Some general considerations on the evaluation of fuzzy rule systems

Siegfried Gottwald
Universität Leipzig, Institut für Logik und Wissenschaftstheorie
Beethovenstr. 15, 04107 Leipzig, Germany
gottwald@uni-leipzig.de

Abstract

The general mathematical problem of fuzzy control is an interpolation problem: a list of fuzzy input-output data, usually provided by a list of linguistic control rules, should be realized as argument-value pairs for a suitably chosen fuzzy function. However, contrary to the usual understanding of interpolation, in the actual approaches this interpolation problem is considered as a global one: one uniformly and globally defined function should realize all the fuzzy input-output data.

In this context the paper discusses some quite general sufficient conditions for the true solution of the interpolation problem, as well as similar conditions for suitably modified data, i.e. for a quite controlled approximation.

1 Introduction

The standard paradigm of fuzzy control is that one supposes to have given, as an incomplete and fuzzy description of a control function Φ from an input space $\mathbf{X}$ to an output space $\mathbf{Y}$, a family

$$\mathcal{D} = (\langle A_i, B_i \rangle)_{1 \le i \le n} \tag{1}$$

of (fuzzy) input-output data pairs to characterize this function Φ.

In the usual approaches such a family of input-output data pairs is provided by a finite list

$$\text{IF } x \text{ is } A_i \quad \text{THEN } y \text{ is } B_i, \qquad i = 1, \dots, n \tag{2}$$

of linguistic control rules, also called fuzzy IF-THEN rules.

The main mathematical problem of fuzzy control, besides the engineering problem to get a suitable list of linguistic control rules for the actual control problem, is therefore the interpolation problem to find a function $\Phi^* : \mathcal{F}(\mathbf{X}) \to \mathcal{F}(\mathbf{Y})$ which interpolates these data, i.e. which satisfies

$$\Phi^*(A_i) = B_i \quad \text{for each } i = 1, \dots, n\,, \tag{3}$$

and which in this way gives a fuzzy representation for the control function Φ.

Actually the standard approach is to look for *one* single function which should interpolate all these data, and which should be globally defined over $\mathcal{F}(\mathbf{X})$.

This "global" interpolation problem, presented by such a finite family (1) of input-output data only, in general has different solutions. However, the main approach toward this global interpolation problem is to search for a solution in a restricted class $\mathcal{IF}$ of functions. And such a restriction of the class of interpolating functions offers also the possibility that within such a class $\mathcal{IF}$ of interpolating functions the interpolation problem becomes unsolvable.

Instead, the global interpolation problem becomes in a natural way intertwined with an *approximation problem*: one may be interested to look for a function $\Psi^* \in \mathcal{IF}$ which does not really interpolate, but which "realizes" the given fuzzy input-output data "suitably well". Such an approximative approach is completely reasonable if one has in mind that even a true solution Φ^* of the interpolation problem (3) only gives a fuzzy representation for the crisp control function Φ.

2 Two standard interpolation strategies

More or less the standard theoretical understanding for the design of a fuzzy controller is the reference to the *compositional rule of inference* (CRI) first discussed by Zadeh [11].

A suitable general context for the structure of the corresponding membership degrees, which at the same time are truth degrees of a corresponding many-valued logic, is a lattice ordered abelian monoid enriched with a further operation $\rightarrowtail$, which is connected with the semigroup operation $*$ by the adjointness condition

$$x * z \leq y \quad \text{iff} \quad z \leq (x \rightarrowtail y)\,.$$

The resulting structure often is called a *residuated lattice*. Its corresponding formalized language has besides the (idempotent) conjunction $\wedge$ which is provided by the lattice meet a further (in general not idempotent) "strong" conjunction &, which has the semigroup operation $*$ as its truth degree function.

For a full formalization one therefore would embed these considerations into the context of the basic fuzzy logic BL or the monoidal t-norm logic MTL, both explained e.g. in [5].

The previously mentioned formalized language may be further enlarged by a suitable class term notation for fuzzy sets by writing $\{x \parallel H(x)\}$ to denote that one fuzzy set A which has as its membership degree $A(a)$ in the point a of the universe of discourse just the truth degree of the formula $H(a)$.

