = \succ$. In 5b, the observation is

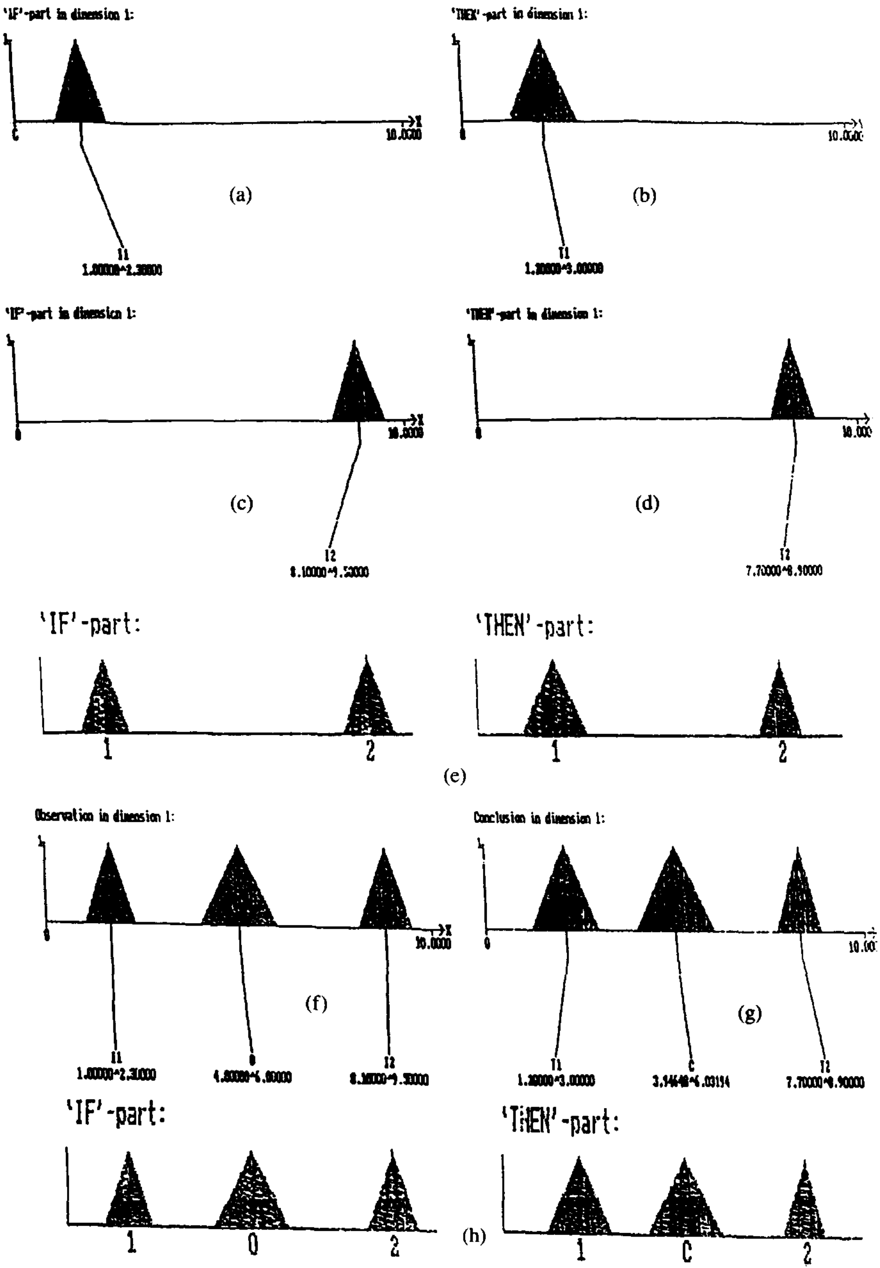
'IF'-part in dimension 1:
'THEN'-part in dimension 1:
(a)
(b)
(c)
(d)
'IF'-part:
'THEN'-part:
1
2
(e)
Observation in dimension 1:
Conclusion in dimension 1:
(f)
(g)
(h)

Figure 4.

