Kybernetika

VOLUME *41* (2005), NUMBER 2

The Journal of the Czech Society for Cybernetics and Information Sciences

Published by:

Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic

Kybernetika is a bi-monthly international journal dedicated for rapid publication of high-quality, peer-reviewed research articles in fields covered by its title.

Kybernetika traditionally publishes research results in the fields of Control Sciences, Information Sciences, System Sciences, Statistical Decision Making, Applied Probability Theory, Random Processes, Fuzziness and Uncertainty Theories, Operations Research and Theoretical Computer Science, as well as in the topics closely related to the above fields.

The Journal has been monitored in the Science Citation Index since 1977 and it is abstracted/indexed in databases of Mathematical Reviews, Current Mathematical Publications, Current Contents ISI Engineering and Computing Technology.

Kybernetika. Volume *41* (2005) ISSN 0023-5954, MK ČR E 4902.

Published bi-monthly by the Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. — Address of the Editor: P. O. Box 18, 182 08 Prague 8, e-mail: kybernetika@utia.cas.cz. — Printed by PV Press, Pod vrstevnicí 5, 140 00 Prague 4. — Orders and subscriptions should be placed with: MYRIS TRADE Ltd., P. O. Box 2, V Štíhlách 1311, 142 01 Prague 4, Czech Republic, e-mail: myris@myris.cz. — Sole agent for all "western" countries: Kubon & Sagner, P. O. Box 34 01 08, D-8 000 München 34, F.R.G.

Published in April 2005.

KYBERNETIKA — VOLUME *41* (2005), NUMBER 2, PAGES 227 – 248

INTERPRETABILITY OF LINGUISTIC VARIABLES: A FORMAL ACCOUNT

Ulrich Bodenhofer and Peter Bauer

This contribution is concerned with the interpretability of fuzzy rule-based systems. While this property is widely considered to be a crucial one in fuzzy rule-based modeling, a more detailed formal investigation of what "interpretability" actually means is not available. So far, interpretability has most often been associated with rather heuristic assumptions about shape and mutual overlapping of fuzzy membership functions. In this paper, we attempt to approach this problem from a more general and formal point of view. First, we clarify what the different aspects of interpretability are in our opinion. Consequently, we propose an axiomatic framework for dealing with the interpretability of linguistic variables (in Zadeh's original sense) which is underlined by examples and application aspects, such as, fuzzy systems design aid, data-driven learning and tuning, and rule base simplification.

Keywords: fuzzy modeling, interpretability, linguistic variable, machine learning

AMS Subject Classification: 94D05, 68T05, 68T35

1. INTRODUCTION

The epoch-making idea of L. A. Zadeh's early work was to utilize what he called "fuzzy sets" as mathematical models of linguistic expressions which cannot be represented in the framework of classical binary logic and set theory in a natural way. The introduction of his seminal article on fuzzy sets [42] contains the following remarkable words:

> *"More often than not, the classes of objects encountered in the real physical world do not have precisely defined criteria of membership. [...] Yet, the fact remains that such imprecisely defined "classes" play an important role in human thinking, particularly in the domains of pattern recognition, communication of information, and abstraction."*

Fuzzy systems became a tremendously successful paradigm – a remarkable triumph which started with well-selling applications in consumer goods implemented by Japanese engineers. The reasons for this development are manifold; however, we are often confronted with the following arguments:

1. The main difference between fuzzy systems and other control or decision support systems is that they are parameterized in an interpretable way – by means

of rules consisting of linguistic expressions. Fuzzy systems, therefore, allow rapid prototyping as well as easy maintenance and adaptation.

2. Fuzzy systems offer completely new opportunities to deal with processes for which only a linguistic description is available. Thereby, they allow to achieve a robust, secure, and reproducible automation of such tasks.

3. Even if conventional control or decision support strategies can be employed, re-formulating a system's actions by means of linguistic rules can lead to a deeper qualitative understanding of its behavior.

We would like to raise the question whether fuzzy systems, as they appear in daily practice, really reflect these undoubtedly nice advantages. One may observe that the *possibility to estimate the system's behavior by reading and understanding the rule base only* is a basic requirement for the validity of the above points. If we adopt the usual wide understanding of fuzzy systems (rule-based systems incorporating vague linguistic expressions), we can see, however, that this property – let us call it *interpretability* – is *not guaranteed by definition*.

In our opinion, interpretability should be *the* key property of fuzzy systems. If it is neglected, one ends up in nothing else than black-box descriptions of input-output relationships and any advantage over neural networks or conventional interpolation methods is lost completely.

The more fuzzy systems became standard tools for engineering applications, the more Zadeh's initial mission became forgotten. In recent years, however, after a relatively long period of ignorance, an increasing awareness of the crucial property of interpretability has emerged [1, 2, 6, 12, 21, 39, 40, 41]. A recent book [11] bundles these forces by presenting a comprehensive overview of research on this topic. So far, the following questions have been identified to have a close connection to interpretability:

1. Does the inference mechanism produce results that are technically and intuitively correct?

2. Is the number of rules still small enough to be comprehensible by a human expert?

3. Is the rule set complete and consistent?

4. Do the fuzzy sets associated to the linguistic expressions really correspond to the human understanding of these expressions?

This paper is solely devoted to the fourth question. So far, there is a kind of shallow understanding that this question is related to shape, ordering, and mutual overlapping of fuzzy membership functions. We intend to approach this question more formally. This is accomplished making the inherent relationships between the linguistic labels explicit by formulating them as (fuzzy) relations. In order to provide a framework that is as general as possible, we consider linguistic variables in their most general form.

Note that the given paper is a revised and extended version of a previously published book chapter [7].

2. PRELIMINARIES

Throughout the whole paper, we do not explicitly distinguish between fuzzy sets and their corresponding membership functions. Consequently, uppercase letters are used for both synonymously. For a given non-empty set X, we denote the set of fuzzy sets on X with $\mathcal{F}(X)$. As usual, a fuzzy set $A \in \mathcal{F}(X)$ is called *normalized* if there exists an $x \in X$ such that $A(x) = 1$.

Triangular norms and conorms [28] are common standard models for fuzzy conjunctions and disjunctions, respectively. In this paper, we will mainly need these two concepts for intersections and unions of fuzzy sets. It is known that couples consisting of a *nilpotent t-norm* and its dual t-conorm [22, 28] are most appropriate choices as soon as fuzzy partitions are concerned [14, 30]. The most important representatives of such operations are the so-called Łukasiewicz operations:

$$T_{\mathbf{L}}(x, y) = \max(x + y - 1, 0)$$
$$S_{\mathbf{L}}(x, y) = \min(x + y, 1)$$

The intersection and union of two arbitrary fuzzy sets $A, B \in \mathcal{F}(X)$ with respect to the Łukasiewicz operations can then be defined as

$$\big(A \cap_{\mathbf{L}} B\big)(x) = T_{\mathbf{L}}\big(A(x), B(x)\big)$$

and

$$\big(A \cup_{\mathbf{L}} B\big)(x) = S_{\mathbf{L}}\big(A(x), B(x)\big),$$

respectively. We restrict to these two standard operations in the following – for the reason of simplicity and the fact that they perfectly fit to the concept of fuzzy partitions due to Ruspini [38]; recall that a family of fuzzy sets $(A_i)_{i \in I} \subseteq \mathcal{F}(X)$ is called *Ruspini partition* if the following equality holds for all $x \in X$:

$$\sum_{i \in I} A_i(x) = 13.$$

Furthermore, recall that a fuzzy set $A \in \mathcal{F}(X)$ is called *convex* if the property

$$x \leq y \leq z \implies A(y) \geq \min\big(A(x), A(z)\big)$$

holds for all $x, y, z \in X$ (given a crisp linear ordering $\leq$ on the domain X) [4, 8, 31, 42].