This context yields for the CRI-based strategy, which was first applied to a control problem by Mamdani/Assilian [8], the following formulation:

From the data (A_i, B_i) one determines a fuzzy relation R in such a way that the approximating function Ψ^*_R for Φ^* becomes "describable" as

$$\Psi^*_R(A)(y) = \sup_{x\in\mathbf{X}} \big(A(x) * R(x,y)\big)\,. \qquad (4)$$

Of course, the most preferable situation would be that the function Ψ^*_R really interpolates the given input-output-data.

In general we shall call functions which can, according to (4), be represented by a fuzzy relation R simply *CRI-representable*.

A closer look at fuzzy control applications shows that one has, besides this approach via CRI-representable functions and a final application of the CRI to fuzzy input data, also a competing approach: the *method of activation degrees* which first was used by Holmblad/Ostergaard [7] in their fuzzy control algorithm for a cement kiln.

This method of activation degrees changes the previous CRI-based approach in the following way:

For each actual input fuzzy set A and each input-output data pair (A_k, B_k) one determines a modification B^*_k of its "local" output B_k, characterized only by (A_k, B_k) and the actual input A, and finally aggregates all these modified "local" outputs into one global output:

$$\Xi^*(A) \;=\; \bigcup_{i=1}^{n} B^*_i\,. \qquad (5)$$

The particular choice of Holmblad/Ostergaard for B^*_k has been

$$B^*_k(y) = \mathsf{hgt}\,(A \cap A_k) \cdot B_k(y)\,. \qquad (6)$$

Here hgt means the supremum of the membership degrees, i.e. of the range of the membership function, and $\cdot$ is the usual product.

In general terms, this modification of the first mentioned approach does not only offer one particular diverging approach toward the general interpolation problem, it also indicates that besides those both CRI-related approaches other ones with different inference and perhaps also with different aggregation operations could be of interest – as long as they are determined by finite lists of input-output data (A_i, B_i) and realize mappings from $\mathcal{F}(\mathbf{X})$ to $\mathcal{F}(\mathbf{Y})$.

This has not been done up to now in sufficient generality. Further on in this paper we shall present some considerations which point in this direction.

3 Interpolation strategies and aggregation operators

There is the well known distinction between FATI and FITA strategies to evaluate systems of linguistic control rules w.r.t. arbitrary fuzzy inputs from $\mathcal{F}(\mathbf{X})$.

The core idea of a FITA strategy is that it is a strategy which **F**irst **I**nfers (by reference to the single rules) and **T**hen **A**ggregates starting from the actual input information A. Contrary to that, a FATI strategy is a strategy which **F**irst **A**ggregates (the information in all the rules into one fuzzy relation) and **T**hen **I**nfers starting from the actual input information A.

From the two standard interpolation strategies of the last section, obviously (4) offers a FATI strategy, and (5) provides a FITA strategy.

Both these strategies use the set theoretic union as their aggregation operator. Furthermore, both of them refer to the compositional rule of inference (CRI) as their core tool of inference.

In general, however, the interpolation operators we intend to consider depend more generally upon some inference operator(s) as well as upon some aggregation operator.

By an *inference operator* we mean here simply a mapping from the fuzzy subsets of the input space to the fuzzy subsets of the output space.[1]

And an *aggregation operator* $\mathbf{A}$, as explained e.g. in [1, 2], is a family $(f^n)_{n\in\mathbb{N}}$ of ("aggregation") operations, each f^n an n-ary one, over some partially ordered set $\mathbf{M}$, with ordering $\leqslant$, with a bottom element $\mathbf{0}$ and a top element $\mathbf{1}$, such that each operation f^n is non-decreasing, maps the bottom to the bottom: $f^n(\mathbf{0},\dots,\mathbf{0}) = \mathbf{0}$, and the top to the top: $f^n(\mathbf{1},\dots,\mathbf{1}) = \mathbf{1}$.

Such an aggregation operator $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$ is *commutative* iff each operation f^n is commutative. And $\mathbf{A}$ is *associative* iff e.g. for $n = k + l$ one always has $f^n(a_1,\dots,a_n) = f^2(f^k(a_1,\dots,a_k), f^l(a_{k+1},\dots,a_n))$ and in general

$$f^n(a_1,\dots,a_n) = f^r(f^{k_1}(a_1,\dots,a_{k_1}),\dots,f^{k_r}(a_{m+1},\dots,a_n))$$

for $n = \sum_{i=1}^{r} k_i$ and $m = \sum_{i=1}^{r-1} k_i$.

