Many-Valued Logic Tools for Granular Modeling

Siegfried Gottwald

Leipzig University, Institute for Logic and Philosophy of Science, Beethovenstr. 15, 04107 Leipzig, Germany
gottwald@uni-leipzig.de

Abstract. A core tool for granular modeling is the use of linguistic rules, e.g. in fuzzy control approaches. We provide the reader with basic mathematical tools to discuss the behavior of system of such linguistic rules.

These mathematical tools range from fuzzy logic and fuzzy set theory, through the consideration of fuzzy relation equations, up to discussions of interpolation strategies and to the use of aggregation operators.

1 Introduction

In their everyday behavior humans quite often reason qualitatively. And they do this rather successful, even in handling complex situations. For the knowledge engineer it is an interesting and important challenge to adopt this type of behavior in his modeling activities.

Fuzzy sets and fuzzy logic have been designed and developed just for this purpose over the last decades. And they offer a rich variety of methods for this purpose: methods, which have quite different mathematical tools as origin and background.

In this chapter we concentrate on tools which are related to logic in the formal mathematical sense of the word, and to a set theoretic – particularly a relation theoretic – background.

The machinery of mathematical fuzzy logic, which is the topic of the Sections 2 and 3, offers a background to model and provides a tool to understand the treatment of vague information.

The machinery of fuzzy relation equations, discussed in Sections 4 to 6, is a core tool from the mathematical background for the understanding of fuzzy control approaches, and a prototypical case for the use of information granules.

All of these tools derive a lot of flexibility from the use of t-norms, i.e. of binary operations in the unit interval which are associative, commutative, and isotonic, and which have 1 as their neutral element.

Some further reflections on the topic of fuzzy relation equations provide additional mathematical tools to understand them: the interpolation method, as well as the use of aggregation operators.

A. Bargiela, W. Pedrycz (Eds.): Human-Centric Information Processing, SCI 182, pp. 153–184.
springerlink.com

2 Fuzzy Sets and Many-Valued Logic

2.1 Membership Degrees as Truth Degrees

A *fuzzy set* A is characterized by its generalized characteristic function $\mu_A : \mathbb{X} \to [0,1]$, called *membership function* of A and defined over some given *universe of discourse* $\mathbb{X}$, i.e. it is a *fuzzy subset* of $\mathbb{X}$.

The essential idea behind this approach was to have the membership degree $\mu_A(a)$ for each point $a \in \mathbb{X}$ as a gradation of its membership with respect to the fuzzy set A. And this degree just is a degree to which the sentence *"a is a member of A"* holds true. Hence it is natural to interpret the membership degrees of fuzzy sets as truth degrees of the membership predicate in some (suitable system of) many-valued logic S.

To do this in a reasonable way one has to accept some minimal conditions concerning the formal language $\mathcal{L}_\mathsf{S}$ of this system.

Disregarding – for simplicity – fuzzy sets of type 2 and of every higher type as well one has, from the set theoretic point of view, fuzzy sets as (generalized) sets of first level over a given class of urelements, the universe of discourse for these fuzzy sets. Therefore the intended language needs besides a (generalized, i.e. graded) binary membership predicate ε e.g. two types of variables: (i) lower case latin letters $a, b, c, \ldots, x, y, z$ for urelements, i.e. for points of the universe of discourse $\mathbb{X}$, and (ii) upper case latin letters $A, B, C, \ldots$ for fuzzy subsets of $\mathbb{X}$. And of course it has some set of connectives and some quantifiers – and thus a suitable notion of well-formed formula.

Having in mind the standard fuzzy sets with membership degrees in the real unit interval $[0,1]$ thus forces to assume that S is an infinitely many-valued logic.

It is not necessary to fix all the details of the language $\mathcal{L}_\mathsf{S}$ in advance. We suppose, for simplicity of notation, that from the context it shall always be clear which objects the individual symbols are to denote.[1] Denoting the truth degree of a well-formed formula H by $[\![H]\!]$, to identify membership degrees with suitable truth degrees then means to put

$$\mu_A(x) = [\![\, x \,\varepsilon\, A \,]\!]. \qquad (1)$$

This type of interpretation proves quite useful: it opens the doors to clarify far reaching analogies between notions and results related to fuzzy sets and those ones related to usual sets, cf. [14, 16].

2.2 Doing Fuzzy Set Theory Using MVL Language

Based on the main idea to look at the membership degrees of fuzzy sets as truth degrees of a suitable membership predicate, it e.g. becomes quite natural to describe fuzzy sets by a (generalized) class term notation, adapting the corresponding notation $\{x \mid H(x)\}$ from traditional set theory and introducing a corresponding notation for fuzzy sets by

[1] Which, formally, means that we assume that a valuation always is determined by the context and has not explicitly to be mentioned.

$$A = \{x \in \mathbb{X} \parallel H(x)\} \ \Leftrightarrow_{\text{def}} \ \mu_A(x) = [\![H(x)]\!] \quad \text{for all } x \in \mathbb{X}, \tag{2}$$

with now H a well-formed formula of the language $\mathcal{L}_{\mathsf{S}}$. As usual, the shorter notation $\{x \parallel H(x)\}$ is also used, and even preferred.

2.2.1 Fuzzy Set Algebra

With this notation the intersection and the cartesian product of fuzzy sets A, B are, even in their t-norm based version, determined as

$$A \cap_{\boldsymbol{t}} B = \{x \parallel x \,\varepsilon\, A \wedge_{\boldsymbol{t}} x \,\varepsilon\, B\},$$
$$A \times_{\boldsymbol{t}} B = \{(x, y) \parallel x \,\varepsilon\, A \wedge_{\boldsymbol{t}} y \,\varepsilon\, B\},$$

and the standard, i.e. min-based form of the *compositional rule of inference*[2] (CRI for short) applied w.r.t. a fuzzy relation R and a fuzzy set A becomes

$$A \circ R = R''A = \{y \parallel \exists x(x \,\varepsilon\, A \wedge_{\boldsymbol{t}} (x, y) \,\varepsilon\, R)\} \tag{3}$$

with $\boldsymbol{t} = \min$. This is the analogue of a formula well known from elementary relation algebra which describes the full image of a set A under a relation R.

Of course, also other choices of the t-norm involved here are possible.

Also for the inclusion relation between fuzzy sets this approach works well. The standard definition of inclusion amounts to

$$A \subset B \quad \Leftrightarrow \quad \mu_A(x) \le \mu_B(x) \ \text{ for all } \ x \in \mathbb{X}$$

which in the language of many-valued logic is the same as

$$A \subset B \quad \Leftrightarrow \quad \models \forall x(x \,\varepsilon\, A \rightarrow_{\boldsymbol{t}} x \,\varepsilon\, B) \tag{4}$$

w.r.t. any one R-implication connective based on a left continuous t-norm.

Obviously this version (4) of inclusion is easily generalized to a "fuzzified", i.e. (truly) many-valued inclusion relation defined as

$$A \subseteqq B =_{\text{def}} \forall x(x \,\varepsilon\, A \rightarrow_{\boldsymbol{t}} x \,\varepsilon\, B). \tag{5}$$

And this many-valued inclusion relation for fuzzy sets has still nice properties, e.g. it is $\boldsymbol{t}$-transitive, i.e. one has:

$$\models \ (A \subseteqq B \wedge_{\boldsymbol{t}} B \subseteqq C \rightarrow_{\boldsymbol{t}} A \subseteqq C).$$

2.2.2 Fuzzy Relation Theory

This natural approach (3) toward the compositional rule of inference is almost the same as the usual definition of the relational product $R \circ S$ of two fuzzy relation, now – even related to some suitable t-norm – to be determined as

[2] This compositional rule of inference is of central importance for the applications of fuzzy sets to fuzzy control and to approximate reasoning, cf. Chapter 1 of the Handbook volume "Fuzzy Sets in Approximate Reasoning and Information Systems" edited by J. Bezdek, D. Dubois and H. Prade.

$$R \circ_{\boldsymbol{t}} S = \{(x,y) \parallel \exists z\big((x,z)\,\varepsilon\, R \wedge_{\boldsymbol{t}} (z,y)\,\varepsilon\, S\big)\}.$$

Relation properties become, in this context, again characterizations which formally read as the corresponding properties of crisp sets. Consider, as an example, transitivity of a fuzzy (binary) relation R in the universe of discourse $\mathbb{X}$ w.r.t. some given t-norm. The usual condition for all $x, y, z \in \mathbb{X}$

$$\boldsymbol{t}(\mu_R(x,y), \mu_R(y,z)) \leq \mu_R(x,z)$$

in the language $\mathcal{L}_{\mathsf{S}}$ of the intended suitable $[0,1]$-valued system for fuzzy set theory becomes the condition

$$\models R(x,y) \wedge_{\boldsymbol{t}} R(y,z) \rightarrow_{\boldsymbol{t}} R(x,z)$$

for all $x, y, z \in \mathbb{X}$ or, even better, becomes

$$\models \forall xyz\big(R(x,y) \wedge_{\boldsymbol{t}} R(y,z) \rightarrow_{\boldsymbol{t}} R(x,z)\big) \tag{6}$$

with the universal quantifier $\forall$ as in the Łukasiewicz systems.

This point of view not only opens the way for a treatment of fuzzy relations quite analogous to the usual discussion of properties of crisp relations, it also opens the way to consider *graded versions* of properties of fuzzy relations, cf. [14]. In the case of transitivity, a graded or "fuzzified" predicate Trans with the intended meaning "is transitive" may be defined as

$$\mathsf{Trans}(R) =_{\mathrm{def}} \forall xyz\big(R(x,y) \wedge_{\boldsymbol{t}} R(y,z) \rightarrow_{\boldsymbol{t}} R(x,z)\big). \tag{7}$$

This point of view has recently been treated in more detail e.g. in [1].

3 T-Norm-Based Mathematical Fuzzy Logics

3.1 Basic Infinite Valued Logics

If one looks for infinite valued logics of the kind which is needed as the underlying logic for a theory of fuzzy sets, one finds three main systems:

- the Łukasiewicz logic L as explained in [31];
- the Gödel logic G from [11];
- the product logic $\mathit{\Pi}$ studied in [23].

In their original presentations, these logics look rather different, regarding their propositional parts. For the first order extensions, however, there is a unique strategy: one adds a universal and an existential quantifier such that quantified formulas get, respectively, as their truth degrees the infimum and the supremum of all the particular cases in the range of the quantifiers.

As a reference for these and also other many-valued logics in general, the reader may consult [16].