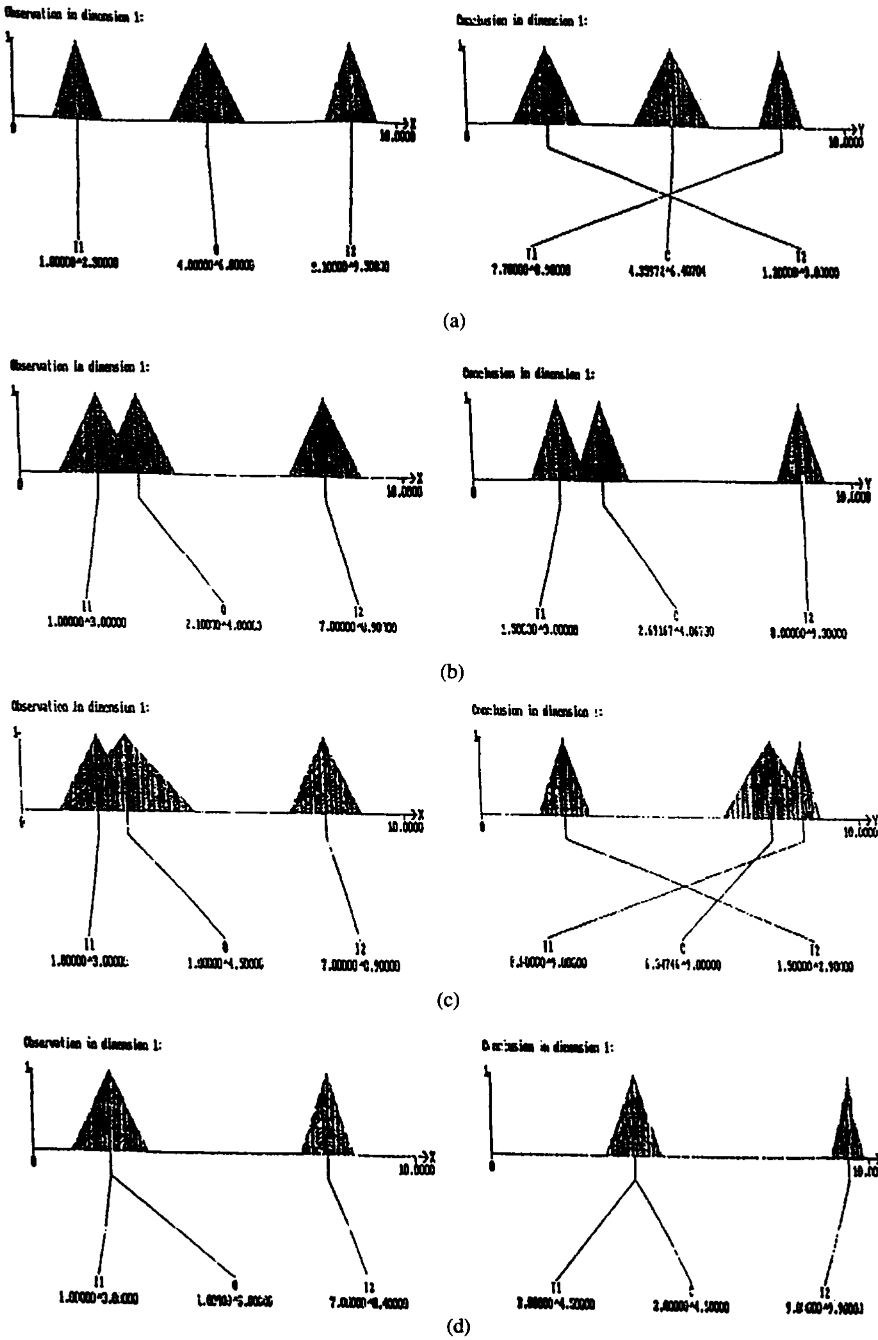

Figure 5.

overlapping with one of the if-parts (I_l), so does the calculated conclusion with the corresponding then-part (T_1). In 5c, one of the boundaries of O is identical with one of the boundaries of I_1, so it is with C and T_1. As here $@ = \succ$, this boundary changes to the maximum in Y from the minimum in X. Finally, in 5d the observation is completely identical with one of the if-parts (I_1), so $C = T_1$. Everywhere it is clear that the dominantly overlapping rule "pulls" the conclusion also near to its consequence part, while an identical if-part in the rule generates an identical then-part, as well.

In the example, one-dimensional X and one-dimensional Y were treated. In real expert systems or control algorithms however usually the rules contain more than one fuzzy variable both in X and Y, i.e., observation and conclusion space are both multidimensional. General type rules have the form

If x_1 *is* A_{1i} *and* x_2 *is* A_{2i} *and... and* x_m *is* A_{mi}
then y_1 *is* B_{1i}, *and* y_2 *is* B_{2i} *and... and* y_n *is* B'_{ni}

We intend to extend the interpolation method for the general case with multidimensional rules and observations (and, as a matter of course, interpolated conclusions). So it is necessary to go back first to the intuitive idea of distance between two fuzzy terms.

If we restrict the examination to an arbitrary α-cut, the distances in every dimension of X can be calculated separately, just like in one dimension. As a result, we obtain m distance pairs: $d_{11}^{\alpha}, d_{12}^{\alpha}, \ldots, d_{1m}^{\alpha}$ and $d_{21}^{\alpha}, d_{22}^{\alpha}, \ldots, d_{2m}^{\alpha}$ (cf. Figure 6).