Lemma 1. [4] Let X be linearly ordered. Then an arbitrary fuzzy set $A \in \mathcal{F}(X)$ is convex if and only if there exists a partition of X into two connected subsets X_1 and X_2 such that, for all $x_1 \in X_1$ and all $x_2 \in X_2$, $x_1 \leq x_2$ holds and such that the membership function of A is non-decreasing over X_1 and non-increasing over X_2.

As a trivial consequence of the previous lemma, a fuzzy set whose membership function is either non-decreasing or non-increasing is convex.

3. FORMAL DEFINITION

Since it has more or less become standard and offers much freedom, in particular with respect to integration of linguistic modifiers and connectives, we closely follow Zadeh's original definition of linguistic variables [43, 44, 45].

Definition 2. A *linguistic variable* V is a quintuple of the form

$$V = (N, G, T, X, S),$$

where N, T, X, G, and S are defined as follows:

1. N is the name of the linguistic variable V;
2. G is a grammar;
3. T is the so-called *term set*, i. e. the set linguistic expressions resulting from G;
4. X is the universe of discourse;
5. S is a $T \to \mathcal{F}(X)$ mapping which defines the semantics – a fuzzy set on X – of each linguistic expression in T.

In this paper, let us assume that the grammar G is always given in *Backus–Naur Form (BNF)* [37].

In our point of view, the ability to interpret the meaning of a rule base qualitatively relies deeply upon an intuitive understanding of the linguistic expressions. Of course, this requires knowledge about inherent relationships between these expressions. Therefore, if qualitative estimations are desired, these relationships need to transfer to the underlying semantics, i. e. the fuzzy sets modeling the labels. In other words, interpretability is strongly connected to the preservation of inherent relationships by the mapping S (according to Definition 2).

The following definition gives an exact mathematical formulation of this property.

Definition 3. Consider a linguistic variable $V = (N, T, X, G, S)$ and an index set I. Let $R = (R_i)_{i \in I}$ be a family of relations on the set of verbal values T, where each relation R_i has a finite arity a_i. Assume that, for every relation R_i, there exists a relation Q_i on the fuzzy power set $\mathcal{F}(X)$ with the same arity.[1] Correspondingly, we abbreviate the family $(Q_i)_{i \in I}$ with Q. Then the linguistic variable V is called *R-Q-interpretable* if and only if the following holds for all $i \in I$ and all $x_1, \ldots, x_{a_i} \in T$:

$$R_i\big(x_1, \ldots, x_{a_i}\big) \implies Q_i\big(S(x_1), \ldots, S(x_{a_i})\big). \tag{1}$$

[1] Q_i is associated with the "semantic counterpart" of R_i, i. e. the relation that models R_i on the semantic level.

Remark 4. The generalization of Definition 3 to fuzzy relations is straightforward. If we admit fuzziness of the relations R_i and Q_i, the implication in (1) has to be replaced by the inequality

$$R_i\big(x_1,\ldots,x_{a_i}\big) \le Q_i\big(S(x_1),\ldots,S(x_{a_i})\big).$$

4. A DETAILED STUDY BY MEANS OF PRACTICAL EXAMPLES

In almost all fuzzy control applications, the domains of the system variables are divided into a certain number of fuzzy sets by means of the underlying ordering – a fact which is typically reflected in expressions like "small", "medium", or "large". We will now discuss a simple example involving orderings to illustrate the concrete meaning of Definition 3.

Let us consider the following linguistic variable:

$$V = (\text{“v1”}, G, T, X, S).$$

The grammatical definition G is given as follows:

$$\begin{array}{lcl}
\bot & := & \langle\text{atomic}\rangle; \\
\langle\text{atomic}\rangle & := & \langle\text{adjective}\rangle \mid \langle\text{adverb}\rangle\ \langle\text{adjective}\rangle; \\
\langle\text{adjective}\rangle & := & \text{“small”} \mid \text{“medium”} \mid \text{“large”}; \\
\langle\text{adverb}\rangle & := & \text{“at least”} \mid \text{“at most”}.
\end{array}$$

Obviously, the following nine-element term set can be derived from G:

$$\begin{aligned}
T = \{&\text{“small”}, \text{“medium”}, \text{“large”}, \\
&\text{“at least small”}, \text{“at least medium”}, \\
&\text{“at least large”}, \text{“at most small”}, \\
&\text{“at most medium”}, \text{“at most large”}.\}
\end{aligned}$$

The universe of discourse is the real interval $X = [0, 100]$.

Taking the "background" or "context" of the variable into account, almost every human has an intuitive understanding of the qualitative meaning of each of the above linguistic expressions, even if absolutely nothing about the quantitative meaning, i. e. the corresponding fuzzy sets, is known. This understanding, to a major part, can be attributed to elementary relationships between the linguistic values. According to Definition 3, let us assume that these inherent relationships are modeled by a family of relations $R = (R_i)_{i\in I}$.

In our opinion, the most obvious relationships in the example term set T are orderings and inclusions. Therefore, we consider the following two binary relations (for convenience, we switch to infix notations here):

$$R = (\preceq, \sqsubseteq). \tag{2}$$

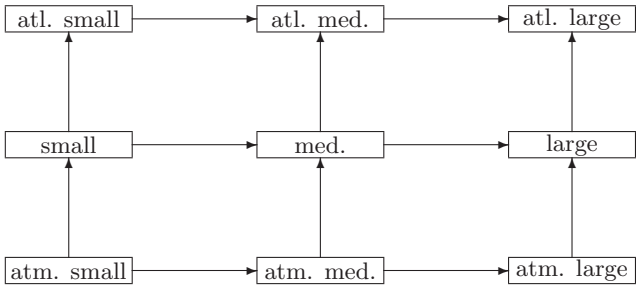

Fig. 1. Hasse diagram of ordering relation $\preceq$.

The first relation $\preceq$ stands for the ordering of the labels, while the second one corresponds to an inclusion relation, e. g. $u \sqsubseteq v$ means that v is a more general term than u.

First of all, one would intuitively expect a proper ordering of the adjectives, i. e.

$$\text{“small”} \preceq \text{“medium”} \preceq \text{“large”}. \tag{3}$$

Moreover, the following conditions of monotonicity seem reasonable for all adjectives u, v (atomic expressions from the set {“small”, “medium”, “large”}):

$$\begin{aligned} u &\sqsubseteq \text{“at least” } u \\ v &\sqsubseteq \text{“at most” } v \\ u \preceq v &\Longrightarrow \text{“at least” } u \preceq \text{“at least” } v \\ u \preceq v &\Longrightarrow \text{“at most” } u \preceq \text{“at most” } v \\ u \preceq v &\Longrightarrow \text{“at least” } v \sqsubseteq \text{“at least” } u \\ u \preceq v &\Longrightarrow \text{“at most” } u \sqsubseteq \text{“at most” } v. \end{aligned}$$

Figures 1 and 2 show Hasse diagrams which fully describe the two relations $\preceq$ and $\sqsubseteq$ (note that both relations are supposed to be reflexive, a fact which, for the sake of simplicity, is not made explicit in the diagrams).

Now we have to define meaningful counterparts of the relations in R on the semantic level, i. e. on $\mathcal{F}(X)$. We start with the usual inclusion of fuzzy sets according to Zadeh [42].