3.1.1 Gödel Logic

The simplest one of these logics is the *Gödel logic* G which has a conjunction $\wedge$ and a disjunction $\vee$ defined by the minimum and the maximum, respectively, of the truth degrees of the constituents:

$$u \wedge v = \min\{u, v\}\,, \qquad u \vee v = \max\{u, v\}\,. \tag{8}$$

For simplicity we denote here and later on the connectives and the corresponding truth degree functions by the same symbol.

The Gödel logic has also a negation $\sim$ and an implication $\to_{\mathsf{G}}$ defined by the truth degree functions

$$\sim u = \begin{cases} 1\,, \text{ if } & u = 0; \\ 0\,, \text{ if } & u > 0. \end{cases} \qquad u \to_{\mathsf{G}} v = \begin{cases} 1\,, \text{ if } & u \le v; \\ v\,, \text{ if } & u > v. \end{cases} \tag{9}$$

3.1.2 Łukasiewicz Logic

The *Łukasiewicz logic* L was originally designed in [31] with only two primitive connectives, an implication $\to_{\mathsf{L}}$ and a negation $\neg$ characterized by the truth degree functions

$$\neg u = 1 - u\,, \qquad u \to_{\mathsf{L}} v = \min\{1, 1 - u + v\}\,. \tag{10}$$

However, it is possible to define further connectives from these primitive ones. With

$$\varphi \,\&\, \psi =_{\mathrm{df}} \neg(\varphi \to_{\mathsf{L}} \neg\psi)\,, \qquad \varphi \,\veebar\, \psi =_{\mathrm{df}} \neg\varphi \to_{\mathsf{L}} \psi \tag{11}$$

one gets a (strong) conjunction and a (strong) disjunction with truth degree functions

$$u \,\&\, v = \max\{u + v - 1, 0\}\,, \qquad u \veebar v = \min\{u + v, 1\}\,, \tag{12}$$

usually called the *Łukasiewicz* (arithmetical) *conjunction* and the *Łukasiewicz* (arithmetical) *disjunction*. It should be mentioned that these connectives are linked together via a De Morgan law using the standard negation of this system:

$$\neg(u \,\&\, v) = \neg u \veebar \neg v\,. \tag{13}$$

With the additional definitions

$$\varphi \wedge \psi =_{\mathrm{df}} \varphi \,\&\, (\varphi \to_{\mathsf{L}} \psi) \qquad \varphi \vee \psi =_{\mathrm{df}} (\varphi \to_{\mathsf{L}} \psi) \to_{\mathsf{L}} \psi \tag{14}$$

one gets another (weak) conjunction $\wedge$ with truth degree function min, and a further (weak) disjunction $\vee$ with max as truth degree function, i.e. one has the conjunction and the disjunction of the Gödel logic also available.

3.1.3 Product Logic

The *product logic* Π, in detail explained in [23], has a fundamental conjunction $\odot$ with the ordinary product of reals as its truth degree function, as well as an implication $\to_{\Pi}$ with truth degree function

$$u \to_{\Pi} v = \begin{cases} 1\,, & \text{if } u \le v; \\ \frac{u}{v}\,, & \text{if } u < v. \end{cases} \tag{15}$$

Additionally it has a truth degree constant $\overline{0}$ to denote the truth degree zero.

In this context, a negation and a further conjunction are defined as

$$\sim \varphi =_{\mathrm{df}} \varphi \to_{\Pi} \overline{0}\,, \qquad \varphi \wedge \psi =_{\mathrm{df}} \varphi \odot (\varphi \to_{\Pi} \psi)\,. \tag{16}$$

Routine calculations show that both connectives coincide with the corresponding ones of the Gödel logic. And also the disjunction $\vee$ of the Gödel logic becomes available, now via the definition

$$\varphi \vee \psi =_{\mathrm{df}} ((\varphi \to_{\Pi} \psi) \to_{\Pi} \psi) \wedge ((\psi \to_{\Pi} \varphi) \to_{\Pi} \varphi)\,. \tag{17}$$

3.2 Standard and Algebraic Semantics

These fundamental infinite valued logics have their *standard semantics* as explained: the real unit interval $[0,1]$ as truth degree set, and the connectives (and quantifiers) as mentioned.

In the standard way, as known from classical logic, one then can introduce for each formula φ the notion of its *validity in a model*, which in these logics means that φ has the truth degree 1 w.r.t. this model. By a *model* we mean either—in the propositional case—an evaluation of the propositional variables by truth degrees, or—in the first-order case—a suitable interpretation of all the non-logical constants together with an assignment of the variables.

Based upon this, one defines *logical validity* of a formula φ as validity of φ in each model, and the *entailment* relation holds between a set Σ of formulas and a formula φ iff each model of Σ is also a model of φ.

In the standard terminology of many-valued logic in general, this means that all the three systems $\mathsf{G}, \mathsf{L}, \Pi$ have the truth degree one as their only designated truth degree.

Besides these standard semantics, all three of these basic infinite valued logics have also *algebraic semantics* determined by suitable classes $\mathcal{K}$ of truth degree structures. The situation is similar here to the case of classical logic: the logically valid formulas in classical logic are also just all those formulas which are valid in all Boolean algebras.

Of course, these structures should have the same signature as the language $\mathcal{L}$ of the corresponding logic. This means that these structures provide for each connective of the language $\mathcal{L}$ an operation of the same arity, and they have to have—in the case that one discusses the corresponding first order logics—suprema and infima for all those subsets which may appear as value sets of formulas. Particularly, hence, they have to be (partially) ordered, or at least pre-ordered.

For each formula φ of the language $\mathcal{L}$ of the corresponding logic, for each such structure $\mathbf{A}$, and for each evaluation e which maps the set of propositional variables of $\mathcal{L}$ into the carrier of $\mathbf{A}$, one has to define a value $e(\varphi)$, and finally one has to define what it means that such a formula φ is *valid in* $\mathbf{A}$. Then a formula φ is *logically valid* w.r.t. this class $\mathcal{K}$ iff φ is valid in all structures from $\mathcal{K}$.

The standard way to arrive at such classes of structures is to start from the Lindenbaum algebra of the corresponding logic, i.e. its algebra of formulas modulo the congruence relation of logical equivalence. For this Lindenbaum algebra one then has to determine a class of similar algebraic structures which—ideally—forms a variety.

For the Gödel logic such a class of structures is, according to the completeness proof of [8], the class of all *Heyting algebras*, i.e. of all relatively pseudo-complemented lattices, which satisfy the prelinearity condition

$$(u \rightarrowtail v) \sqcup (v \rightarrowtail u) = \mathbf{1}\,. \tag{18}$$

Here $\sqcup$ is the lattice join and $\rightarrowtail$ the relative pseudo-complement.

For the Łukasiewicz logic the corresponding class of structures is the class of all *MV-algebras*, first introduced again within a completeness proof in [3], and extensively studied in [5].

And for the product logic the authors of [23] introduce a class of lattice ordered semigroups which they call *product algebras*.

It is interesting to recognize that all these structures—pre-linear Heyting algebras, MV-algebras, and product algebras—are abelian lattice ordered semigroups with an additional "residuation" operation.

3.3 Logics with t-Norm Based Connectives

The fundamental infinite valued logics from Section 3.1 look quite different if one has in mind the form in which they first were presented.

Fortunately, however, there is a common generalization which allows to present all these three logics in a uniform way. In this uniform presentation one of the conjunction connectives becomes a core role: $\wedge$ in the system G, & in the system L, and $\odot$ in the system $\mathit{\Pi}$.

But this uniform generalization covers a much larger class of infinite valued logics over $[0,1]$: the core conjunction connective—which shall now in general be denoted &—has only to have a truth degree function $\otimes$ which, as a binary operation in the real unit interval, should be an associative, commutative, and isotonic operation which has 1 as a neutral element, i.e. should satisfy for arbitrary $x, y, z \in [0,1]$:

(T1) $x \otimes (y \otimes z) = (x \otimes y) \otimes z$,
(T2) $x \otimes y = y \otimes x$,
(T3) if $x \leq y$ then $x \otimes z \leq y \otimes z$,
(T4) $x \otimes 1 = x$.

Such binary operations are known as *t-norms* and have been used in the context of probabilistic metric spaces, cf. e.g. [29]. At the same time they are considered as natural candidates for truth degree functions of conjunction connectives. And from such a t-norm one is able to derive (essentially) all the other truth degree functions for further connectives.

The minimum operation $u \wedge v$ from (8), the Łukasiewicz arithmetic conjunction $u \,\&\, v$ from (12), and the ordinary product are the best known examples of t-norms.

In algebraic terms, such a t-norm $\otimes$ makes the real unit interval into an *ordered monoid*, i.e. into an abelian semigroup with unit element. And this ordered monoid is even *integral*, i.e. its unit element is at the same time the universal upper bound of the ordering. Additionally this monoid has because of

$$0 \otimes x \leq 0 \otimes 1 = 0 \tag{19}$$

the number 0 as an *annihilator*.

Starting from a t-norm $\otimes$ one finds a truth degree function $\rightarrowtail$ for an implication connective via the *adjointness condition*

$$x \otimes z \leq y \quad\Longleftrightarrow\quad z \leq (x \rightarrowtail y)\,. \tag{20}$$

However, to guarantee that this adjointness condition (20) determines the operation $\rightarrowtail$ uniquely, one has to assume that the t-norm $\otimes$ is a *left continuous* function in both arguments. Indeed, the adjointness condition (20) is equivalent to the condition that $\otimes$ is left continuous in both arguments, cf. [16].

Instead of this adjointness condition (20) one could equivalently either give the direct definition

$$x \rightarrowtail y = \sup\{z \mid x \otimes z \leq y\} \tag{21}$$

of the *residuation operation* $\rightarrowtail$, or one could force the t-norm $\otimes$ to have the *sup-preservation property*

$$\sup_{i\to\infty} (x_i \otimes y) = (\sup_{i\to\infty} x_i) \otimes y \tag{22}$$

for each $y \in [0,1]$ and each non-decreasing sequence $(x_i)_{i\to\infty}$ from the real unit interval.

In this framework one additionally introduces a further unary operation $-$ by

$$-x =_{\mathrm{df}} x \rightarrowtail 0\,, \tag{23}$$

and considers this as the truth degree function of a negation connective. That this works also in the formalized language of the corresponding system of logic forces to introduce into this language a truth degree constant $\overline{0}$ to denote the truth degree zero.