For the final interpolation, however, a pair of single weighting factors is necessary that must somehow accumulate the information in all these distance pairs. In multidimensional spaces distance is always understood as the length of the vector defined by the two endpoints. In order to calcu-

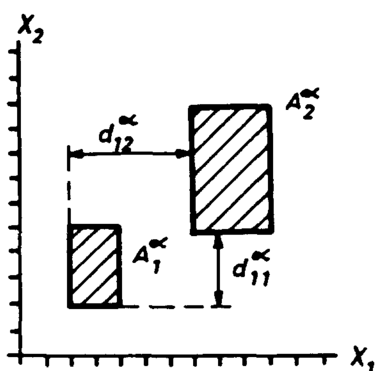

Figure 6.

late the length, it is necessary that a uniform metric is introduced over the whole space. This is achieved by the normalization, as we mentioned it earlier.

Let us highlight the main problem by a simple example. If we define $X_1 = \{x_1 \dots x_{10}\}$ and $X_2 = \{x_1 \dots x_{1000}\}$, $|i - j|$ will be a "large distance" in X_1 if e.g., $i = 1$, $j = 5$; and the same will be a "very small distance" in X_2. It is necessary to normalize the distance in every dimension by either applying the same discretization in every dimension (both X and Y), or by defining the metric as

$$d(x_i, x_j) = |i - j| / (\max\{k | x_k \in X_n\} - \min\{k | x_k \in X_n\}).$$

Having this uniform metric, we suggest the use of the reciprocal values of the lengths of the distance vectors as weighting factors, in accordance with the previous formulas.

Statement 2*b*:

In this case, the weighting factors for the extended linear interpolation are

$$w_{1L}^{\alpha} = \left(\sqrt{(d_{L11}^{\alpha})^2 + (d_{L12}^{\alpha})^2 + \cdots + (d_{L1m}^{\alpha})^2}\right)^{-1}$$

$$w_{1U}^{\alpha} = \left(\sqrt{(d_{U11}^{\alpha})^2 + (d_{U12}^{\alpha})^2 + \cdots + (d_{U1m}^{\alpha})^2}\right)^{-1}$$

and similarly for w_{2L}^{α} *and* w_{U2}^{α}.

If Y is multidimensional, these weights must be applied in every dimension of Y, the same forms must be applied for every dimension of Y separately.

Statement 4:

If R_1 *and* R_2, *further on* O *contain only triangular membership functions in* X *and* Y, *the conclusion obtained by linear interpolation is also triangular in every dimension of* Y.

The proof is similar to that of the one-dimensional case as the combination of straight lines results into linear surfaces (hyperplanes) in multiple dimensions.

The illustrative program is able to treat also multidimensional rules (both in X and Y). One example is shown in Figure 7. Here, both X and Y have five dimensions (it is not necessary that these dimensionalities are the same). In 7*a* the two rules are shown: a triangular membership function in every component of X and Y. The observation has also five dimensions, the calculated conclusion has a similar structure in every