Our aggregation operators further on are supposed to be commutative as well as associative ones.[2]

Observe that an associative aggregation operator $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$ is essentially determined by its binary aggregation function f^2; more precisely: by its subfamily $(f^n)_{n\leq 2}$.

Additionally we call an aggregation operator $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$

$$\begin{aligned} \textit{additive} \quad &\text{iff} \quad \text{always} \quad b \leqslant f^2(b,c)\,, \\ \textit{multiplicative} \quad &\text{iff} \quad \text{always} \quad f^2(b,c) \leqslant b\,, \\ \textit{idempotent} \quad &\text{iff} \quad \text{always} \quad b = f^2(b,b)\,. \end{aligned}$$

Corollary 1 *Let $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$ be an aggregation operator. Then one has*

(i) for idempotent $\mathbf{A}$ always $f^2(\mathbf{0},b) \leqslant b$;

(ii) for additive $\mathbf{A}$ always $b \leqslant f^2(\mathbf{0},b)$;

(iii) for multiplicative $\mathbf{A}$ always $f^2(\mathbf{0},b) = \mathbf{0}$.

If we now consider interpolation operators Φ of FITA-type and interpolation operators Ψ of FATI-type then they have the abstract forms

$$\Psi_{\mathcal{D}}(A) = \mathbf{A}(\theta_1(A),\dots,\theta_n(A))\,, \qquad (7)$$

$$\Xi_{\mathcal{D}}(A) = \widehat{\mathbf{A}}(\theta_1,\dots,\theta_n)(A)\,. \qquad (8)$$

Here we assume that each one of the "local" inference operators θ_i is determined by the single input-output pair $\langle A_i, B_i\rangle$. Therefore we also shall write $\theta_{\langle A_i,B_i\rangle}$ instead of θ_i only. And we have to assume that the aggregation operator $\mathbf{A}$ operates on fuzzy sets, and that the aggregation operator $\widehat{\mathbf{A}}$ operates on inference operators.

With this extended notation the formulas (7), (8) become

$$\Psi_{\mathcal{D}}(A) = \mathbf{A}(\theta_{\langle A_1,B_1\rangle}(A),\dots,\theta_{\langle A_n,B_n\rangle}(A)) \qquad (9)$$

$$\Xi_{\mathcal{D}}(A) = \widehat{\mathbf{A}}(\theta_{\langle A_1,B_1\rangle},\dots,\theta_{\langle A_n,B_n\rangle})(A)\,. \qquad (10)$$

[1] This terminology has its historical roots in the fuzzy control community. There is no relationship at all with the logical notion of inference intended and supposed here; but–of course–also not ruled out.

[2] It seems that this is a rather restrictive choice from a theoretical point of view. However, in all the usual cases these restrictions are satisfied.

In the previous examples (4) of a FATI and (5) of a FITA strategy, both aggregation operators $\mathbf{A}, \widehat{\mathbf{A}}$ have been the set theoretic union (of fuzzy sets, and of fuzzy relations, respectively).

4 Some particular examples

Some particular cases of these interpolation procedures have been discussed in [9]. These authors consider four different cases. First they look at the FITA-type interpolation

$$\Psi^1_{\mathcal{D}}(A) = \bigcap_i \left(A \circ (A_i \rhd B_i)\right), \qquad (11)$$

using as in [4] the notation $A_i \rhd B_i$ to denote the fuzzy relation with membership function

$$(A_i \rhd B_i)(x, y) = A_i(x) \rightarrowtail B_i(y) .$$

Their second example discusses a FATI-type approach given by

$$\Xi^2_{\mathcal{D}}(A) = A \circ \bigcap_i \left((A_i \rhd B_i)\right), \qquad (12)$$

and is thus just the common CRI-based strategy of the S-pseudo-solution, used in this general form already in [3], cf. also [4].