And finally one likes to have the weak conjunction and disjunction connectives $\wedge, \vee$ available. These connectives should also be added to the vocabulary. However, it suffices to add only the min-conjunction $\wedge$, because then for each left continuous t-norm $\otimes$ and its residuated implication $\rightarrowtail$ one has, completely similar to the situation (17) in the product logic,

$$u \vee v = ((u \rightarrowtail v) \rightarrowtail v) \wedge ((v \rightarrowtail u) \rightarrowtail u)\,. \tag{24}$$

All these considerations lead in a natural way to algebraic structures which, starting from the unit interval, consider a left continuous t-norm $\otimes$ together with its residuation operation $\rightarrowtail$, with the minimum-operation $\wedge$, and the maximum operation $\vee$ as basic operations of such an algebraic structure, and with the particular truth degrees $0, 1$ as fixed objects (i.e. as nullary operations) of the structure. Such an algebraic structure

$$\langle [0,1], \wedge, \vee, \otimes, \rightarrowtail, 0, 1 \rangle \tag{25}$$

shall be coined to be a *t-norm algebra.*

3.4 Continuous t-Norms

Among the large class of all t-norms the continuous ones are best understood. A t-norm is continuous iff it is continuous as a real function of two variables, or equivalently, iff it is continuous in each argument (with the other one as a parameter), cf. [16, 29].

Furthermore, all continuous t-norms are ordinal sums of only three of them: the Łukasiewicz arithmetic t-norm $u \,\&\, v$ from (12), the ordinary product t-norm, and the minimum operation $u \wedge v$. The definition of an ordinal sum of t-norms is the following one.

Definition 1. *Suppose that* $([a_i, b_i])_{i \in I}$ *is a countable family of non-overlapping proper subintervals of the unit interval* $[0,1]$*, let* $(\boldsymbol{t}_i)_{i \in I}$ *be a family of t-norms, and let* $(\varphi_i)_{i \in I}$ *be a family of mappings such that each* φ_i *is an order isomorphism from* $[a_i, b_i]$ *onto* $[0,1]$*. Then the* (*generalized*) ordinal sum *of the combined family* $(([a_i, b_i], \boldsymbol{t}_i, \varphi_i))_{i \in I}$ *is the binary function* $T : [0,1]^2 \to [0,1]$ *characterized by*

$$T(u,v) = \begin{cases} {\varphi_k}^{-1}(\boldsymbol{t}_k(\varphi_k(u), \varphi_k(v)), & \text{if } u, v \in [a_k, b_k] \\ \min\{u, v\} & \text{otherwise.} \end{cases} \tag{26}$$

It is easy to see that an order isomorphic copy of the minimum t-norm is again the minimum operation. Thus the whole construction of ordinal sums of t-norms even allows to assume that the summands are formed from t-norms different from the minimum t-norm. This detail, however, shall be inessential for the present considerations.

But it should be mentioned that all the endpoints a_i, b_i of the interval family $([a_i, b_i])_{i \in I}$ give *idempotents* of the resulting ordinal sum t-norm T:

$$T(a_i, a_i) = a_i\,, \quad T(b_i, b_i) = b_i \quad \text{for all } i \in I.$$

Conversely, if one knows all the idempotents of a given continuous t-norm $\boldsymbol{t}$, i.e. all $u \in [0,1]$ with $\boldsymbol{t}(u,u) = u$, then one is able to give a representation of $\boldsymbol{t}$ as an ordinal sum, as explained again in [29].

The general result, given e.g. in [16, 29], reads as follows.

Theorem 1. *Each continuous t-norm* $\boldsymbol{t}$ *is the* (*generalized*) *ordinal sum of* (*isomorphic*) *copies of the Łukasiewicz t-norm, the product t-norm, and the minimum t-norm.*

As was mentioned in Section 3.1, the t-norm based logics which are determined by these three t-norms are well known and adequately axiomatized.

Therefore one is interested to find adequate axiomatizations also for further continuous t-norms. A global solution of this problem, i.e. a solution which did not only cover some few particular cases, appeared as quite difficult. Therefore, instead, one first has been interested to find all those formulas of the language of t-norm based systems which are logically valid in each one of these logics.

There seems to be a natural way to get an algebraic semantics for these considerations: the class of all t-norm algebras with a continuous t-norm should either form such an algebraic semantics, or should be a constitutive part—preferably a generating set—of a variety of algebraic structures which form such an algebraic semantics.

However, there seems to be an inadequacy in the description of this algebraic semantics: on the one hand the notion of t-norm algebra is a purely algebraic notion, the notion of continuity of a t-norm on the other hand is an analytical one. Fortunately, there is a possibility to give an algebraic characterization for the continuity of t-norms. It needs a further notion.

Definition 2. *A t-norm algebra* $\langle [0,1], \wedge, \vee, \otimes, \rightarrowtail, 0, 1 \rangle$ *is* divisible *iff one has for all* $a, b \in L$*:*

$$a \wedge b = a \otimes (a \rightarrowtail b)\,. \tag{27}$$

And this notion gives the algebraic counterpart for the continuity, as shown e.g. in [16, 29].

Proposition 1. *A t-norm algebra* $\langle [0,1], \wedge, \vee, \otimes, \rightarrowtail, 0, 1 \rangle$ *is divisible iff the t-norm* $\otimes$ *is continuous.*

3.5 The Logic of Continuous t-Norms

Instead of considering for each particular t-norm $\boldsymbol{t}$ the $\boldsymbol{t}$-based logic, it seems preferable and more interesting to consider the common logic of all t-norms of some kind. This was first realized for the class of all continuous t-norms by Hájek [22]. This logic should have as a natural *standard semantics* the class of all t-norm algebras with a continuous t-norm.

However, to built up a logic with an algebraic semantics determined by a class $\mathcal{K}$ of algebraic structures becomes quite natural in the cases that this class $\mathcal{K}$ is a variety: i.e. a class which is equationally definable—or equivalently, in more algebraic terms, which is closed under forming direct products, substructures, and homomorphic images.

Unfortunately, the class of t-norm algebras (with a continuous t-norm or not) is not a variety: it is not closed under direct products because each t-norm algebra is linearly ordered, but the direct products of linearly ordered structures are not linearly ordered, in general. Hence one may expect that it would be helpful for the development of a logic of continuous t-norms to extend the class of all divisible t-norm algebras in a moderate way to get a variety. And indeed this idea works, and is in detail explained in [22].

The core points are that one considers instead of the divisible t-norm algebras, which are linearly ordered integral monoids as mentioned previously, now lattice ordered integral monoids which are divisible, which have an additional residuation operation connected with the semigroup operation via an adjointness condition like (20), and which satisfy a pre-linearity condition like (18). These structures have been called BL-algebras; they are completely defined in the following way.

Definition 3. *A* BL-algebra $\mathbf{L} = \langle L, \vee, \wedge, *, \to, \mathbf{0}, \mathbf{1}\rangle$ *is an algebraic structure such that*

(*i*) $(L, \vee, \wedge, \mathbf{0}, \mathbf{1})$ *is a bounded lattice, i.e. has* $\mathbf{0}$ *and* $\mathbf{1}$ *as the universal lower and upper bounds w.r.t. the lattice ordering* $\leq$,

(*ii*) $(L, *, \mathbf{1})$ *is an abelian monoid, i.e. a commutative semigroup with unit* $\mathbf{1}$ *such that the multiplication* $*$ *is associative, commutative and satisfies* $\mathbf{1} * x = x$ *for all* $x \in L$,

(*iii*) *the binary operations* $*$ *and* $\to$ *form an adjoint pair, i.e. satisfy for all* $x, y, z \in L$ *the* adjointness *condition*

$$z \leq (x \to y) \iff x * z \leq y, \tag{28}$$

(*iv*) *and moreover, for all* $x, y \in L$ *one has satisfied the* pre-linearity *condition*

$$(x \to y) \vee (y \to x) = \mathbf{1} \tag{29}$$

as well as the divisibility *condition*

$$x * (x \to y) = x \wedge y\,. \tag{30}$$

The axiomatization of Hájek [22] for the *basic t-norm logic* BL (in [16] denoted BTL), i.e. for the class of all well-formed formulas which are valid in all BL-algebras, is given in a language $\mathcal{L}_T$ which has as basic vocabulary the connectives $\to, \&$ and the truth degree constant $\overline{0}$, taken in each BL-algebra $\langle L, \cap, \cup, *, \rightarrowtail, 0, 1\rangle$ as the operations $\rightarrowtail, *$ and the element 0. Then this t-norm based logic has as axiom system $\mathsf{Ax}_{\mathsf{BL}}$ the following schemata:

($\mathsf{Ax}_{\mathsf{BL}}1$) $(\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))\,,$
($\mathsf{Ax}_{\mathsf{BL}}2$) $\varphi \,\&\, \psi \to \varphi\,,$
($\mathsf{Ax}_{\mathsf{BL}}3$) $\varphi \,\&\, \psi \to \psi \,\&\, \varphi\,,$
($\mathsf{Ax}_{\mathsf{BL}}4$) $(\varphi \to (\psi \to \chi)) \to (\varphi \,\&\, \psi \to \chi)\,,$
($\mathsf{Ax}_{\mathsf{BL}}5$) $(\varphi \,\&\, \psi \to \chi) \to (\varphi \to (\psi \to \chi))\,,$
($\mathsf{Ax}_{\mathsf{BL}}6$) $\varphi \,\&\, (\varphi \to \psi) \to \psi \,\&\, (\psi \to \varphi)\,,$
($\mathsf{Ax}_{\mathsf{BL}}7$) $((\varphi \to \psi) \to \chi) \to (((\psi \to \varphi) \to \chi) \to \chi)\,,$
($\mathsf{Ax}_{\mathsf{BL}}8$) $\overline{0} \to \varphi\,,$

and has as its (only) inference rule the rule of detachment, or: modus ponens (w.r.t. the implication connective $\to$).

The logical calculus which is constituted by this axiom system and its inference rule, and which has the standard notion of derivation, shall be denoted by $\mathbb{K}_{\mathsf{BL}}$ or just by BL. (Similarly in other cases.)

Starting from the primitive connectives $\to, \&$ and the truth degree constant $\overline{0}$, the language $\mathcal{L}_T$ of BL is extended by definitions of additional connectives $\wedge, \vee, \neg$:

$$\varphi \wedge \psi =_{\text{df}} \varphi \,\&\, (\varphi \to \psi)\,, \tag{31}$$
$$\varphi \vee \psi =_{\text{df}} ((\varphi \to \psi) \to \psi) \wedge ((\psi \to \varphi) \to \varphi)\,, \tag{32}$$
$$\neg\varphi =_{\text{df}} \varphi \to \overline{0}\,. \tag{33}$$

These additional connectives $\wedge, \vee$ just have the lattice operations $\cap, \cup$ as their truth degree functions.