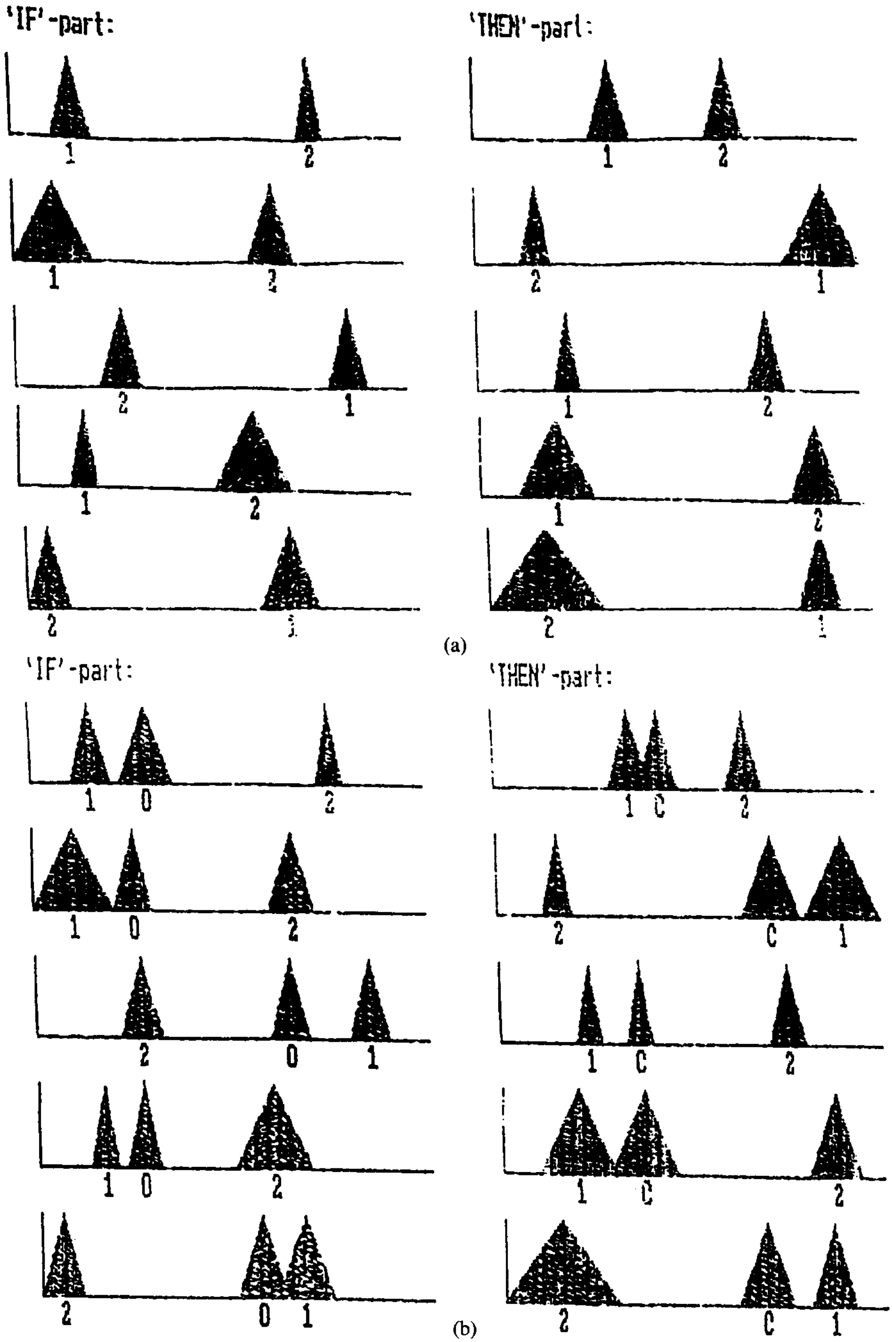

Figure 7.

dimension of Y. In the Figure, it can be seen clearly that O is much nearer to the if-part of R_1 than to R_2, so is the conclusion to the components of T_2.

It is necessary to deal with some extremal cases. The conditions of flanking can be made somewhat looser if we allow that $\prec$ is replaced by $\leq$ (and $\succ$ by $\geq$, if necessary). This means that some distances may be equal to 0. It is expected that in such a case, the rule to which a full or partial identity of the observation is true, becomes dominant in the conclusion.

We extend the weights to possibly $1/0 = \infty$ and we understand $\infty/\infty = 1$, $c/\infty = 0$, $c + \infty = \infty$.

Statement 5:

If observation O is identical with I_i in rule R_i then the conclusion will be identical with T_i of the same rule.

The proof is obvious from the way of extending the weights.

Another extremal case is if the triangular membership function has a positive value at the "end" of X_i. Then, an extension of X_i and an extrapolation of the membership function is advisable (including even "negative subscripts" if necessary), as so Statement $2b$ concerning multidimensional triangular-shaped membership functions can be applied, and calculations can be restricted to two α's (e.g., 0 and 1).

It is an interesting problem, what happens if X is infinite and the membership functions of the if-parts of the rules are not bounded. (A typical example for that is the S-shaped membership function that is very common in practical applications.) Here formally, the minimum or maximum points of the α-cuts are in $+\infty$ or $-\infty$. If in the interpolation $\pm\infty$ is "linearly combined" with some finite value, the distance $\pm\infty$ is also $\pm\infty$ from any finite point and so the weight is $1/\infty = 0$. If however, one end of the membership functions is in $\pm\infty$ for both rules and the observation as well, interpolation on the other end must be done and the interpolated conclusion will have also an "infinite end." Details of such extremal cases must be worked out for every concrete algorithm depending on the type of membership functions to be expected in that particular field of application.