Their third example is again of FITA-type and determined by

$$\Psi^3_{\mathcal{D}}(A) = \bigcap_i \{y \,\|\, \delta(A, A_i) \to B_i(y)\}, \qquad (13)$$

using besides the previously mentioned class term notation for fuzzy sets the activation degree

$$\delta(A, A_i) = \bigwedge_{x \in \mathbf{X}} (A(x) \to A_i(x)) \qquad (14)$$

which is a degree of subsethood of the actual input fuzzy set A w.r.t. the i-th rule input A_i.

And the fourth one is a modification of the third one, determined for $N = \{1, 2, \ldots, n\}$ by

$$\Psi^4_{\mathcal{D}}(A) = \bigcap_{\emptyset \neq J \subseteq N} \{y \,\|\, \delta(A, \bigcup_{j \in J} A_j) \to \bigcup_{j \in J} B_i(y)\} . \qquad (15)$$

In these examples the main aggregation operators are the set theoretic union and the set theoretic intersection. Both are obviously associative, commutative, and idempotent. Additionally the union is an additive, and the intersection a multiplicative aggregation operator.

5 Stability conditions for the given data

If $\Theta_{\mathcal{D}}$ is a fuzzy inference operator of one of the types (9), (10), then the interpolation property one likes to have realized is that one has

$$\Theta_{\mathcal{D}}(A_i) = B_i \qquad (16)$$

for all the data pairs $\langle A_i, B_i \rangle$. In the particular case that the operator $\Theta_{\mathcal{D}}$ is given by (4), this is just the problem to solve the system (16) of fuzzy relation equations.

Definition 1 *In the present generalized context let us call the property (16) the* $\mathcal{D}$-stability *of the fuzzy inference operator* $\Theta_{\mathcal{D}}$*.*

To find $\mathcal{D}$-stability conditions on this abstract level seems to be rather difficult in general. However, the restriction to fuzzy inference operators of FITA-type makes things easier.

It is necessary to have a closer look at the aggregation operator $\mathbf{A} = (f^n)_{n \in \mathbb{N}}$ involved in (7) which operates on $\mathcal{F}(\mathbf{Y})$, of course with inclusion as partial ordering.

Definition 2 *Having* $B, C \in \mathcal{F}(\mathbf{Y})$ *we say that* C *is* $\mathbf{A}$-negligible w.r.t. B *iff* $f^2(B, C) = f^1(B)$ *holds true.*

The core idea here is that in any aggregation by $\mathbf{A}$ the presence of the fuzzy set B among the aggregated fuzzy sets makes any presence of C superfluous.

Examples:

1. C is $\bigcup$-negligible w.r.t. B iff $C \subseteq B$; and this holds similarly true for all idempotent and additive aggregation operators.

2. C is $\bigcap$-negligible w.r.t. B iff $C \supseteq B$; and this holds similarly true for all idempotent and multiplicative aggregation operators.

3. The bottom element $C = \mathbf{0}$ in the domain of an additive and idempotent aggregation operator $\mathbf{A}$ is $\mathbf{A}$-negligible w.r.t. any other element of that domain.

Proposition 2 *Consider a fuzzy inference operator* $\Psi_{\mathcal{D}} = \mathbf{A}(\theta_{\langle A_1,B_1\rangle}, \dots, \theta_{\langle A_n,B_n\rangle})$ *of FITA-type. It is sufficient for the* $\mathcal{D}$*-stability of* $\Psi_{\mathcal{D}}$*, that one always has*

$$\theta_{\langle A_k,B_k\rangle}(A_k) = B_k$$

and additionally that for each $i \neq k$ *the fuzzy set*

$$\theta_{\langle A_k,B_k\rangle}(A_i) \ \text{ is } \mathbf{A}\text{-negligible w.r.t. } \theta_{\langle A_k,B_k\rangle}(A_k)\,.$$

The proof follows immediately from the corresponding definitions. And this result has two interesting specializations which generalize well known results about fuzzy relation equations.