It is a routine matter, but a bit tedious, to check that this logical calculus $\mathbb{K}_{\mathsf{BL}}$, usually called the axiomatic system BL, is sound, i.e. derives only such formulas which are valid in all BL-algebras. A proof is given in [22], together with a proof of a corresponding completeness theorem.

Corollary 1. *The Lindenbaum algebra of the axiomatic system* BL *is a BL-algebra.*

Theorem 2 (General Completeness). *A formula φ of the language $\mathcal{L}_T$ is derivable within the axiomatic system* BL *iff φ is valid in all BL-algebras.*

The proof method yields that each BL-algebra is (isomorphic to) a subdirect product of linearly ordered BL-algebras, i.e. of BL-chains. Thus it allows a nice modification of the previous result.

Corollary 2 (General Completeness; Version 2). *A formula φ of $\mathcal{L}_T$ is derivable within the axiomatic system* BL *iff φ is valid in all BL-chains.*

But even more is provable and leads back to the starting point of the whole approach: the logical calculus $\mathbb{K}_{\mathsf{BL}}$ characterizes just those formulas which hold true w.r.t. all divisible t-norm algebras. This was proved in [4].

Theorem 3 (Standard Completeness). *The class of all formula which are provable in the system* BL *coincides with the class of all formulas which are logically valid in all t-norm algebras with a continuous t-norm.*

And another generalization of Theorem 2 deserves to be mentioned. To state it, let us call *schematic extension* of BL every extension which consists in an addition of finitely many axiom schemata to the axiom schemata of BL. And let us denote such an extension by $\mathsf{BL}(\mathcal{C})$. And call BL($\mathcal{C}$)-algebra each BL-algebra $\mathbf{A}$ which makes $\mathbf{A}$-valid all formulas of $\mathcal{C}$.

Then one can prove, as done in [22], an even more general completeness result.

Theorem 4 (Extended General Completeness). *For each finite set $\mathcal{C}$ of axiom schemata and any formula φ of $\mathcal{L}_T$ there are equivalent:*

(i) φ is derivable within $\mathsf{BL}(\mathcal{C})$*;*
(ii) φ is valid in all $\mathsf{BL}(\mathcal{C})$*-algebras;*
(iii) φ is valid in all $\mathsf{BL}(\mathcal{C})$*-chains.*

The extension of these considerations to the first-order case is also given in [22], but shall not be discussed here.

But the algebraic machinery allows even deeper insights. After some particular results e.g. in [24, 25], the study of such subvarieties of the variety of all BL-algebras which are generated by single t-norm algebras of the form $\langle [0,1], \wedge, \vee, \otimes, \rightarrowtail, 0, 1\rangle$ with a continuous t-norm $\otimes$ led to (finite) axiomatizations of those t-norm based logics which have a standard semantics determined just by this continuous t-norm algebra. These results have been presented in [10].

3.6 The Logic of Left Continuous t-Norms

The guess of Esteva/Godo [9] has been that one should arrive at the logic of left continuous t-norms if one starts from the logic of continuous t-norms and deletes the continuity condition, i.e. the divisibility condition (27).

The algebraic approach needs only a small modification: in the Definition 3 of BL-algebras one has simply to delete the divisibility condition (30). The resulting algebraic structures have been called *MTL-algebras*. They again form a variety.

Following this idea, one has to modify the previous axiom system in a suitable way. And one has to delete the definition (31) of the connective $\wedge$, because this definition (together with suitable axioms) essentially codes the divisibility condition. The definition (32) of the connective $\vee$ remains unchanged.

As a result one now considers a new system MTL of mathematical fuzzy logic, known as *monoidal t-norm logic*, characterized semantically by the class of all MTL-algebras. It is connected with the axiom system

$$
\begin{array}{ll}
(\mathsf{Ax}_{\mathsf{MTL}}1) & (\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}2) & \varphi \,\&\, \psi \to \varphi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}3) & \varphi \,\&\, \psi \to \psi \,\&\, \varphi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}4) & (\varphi \to (\psi \to \chi)) \to (\varphi \,\&\, \psi \to \chi)\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}5) & (\varphi \,\&\, \psi \to \chi) \to (\varphi \to (\psi \to \chi))\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}6) & \varphi \wedge \psi \to \varphi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}7) & \varphi \wedge \psi \to \psi \wedge \varphi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}8) & \varphi \,\&\, (\varphi \to \psi) \to \varphi \wedge \psi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}9) & \overline{0} \to \varphi\,, \\
(\mathsf{Ax}_{\mathsf{MTL}}10) & ((\varphi \to \psi) \to \chi) \to (((\psi \to \varphi) \to \chi) \to \chi)\,,
\end{array}
$$

together with the rule of detachment (w.r.t. the implication connective $\to$) as (the only) inference rule.

It is a routine matter, but again tedious, to check that this logical calculus $\mathbb{K}_{\mathsf{MTL}}$ is sound, i.e. derives only such formulas which are valid in all MTL-algebras.

Corollary 3. *The Lindenbaum algebra of the logical calculus* $\mathbb{K}_{\mathsf{MTL}}$ *is an MTL-algebra.*

Proofs of this result and also of the following completeness theorem are given in [9].

Theorem 5 (General Completeness). *A formula φ of the language $\mathcal{L}_T$ is derivable within the logical calculus $\mathbb{K}_{\mathsf{MTL}}$ iff φ is valid in all MTL-algebras.*

Again the proof method yields that each MTL-algebra is (isomorphic to) a subdirect product of linearly ordered MTL-algebras, i.e. of MTL-chains.

Corollary 4 (General Completeness; Version 2). *A formula φ of $\mathcal{L}_T$ is derivable within the axiomatic system* MTL *iff φ is valid in all MTL-chains.*

And again, similar as for the BL-case, even more is provable: the logical calculus $\mathbb{K}_{\mathsf{MTL}}$ characterizes just these formulas which hold true w.r.t. all those t-norm based logics which are determined by a left continuous t-norm. A proof is given in [27].

Theorem 6 (Standard Completeness). *The class of all formulas which are provable in the logical calculus $\mathbb{K}_{\mathsf{MTL}}$ coincides with the class of all formulas which are logically valid in all t-norm algebras with a left continuous t-norm.*

This result again means, as the similar one for the logic of continuous t-norms, that the variety of all MTL-algebras is the smallest variety which contains all t-norm algebras with a left continuous t-norm.

Because of the fact that the BL-algebras are the divisible MTL-algebras, one gets another adequate axiomatization of the basic t-norm logic BL if one extends the axiom system $\mathbb{K}_{\mathsf{MTL}}$ with the additional axiom schema

$$\varphi \wedge \psi \to \varphi \& (\varphi \to \psi) \,. \tag{34}$$

The simplest way to prove that this implication is sufficient is to show that the inequality $x * (x \rightarrowtail y) \leq x \cap y$, which corresponds to the converse implication, holds true in each MTL-algebra. Similar remarks apply to further extensions of MTL we are going to mention.

Also for MTL an extended completeness theorem similar to Theorem 4 remains true.

Theorem 7 (Extended General Completeness). *For each finite set $\mathcal{C}$ of axiom schemata and any formula φ of $\mathcal{L}_T$ the following are equivalent:*

(i) φ is derivable within the logical calculus $\mathbb{K}_{\mathsf{MTL}} + \mathcal{C}$;
(ii) φ is valid in all MTL($\mathcal{C}$)-algebras;
(iii) φ is valid in all MTL($\mathcal{C}$)-chains.

Again the extension to the first-order case is similar to the treatment in [22] for BL and shall not be discussed here.

The core point is that the formal language has to use predicate symbols to introduce atomic formulas, and that the logical apparatus has to be extended by quantifiers: and these are usually a generalization $\forall$ and a particularization $\exists$. As semantic interpretations of these quantifiers one uses in case of $\forall$ the infimum of the set of truth degrees of all the instances, and in case of $\exists$ the corresponding supremum. The semantic models $(\mathbf{M}, L)$ for this first-order language are determined by a nonempty universe M, a truth degree lattice L, and for each n-ary predicate symbol P some n-ary L-valued fuzzy relation in M.

This forces either to assume that the truth degree lattices L are complete lattices, or weaker and more preferably that the model $(\mathbf{M}, L)$ has to be a *safe* one, i.e. one for which all those subsets of L have suprema and infima which may occur as truth degree sets of instances of quantified formulas of the language.

4 Linguistic Control and Fuzzy Relational Equations

4.1 The Standard Paradigm

The standard paradigm of rule based fuzzy control is that one supposes to have given, in a granular way, an incomplete and fuzzy description of a control function Φ from an input space $\mathbb{X}$ to an output space $\mathbb{Y}$, realized by a finite family

$$\mathcal{D} = (\langle A_i, B_i \rangle)_{1 \leq i \leq n} \tag{35}$$

of (fuzzy) input-output data pairs. These granular data are supposed to characterize this function Φ sufficiently well.

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

$$\text{IF } \alpha \text{ is } A_i, \text{ THEN } \beta \text{ is } B_i, \quad i = 1, \ldots, n\,, \tag{36}$$

of linguistic control rules, also called fuzzy IF-THEN rules, describing some control procedure with input variable α and output variable β.

Mainly in engineering papers one often considers also the case of different input variables $\alpha_1, \ldots, \alpha_m$. In this case the linguistic control rules become the form

$$\text{IF } \alpha_1 \text{ IS } A_i^1, \text{ AND } \ldots \text{ AND } \alpha_m \text{ IS } A_i^m, \quad \text{THEN } \beta \text{ IS } B_i, \quad i = 1, \ldots, n\,. \tag{37}$$

But from a mathematical point of view such rules are subsumed among the former ones and cover only a restricted class of cases. To see this one simply has to allow as the input universe for α the cartesian product $\mathbb{X} = \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$ of the input universes $\mathbb{X}_i$ of α_i, $i = 1, \ldots, m$. This yields for a given list $A_1, \ldots, A_m$ of input sets for the variables $\alpha_1, \ldots, \alpha_m$ the particular fuzzy input set $A = A_1 \times \cdots \times A_m$ for the combined variable α. The above mentioned restriction comes from the fact that not all fuzzy subsets of $\mathbb{X}$ have this form of a fuzzy cartesian product of fuzzy subsets of the universes $\mathbb{X}_i$.