7. INTERPOLATION OF $2k$ RULES IN GENERAL

In the previous section, we discussed only the interpolation on the basis of two rules flanking the observation. The fundamental idea described here can be extended for $2k$ rules if k if-parts on the left and k if-parts on the right side of the observation flank it in the same sense as the single pair did it in the simple interpolation case:

$$\min\{supp(I_i)\} \le \min\{supp(O)\} \le \min\{supp(I_j)\} \quad \text{and}$$
$$\max\{supp(I_i)\} \le \max\{supp(O)\} \le \max\{supp(I_j)\} \text{ if}$$
$$\min\{supp(T_i)\} @ \min\{supp(T_j)\} \quad \text{and}$$
$$\max\{supp(T_i)\} @ \max\{supp(T_j)\}, \text{ where @ is either} \le \text{ or } \ge .$$

$$i \in \{1,\dots,k\} \quad \text{and} \quad j \in \{k+1,\dots,2k\}$$

Then, the $2k$-rules interpolation of $R_1,\dots,R_{2k}$ and O fulfilling the general flanking conditions is given by the formulas

$$d^0_{Li} = \min\{\text{supp}(O)\} - \min\{\text{supp}(I_i)\} \qquad \text{for } i \in \{1,\dots,k\} \quad \text{and}$$

$$d^0_{Lj} = \min\{\text{supp}(I_j)\} - \min\{\text{supp}(O)\} \qquad \text{for } j \in \{k+1,\dots,2k\}$$

further on

$$d^0_{Ui} = \max\{\text{supp}(0)\} - \max\{\text{supp}(I_i)\} \qquad \text{for } i \in \{1,\dots,k\} \quad \text{and}$$

$$d^0_{Uj} = \max\{\text{supp}I_j)\} - \max\{\text{supp}(O)\}$$

for $j \in \{k+1,\dots,2k\}$. Here, the extended linear weighting factors are

$$w_{Li} = 1/d_{Li} \quad \text{and} \quad w_{Ui} = 1/d_{Ui} i \in \{1,\dots,2k\}$$

further on, similarly

$$d^\alpha_{Li} = \min\{O_\alpha\} - \min\{I_{i\alpha}\} \qquad \text{for } i \in \{1,\dots,k\} \quad \text{and}$$

$$d^\alpha_{Lj} = \min\{I_{j\alpha}\} - \min\{O_\alpha\} \qquad \text{for } j \in \{k+1,\dots,2k\}$$

$$d^\alpha_{Ui} = \max\{O_\alpha\} - \max\{I_{i\alpha}\} \qquad \text{for } i \in \{1,\dots,k\} \quad \text{and}$$

$$d^\alpha_{Uj} = \max\{I_{j\alpha}\} - \max\{O_\alpha\} \qquad \text{for } j \in \{k+1,\dots,2k\}.$$

In multiple dimensions, d-s are calculated in every (normalized) dimension separately, and the resulting distance is taken e.g., in the Minkowski sense. The weights are

$$w^\alpha_{Li} = 1/d^\alpha_{Li} \quad \text{and} \quad w^\alpha_{Ui} = 1/d^\alpha_{Ui} i \in \{1,\dots,2k\}$$

and the interpolated conclusion is

$$\min_{@}\{C_\alpha\} = \frac{\sum_{i=1}^{2k} w_{iL}^{\alpha} \min_{@}\{T_{i\alpha}\}}{\sum_{i=1}^{2k} w_i^{\alpha}} \quad \text{and}$$

$$\max_{@}\{C_\alpha\} = \frac{\sum_{i=1}^{2k} w_{iL}^{\alpha} \max_{@}\{T_{i\alpha}\}}{\sum_{i=1}^{2k} w_i^{\alpha}},$$

where @ is either $\leq$ or $\geq$ as in the flanking conditions.

The usability of this definition is illustrated by the following Statement:

Statement 6:

If there is an isomorphism ι

$$\iota: \mathscr{F}(X) \to \mathscr{F}(Y),$$

where $\mathscr{F}(S)$ *is the set of fuzzy sets of* S, *which isomorphism is invariant for the metrics in* X *and* Y, *and the rule system is* $R = \{R_1, \ldots, R_{2k}\}$, *the observation is* O *and*

$$T_i(y) = \iota(I_i(x)) \qquad \textit{for } i \in \{1, \ldots, 2k\},$$

then

$$C(y) = \iota(O(x))$$

Idea of the proof.

We prove that for every α

$$C_\alpha(y) = \iota_\alpha(O(x))$$

The above is expressed by min and max of C_α and for the simplicity Y is transformed into X. Then w_{Li}^{α} and w_{Ui}^{α} are expressed by min{supp(O)} − min{supp(I_i)}, etc. by the flanking conditions, the equation to prove is reduced to

$$\sum_{i=1}^{k} \frac{z_i}{z_i - c} + \sum_{i=k+1}^{2k} \frac{z_i}{c - z_1} = \sum_{i=1}^{k} \frac{c}{c - z_i} + \sum_{i=k+1}^{2} \frac{c}{z_i - c},$$

where z_i stands for min{$T_{i\alpha}$} or max{$T_{i\alpha}$}, i.e., also for min{$\iota_\alpha(I_i)$} or max{$\iota_\alpha(I_i)$} and c for min{C_α} or max{C_α} and also for min{$\iota_\alpha(C)$} $=^?$ min {O_α} or max{$\iota_\alpha(C)$} $=^?$ max{O_α}. This is however, identically true as it is reducible to $k = k$.