Definition 5. Consider two fuzzy sets of $A, B \in \mathcal{F}(X)$. A is called a *subset* of B, short $A \subseteq B$, if and only if, for all $x \in X$, $A(x) \leq B(x)$. Consequently, in this case, B is called a *superset* of A.

For defining a meaningful counterpart of the ordering relation $\preceq$, we adopt a simple variant of the general framework for ordering fuzzy sets proposed in [4, 5], which includes well-known orderings of fuzzy numbers based on the extension principle [27, 29].

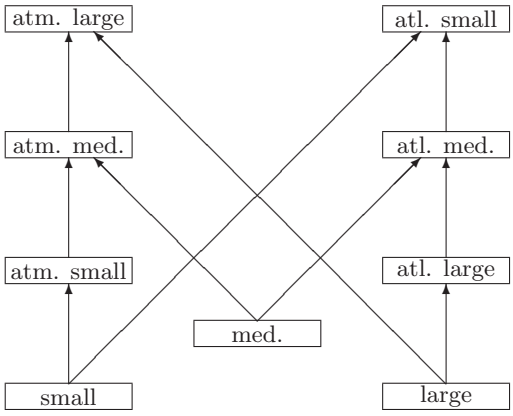

Fig. 2. Hasse diagram of inclusion relation $\sqsubseteq$.

Definition 6. Suppose that a universe X is equipped with a crisp linear ordering $\leq$. Then a preordering $\lesssim$ of fuzzy sets can be defined by

$$A \lesssim B \iff \big(\mathrm{ATL}(B) \subseteq \mathrm{ATL}(A) \text{ and } \mathrm{ATM}(A) \subseteq \mathrm{ATM}(B)\big),$$

where the operators ATL and ATM are defined as follows:

$$\mathrm{ATL}(A)(x) = \sup\{A(y) \mid y \leq x\}$$
$$\mathrm{ATM}(A)(x) = \sup\{A(y) \mid y \geq x\}.$$

Figure 3 shows an example what the operators ATL and ATM give for a non-trivial fuzzy set. It is easy to see that ATL always yields the smallest superset with non-decreasing membership function, while ATM yields the smallest superset with non-increasing membership function. For more details about the particular properties of the ordering relation $\lesssim$ and the two operators ATL and ATM, see [4, 5].

Summarizing, the set of counterpart relations Q looks as follows (with the relations from Definitions 5 and 6):

$$Q = (\lesssim, \subseteq). \tag{4}$$

Now R-Q-interpretability of linguistic variable V (with definitions of R and Q according to (2) and (4), respectively) specifically means that the following two implications hold for all $u, v \in T$:

$$u \preceq v \implies S(u) \lesssim S(v) \tag{5}$$
$$u \sqsubseteq v \implies S(u) \subseteq S(v). \tag{6}$$

This means that the mapping S plays the crucial role in terms of interpretability. In this particular case, R-Q-interpretability is the property that an ordering or inclusion

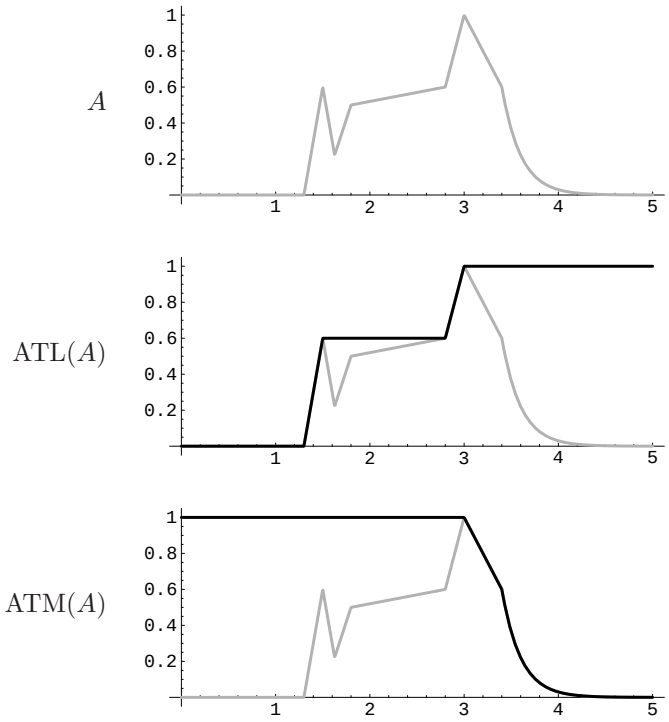

Fig. 3. A fuzzy set $A \in \mathcal{F}(\mathbb{R})$ and the results which are obtained when applying the operators ATL and ATM.

relationship between two linguistic terms is never violated by the two corresponding fuzzy sets. From (3) and (5), we can deduce the first basic necessary condition for the fulfillment of R-Q-interpretability – that the fuzzy sets associated with the three adjectives must be in proper order:

$$S(\text{“small”}) \lesssim S(\text{“medium”}) \lesssim S(\text{“large”}). \tag{7}$$

It is easy to observe that this basic ordering requirement is violated by the example shown in Figure 4, while it is fulfilled by the fuzzy sets in Figure 5.

In order to fully check R-Q-interpretability of V, the semantics of linguistic expressions containing an adverb (“at least” or “at most”) have to be considered as well. The definition of linguistic variables does not explicitly contain any hint how to deal with the semantics of such expressions. From a pragmatic viewpoint, two different ways are possible: one simple variant is to define a separate fuzzy set for each expression, regardless whether they contain an adverb or not. As a second traditional variant, we could use fuzzy modifiers – $\mathcal{F}(X) \to \mathcal{F}(X)$ functions – for modeling the semantics of adverbs. In this example, it is straightforward to use the

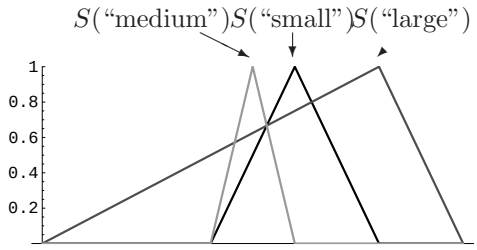

Fig. 4. A non-interpretable setting.

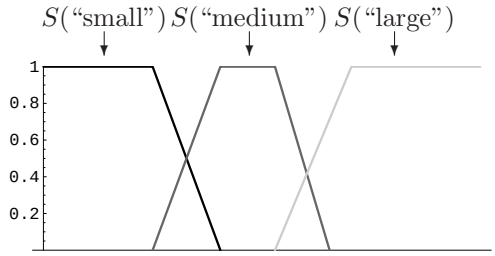

Fig. 5. An example of an interpretable setting.

fuzzy modifiers introduced in Definition 6 (see [4, 3, 15] for a detailed justification):

$$S(\text{“at least”}\,A) = \mathrm{ATL}\big(S(A)\big)$$
$$S(\text{“at most”}\,A) = \mathrm{ATM}\big(S(A)\big).$$

Since it is by far simpler and easier to handle with respect to interpretability, we strongly suggest the second variant.

In case that we use the above fuzzy modifiers for modeling the two adverbs “at least” and “at most”, we are now able to formulate a necessary condition for the fulfillment of R-Q-interpretability in our example.

Theorem 7. Consider the linguistic variable V and the two relation families R and Q as defined above. Provided that the mapping S always yields a normalized fuzzy set, the following two statements are equivalent:

(i) V is R-Q-interpretable

(ii) $S(\text{“small”}) \lesssim S(\text{“medium”}) \lesssim S(\text{“large”})$.