Let us assume for simplicity that all the input data A_i are *normal*, i.e. that for each i there is a point x_0^i in the universe of discourse with $A_i(x_0^i) = 1$. Sometimes even weak normality would suffice, i.e. that the supremum over all the membership degrees of the A_i equals one; but we do not indent to discuss this in detail.

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 $\Phi^* : \mathbb{F}(\mathbb{X}) \to \mathbb{F}(\mathbb{Y})$ which interpolates these data, i.e. which satisfies

$$\Phi^*(A_i) = B_i \quad \text{for each } i = 1, \ldots, n\,, \tag{38}$$

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, more precisely: for some uniformly defined function, which should interpolate all these data, and which should be globally defined over the class $\mathbb{F}(\mathbb{X})$ of all fuzzy subsets of $\mathbb{X}$, or at least over a suitably chosen sufficiently large subclass of $\mathbb{F}(\mathbb{X})$.

Following Zadeh [36], this idea is formally realized by a fuzzy relation R, which connects fuzzy input information A with fuzzy output information $B = A \circ R$ via the compositional rule of inference (3). Therefore, applying this idea to the linguistic control rules themselves, transforms these rules in a natural way into a system of fuzzy relation equations

$$A_i \circ R = B_i, \qquad \text{for} \quad i = 1, \ldots, n\,. \tag{39}$$

The problem, to determine a fuzzy relation R which realizes via (3) such a list (36) of linguistic control rules, becomes the problem to determine a solution of the corresponding system (39) of relation equations.

This problem proves to be a rather difficult one: it often happens that a given system (39) of relation equations is unsolvable. This is already the case in the more specific situation that the membership degrees belong to a Boolean algebra, as discussed (as a problem for Boolean matrices) e.g. in [30].

Nice solvability criteria are still largely unknown. Thus the investigation of the structure of the solution space for (39) was one of the problems discussed rather early. One essentially has that this space is an upper semilattice under the simple set union determined by the maximum of the membership degrees; cf. e.g. [6].

And this semilattice has, if it is nonempty, a universal upper bound

$$\widehat{R} = \bigcap_{i=1}^{n} \{(x,y) \parallel A_i(x) \to B_i(y)\} \tag{40}$$

as explained by the following result.

Theorem 8. *The system (39) of relation equations is solvable iff the fuzzy relation $\widehat{R}$ is a solution of it. And in the case of solvability, $\widehat{R}$ is always the largest solution of the system (39) of relation equations.*

This result was first stated in [35] for the particular case of the min-based Gödel implication $\to$ in (40), and generalized to the case of the residuated implications based upon arbitrary left continuous t-norms—and hence to the present situation—by this author in [13]; cf. also his [14].

Besides the reference to the CRI in this type of approach toward fuzzy control, the crucial point is to determine a fuzzy relation out of a list of linguistic control rules.

The fuzzy relation $\widehat{R}$ can be seen as a formalization of the idea that the list (36) of control rules has to be read as:

$$\begin{aligned}&\text{IF input is } A_1 \text{ THEN output is } B_1\\&\text{AND}\\&\dots\\&\text{AND}\\&\text{IF input is } A_n \text{ THEN output is } B_n\,.\end{aligned}$$

Having in mind such a formalization of the list (36) of control rules, there is immediately also another way how to read this list: substitute an OR for the AND to combine the single rules.

It is this understanding of the list of linguistic control rules as a (rough) description of a *fuzzy function* which characterizes the approach of Mamdani/Assilian [32]. Therefore they consider instead of $\widehat{R}$ the fuzzy relation

$$R_{\mathrm{MA}} = \bigcup_{i=1}^{n} (A_i \times B_i)\,,$$

again combined with the compositional rule of inference.

4.2 Solutions and Pseudo-solutions of Fuzzy Relational Equations

Linguistic control rules serve as a tool for rough model building, and their translation into fuzzy relational equations does so too. So it may not really be necessary to solve systems of fuzzy relational equations (39) in the standard sense, but some kind of "approximation" of solutions may do as well.

There is, however, a a quite fundamental problem to understand this remark: an approximation is always an approximation of something. So an approximate solution of a system (39) should (normally) be an approximation of a true solution of this system. But what, if the system is not solvable at all?

A way out is offered by another type of approximation: by an approximation of the "behavior" of the functional which is given by the left hand sides of the equations from (39). In terms of the linguistic control rules which constituted (39) this means that one looks for an approximate realization of their behavior.

This way of doing, originating from [32] and the fuzzy relation R_{MA} introduced there, can also be seen as to "fake" something similar to a solution – and to work with it like a solution.

Such "faked" solutions shall be coined *pseudo-solutions*, following [21]. The best known pseudo-solutions for a system (39) are $\widehat{R}$ and R_{MA}.

From Theorem 8 it is clear that $\widehat{R}$ is a pseudo-solution of a system (39) just in the case that the system (39) is not solvable. So one is immediately lead to the

Problem: Under which conditions is the pseudo-solution R_{MA} really a solution of the corresponding system (39) of relation equations.

This problem is discussed in [28]. And one of the main results is the next theorem.

Theorem 9. *Let all the input sets A_i be normal. Then the fuzzy relation R_{MA} is a solution of the corresponding system of fuzzy relation equations iff for all $i, j = 1, \dots, n$ one has*

$$\models \quad \exists x(A_i(x) \;\&\; A_j(x)) \rightarrow B_i \equiv^* B_j \,. \tag{41}$$

This MA-solvability criterion (41) is a kind of functionality of the list of linguistic control rules, at least in the case of the presence of an involutive negation: because in such a case one has

$$\models \quad \exists x(A_i(x) \;\&\; A_j(x)) \;\leftrightarrow\; A_i \cap_{\boldsymbol{t}} A_j \not\equiv^* \emptyset \,,$$

and thus condition (41) becomes

$$\models \quad A_i \cap_{\boldsymbol{t}} A_j \not\equiv^* \emptyset \;\rightarrow\; B_i \equiv^* B_j \,. \tag{42}$$

And this can be understood as a fuzzification of the idea *"if A_i and A_j coincide to some degree, than also B_i and B_j should coincide to a certain degree"*.

Of course, this fuzzification is neither obvious nor completely natural, because it translates "degree of coincidence" in two different ways.

This leads back to the well known result, explained e.g. in [14], that the system of relation equations is solvable in the case that all the input fuzzy sets A_i are pairwise $\boldsymbol{t}$-disjoint:

$$A_i \cap_{\boldsymbol{t}} A_j = \emptyset \quad \text{for all } i \neq j.$$

Here the reader should have in mind that this $\boldsymbol{t}$-disjointness is, in general, weaker than the standard min-disjointness: it does not force the disjointness of the supports of the fuzzy sets A_i, A_j, and even allows a height $\mathsf{hgt}(A_i \cap A_j) = 0.5$ for the case that $\boldsymbol{t}$ is the Łukasiewicz t-norm.

However, it may happen that the system of relation equations is solvable, i.e. has $\widehat{R}$ as a solution, without having the fuzzy relation R_{MA} as a solution.

An example is given in [21].

Therefore condition (41) is only a sufficient one for the solvability of the system (39) of relation equations.

Hence one has as a new problem to give additional assumptions, besides the solvability of the system (39) of relation equations, which are sufficient to guarantee that R_{MA} is a solution of (39).

As in [14] and already in [12], we subdivide the problem whether a fuzzy relation R is a solution of the system of relation equations into two cases: (i) whether one has satisfied the *subset property* w.r.t. a system (39), i.e. whether one has

$$A_i \circ R \subset B_i, \quad \text{for} \quad i = 1, \dots, n \,, \tag{43}$$

and (ii) whether one has the *superset property* w.r.t. (39), i.e. whether one has one has

$$A_i \circ R \supset B_i, \quad \text{for} \quad i = 1, \dots, n \,. \tag{44}$$

The core result is the following theorem.

Theorem 10. *If the input set A_k is normal then* $A_k \circ \widehat{R} \subset B_k \subset A_k \circ R_{\mathrm{MA}}$.

So we know that with normal input sets the fuzzy outputs $A_i \circ \widehat{R}$ always are fuzzy subsets of $A_i \circ R_{\mathrm{MA}}$.

Furthermore we immediately have the following global result.

Proposition 2. *If all the input sets A_i of the system of relation equations are normal and if one also has $R_{\mathrm{MA}} \subset \widehat{R}$, then the system of relation equations is solvable, and R_{MA} is a solution.*

Hence the pseudo-solutions R_{MA} and $\widehat{R}$ are upper and lower approximations for the realizations of the linguistic control rules.

Now one may equally well look for new pseudo-solutions, e.g. by some *iteration* of these pseudo-solutions in the way, that for the *next iteration step* in such an iteration process the system of relation equations is changed such that its (new) output sets become the *real output* of the *former* iteration step. This has been done in [21].

To formulate the dependence of the pseudo-solutions R_{MA} and $\widehat{R}$ from the input and output data, we denote the "original" pseudo-solutions with the input-output data (A_i, B_i) in another way and write

$$R_{\mathrm{MA}}[B_k] \text{ for } R_{\mathrm{MA}}\,, \qquad \widehat{R}[B_k] \text{ for } \widehat{R}\,.$$

Theorem 11. *One has always*

$$A_i \circ \widehat{R}[B_k] \subset A_i \circ R_{\mathrm{MA}}[A_k \circ \widehat{R}[B_k]] \subset A_i \circ R_{\mathrm{MA}}[B_k]\,.$$

Thus the iterated relation $R_{\mathrm{MA}}[A_k \circ \widehat{R}]$ is a pseudo-solution which somehow better approximates the intended behavior of the linguistic control rules as each one of R_{MA} and $\widehat{R}$. For details cf. again [21].

4.3 Invoking More Formal Logic

The languages of the first-order versions of any one of the standard fuzzy logics discussed in Section 3 can be used to formalize the main ideas behind the use of linguistic control rules in fuzzy control matters. This offers, besides the above presented reference to fuzzy set and fuzzy relation matters, a second way for a mathematical analysis of this rough modeling strategy.

We shall use here the logic $\mathsf{BL}\forall$, i.e. the first-order version of the logic BL of continuous t-norms as the basic reference logic for this analysis. And we follow in the presentation of the material closely the paper [33].

A formula of this logic, with n free variables, describes w.r.t. each $\mathsf{BL}\forall$-model $(\mathbf{M}, L)$ some n-ary L-fuzzy relation in the universe $M = |\mathbf{M}|$. This is the obvious specification of the approach of Section 2 to describe fuzzy sets by formulas of a suitable language.

A BL-theory $\mathcal{T}$ is just a crisp, i.e. classical set of formulas of the language of $\mathsf{BL}\forall$. The notation $\mathcal{T} \vdash A$ means that the formula A is provable in a theory $\mathcal{T}$.