8. AN OUTLOOK TO FURTHER APPROXIMATION TECHNIQUES

Interpolation in the previous sense can be extended to extrapolation: any approximation of $\mathscr{R}(x)$ *between the rules* can be continued also *outside* of the area covered by the set of rules. Reliability of the approximation obtained in this way depends on the distance of the approximated conclusion from the nearest rule and the "goodness" of the approximation between the rules.

In order to extend the idea of interpolation that is based on the fuzzy distance we introduced earlier, the geometrically interpretable fuzzy distance will be extended to a more general signed distance concept. Using the signed distance (allowing that an observation located completely outside of the scope of the known rules), the following extended forms will be used for determining the approximation function:

$$\begin{aligned} inf_{\prec}\{B_\alpha^*\} &= \left[\sum_{i=1}^{k}(A_{\alpha L}^* - A_{i,\alpha L})^{-1} inf_{\prec}\{B_{i,\alpha L}\} \right. \\ &\quad \left. + \sum_{i=k+1}^{2k}(A_{i,\alpha L} - A_{\alpha l}^*)^{-1} inf_{\prec}\{B_{i,\alpha L}\}\right] \\ &\quad \times \left[\sum_{i=1}^{k}(A_{\alpha L}^* - A_{i,\alpha L})^{-1} + \sum_{i=k+1}^{2K}(A_{i,\alpha L} - A_{\alpha L}^*)^{-1}\right]^{-1} \end{aligned}$$

etc. If for arbitrary rules and observation the distances are taken with alternating signs (hypothetical left and right sides of the observation), a smooth inter/extrapolation curve is obtained. $\mathscr{R}(x)$ is approximated from six rules on Figure 8.

A demonstration software has been developed for illustrating this technique. A rule base and an approximation function are shown in Figure 9.

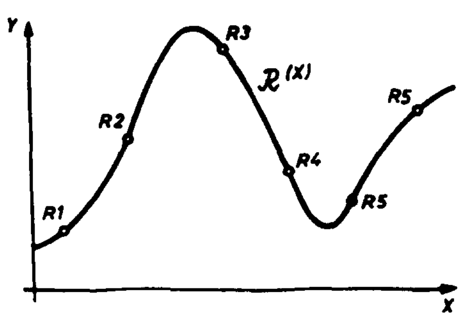

Figure 8.

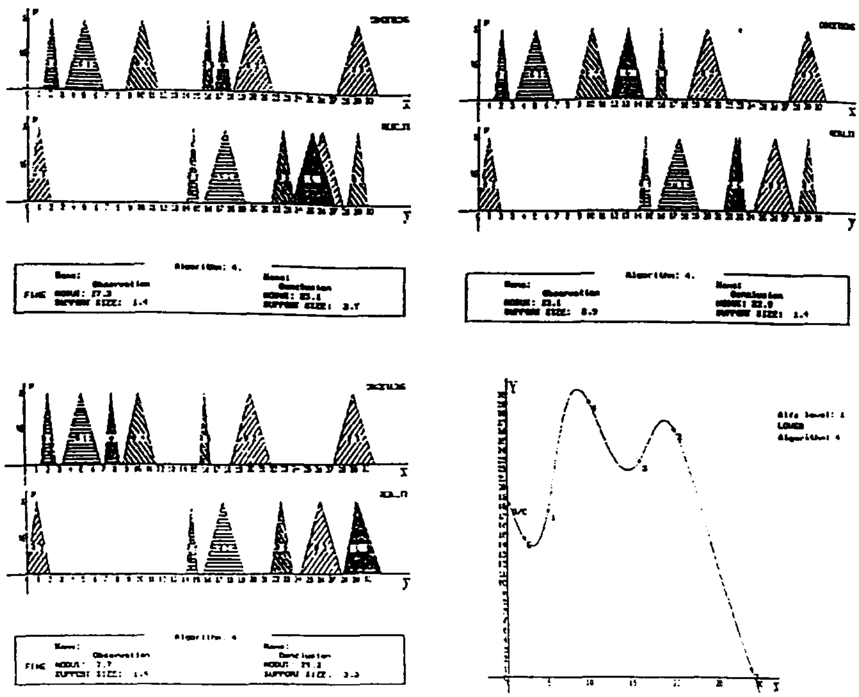

Figure 9.