Corollary 3 *It is sufficient for the* $\mathcal{D}$*-stability of a fuzzy inference operator* $\Psi_{\mathcal{D}}$ *of FITA-type that one has* $\Psi_{\mathcal{D}}(A_i) = B_i$ *for all* $1 \leq i \leq n$ *and that always* $\theta_{\langle A_i,B_i\rangle}(A_j)$ *is* $\mathbf{A}$*-negligible w.r.t.* $\theta_{\langle A_i,B_i\rangle}(A_i)$*.*

Corollary 4 *It is sufficient for the* $\mathcal{D}$*-stability of a fuzzy inference operator* $\Psi_{\mathcal{D}}$ *of FITA-type, which is based upon an additive and idempotent aggregation operator, that one has*

$$\Psi_{\mathcal{D}}(A_i) = B_i \qquad \text{for all } 1 \leq i \leq n$$

and that always $\theta_{\langle A_i,B_i\rangle}(A_j)$ *is the bottom element in the domain of the aggregation operator* $\mathbf{A}$*.*

Obviously this is a direct generalization of the fact that systems of fuzzy relation equations are solvable if their input data form a pairwise disjoint family (w.r.t. the corresponding t-norm based intersection).

To extend these considerations from inference operators (7) of the FITA type to those ones of the FATI type (8) let us consider the following notion.

Definition 3 *Suppose that* $\widehat{\mathbf{A}}$ *is an aggregation operator for inference operators, and that* $\mathbf{A}$ *is an aggregation operator for fuzzy sets. Then* $(\widehat{\mathbf{A}}, \mathbf{A})$ *is an* application distributive *pair of aggregation operators iff*

$$\widehat{\mathbf{A}}(\theta_1, \dots, \theta_n)(X) = \mathbf{A}(\theta_1(X), \dots, \theta_n(X)) \quad (17)$$

holds true for arbitrary inference operators $\theta_1, \dots, \theta_n$ *and fuzzy sets* X*.*

Using this notion it is easy to see that one has on the left hand side of (17) a FATI type inference operator, and on the right hand side an associated FITA type inference operator. So one is able to give a reduction of the FATI case to the FITA case.

Proposition 5 *Suppose that* $(\widehat{\mathbf{A}}, \mathbf{A})$ *is an application distributive pair of aggregation operators. Then a fuzzy inference operator* $\Xi_{\mathcal{D}}$ *of FATI-type is* $\mathcal{D}$*-stable iff its associated fuzzy inference operator* $\Psi_{\mathcal{D}}$ *of FITA-type is* $\mathcal{D}$*-stable.*

6 Stability conditions for modified data

The combined approximation and interpolation problem, as previously explained, sheds new light on the standard approaches toward fuzzy control via CRI-representable functions originating from the works of Mamdani/Assilian [8] and Sanchez [10] particularly for the case that neither the Mamdani/Assilian relation R_{MA}, determined by the membership degrees

$$R_{\mathrm{MA}}(x, y) = \bigvee_{i=1}^{n} A_i(x) * B_i(y)\,, \qquad (18)$$

nor the Sanchez relation $\widehat{R}$, determined by the membership degrees

$$\widehat{R}(x, y) = \bigwedge_{i=1}^{n} (A_i(x) \rightarrowtail B_i(y))\,, \qquad (19)$$

offer a solution for the system of fuzzy relation equations.

As is well known and explained e.g. in [4], the approximating interpolation function CRI-represented by $\widehat{R}$ always gives a lower approximation, and that one CRI-represented by R_{MA} gives an upper approximation for normal input data.

Extending these results, in [6] the iterative combination of these methods has been discussed to get better approximation results. For the iterations there, always the next iteration step consisted in an application of a predetermined one of the two approximation methods to the data family with the original input data and the real, approximating output data which resulted from the application of the former approximation method. A similar iteration idea was also discussed in [9], however restricted always to the iteration of only one of the approximation methods explained in (11), (12), (13), and (15).

Therefore now we discuss the $\mathcal{D}$-stability for a modified operator $\Theta^*_{\mathcal{D}}$ which is determined by the kind of iteration of $\Theta_{\mathcal{D}}$ just explained.

Let the $\Theta_{\mathcal{D}}$-*modified* data set $\mathcal{D}^*$ be given as

$$\mathcal{D}^* = (\langle A_i, \Theta_{\mathcal{D}}(A_i)\rangle)_{1\leq i\leq n}\,, \tag{20}$$

and define the modified fuzzy inference operator $\Theta^*_{\mathcal{D}}$ as

$$\Theta^*_{\mathcal{D}} = \Theta_{\mathcal{D}^*}\,. \tag{21}$$

For these modifications, the problem of stability reappears. But it becomes a simpler one in the sense that the stability criteria now refer only to the input data A_i of the data set $\mathcal{D} = (\langle A_i, B_i\rangle)_{1\leq i\leq n}$.