In what follows, we will use the letters $A, B, \dots$ formulas as well as for L-fuzzy sets in some $\mathsf{BL}\forall$-model $(\mathbf{M}, L)$.

This yields that e.g. the result (3) of a – now t-norm based and not only min-based – CRI-application to a fuzzy set A and a fuzzy relation R is described by the formula

$$(\exists x)(A(x) \mathbin{\&} R(x,y)) \,. \tag{45}$$

Furthermore, if $\mathcal{T}$ is a BL-theory and $\mathcal{T} \vdash (\exists x)A(x)$ then this means that the fuzzy set described by A is normal.

Let us additionally assume that $\mathcal{T}$ is a consistent $\mathsf{BL}\forall$-theory which formalizes some additional assumptions which may be made for the considerations of a fuzzy control problem, e.g. that the input fuzzy sets may be normal ones.

On this basis, the results in fuzzy relation equations can be succinctly and transparently formulated in the syntax of $\mathsf{BL}\forall$. So let us start from a system (36) of linguistic control rules with input fuzzy sets A_i over some m-dimensional universe $\mathbb{X} = \mathbb{X}_1 \times \cdots \times \mathbb{X}_m$.

With our notational convention this means that we have given formulas $A_i(x_1, \dots, x_m)$ with free variables $x_1, \dots, x_m$ which describe these input sets of the single control rules. For simplicity, let us write $\overrightarrow{x} = (x_1, \dots, x_m)$ for this list of (free) variables.

The problem which corresponds to the solvability problem of a system (39) of relation equations now is to find a formula $R(\overrightarrow{x}, y)$ with $m+1$ free variables such that

$$\mathcal{T} \vdash (\exists x_1) \cdots (\exists x_n)(A_i(\overrightarrow{x}) \mathbin{\&} R(\overrightarrow{x}, y)) \leftrightarrow B_i(y) \tag{46}$$

holds for every $i = 1, \dots, n$. If such a formula R exists then we say that the *system of fuzzy relation equations* (46) is *solvable in* $\mathcal{T}$ and $R(\overrightarrow{x}, y)$ is its solution.

Lemma 1. *The following is* $\mathsf{BL}\forall$*-provable:*

$$\mathcal{T} \vdash (\exists x_1) \cdots (\exists x_m) \left(A_i(\overrightarrow{x}) \mathbin{\&} \bigwedge_{j=1}^{n} (A_j(\overrightarrow{x}) \to B_j(y)) \right) \to B_i(y).$$

The following theorem presents in a purely syntactical way the well known fundamental result on the solution of the fuzzy relation equations.

Theorem 12. *The system of fuzzy relation equations (46) is solvable in* $\mathcal{T}$ *iff*

$$\mathcal{T} \vdash (\exists x_1) \cdots (\exists x_m) \left(A_i(\overrightarrow{x}) \mathbin{\&} \bigwedge_{j=1}^{n} (A_j(\overrightarrow{x}) \to B_j(y)) \right) \leftrightarrow B_i(y) \tag{47}$$

holds for all $i = 1, \dots, n$.

Corollary 5. *If (46) holds for some* $R(\overrightarrow{x}, y)$ *then*

$$\mathcal{T} \vdash R(\overrightarrow{x}, y) \to \bigwedge_{j=1}^{n} (A_j(\overrightarrow{x}) \to B_j(y)) \,.$$

This means, in the present terminology, that the solution (47) is maximal.

To simplify the formulas to come let us assume for the rest of this subsection that $m = 1$. This corresponds completely to the former subsumption of the case (37) of a system of linguistic control rules with m input variables under the case (36) with only one input variable.

We will work with the following two special kinds of formulas. The first one has the form of a conjunction of implications from Theorem 12

$$R^{Imp}(x, y) =_{\mathrm{def}} \bigwedge_{j=1}^{m} (A_j(x) \to B_j(y))) . \tag{48}$$

The second one, the so called Mamdani-Assilian formula, shall be

$$R^{MA}(x, y) = \bigvee_{i=1}^{m} (A_i(x) \mathbin{\&} B_i(y)) . \tag{49}$$

Lemma 2. *Let $\mathcal{T} \vdash (\exists x)A_i$ for all $i = 1, \ldots, n$. Then*

$$\mathcal{T} \vdash B_i(y) \to (\exists x)(A_i(x) \mathbin{\&} \bigvee_{j=1}^{m} (A_j(x) \mathbin{\&} B_j(y))). \tag{50}$$

Joining Lemmas 1 and 2, we get get following theorem, which says that our conjunction (48) of implications gives the lower and the Mamdani-Assilian formula (49) the upper bound for the solutions of the system (46).

Theorem 13. *Let $\mathcal{T} \vdash (\exists x)A_i$ for $i = 1, \ldots, n$. Then the following is provable:*

$$\mathcal{T} \vdash (\exists x)(A_i(x) \mathbin{\&} R^{Impl}(x, y)) \to B_i(y) ,$$
$$\mathcal{T} \vdash B_i(y) \to (\exists x)(A_i(x) \mathbin{\&} R^{MA})$$

for each $i = 1, \ldots, n$.

The following theorem has a semantical counterpart proved with some more restrictive assumptions by Perfilieva/Tonis [34].

Theorem 14. *Let $A(x)$ be an arbitrary formula. The following is provable for each $i = 1, \ldots, n$:*

$$\mathcal{T} \vdash (A(x) \leftrightarrow A_i(x)) \mathbin{\&} (B_i(y) \to (\exists x)(A_i(x) \mathbin{\&} R^{Impl}(x, y))) \to ((\exists x)(A(x) \mathbin{\&} R^{Impl}(x, y)) \leftrightarrow B_i(y)) . \tag{51}$$

This theorem suggests that the formula

$$\xi(y) =_{\mathrm{def}} (B_i(y) \to (\exists x)(A_i(x) \mathbin{\&} R^{Impl}(x, y)))$$

can be the basis of a *solvability sign* similar to a more general solvability index discussed by this author e.g. in [14, 15] and defined as $\xi(y) =_{\mathrm{def}} \bigwedge_{i=1}^{n} \xi_i(y)$.

Then from Theorem 14 it follows:

Corollary 6

$$\mathcal{T} \vdash \xi(y) \to ((\exists x)(A_i(x) \,\&\, R^{Impl}(x,y)) \leftrightarrow B_i(y))$$

for each $i = 1, \ldots, n$.

It follows from this corollary that in a model of $\mathcal{T}$, the solvability sign can be interpreted as a degree, in which the system (46) is solvable in $\mathcal{T}$.

Theorem 15. *The system (46) is solvable in* $\mathcal{T}$ *iff*

$$\mathcal{T} \vdash (\forall y)\xi(y).$$

Theorem 16. *If the system (46) is solvable in* $\mathcal{T}$ *then*

$$\mathcal{T} \vdash (\forall x)(A_i(x) \leftrightarrow A_j(x)) \to (\forall y)(B_i(y) \leftrightarrow B_j(y)) \tag{52}$$

for all $i, j = 1, \ldots, n$.

Klawonn [28] gave a necessary and sufficient condition that the Mamdani-Assilian fuzzy relation is a solution of a system of fuzzy relation equations. The formalized version of this result can again be BL∀-proved.

Theorem 17. *If* $T \vdash (\exists x)A_i(x)$, $i = 1, \ldots, m$. *Then* $R^{MA}(x,y)$ *is a solution of (46) iff*

$$\mathcal{T} \vdash (\exists x)(A_i(x) \,\&\, A_j(x)) \to (\forall y)(B_i(y) \leftrightarrow B_j(y)), \qquad i = 1, \ldots, n. \tag{53}$$

The following theorem is a corollary of the previous results.

Corollary 7. (i) *If there is an index* k *such that*

$$\mathcal{T} \nvdash (\exists y)B_k(y) \to (\exists x)A_k(x)\,, \tag{54}$$

then the system (46) is not solvable.

(ii) *Assume*

$$\mathcal{T} \nvdash (\exists x)(\exists y) \bigvee_{j=1}^{n} (A_i(x) \,\&\, B_i(y)) \to (\exists x)(\exists y)R(x,y)\,.$$

Then $R(x,y)$ *is not a solution of (46).*

5 Approximation and Interpolation

The standard mathematical understanding of *approximation* is that by an approximation process some mathematical object A, e.g. some function, is approximated, i.e. determined within some (usually previously unspecified) error bounds.

Additionally one assumes that the approximating object B for A is of some predetermined, usually "simpler" kind, e.g. a polynomial function.

So one may approximate some transcendental function, e.g. the trajectory of some non-linear process, by a piecewise linear function, or by a polynomial function of some bounded degree. Similarly one approximates e.g. in the Runge-Kutta methods the solution of a differential equation by a piecewise linear function, or one uses splines to approximate a difficult surface in 3-space by planar pieces.

The standard mathematical understanding of *interpolation* is that a function f is only partially given by its values *at some points* of the domain of the function, the interpolation nodes.

The problem then is to determine "the" values of f for all the other points of the domain (usually) *between* the interpolation nodes – sometimes also outside these interpolation nodes (extrapolation).

And this is usually done in such a way that one considers groups of neighboring interpolation nodes which uniquely determine an interpolating function of some predetermined type within their convex hull (or something like): a function which has the interpolation nodes of the actual group as argument-value pairs – and which in this sense *locally* approximates the function f.

In the *standard fuzzy control approach* the input-output data pairs of the linguistic control rules just provide interpolation nodes.

However, what is lacking – at least up to now – that is the idea of a *local approximation* of the intended crisp control function by some fuzzy function. Instead, in the standard contexts one always asks for something like a *global* interpolation, i.e. one is interested to interpolate *all* nodes by *only one* interpolation function.

To get a local approximation of the intended crisp control function Φ, one needs some notion of "nearness" or of "neighboring" for fuzzy data granules. Such a notion is lacking in general.

For the particular case of a linearly ordered input universe $\mathbb{X}$, and the additional assumption that the fuzzy input data are unimodal, one gets in a natural way from this crisp background a notion of neighboring interpolation nodes: fuzzy nodes are neighboring if their kernel points are.

In general, however, it seems most appropriate to suppose that one may be able to infer from the control problem a—perhaps itself fuzzy—partitioning of the whole input space (or similarly of the output space). Then one will be in a position to split in a natural way the data set (35), or correspondingly the list (36) of control rules, into different groups—and to consider the localized interpolation problems separately for these groups.