Proof.

(i) $\Rightarrow$ (ii): Trivial (see above).

(ii) ⇒ (i): The following basic properties hold for all normalized fuzzy sets $A, B \in \mathcal{F}(X)$ [4, 8]:

$$A \subseteq \mathrm{ATL}(A) \tag{8}$$
$$A \subseteq \mathrm{ATM}(A) \tag{9}$$
$$\mathrm{ATL}\big(\mathrm{ATL}(A)\big) = \mathrm{ATL}(A) \tag{10}$$
$$\mathrm{ATM}\big(\mathrm{ATM}(A)\big) = \mathrm{ATM}(A) \tag{11}$$
$$\mathrm{ATL}\big(\mathrm{ATM}(A)\big) = \mathrm{ATM}\big(\mathrm{ATL}(A)\big) = X \tag{12}$$
$$A \subseteq B \implies \mathrm{ATL}(A) \subseteq \mathrm{ATL}(B) \tag{13}$$
$$A \subseteq B \implies \mathrm{ATM}(A) \subseteq \mathrm{ATM}(B). \tag{14}$$

Since the relations $\subseteq$ and $\lesssim$ are reflexive and transitive [4, 5], it is sufficient to prove the relations indicated by arrows in the two Hasse diagrams (see Figures 1 and 2).

Let us start with the ordering relation. The validity of the relations in the middle row is exactly assumption (ii). The relations in the two other rows follow directly from the following two relationships which can be proved easily using (10), (11), and (12):

$$A \lesssim B \implies \mathrm{ATL}(A) \lesssim \mathrm{ATL}(B)$$
$$A \lesssim B \implies \mathrm{ATM}(A) \lesssim \mathrm{ATM}(B).$$

The three vertical relationships in Figure 1 follow directly from

$$\mathrm{ATM}(A) \lesssim A \lesssim \mathrm{ATL}(A)$$

which can be shown using (8), (9), (13), and (14).

The relations in the Hasse diagram in Figure 2 follow from (8), (9), and the definition of the preordering $\lesssim$ (cf. Definition 6). □

Obviously, interpretability of V in this example (with respect to the families R and Q) does not fully correspond to an intuitive human understanding of interpretability. For instance, all three expressions “small”, “medium”, and “large” could be mapped to the same fuzzy set without violating R-Q-interpretability. The intention was to give an example which is just expressive enough to illustrate the concrete meaning and practical relevance of Definition 3.

In order to formulate an example in which R-Q-interpretability is much closer to a human-like understanding of interpretability (e. g. including separation constraints), we have to consider an extended linguistic variable

$$V' = (\text{“v2”}, G', T', X', S').$$

The extended grammar G' is given as follows:

$$
\begin{aligned}
\bot \quad &:= \quad \langle \text{exp} \rangle \mid \langle \text{bounds} \rangle; \\
\langle \text{exp} \rangle \quad &:= \quad \langle \text{atomic} \rangle \mid \langle \text{atomic} \rangle\ \langle \text{binary} \rangle\ \langle \text{atomic} \rangle; \\
\langle \text{atomic} \rangle \quad &:= \quad \langle \text{adjective} \rangle \mid \langle \text{adverb} \rangle\ \langle \text{adjective} \rangle; \\
\langle \text{adjective} \rangle \quad &:= \quad \text{“small”} \mid \text{“medium”} \mid \text{“large”}; \\
\langle \text{adverb} \rangle \quad &:= \quad \text{“at least”} \mid \text{“at most”}; \\
\langle \text{binary} \rangle \quad &:= \quad \text{“and”} \mid \text{“or”}; \\
\langle \text{bounds} \rangle \quad &:= \quad \text{“empty”} \mid \text{“anything”}.
\end{aligned}
$$

It is easy to see that the corresponding term set T' has the following elements: the grammar admits nine atomic expressions (three adjectives plus two adverbs times three adjectives; note that this subset coincides with T from the previous example). Hence, there are $9 + 2 \cdot 9^2 = 171$ expressions of type $\langle \text{exp} \rangle$. Finally adding the two expressions of type $\langle \text{bounds} \rangle$, the term set T' has a total number of 173 elements.

As in the previous example, we would like to use an inclusion and an ordering relation. Since the two relations $\lesssim$ and $\subseteq$ are defined for arbitrary fuzzy sets, we can keep the relation family Q as it is. If we took R as defined above, R-Q-interpretability would be satisfied under the same conditions as in Theorem 7. This example, however, is intended to demonstrate that partition and convexity constraints can be formulated level of linguistic expressions, too. Therefore, we extend the inclusion relation $\sqsubseteq$ as follows. Let us consider a binary relation $\boxminus$ on T'. First of all, we require that $\boxminus$ coincides with $\sqsubseteq$ on the set of atomic expressions ($u, v \in T$):

$$(u \sqsubseteq v) \Longrightarrow (u \boxminus v). \tag{15}$$

Of course, we assume that the two binary connectives are non-decreasing with respect to inclusion, commutative, and that the “and” connective yields subsets and the “or” connective yields supersets (for all $u, v, w \in T$):

$$(v \boxminus w) \Longrightarrow (u \text{ “and” } v) \boxminus (u \text{ “and” } w) \tag{16}$$

$$(v \boxminus w) \Longrightarrow (u \text{ “or” } v) \boxminus (u \text{ “or” } w) \tag{17}$$

$$(u \text{ “and” } v) \boxminus (v \text{ “and” } u) \tag{18}$$

$$(u \text{ “or” } v) \boxminus (v \text{ “or” } u) \tag{19}$$

$$(u \text{ “and” } v) \boxminus u \tag{20}$$

$$u \boxminus (u \text{ “or” } v). \tag{21}$$

Next, let us suppose that “anything” is the most general and that “empty” is the least general expression, i. e., for all $u \in T'$,

$$u \boxminus \text{“anything”} \quad \text{and} \quad \text{“empty”} \boxminus u. \tag{22}$$

Now we can impose reasonable disjointness constraints like

$$\text{“small and at least medium”} \boxminus \text{“empty”} \tag{23}$$

$$\text{“at most medium and large”} \boxminus \text{“empty”} \tag{24}$$

and coverage properties:

$$\text{“small or medium”} \mathrel{\dot{\sqsubseteq}} \text{“at most medium”} \tag{25}$$
$$\text{“at most medium”} \mathrel{\dot{\sqsubseteq}} \text{“small or medium”} \tag{26}$$
$$\text{“anything”} \mathrel{\dot{\sqsubseteq}} \text{“at most medium or large”} \tag{27}$$
$$\text{“medium or large”} \mathrel{\dot{\sqsubseteq}} \text{“at least medium”} \tag{28}$$
$$\text{“at least medium”} \mathrel{\dot{\sqsubseteq}} \text{“medium or large”} \tag{29}$$
$$\text{“small or at least medium”} \mathrel{\dot{\sqsubseteq}} \text{“anything”}. \tag{30}$$

Finally, let us assume that “small” and “large” are the two boundaries with respect to the ordering of the labels:

$$\text{“at most small”} \mathrel{\dot{\sqsubseteq}} \text{“small”} \tag{31}$$
$$\text{“anything”} \mathrel{\dot{\sqsubseteq}} \text{“at least small”} \tag{32}$$
$$\text{“at least large”} \mathrel{\dot{\sqsubseteq}} \text{“large”} \tag{33}$$
$$\text{“anything”} \mathrel{\dot{\sqsubseteq}} \text{“at most large”}. \tag{34}$$