How to estimate $\mathscr{R}(x)$ when the rules contain internal conflicts? If there are rules with (partially) identical antecedents and different consequents, interpolation and extrapolation are ambiguous. In this case, it might occur that the approximation curve has a singularity area (approximated conclusions lie out of the range of Y), where clearly, the function is a too rough estimation, a result of the inexactness in the information contained in the conflicting rule base. Such a case is shown in Figure 10. Here, $\mathscr{R}(x)$ and the conclusion should be estimated by some compromise with simultaneous consideration of the conflicting rules.

A possible solution for this problem is the use of (linear) regression. For every α-cut, the best fitting straight line is computed to the points representing the rules, by the least square method. In 1 + 1 dimensions, this is defined by

$$y = ax + b = \left(\sum x_i y_i - \sum x_i \sum y_i / r\right)\left(\sum x_i^2 - \sum x_i^2 / r\right)^{-1} x + \left(\sum y_i / r - a \sum x_i / r\right)$$

A rule base with overstressedly conflicting 12 rules is shown in Figure 11*a*, $\mathscr{R}(x)$ is estimated by a straight regression line. This solution can be

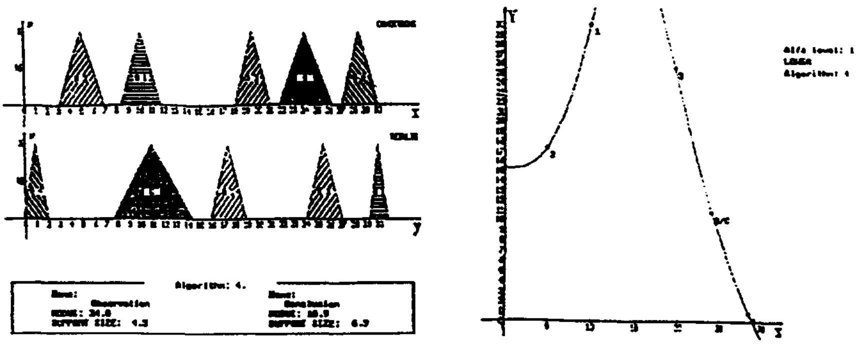

Figure 10.

interpreted as a compromise between two clearly different tendencies in the knowledge.

For compound variables (X has k_1 and Y has k_2 components), the estimation is a $k_1 \times k_2$ dimensional hyperplane. The estimation can always be decomposed into k_2 $(k_1 + 1)$-dimensional cases. Y_i is approximated by $\sum_{j=1}^{k_1} a_{ij}x_j + b_i$. The best fitting hyperplane is given by

$$a = [a_i] \quad \text{and} \quad a = \sum_i y_i/r - a^T\left[\sum_j x_{ij}\right] \text{ where}$$

$$a = \left(\left[x_{ij} - \sum_j x_{ij}/r\right]^T\left[x_{ij} - \sum_j x_{ij}/r\right]\right)^{-1}$$
$$\times\left[x_{ij} - \sum_j x_{ij}/r\right]^T\left[y_i - \sum_i y_i/r\right]$$

$i = 1 \ldots r$, $j = 1 \ldots k_1$, [ ] indicates a matrix, ^{T} is the transposed.

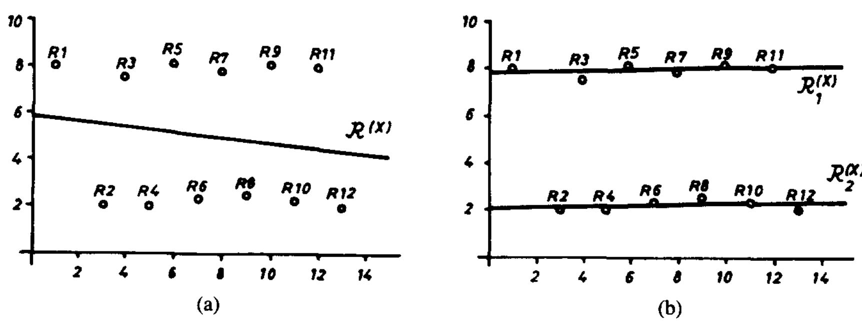

Figure 11.

As a matter of course, the linear regression gives a very rough approximation of $\mathscr{R}(x)$ if it is not linear, so it is reasonable to calculate the regression lines only locally (in a "window" around O). Then, $\mathscr{R}$ is estimated by a partially linear function.

Disadvantage of the "window" approximation is that the functions obtained in this manner are not continuous. It is possible to extend this method to fuzzy windows where smoothness of the approximation is guaranteed (see [20]).