This offers obviously better chances for finding interpolating functions, particularly for getting solvable systems of fuzzy relation equations. However, one has to be aware that one should additionally take care that the different local interpolation functions fit together somehow smoothly—again an open problem that needs a separate discussion. It is a problem that is more complicated for fuzzy interpolation than for the crisp counterpart because the fuzzy interpolating functions may realize the fuzzy interpolation nodes only approximately.

However, one may start from ideas like these to speculate about fuzzy versions of the standard *spline interpolation* methodology.

In the context of fuzzy control the control function Φ, which has to be determined, is described only roughly, i.e. given only by its behavior in some (fuzzy) points of the state space.

The standard way to roughly describe the control function is to give a list (36) of linguistic control rules connecting fuzzy subsets A_i of the input space $\mathbb{X}$ with fuzzy subsets B_i of the output space $\mathbb{Y}$ indicating that one likes to have

$$\Phi^*(A_i) = B_i\,, \quad i = 1, \dots, n \tag{55}$$

for a suitable "fuzzified" version $\Phi^* : \mathbb{F}(\mathbb{X}) \to \mathbb{F}(\mathbb{Y})$ of the control function $\Phi : \mathbb{X} \to \mathbb{Y}$.

The additional approximation idea of the CRI is to approximate Φ^* by a fuzzy function $\Psi^* : \mathbb{F}(\mathbb{X}) \to \mathbb{F}(\mathbb{Y})$ determined for all $A \in \mathbb{F}(\mathbb{X})$ by

$$\Psi^*(A) = A \circ R \tag{56}$$

which refers to some suitable fuzzy relation $R \in \mathbb{F}(\mathbb{X} \times \mathbb{Y})$, and understands $\circ$ as sup-t-composition.

Formally thus the equations (55) become transformed into some well known system (39) of relation equations

$$A_i \circ R = B_i, \qquad i = 1, \dots, n$$

to be solved for the unknown fuzzy relation R.

This approximation idea fits well with the fact that one often is satisfied with pseudo-solutions of (39), and particularly with the MA-pseudo-solution R_{MA}, or the S-pseudo-solution $\widehat{R}$. Both of them determine approximations Ψ^* to the (fuzzified) control function Φ^*.

What remains open in this discussion up to now are quality considerations for such approximations via pseudo-solutions of systems (39) of fuzzy relational equations. This is a topic which has been discussed only scarcely. Qualitative approaches referring to the ideas of upper and lower approximations w.r.t. fuzzy inclusion, as considered here in Subsection 4.2, have been considered in [17] and extensively reported in [20]. The interested reader may consult these sources. More quantitative approaches, based e.g. upon the consideration of metrics or pseudo-metrics in spaces of fuzzy sets, are missing up to now – as far as this author is aware of.

6 Aggregation Operations and Fuzzy Control Strategies

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 $\mathbb{F}(\mathbb{X})$.

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.

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

In general, however, the interpolation operators may 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.[3]

And an *aggregation operator* $\mathbf{A}$, as explained e.g. in [2, 7], 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 a *commutative* one iff each operation f^n is commutative. And $\mathbf{A}$ is an *associative* aggregation operator iff

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

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.[4]

As in [18], we now consider interpolation operators Φ of FITA-type and interpolation operators Ψ of FATI-type which have the abstract forms

$$\Psi_{\mathcal{D}}(A) = \mathbf{A}(\theta_{\langle A_1,B_1\rangle}(A),\dots,\theta_{\langle A_n,B_n\rangle}(A))\,, \tag{57}$$

$$\Xi_{\mathcal{D}}(A) = \widehat{\mathbf{A}}(\theta_{\langle A_1,B_1\rangle},\dots,\theta_{\langle A_n,B_n\rangle})(A)\,. \tag{58}$$

Here we assume that each one of the "local" inference operators θ_i is determined by the single input-output pair $\langle A_i, B_i\rangle$. This restriction is sufficient for the present purpose.

In this Section we survey the main notions and results. The interested reader gets more details from [19] and also from [20].

6.1 Stability Conditions

If $\Theta_{\mathcal{D}}$ is a fuzzy inference operator of one of the types (57), (58), then the interpolation property one likes to have realized is that one has

$$\Theta_{\mathcal{D}}(A_i) = B_i \tag{59}$$

for all the data pairs $\langle A_i, B_i\rangle$. In the particular case that the operator $\Theta_{\mathcal{D}}$ is given by (3), this is just the problem to solve the system (59) of fuzzy relation equations.

[3] 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.

[4] 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.

Definition 4. *In the present generalized context let us call the property (59) 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 (57) which operates on $\mathbb{F}(\mathbb{Y})$, of course with the inclusion relation for fuzzy sets as partial ordering.

Definition 5. *Having* $B, C \in \mathbb{F}(\mathbb{Y})$ *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.

Hence one e.g. has that C is $\bigcup$-negligible w.r.t. B iff $C \subset B$; and C is $\bigcap$-negligible w.r.t. B iff $C \supset B$.

Proposition 3. *Consider a fuzzy inference operator* $\Psi_{\mathcal{D}}$ *of FITA-type (57). It is sufficient for the* $\mathcal{D}$*-stability of* $\Psi_{\mathcal{D}}$*, i.e. to have*

$$\Psi_{\mathcal{D}}(A_k) = B_k \qquad \text{for all } k = 1, \dots, n$$

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*

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

This result has two quite interesting specializations which themselves generalize well known results about fuzzy relation equations.

Corollary 8. *It is sufficient for the* $\mathcal{D}$*-stability of a fuzzy inference operator* $\Psi_{\mathcal{D}}$ *of FITA-type that one has*

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

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

To state the second one of these results, call an aggregation operator $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$ *additive* iff always $b \leqslant f^2(b, c)$, and call it *idempotent* iff always $b = f^2(b, b)$.

Corollary 9. *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*

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

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 $\cap$ and cartesian product $\times$) because in this case one has usually:

$$\begin{aligned}\theta_{\langle A_i,B_i\rangle}(A_j) &= A_j \circ (A_i \times B_i) = \{y \parallel \exists x (x \,\varepsilon\, A_j \,\&\, (x,y) \,\varepsilon\, A_i \times B_i)\} \\ &= \{y \parallel \exists x (x \,\varepsilon\, A_j \cap A_i \,\&\, y \,\varepsilon\, B_i)\} \,.\end{aligned}$$

To extend these considerations from inference operators (57) of the FITA type to those ones of the FATI type (58) let us consider the following notion.

Definition 6. *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*

$$\widehat{\mathbf{A}}(\theta_1,\ldots,\theta_n)(X) = \mathbf{A}(\theta_1(X),\ldots,\theta_n(X)) \tag{60}$$

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

Using this notion it is easy to see that one has on the left hand side of (60) 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, assuming that such application distributive pairs of aggregation operators exist.

Proposition 4. *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.*

The general approach of this Section can also be applied to the problem of $\mathcal{D}$-stability for a modified operator $\Theta^*_{\mathcal{D}}$ which is determined by the kind of iteration of $\Theta_{\mathcal{D}}$ which previously led to Theorem 11. To do this, let us consider the $\Theta_{\mathcal{D}}$-*modified* data set $\mathcal{D}^*$ given as

$$\mathcal{D}^* = (\langle A_i, \Theta_{\mathcal{D}}(A_i)\rangle)_{1\le i\le n}\,, \tag{61}$$

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

$$\Theta^*_{\mathcal{D}} = \Theta_{\mathcal{D}^*}\,. \tag{62}$$

For these modifications, the problem of stability reappears. Of course, the new situation here is only a particular case of the former. And 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 $\mathcal{D} = (\langle A_i, B_i\rangle)_{1\le i\le n}$.

Proposition 5. *It is sufficient for the* $\mathcal{D}^*$*-stability of a fuzzy inference operator* $\Psi^*_{\mathcal{D}}$ *of FITA-type that one has*

$$\Psi^*_{\mathcal{D}}(A_i) = \Psi_{\mathcal{D}^*}(A_i) = \Psi_{\mathcal{D}}(A_i) \qquad \text{for all } \ 1\le i\le n \tag{63}$$

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 (63) and at the negligibility conditions.

Corollary 10. *The conditions (63) are always satisfied if the inference 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 [21].*

Corollary 11. *In the case that the aggregation operator is the set theoretic union, i.e.* $A = \bigcup$*, the conditions (63) together with the inclusion relationships*

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

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

Again one is able to transfer this result to FATI-type fuzzy inference operators.

Corollary 12. *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.*

6.2 Application Distributivity

Based upon the notion of application distributive pair of aggregation operators the property of $\mathcal{D}$-stability can be transferred back and forth between two inference operators of FATI-type and of FITA-type if they are based upon a pair of application distributive aggregation operators.

What has not been discussed previously was the existence and the uniqueness of such pairs. Here are some results concerning these problems.

The uniqueness problem has a simple solution.

Proposition 6. *If* $(\widehat{\mathbf{A}}, \mathbf{A})$ *is an application distributive pair of aggregation operators then* $\widehat{\mathbf{A}}$ *is uniquely determined by* $\mathbf{A}$*, and conversely also* $\mathbf{A}$ *is uniquely determined by* $\widehat{\mathbf{A}}$*.*

And for the existence problem we have a nice reduction to the two-argument case.

Theorem 18. *Suppose that* $\mathbf{A}$ *is a commutative and associative aggregation operator. For the case that there exists an aggregation operator* $\widehat{\mathbf{A}}$ *such that* $(\widehat{\mathbf{A}}, \mathbf{A})$ *form an application distributive pair of aggregation operators it is necessary and sufficient that there exists some operation* G *for fuzzy inference operators satisfying*

$$\mathbf{A}(\theta_1(X), \theta_2(X)) = G(\theta_1, \theta_2)(X) \tag{64}$$

for all fuzzy inference operators θ_1, θ_2 *and all fuzzy sets* X*.*

However, there is an important restriction concerning the existence of such pairs of application distributive aggregation operators, at least for an interesting particular case.

Definition 7. *An aggregation operator* $\mathbf{A} = (f^n)_{n\in\mathbb{N}}$ *for fuzzy subsets of a common universe of discourse* $\mathbb{X}$ *is* pointwise defined *iff for each* $n \in \mathbb{N}$ *there exists a function* $g_n : [0,1]^n \to [0,1]$ *such that for all* $A_1, \dots, A_n \in \mathcal{F}$ *and all* $x \in \mathbf{X}$ *there hold*

$$f^n(A_1, \dots, A_n)(x) = g_n(A_1(x), \dots, A_n(x)) . \tag{65}$$

And an aggregation operator $\widehat{\mathbf{A}}$ *for inference operators is* pointwise defined *iff it can be reduced to a pointwise defined aggregation operator for fuzzy relations.*

The restrictive result, proved in [26], now reads as follows.