If we denote the reflexive and transitive closure of $\mathrel{\dot{\sqsubseteq}}$ with $\sqsubseteq'$, we can finally write down the desired family of relations:

$$R' = (\preceq, \sqsubseteq'). \tag{35}$$

In order to study the R'-Q-interpretability of V', we need to define the semantics of those expressions that have not been contained in T. Of course, for the expressions in T, we use the same semantics as in the previous example, i. e., for all $u \in T$, $S'(u) = S(u)$. Further, let us make the convention that the two expressions of type $\langle\text{bounds}\rangle$ are always mapped to the empty set and the whole universe, respectively:

$$S'(\text{“empty”}) = \emptyset \qquad S'(\text{“anything”}) = X.$$

The “and” and the “or” connective are supposed to be “implemented” by the intersection and union with respect to the Łukasiewicz t-norm and its dual t-conorm (for all $u, v \in T$):

$$S'(u \text{ “and” } v) = S'(u) \cap_{\mathbf{L}} S'(v)$$
$$S'(u \text{ “or” } v) = S'(u) \cup_{\mathbf{L}} S'(v).$$

Now we are able to fully characterize R'-Q-interpretability for the given example (the linguistic variable V').

Theorem 8. Provided that S' yields a normalized fuzzy set for each adjective, V' is R'-Q-interpretable if and only if the following three properties hold together:

1. $S'(\text{“small”}) \lesssim S'(\text{“medium”}) \lesssim S'(\text{“large”})$;
2. $S'(\text{“small”})$, $S'(\text{“medium”})$, and $S'(\text{“large”})$ are convex;
3. $S'(\text{“small”})$, $S'(\text{“medium”})$, and $S'(\text{“large”})$ form a Ruspini partition.

P r o o f . First of all, let us assume that V' is R'-Q-interpretable. The first property follows trivially as in the proof of Theorem 7. Now, taking (23) into account, we obtain from R'-Q-interpretability that

$$\begin{aligned} 0 &\geq T_{\mathbf{L}}\big(S'(\text{“small”})(x), S'(\text{“at least medium”})(x)\big) \\ &= T_{\mathbf{L}}\big(S'(\text{“small”})(x), \mathrm{ATL}(S'(\text{“medium”}))(x)\big) \\ &\geq T_{\mathbf{L}}\big(S'(\text{“small”})(x), S'(\text{“medium”})(x)\big), \end{aligned}$$

i. e. that the $T_{\mathbf{L}}$-intersection of $S'(\text{“small”})$ and $S'(\text{“medium”})$ is empty. Analogously, we are able to show that the $T_{\mathbf{L}}$-intersection of $S'(\text{“medium”})$ and $S'(\text{“large”})$ is empty, too. Since $S'(\text{“medium”}) \lesssim S'(\text{“large”})$ implies that

$$\mathrm{ATL}\big(S'(\text{“large”})\big) \subseteq \mathrm{ATL}\big(S'(\text{“medium”})\big),$$

it follows that, by the same argument as above, that the $T_{\mathbf{L}}$-intersection of $S'(\text{“small”})$ and $S'(\text{“large”})$ is empty as well. Now consider (25) and (26). R'-Q-interpretability then implies the following (for all $x \in X$):

$$\begin{aligned} S_{\mathbf{L}}\big(S'(\text{“small”})(x), S'(\text{“medium”})(x)\big) &= S'(\text{“at most medium”})(x) \\ &= \mathrm{ATL}\big(S'(\text{“medium”})\big)(x). \end{aligned}$$

Taking (27) into account as well, we finally obtain

$$\begin{aligned} 1 &\leq S_{\mathbf{L}}\big(S'(\text{“at most medium”})(x), S'(\text{“large”})(x)\big) \\ &= S_{\mathbf{L}}\big(S'(\text{“small”})(x), S'(\text{“medium”})(x), S'(\text{“large”})(x)\big) \end{aligned}$$

which proves that the $S_{\mathbf{L}}$-union of all fuzzy sets associated to the three adjectives yields the whole universe X. Since all three fuzzy sets are normalized and properly ordered, not more than two can have a membership degree greater than zero at a given point $x \in X$. This implies that the three fuzzy sets form a Ruspini partition [30].

From (31) and (33) and R'-Q-interpretability, we can infer that

$$\begin{aligned} \mathrm{ATM}\big(S'(\text{“small”})\big) &= S'(\text{“small”}), \\ \mathrm{ATL}\big(S'(\text{“large”})\big) &= S'(\text{“large”}), \end{aligned}$$

hence, $S'(\text{“small”})$ has a non-increasing membership function and $S'(\text{“large”})$ has a non-decreasing membership function. Both fuzzy sets, therefore, are convex. Since the three fuzzy sets $S'(\text{“small”})$, $S'(\text{“medium”})$, and $S'(\text{“large”})$ form a Ruspini partition, while only two can overlap to a positive degree, we have that $S'(\text{“medium”})$ is non-decreasing to the left of any value x for which $S'(\text{“medium”})(x) = 1$ and non-increasing to the right. Therefore, by Lemma 1, $S'(\text{“medium”})$ is convex, too.

Now let us prove the reverse direction, i. e. we assume all three properties and show that R'-Q-interpretability must hold. By (15) and the first property from Theorem 8, we can rely on the fact that all correspondences remain preserved for cases that are already covered by Theorem 7. Therefore, it is sufficient to show that S'

preserves all relationships (16) – (34). Clearly, the preservation of (16) – (21) follows directly from elementary properties of t-norms and t-conorms [28]. The inclusions (22) are trivially maintained, since S' is supposed to map "empty" to the empty set and "anything" to the universe X. The preservation of the two disjointness conditions (23) and (24) follows from the fact that $S'(\text{“small”})$, $S'(\text{“medium”})$, and $S'(\text{“large”})$ form a Ruspini partition and that the first property holds. The same is true for the six coverage properties (25) – (30). Since all three fuzzy sets are convex and form a Ruspini partition, $S'(\text{“small”})$ must have a non-increasing membership function and $S'(\text{“large”})$ must have a non-decreasing membership function. Therefore, the following needs to hold:

$$\begin{aligned} \mathrm{ATM}\big(S'(\text{“small”})\big) &= S'(\text{“small”}) \\ \mathrm{ATL}\big(S'(\text{“large”})\big) &= S'(\text{“large”}). \end{aligned}$$

This is a sufficient condition for the preservation of inclusions (31) and (33). Then the preservation of (32) and (34) follows from the fact, that $\mathrm{ATL}(\mathrm{ATM}(A)) = X$ for any normalized fuzzy set A (cf. (12)). Since $\sqsubseteq'$ and $\subseteq$ are both supposed to be transitive ($\sqsubseteq'$ being the transitive closure of the intermediate relation $\sqsubset$), the preservation of all other relationships follows instantly. □

At first glance, this example might seem unnecessarily complicated, since the final result is nothing else than exactly those common sense assumptions – proper ordering, convexity, partition constraints – that have been identified as crucial for interpretability before in several recent publications (see [10] for an overview). However, we must take into account that they are not just heuristic assumptions here, but necessary conditions that are enforced by intuitive requirements on the level of linguistic expressions. From this point of view, this example provides a sound justification for exactly those three crucial assumptions.

Now the question arises how the three properties can be satisfied in practice. In particular, it is desirable to have a constructive characterization of the constraints implied by requiring R'-Q-interpretability. The following theorem provides a unique characterization of R'-Q-interpretability under the assumption that we are considering real numbers and fuzzy sets with continuous membership functions – both are no serious restrictions from the practical point of view. Fortunately, we obtain a parameterized representation of all mappings S' that maintain R'-Q-interpretability.