Proposition 6 *It is sufficient for the $\mathcal{D}^*$-stability of a fuzzy inference operator $\Psi^*_{\mathcal{D}}$ of FITA-type that one has for all $1 \leq i \leq n$:*

$$\Psi^*_{\mathcal{D}}(A_i) = \Psi_{\mathcal{D}^*}(A_i) = \Psi_{\mathcal{D}}(A_i) \tag{22}$$

and that always $\theta_{\langle A_i,\Psi_{\mathcal{D}}(A_i)\rangle}(A_j)$ is $\mathbf{A}$-negligible w.r.t. $\theta_{\langle A_i,\Psi_{\mathcal{D}}(A_i)\rangle}(A_i)$.

Let us look separately at the conditions (22) and at the negligibility conditions.

Corollary 7 *The conditions* (22) *are always satisfied if the operator $\Psi^*_{\mathcal{D}}$ is determined by the standard output-modified system of relation equations $A_i \circ R[A_k \circ R] = B_i$ in the notation of* [6].

Corollary 8 *In the case $A = \bigcup$ the conditions* (22) *together with the inclusion relationships*

$$\theta_{\langle A_i,\Psi_{\mathcal{D}}(A_i)\rangle}(A_j) \subseteq \theta_{\langle A_i,\Psi_{\mathcal{D}}(A_i)\rangle}(A_i)$$

are sufficient for the $\mathcal{D}^*$-stability of $\Psi^*_{\mathcal{D}}$.

As in Section 5 one is able to transfer this result to FATI-type fuzzy inference operators.

Corollary 9 *Suppose that $(\widehat{\mathbf{A}}, \mathbf{A})$ is an application distributive pair of aggregation operators. Then a fuzzy inference operator $\Phi^*_{\mathcal{D}}$ of FATI-type is $\mathcal{D}^*$-stable iff its associated fuzzy inference operator $\Psi^*_{\mathcal{D}}$ of FITA-type is $\mathcal{D}^*$-stable.*

References

[1] T. Calvo, G. Mayor, R. Mesiar (eds.): *Aggregation Operators: New Trends and Applications*, Heidelberg, 2002.

[2] D. Dubois, H. Prade: On the use of aggregation operations in information fusion processes, *Fuzzy Sets Systems* **142** (2004), 143–161.

[3] S. Gottwald: Characterizations of the solvability of fuzzy equations. *Elektron. Informationsverarb. Kybernet.* **22** (1986) 67–91.

[4] S. Gottwald: *Fuzzy Sets and Fuzzy Logic.* Braunschweig/Wiesbaden and Toulouse, 1993.

[5] S. Gottwald, P. Hájek: T-norm based mathematical fuzzy logics. In: *Logical, Algebraic, Analytic, and Probabilistic Aspects of Triangular Norms* (E.P. Klement and R. Mesiar, eds.), Dordrecht, 2005, 275–299.

[6] S. Gottwald, V. Novák, I. Perfilieva: Fuzzy control and t-norm-based fuzzy logic. Some recent results, in: *Proc.* IPMU'2002, Annecy, 2002, 1087–1094.

[7] L.P. Holmblad, J.J. Ostergaard: Control of a cement kiln by fuzzy logic, in: M.M. Gupta/E. Sanchez (eds.), *Fuzzy Information and Decision Processes.* Amsterdam, 1982, 389–399.

[8] A. Mamdani, S. Assilian: An experiment in linguistic synthesis with a fuzzy logic controller, *Internat. J. Man-Mach. Studies* **7** (1975) 1–13.

[9] N.N. Morsi, A.A. Fahmy: On generalized modus ponens with multiple rules and a residuated implication, *Fuzzy Sets Systems* **129** (2002) 267–274.

[10] E. Sanchez: Resolution of composite fuzzy relation equations, *Information and Control*, **30** (1976) 38–48.

[11] L.A. Zadeh: Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Systems, Man and Cybernet.* **SMC-3** (1973) 28–44.