Proposition 7. *Among the commutative, associative, and pointwise defined aggregation operators is* $(\bigcup, \bigcup)$ *the only application distributive pair.*

6.3 Invoking a Defuzzification Strategy

In a lot of practical applications of the fuzzy control strategies which form the starting point for the previous general considerations, the fuzzy model—e.g. determined by a list (36) of linguistic IF-THEN-rules—is realized in the context of a further *defuzzification strategy*, which is nothing but a mapping $F : \mathbb{F}(\mathbb{Y}) \to \mathbb{Y}$ for fuzzy subsets of the output space $\mathbb{Y}$.

Having this in mind, it seems to be reasonable to consider the following modification of the $\mathcal{D}$-stability condition, which is a formalization of the idea to have "stability modulo defuzzification".

Definition 8. *A fuzzy inference operator* $\Theta_{\mathcal{D}}$ *is* $(F, \mathcal{D})$*-stable w.r.t. a defuzzification strategy* $F : \mathbb{F}(\mathbb{Y}) \to \mathbb{Y}$ *iff one has*

$$F(\Theta_{\mathcal{D}}(A_i)) = F(B_i) \tag{66}$$

for all the data pairs $\langle A_i, B_i \rangle$ *from* $\mathcal{D}$.

For the fuzzy modeling process which is manifested in the data set $\mathcal{D}$ this condition (66) is supposed to fit well with the control behavior one is interested to implement. If for some application this condition (66) seems to be unreasonable, this indicates that either the data set $\mathcal{D}$ or the chosen defuzzification strategy F are unsuitable.

As a first, and rather restricted stability result for this modified situation, the following Proposition shall be mentioned.

Proposition 8. *Suppose that* $\Theta_{\mathcal{D}}$ *is a fuzzy inference operator of FITA-type (57), that the aggregation is union* $\mathbf{A} = \bigcup$ *as e.g. in the fuzzy inference operator for the Mamdani–Assilian case, and that the defuzzification strategy* F *is the "mean of max" method. Then it is sufficient for the* $(F, \mathcal{D})$*-stability of* $\Theta_{\mathcal{D}}$ *to have satisfied*

$$\mathsf{hgt}\Big(\bigcup_{j=1, j\neq k}^{n} \theta_k(A_j)\Big) < \mathsf{hgt}(\theta_k(A_k)) \tag{67}$$

for all $k = 1, \dots, n$.

The proof follows from the corresponding definitions by straightforward routine calculations, and hgt means the “height” of a fuzzy set, i.e. the supremum of its membership degrees.

Further investigations into this topic are necessary.

7 Conclusion

Abstract mathematical tools like mathematical fuzzy logics, like fuzzy set theory and fuzzy relational equations, but also like interpolation strategies, and like operator equations which use the more recent topic of aggregation operators shed interesting light on the formal properties of granular modeling approaches which use the technique of linguistic variables and linguistic rules.

This point of view has been exemplified using considerations upon the fuzzy control related topic of the realizability behavior of systems of linguistic rules, mostly with the background idea to refer to the compositional rule of inference in implementing these rules.

Our reference to the mathematical fuzzy logic BL, with its algebraic semantics determined by the class of all prelinear, divisible, and integral residuated lattices, also opens the way for further generalizations of our considerations toward more general membership degree structures for fuzzy sets: like the type-2 or interval type-2 cases. And the reader should also have in mind that there is a close relationship between rough sets and L-fuzzy sets over a suitable 3-element lattice.

Additionally it is easy to recognize that the discussions of the interpolating behavior in Section 5 as well as the operator oriented considerations of Section 6 are essentially independent of the particular choice of the membership degree structure. Hence they too may be generalized to other membership degree setting discussed in this book.

References

1. Běhounek, L., Bodenhofer, U., Cintula, P.: Relations in fuzzy class theory: initial steps. Fuzzy Sets and Systems 159, 1729–1772 (2008)
2. Calvo, T., Mayor, G., Mesiar, R. (eds.): Aggregation Operators: New Trends and Applications. Physica-Verlag, Heidelberg (2002)
3. Chang, C.C.: Algebraic analysis of many valued logics. Transactions American Mathematical Society 88, 476–490 (1958)
4. Cignoli, R., Esteva, F., Godo, L., Torrens, A.: Basic fuzzy logic is the logic of continuous t-norms and their residua. Soft Computing 4, 106–112 (2000)
5. Cignoli, R.L.O., D’Ottaviano, I.M.L., Mundici, D.: Algebraic foundations of many-valued reasoning. Trends in Logic–Studia Logica Library, vol. 7. Kluwer Academic Publishers, Dordrecht (2000)
6. Nola, A.D., Sessa, S., Pedrycz, W., Sanchez, E.: Fuzzy relation equations and their applications to knowledge engineering. Theory and Decision Library. Series D: System Theory, Knowledge Engineering and Problem Solving, vol. 3. Kluwer Academic Publishers Group, Dordrecht (1989)

7. Dubois, D., Prade, H.: On the use of aggregation operations in information fusion processes. Fuzzy Sets and Systems 142, 143–161 (2004)
8. Dummett, M.: A propositional calculus with denumerable matrix. Journal Symbolic Logic 24, 97–106 (1959)
9. Esteva, F., Godo, L.: Monoidal t-norm based logic: toward a logic for left-continuous t-norms. Fuzzy Sets and Systems 124, 271–288 (2001)
10. Esteva, F., Godo, L., Montagna, F.: Equational characterization of the subvarieties of BL generated by t-norm algebras. Studia Logica 76, 161–200 (2004)
11. Gödel, K.: Zum intuitionistischen Aussagenkalkül. Anzeiger Akad. Wiss. Wien, math.-naturwiss. Kl. 69, 65–66 (1932)
12. Gottwald, S.: Criteria for noninteractivity of fuzzy logic controller rules. In: Straszak, A. (ed.) Large Scale Systems: Theory and Applications 1983 (Warsaw, 1983), pp. 229–233. Pergamon Press, Oxford (1984)
13. Gottwald, S.: Characterizations of the solvability of fuzzy equations. Elektron. Informationsverarbeitung Kybernetik 22, 67–91 (1986)
14. Gottwald, S.: Fuzzy Sets and Fuzzy Logic. Artificial Intelligence. Vieweg, Braunschweig/Wiesbaden, and Tecnea, Toulouse (1993)
15. Gottwald, S.: Generalized solvability behaviour for systems of fuzzy equations. In: Novák, V., Perfilieva, I. (eds.) Discovering the World with Fuzzy Logic, pp. 401–430. Physica-Verlag, Heidelberg (2000)
16. Gottwald, S.: A Treatise on Many-Valued Logics. Studies in Logic and Computation, vol. 9. Research Studies Press, Baldock (2001)
17. Gottwald, S.: Mathematical fuzzy control. A survey of some recent results. Logic Journal of the IGPL 13, 525–541 (2005)
18. Gottwald, S.: On a generalization of fuzzy relation equations. In: Proc. 11th Internat. Conf. IPMU, Paris, vol. 2, pp. 2572–2577, Edition EDK (2006)
19. Gottwald, S.: An abstract approach toward the evaluation of fuzzy rule systems. In: Wang, P.P., Ruan, D., Kerre, E.E. (eds.) Fuzzy Logic. A Spectrum of Theoretical & Practical Issues. Studies in Fuzziness and Soft Computing, vol. 215, pp. 281–297. Springer, Berlin (2007)
20. Gottwald, S.: Calculi of information granules: Fuzzy relational equations. In: Pedrycz, W., Skowron, A., Kreinovich, V. (eds.) Handbook of Granular Computing, ch. 11, pp. 225–248. J. Wiley & Sons, Chichester (2008)
21. Gottwald, S., Novák, V., Perfilieva, I.: Fuzzy control and t-norm-based fuzzy logic. Some recent results. In: Proc. 9th Internat. Conf. IPMU, Annecy, ESIA – Université de Savoie, pp. 1087–1094 (2002)
22. Hájek, P.: Metamathematics of Fuzzy Logic. Trends in Logic, vol. 4. Kluwer Academic Publishers, Dordrecht (1998)
23. Hájek, P., Godo, L., Esteva, F.: A complete many-valued logic with product-conjunction. Archive Math. Logic 35, 191–208 (1996)
24. Haniková, Z.: Standard algebras for fuzzy propositional calculi. Fuzzy Sets and Systems 124, 309–320 (2001)
25. Haniková, Z.: Mathematical and Metamathematical Properties of Fuzzy Logic. PhD thesis, Charles University, Prague, Czech Republic (2003)
26. Hüsam, P.: FATI versus FITA. Ein abstrakter Vergleich gängiger Strategien bei "fuzzy control"-Systemen. Diplomarbeit, Mathematisches Institut, Universität Leipzig (2008)
27. Jenei, S., Montagna, F.: A proof of standard completeness for Esteva and Godo's logic MTL. Studia Logica 70, 183–192 (2002)

28. Klawonn, F.: Fuzzy points, fuzzy relations and fuzzy functions. In: Discovering the world with fuzzy logic. Stud. Fuzziness Soft Comput, vol. 57, pp. 431–453. Physica, Heidelberg (2000)
29. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
30. Duncan Luce, R.: A note on Boolean matrix theory. Proc. Amer. Math. Soc. 3, 382–388 (1952)
31. Łukasiewicz, J., Tarski, A.: Untersuchungen über den Aussagenkalkül. Comptes Rendus Séances Soc. Sciences et Lettres Varsovie, cl. III 23, 30–50 (1930)
32. Mamdani, A., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. Internat. J. Man-Machine Studies 7, 1–13 (1975)
33. Novák, V., Perfilieva, I., Gottwald, S.: Fuzzy relation equations via basic predicate fuzzy logic. In: Da Ruan, P., Kerre, E.E. (eds.) Computational Intelligent Systems for Applied Research, pp. 81–88. World Scientific, New Jersey (2002)
34. Perfilieva, I., Tonis, A.: Compatibility of systems of fuzzy relation equations. Internat. J. of General Systems 29, 511–528 (2000)
35. Sanchez, E.: Resolution of composite fuzzy relation equations. Information and Control 30, 38–48 (1976)
36. Zadeh, L.A.: Outline of new approach to the analysis of complex systems and decision processes. IEEE Trans. Systems, Man, and Cybernet. SMC-3, 28–44 (1973)