Theorem 9. Assume that X is a connected subset of the real line and that S', for each adjective, yields a normalized fuzzy set with continuous membership function. Then the three properties from Theorem 8 are fulfilled if and only if there exist four values $a, b, c, d \in X$ satisfying $a < b \leq c < d$ and two continuous non-decreasing $[0,1] \to [0,1]$ functions f_1, f_2 fulfilling $f_1(0) = f_2(0) = 0$ and $f_1(1) = f_2(1) = 1$ such

that the semantics of the three adjectives are defined as follows:

$$S'(\text{“small”})(x) = \begin{cases} 1 & \text{if } x \leq a \\ 1 - f_1\left(\frac{x-a}{b-a}\right) & \text{if } a < x < b \\ 0 & \text{if } x \geq b \end{cases} \tag{36}$$

(37)

$$S'(\text{“medium”})(x) = \begin{cases} 0 & \text{if } x \leq a \\ f_1\left(\frac{x-a}{b-a}\right) & \text{if } a < x < b \\ 1 & \text{if } b \leq x \leq c \\ 1 - f_2\left(\frac{x-c}{d-c}\right) & \text{if } c < x < d \\ 0 & \text{if } x \geq d \end{cases} \tag{38}$$

$$S'(\text{“large”})(x) = \begin{cases} 0 & \text{if } x \leq c \\ f_2\left(\frac{x-c}{d-c}\right) & \text{if } c < x < d \\ 1 & \text{if } x \geq d. \end{cases} \tag{39}$$

P r o o f . It is a straightforward, yet tedious, task to show that the fuzzy sets defined as above fulfill the three properties from Theorem 8. Under the assumption that these three properties are satisfied, we make the following definitions:

$$\begin{aligned} a &= \sup\{x \mid S'(\text{“small”})(x) = 1\} \\ b &= \inf\{x \mid S'(\text{“medium”})(x) = 1\} \\ c &= \sup\{x \mid S'(\text{“medium”})(x) = 1\} \\ d &= \inf\{x \mid S'(\text{“large’})(x) = 1\}. \end{aligned}$$

The two functions f_1, f_2 can be defined as follows:

$$\begin{aligned} f_1(x) &= S'(\text{“medium”})\big(a + x \cdot (b-a)\big) \\ f_2(x) &= S'(\text{“large”})\big(c + x \cdot (d-c)\big). \end{aligned}$$

Since all membership functions associated with adjectives are continuous, the functions f_1 and f_2 are continuous. Taking the continuity and the fact that the three fuzzy sets associated to the adjectives form a Ruspini partition into account, it is clear that the following holds:

$$\begin{aligned} S'(\text{“small”})(a) &= 1 \\ S'(\text{“small”})(b) &= S'(\text{“small”})(c) = S'(\text{“small”})(d) = 0 \\ S'(\text{“medium”})(b) &= S'(\text{“medium”})(c) = 1 \\ S'(\text{“medium”})(a) &= S'(\text{“medium”})(d) = 0 \\ S'(\text{“large”})(d) &= 1 \\ S'(\text{“large”})(a) &= S'(\text{“large”})(b) = S'(\text{“large”})(c) = 0. \end{aligned}$$

These equalities particularly imply that $f_1(0) = f_2(0) = 0$ and $f_1(1) = f_2(1) = 1$ holds. As a consequence of convexity and Lemma 1, we know that the membership function of S'("medium") to the left of b is non-decreasing. Analogously, we can infer that the same is true for S'("large") to the left of d. Therefore, f_1 and f_2 are non-decreasing. To show that the three representations (36) – (39) hold is a routine matter. □

5. APPLICATIONS

5.1. Design aid

As long as the top-down construction of small fuzzy systems (e. g. two-input single-output fuzzy controllers) is concerned, interpretability is usually not such an important issue, since the system is simple enough that a conscious user will refrain from making settings which contradict his/her intuition.

In the design of complex fuzzy systems with a large number of variables and rules, however, interpretability is a most crucial point. Integrating tools which guide the user through the design of a large fuzzy system by preventing him/her from making non-interpretable settings accidentally are extremely helpful. As a matter of fact, debugging of large fuzzy systems becomes a tedious task if it is not guaranteed that the intuitive meanings of the labels used in the rule base are reflected in their corresponding semantics.

To be more precise, our goal is not to bother the user with additional theoretical aspects. Instead, the idea is to integrate these aspects into software tools for fuzzy systems design, but not necessarily transparent for the user, with the aim that he/she can build interpretable fuzzy systems in an even easier way than with today's software tools. Theorem 9 gives a clue how this could be accomplished. This result, for one particular example, clearly identifies how much freedom one has in choosing interpretable settings. The example is not quite representative, since three linguistic expressions are a quite restrictive assumption. However, the extension to an arbitrary finite number of such expressions is straightforward, no matter whether we consider such a typical "small"-"medium"-"large" example or a kind of symmetric setting (e. g. "neg. large", "neg. medium", "neg. small", "approx. zero", "pos. small", "pos. medium", "pos. large") as it is common in many fuzzy control applications. In all these cases, the requirements for interpretability are similar and, by Theorem 9, the resulting set of degrees of freedom is an increasing chain of values that mark the beginning/ending of the kernels of the fuzzy sets and a set of continuous non-decreasing functions that control the shape of the transitions between two neighboring fuzzy sets. While linear transitions are common and easy to handle, smooth transitions by means of polynomial functions with higher degree may be beneficial in some applications as well.

As simple examples, the following three polynomial $[0, 1] \to [0, 1]$ functions of degrees 1, 3, and 5 perfectly serve as transition functions in the sense of Theorem 9. They produce membership functions that are continuous (p_1), differentiable (p_3),

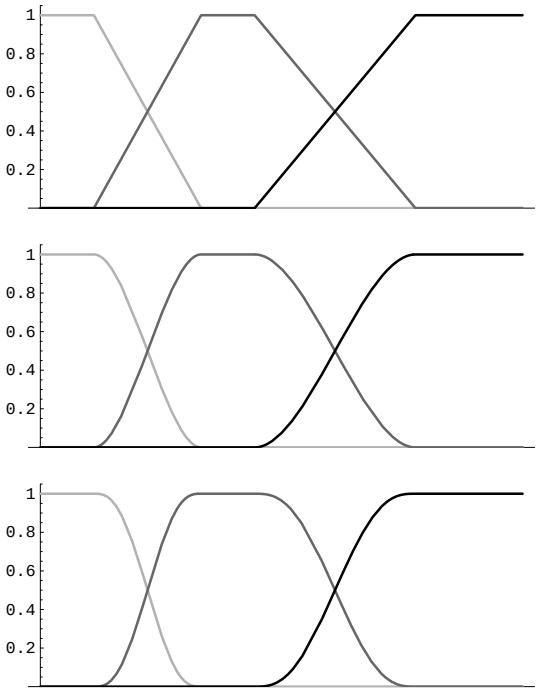

Fig. 6. Three interpretable fuzzy partitions
with polynomial transitions of degree 1 (top), 3 (middle), and 5 (bottom).

and twice differentiable (p_5), respectively:

$$p_1(x) = x$$
$$p_3(x) = -2x^3 + 3x^2$$
$$p_5(x) = 6x^5 - 15x^4 + 10x^3.$$

Figure 6 shows examples of interpretable fuzzy partitions with three fuzzy sets using the transition functions p_1, p_3, and p_5.