It is questionable if the compromise estimation as in the Figure is the right way to treat a conflicting rule base. It might be more informative to detect different tendencies simultaneously and consider several alternatives for the conclusion. A promising way is offered by the use of *edge detection techniques* e.g. to recognize parallel tendencies. A fuzzy way of edge detection is proposed in [21]; here we indicate that the same technique can be applied for inference, as well. Figure $11b$ depicts the same rule base as Figure $11a$, however, the fuzzy Hough-transformation is applied, and two markant parallel tendencies of rules are detected.

There are many more possibilities to extend the idea of rule interpolation in the direction of various function approximations. The authors intend to investigate this matter further.

The use of rule interpolation and general rule approximation opens some new possibilities to the application of fuzzy expert systems and control.

References

1. Zadeh, L. A., Fuzzy algorithms. *Information and Control*, 12, 94–102, 1968.
2. Zadeh, L. A., Outline of a new approach to the analysis of complex systems. *IEEE Trans. Syst. Man Cybern.* 28–44, 1973.
3. Mamdani, E. H., Application of fuzzy algorithms for the control of a dynamic plant. *Proc. IEE* 121(12), 1585–1588, 1974.
4. Mamdani, E. H., and Assilian, S., An experiment in linguistic synthesis with a fuzzy logic controller. In *Fuzzy Reasoning and Its Applications*, (E. H. Mamdani and B. R. Gaines, Eds.), Academic Press, London, 311–323, 1981.
5. Sugeno, M., and Nishida, M., Fuzzy control of model car. *Fuzzy Sets and Systems*, 16, 103–113, 1985.
6. Hirota, K., Arai, Y., and Hachisu, Sh., Real time fuzzy pattern recognition and fuzzy controlled robot arm. *Preprints of Second IFSA Congress, Tokyo*, 274–277, 1987.
7. Kóczy, L. T., and Hirota, K., Fuzzy inference by compact rules. *Proc. Int. Conf. Fuzzy Logic & Neural Networks IIZUKA '90, Iizuka, Fukuoka*, 307–310, 1990.

8. Kóczy, L. T., Complexity of bounded compact rule based fuzzy inference. *Proc. Third Joint IFSA-EC and EURO-WG Workshop on Fuzzy Sets, Visegrád*, 59–60, 1990.

9. Kóczy, L. T., On the computational complexity of rule based fuzzy inference. *NAFIPS-91*, Columbia, Missouri, 87–91, 1991.

10. Kóczy, L. T., Computational complexity of various fuzzy inference algorithms, *Annales Univ. Sci Budapest, Sect. Comp.* 12, 151–158, 1991.

11. Mizumoto, M., and Zimmermann, H. J., Comparison of fuzzy reasoning methods. *Fuzzy Sets and Systems*, 8, 253–283, 1982.

12. Zimmermann, H. J., Fuzzy set theory—and inference mechanism. *NATO ASI Series F48, Mathematical Models for Decision Support*, (G. Mitra, Ed.), Springer, Berlin–Heidelberg, 1988, pp. 727–741.

13. Dubois, D., and Prade, H., Gradual rules in approximate reasoning. To appear in *Information Science*, 1993.

14. Dubois, D., Grabisch, M., and Prade, H., Gradual rules and the approximation of functions, *2nd Int. Conf. Fuzzy Logic and Neural Networks, Iizuka '92, Iizuka*, 629–632, 1992.

15. Türkşen, I. B., and Zhong, Z., An approximate analogical reasoning approach based on similarity measures. *IEEE Trans. Syst. Man Cybern.*, 1049–1056, 1988.

16. Mukaidono, M., Ding, L., and Shen, Z., Approximate reasoning based on Revision Principle, *NAFIPS '90*, Toronto, 94–97, 1990.

17. Ding, L., Shen, Z., and Mukaidono, M., Revision Principle for approximate reasoning—based on linear revising method, *2nd Int. Conf. Fuzzy Logic and Neural Networks, Iizuka '92, Iizuka*, 305–308, 1992.

18. Zadeh, L. A., Interpolative reasoning in fuzzy logic and neural network theory (plenary paper). *IEEE Int. Conf. Fuzzy Systems*, San Diego, 1992.

19. Kóczy, L. T., and Hirota, K., Ordering, distance and closeness of fuzzy sets. Submitted to *Fuzzy Sets and Systems*, special issue on Fuzzy Data Analysis.

20. Kóczy, L. T., Inference in fuzzy rule bases with conflicting evidence. *NAFIPS '92, NASA Conf. Publ. 10112, Vol. 2*, Puerto Vallarta, 608–616, 1992.

21. Han, J. H., Kóczy, L. T., and Poston, T., Fuzzy Hough transform for sparse data. *FUZZ-IEEE '93*, San Francisco, 803–808, 1993.