5.2. Data–driven learning and tuning

Automatic design and tuning of fuzzy systems has become a central issue in machine learning, data analysis, and the identification of functional dependencies in the analysis of complex systems. In the last years, a vast number of scientific publications dealt with this problem. Most of them, however, disregarded the importance of interpretability – leading to results which are actually black-box functions that do not provide any meaningful linguistic information (typical pictures like in Figure 4 can be found in an enormous number of papers).

One may argue that proper input-output behavior is the central goal of automatic

tuning. To some extent, this is true; however, as stated already in Section 1, this is not the primary mission of fuzzy systems.

Again, Theorem 9 gives a clear indication how the space of possible solutions among interpretable settings may be parameterized – by an ascending chain of transition points (given a set of transition functions). Note that this kind of parametrization even leads to a reduction of the search space. Parameterizing three trapezoid fuzzy sets independently requires a total number of twelve parameters and most probably leads to difficulty interpretable results. Requiring interpretability (as described in the previous section) leads to a set of only four parameters. It is true that such a setting is much more restrictive. However, in our opinion, it is not necessarily the case that requiring interpretability automatically leads to a painful loss of accuracy. The requirement of interpretability implies more constraints that have to be taken into account and, therefore, is more difficult to handle for many tuning algorithms, no matter whether we consider genetic algorithms, fuzzy-neuro methods, or numerical optimization. As a recent investigation has shown, it is indeed possible to require interpretability while, at the same time, maintaining high accuracy and robustness [26]. Many other studies have also come up with tuning algorithms that produce interpretable *and* accurate results [2, 9, 16, 17, 21, 34, 40].

5.3. Rule base simplification

The examples in Section 4 used an inclusion relation $\sqsubseteq$ and its counterpart on the semantic side – the inclusion relation $\subseteq$. Both relations are preorderings which particularly implies that their symmetric kernels are equivalence relations. As easy to see, the relation $\subseteq$ is even an ordering on $\mathcal{F}(X)$, which implies that the symmetric kernel is the crisp equality relation, i. e. $A \subseteq B$ and $B \subseteq A$ hold together if and only if the two membership functions coincide exactly.

Now let us assume that we are given a linguistic variable V and two relation families R and Q, where R contains an inclusion relation $\sqsubseteq$ and Q contains its counterpart $\subseteq$. If we have two linguistic labels u and v for which $u \sqsubseteq v$ and $v \sqsubseteq v$ hold, R-Q-interpretability guarantees that the inclusions $S(u) \subseteq S(v)$ and $S(v) \subseteq S(u)$ hold, i. e. $u \sqsubseteq v$ and $v \sqsubseteq v$ are sufficient conditions that the membership functions of $S(u)$ and $S(v)$ are equal. This means that the equivalence relation defined as (for two $u, v \in T$)

$$(u \equiv v) \iff (u \sqsubseteq v \text{ and } v \sqsubseteq u)$$

may be considered as a set of *simplification rules*, while R-Q-interpretability corresponds to the validity of these rules on the semantic side.

Let us recall the second example (linguistic variable V' defined as in Section 4). The two inequalities (25) and (26) together imply

$$\text{“small or medium”} \equiv \text{“at most medium”}.$$

This could be read as a replacement rule

$$\text{“small or medium”} \longrightarrow \text{“at most medium”},$$

with the meaning that, in any linguistic expression, "small or medium" can be replaced by "at most medium". If we assume interpretability, we can be sure that this replacement is also semantically correct. If we incorporate as many reasonable relationships in the relation $\sqsubseteq$ as possible such that interpretability can still be fulfilled, we are able to provide a powerful set of simplification rules.

Of course, very simple grammars do not necessitate any simplification. However, if we have to consider very complex rule bases like they appear in grammar-based rule base optimization methods (e. g. inductive learning [32, 33, 35, 36] or fuzzy genetic programming [23, 24]), simplification is a highly important concern. The methodology presented here allows to deal with simplification in a symbolic fashion – assuming interpretability – without the need to consider the concrete semantics of the expressions anymore.

6. CONCLUSION

This paper has been devoted to the interpretability of linguistic variables. In order to approach this key property in a systematic and mathematically exact way, we have proposed to make implicit relationships between the linguistic labels explicit by formulating them as (fuzzy) relations. Then interpretability corresponds to the preservation of this relationships by the associated meaning. This idea has been illustrated by means of two extensive examples. These case studies have demonstrated that well-known common sense assumptions about the membership functions, such as, ordering, convexity, or partition constraints, have a sound justification also from a formal linguistic point of view. In contrast to other investigations, the model proposed in this paper cannot just be applied to simple fuzzy sets, but also allows smooth integration of connectives and ordering-based modifiers.

By characterizing parameterizations which ensure interpretability, we have been able to provide hints for the design and tuning of fuzzy systems with interpretable linguistic variables. Finally, we have seen that interpretability even corresponds to the fact that symbolic simplification rules on the side of linguistic expressions still remain valid on the semantic side.

Looking back on the questions posed in Section 1, this paper has been concerned with the fourth one. However, the ideas presented in this paper also have strong influence on the way the Questions 1 – 3 can be handled. Partitions constraints, as we derived them, enforce semantic properties of the expressions used in the rule antecedents that ease to investigate/guarantee completeness and/or consistency (cf. Question 3). As we use linguistic variables in their most general form, high-level language elements, such as linguistic modifiers or more advanced connectives, can smoothly be integrated, which directly lead to more compact rule sets (cf. Question 2). The ideas stated in Subsection 5.3. strongly support this viewpoint, too. The first question is a slightly different matter which should rather be approached from the side of approximate reasoning [18, 19, 20] and relational equations [13, 25, 30]. However, even following these lines, partition constraints play a crucial role. Therefore, we dare to conclude that interpretability of linguistic variables is a most basic requirement for any study of aspects of interpretability of fuzzy

systems.

ACKNOWLEDGEMENTS

Ulrich Bodenhofer gratefully acknowledges support by the Austrian Government, the State of Upper Austria, and the Johannes Kepler University Linz in the framework of the K*plus* Competence Center Program.

(Received June 20, 2004.)

REFERENCES

[1] R. Babuška: Construction of fuzzy systems – interplay between precision and transparency. In: Proc. Europ. Symp. on Intelligent Techniques, Aachen 2000, pp. 445–452.

[2] M. Bikdash: A highly interpretable form of Sugeno inference systems. IEEE Trans. Fuzzy Systems *7* (1999), 686–696.

[3] U. Bodenhofer: The construction of ordering-based modifiers. In: Fuzzy-Neuro Systems '99 (G. Brewka, R. Der, S. Gottwald and A. Schierwagen, eds.), Leipziger Universitätsverlag 1999, pp. 55–62.

[4] U. Bodenhofer: A Similarity-Based Generalization of Fuzzy Orderings. (Schriftenreihe der Johannes-Kepler-Universität Linz *C26.*) Universitätsverlag Rudolf Trauner, Linz 1999.

[5] U. Bodenhofer: A general framework for ordering fuzzy sets. In: Technologies for Constructing Intelligent Systems 1: Tasks, (B. Bouchon-Meunier, J. Guitiérrez-Ríoz, L. Magdalena, and R. R. Yager, eds., Studies in Fuzziness and Soft Computing 89), Physica–Verlag, Heidelberg 2002, pp. 213–224.

[6] U. Bodenhofer and P. Bauer: Towards an axiomatic treatment of "interpretability". In: Proc. 6th Internat. Conference on Soft Computing, Iizuka 2000, pp. 334–339.

[7] U. Bodenhofer and P. Bauer: A formal model of interpretability of linguistic variables. In: Interpretability Issues in Fuzzy Modeling (J. Casillas, O. Cordón, F. Herrera and L. Magdalena, eds.), Studies in Fuzziness and Soft Computing 128), Springer, Berlin 2003, pp. 524–545.

[8] U. Bodenhofer, M. De Cock, and E. E. Kerre: Openings and closures of fuzzy preorderings: Theoretical basics and applications to fuzzy rule-based systems. Internat. J. General Systems *4* (2003), 343–360.

[9] U. Bodenhofer and E. P. Klement: Genetic optimization of fuzzy classification systems – a case study. In: Computational Intelligence in Theory and Practice (B. Reusch and K.-H. Temme, eds., Advances in Soft Computing), Physica–Verlag, Heidelberg 2001, pp. 183–200.

[10] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena: Interpretability improvements to find the balance interpretability-accuracy in fuzzy modeling: an overview. In: Interpretability Issues in Fuzzy Modeling (J. Casillas, O. Cordón, F. Herrera and L. Magdalena, eds., Studies in Fuzziness and Soft Computing 128), Springer–Verlag, Berlin 2003, pp. 3–24.

[11] J. Casillas, O. Cordón, F. Herrera, and L. Magdalena (eds.): Interpretability Issues in Fuzzy Modeling (Studies in Fuzziness and Soft Computing 128). Springer–Verlag, Berlin 2003.

[12] O. Cordón and F. Herrera: A proposal for improving the accuracy of linguistic modeling. IEEE Trans. Fuzzy Systems *8* (2000), 335–344.

[13] B. De Baets: Analytical solution methods for fuzzy relational equations. In: Fundamentals of Fuzzy Sets (D. Dubois and H. Prade, eds., The Handbooks of Fuzzy Sets 7), Kluwer Academic Publishers, Boston 2000, pp. 291–340.
[14] B. De Baets and R. Mesiar: T-partitions. Fuzzy Sets and Systems *97* (1998), 211–223.
[15] M. De Cock, U. Bodenhofer, and E. E. Kerre: Modelling linguistic expressions using fuzzy relations. In: Proc. 6th Internat. Conference on Soft Computing, Iizuka 2000, pp. 353–360.
[16] M. Drobics and U. Bodenhofer: Fuzzy modeling with decision trees. In: Proc. 2002 IEEE Inernat. Conference on Systems, Man and Cybernetics, Hammamet 2002.
[17] M. Drobics, U. Bodenhofer, and E. P. Klement: FS-FOIL: An inductive learning method for extracting interpretable fuzzy descriptions. Internat. J. Approx. Reason. *32* (2003), 131–152.
[18] D. Dubois and H. Prade: What are fuzzy rules and how to use them. Fuzzy Sets and Systems *84* (1996), 169–185.
[19] D. Dubois, H. Prade, and L. Ughetto: Checking the coherence and redundancy of fuzzy knowledge bases. IEEE Trans. Fuzzy Systems *5* (1997), 398–417.
[20] D. Dubois, H. Prade, and L. Ughetto: Fuzzy logic, control engineering and artificial intelligence. In: Fuzzy Algorithms for Control (H. B. Verbruggen, H.-J. Zimmermann, and R. Babuška, eds., International Series in Intelligent Technologies), Kluwer Academic Publishers, Boston 1999, pp. 17–57.
[21] J. Espinosa and J. Vandewalle: Constructing fuzzy models with linguistic integrity from numerical data – AFRELI algorithm. IEEE Trans. Fuzzy Systems *8* (2000), 591–600.
[22] J. Fodor and M. Roubens: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht 1994.
[23] A. Geyer–Schulz: Fuzzy Rule-based Expert Systems and Genetic Machine Learning. (Studies in Fuzziness *3*.) Physica–Verlag, Heidelberg 1995.
[24] A. Geyer–Schulz: The MIT beer distribution game revisited: Genetic machine learning and managerial behavior in a dynamic decision making experiment. In: Genetic Algorithms and Soft Computing (F. Herrera and J. L. Verdegay, eds.), Studies in Fuzziness and Soft Computing *8*, Physica–Verlag, Heidelberg 1996, pp. 658–682.
[25] S. Gottwald: Fuzzy Sets and Fuzzy Logic. Vieweg, Braunschweig 1993.
[26] J. Haslinger, U. Bodenhofer, and M. Burger: Data-driven construction of Sugeno controllers: Analytical aspects and new numerical methods. In: Proc. Joint 9th IFSA World Congress and 20th NAFIPS Internat. Conference, Vancouver 2001, pp. 239–244.
[27] E. E. Kerre, M. Mareš, and R. Mesiar: On the orderings of generated fuzzy quantities. In: Proc. 7th Internat. Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris 1998, pp. 250–253.
[28] E. P. Klement, R. Mesiar, and E. Pap: Triangular Norms (Trends in Logic 8). Kluwer Academic Publishers, Dordrecht 2000.
[29] L. T. Kóczy and K. Hirota: Ordering, distance and closeness of fuzzy sets. Fuzzy Sets and Systems *59* (1993), 281–293.
[30] R. Kruse, J. Gebhardt, and F. Klawonn: Foundations of Fuzzy Systems. Wiley, New York 1994.
[31] R. Lowen: Convex fuzzy sets. Fuzzy Sets and Systems *3* (1980), 291–310.
[32] R. S. Michalski, I. Bratko, and M. Kubat: Machine Learning and Data Mining. Wiley, Chichester 1998.
[33] S. Muggleton and L. De Raedt: Inductive logic programming: Theory and methods. J. Logic Program. *19/20* (1994), 629–680.
[34] W. Pedrycz and Z. A. Sosnowski: Designing decision trees with the use of fuzzy granulation. IEEE Trans. Systems Man Cybernet. A *30* (2000), 151–159.

[35] J. R. Quinlan: Induction of decision trees. Mach. Learning *1* (1986), 81–106.
[36] J. R. Quinlan: Learning logical definitions from relations. Mach. Learning *5* (1990), 239–266.
[37] A. Ralston, E. D. Reilly, and D. Hemmendinger (eds.): Encyclopedia of Computer Science. Fourth edition. Groves Dictionaries, Williston 2000.
[38] E. H. Ruspini: A new approach to clustering. Inform. and Control *15* (1969), 22–32.
[39] M. Setnes, R. Babuška, and H. B. Verbruggen: Rule-based modeling: Precision and transparency. IEEE Trans. Systems Man Cybernet. C *28* (1998), 165–169.
[40] M. Setnes and H. Roubos: GA-fuzzy modeling and classification: Complexity and performance. IEEE Trans. Fuzzy Systems *8* (2000), 509–522.
[41] J. Yen, L. Wang, and C. W. Gillespie: Improving the interpretability of TSK fuzzy models by combining global learning and local learning. IEEE Trans. Fuzzy Systems *6* (1998), 530–537.
[42] L. A. Zadeh: Fuzzy sets. Inform. and Control *8* (1965), 338–353.
[43] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning I. Inform. Sci. *8* (1975), 199–250.
[44] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning II. Inform. Sci. *8* (1975), 301–357.
[45] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning III. Inform. Sci. *9* (1975), 43–80.

Ulrich Bodenhofer, Software Competence Center Hagenberg, A-4232 Hagenberg. Austria.
e-mail: ulrich.bodenhofer@scch.at

Peter Bauer, COMNEON electronic technology, A-4040 Linz. Austria.
e-mail: peter.bauer@comneon.com