On Perception-based Logical Deduction and Its Variants

Martin Štěpnička, Antonín Dvořák

Institute for Research and Applications of Fuzzy Modeling, CE IT4Innovations, University of Ostrava, 30. dubna 22, 701 03 Ostrava, Czech Republic

martin.stepnicka@osu.cz, antonin.dvorak@osu.cz

Abstract

We present and analyze inference method called Perception-based Logical Deduction (PbLD) aimed at the treatment of fuzzy IF-THEN rules as linguistically expressed genuine logical implications. Besides the original PbLD, we propose a new balancing variant of PbLD, introduce both variants with fuzzy inputs and study them from the point of view of the interpolativity property.

Keywords: Perception-based Logical Deduction, Fuzzy interpolation, Inference mechanism

1. Introduction

1.1. Fuzzy Rules and Inference Mechanisms

Consider a finite set of fuzzy IF-THEN rules (fuzzy rules), which is called a fuzzy rule base:

$$\mathcal{R}_i := \mathsf{IF} \ X \text{ is } \mathcal{A}_i \ \mathsf{THEN} \ Y \text{ is } \mathcal{B}_i, \quad i = 1, \dots, n$$

where X,Y are linguistic variables [1] and $\mathcal{A}_i,\mathcal{B}_i$ are linguistic labels (e.g., "small", "around ten" etc.). The interpretation of expressions $\mathcal{A}_i,\mathcal{B}_i$ is modeled in a standard manner by appropriate antecedent and consequent fuzzy sets A_i,B_i , which are defined on some non-empty universes U,V, respectively. For our purposes, we can restrict our choice of U,V to closed real intervals. There exist two standard approaches to modelling a given fuzzy rule base by an appropriate fuzzy relation $R \in \mathcal{F}(U \times V)$ [2]. The first approach constructs $\hat{R} \in \mathcal{F}(U \times V)$:

$$\hat{R}(u,v) = \bigwedge_{i=1}^{n} (A_i(u) \to B_i(v))$$
 (2)

where \rightarrow is a fuzzy implication. In this *implicative* approach, each fuzzy rule is viewed as a constraint. Hence, all constraints are merged in a conjunctive way (the more rules, the more constraints and the less possible values to satisfy them) [2]. Therefore, \hat{R} based on a residual implication seems to be appropriate to model the conditional form of the rules.

The second approach, initiated by a successful experimental application by Mamdani and Assilian [3], consists in constructing the fuzzy relation

 $\check{R} \in \mathcal{F}(U \times V)$ defined by

$$\check{R}(u,v) = \bigvee_{i=1}^{n} (A_i(u) * B_i(v)), \qquad (3)$$

where * is a t-norm. Obviously, \check{R} can hardly be considered as a model of fuzzy rule base (1). The fuzzy rules modelled by (3) are not viewed as constraints but are considered as pieces of data. The maximum in (3) then expresses accumulation of data [2]. For further sources related to the interpretability and interpretation of such systems, we refer to [4, 5, 6].

There are many inference mechanisms that, with the help of fuzzy rules, deduce an appropriate output $B_0 \in \mathcal{F}(V)$ based on a given observation (input) $A_0 \in \mathcal{F}(U)$. Let us restrict our focus on the fuzzy relational inference systems that directly use a fuzzy relation $R \in \mathcal{F}(U \times V)$ as a model of fuzzy rules, and an image of a fuzzy set under the fuzzy relation as a model of the inference mechanism. Most often, the direct image (sup-* composition) is used,

$$B_0 = A_0 \circ R \,, \tag{4}$$

also called Compositional Rule of Inference (CRI) [7], which is defined, for all $v \in V$, by

$$(A_0 \circ R)(v) = \bigvee_{u \in U} (A_0(u) * R(u, v)).$$
 (5)

Another alternative is the *Bandler-Kohout sub-product* (BK-subproduct):

$$B_0 = A_0 \lhd R \,, \tag{6}$$

which is defined, for all $v \in V$, by

$$(A_0 \triangleleft R)(v) = \bigwedge_{u \in U} (A_0(u) \to R(u, v)). \tag{7}$$

The BK-subproduct was firstly suggested as an inference mechanism in [8] and later on, in [9], it was shown that both inference mechanisms are equally good. Particularly, the advantages, which may be obtained from using one or the other inference mechanism, do not come from the mechanism itself, but from a proper combination of the mechanism and the model of fuzzy rules. If an inference mechanism provides some advantage when connected with the model \check{R} , the same holds for

the other inference mechanism connected to \hat{R} and, vice-versa, the same holds for disadvantages. There is no advantage of any of the two above mentioned inferences that would hold independently on the choice of the fuzzy rule base model, see [9].

The only exception in favor of the implicative approach represented by \hat{R} is an existence of a sort of "built-in" consistency checking mechanism [10], which uses the definition of the notion of coherence [11]. This only underlines the fact that implicative rules are not only more appropriate semantical model with respect to the syntactical form of the IF-THEN rules (1), but at least equally or even more practical than the Mamdani-Assilian rules that are used significantly more often [12]. Therefore, we will focus mainly on the implicative type in this paper.

1.2. Motivation

Models of fuzzy rule bases and inference mechanisms described in the previous subsection based on CRI or BK-subproduct are well suited for the purpose of approximation of an unknown function characterized imprecisely by fuzzy rules. The fuzzy sets A_i , B_i , which interpret linguistic labels A_i , B_i , are usually of one of standard shapes, e.g., triangular, trapezoidal etc. Further, they are usually uniformly distributed along the intervals of real numbers U and V, forming a fuzzy partition.

If we are interested in capturing the meaning of linguistic labels, which are used by humans most often, in a way that is in accordance with their intuitive understanding, we have to search for another model of these labels. Consider, for example, linguistic expressions "extremely small" and "small". We argue that, according to intuition shared by humans, if something is extremely small, it is, at the same time, small. Using the language of fuzzy set theory, the interpretation of "extremely small" should be a fuzzy subset of the interpretation of "small". Typical shapes of interpretations of these linguistic expressions are depicted in Fig. 1. If we accept that a model of these linguistic expressions should possess this property, we find that inference mechanisms and fuzzy rule base models described above are not well suited for it.

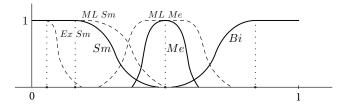
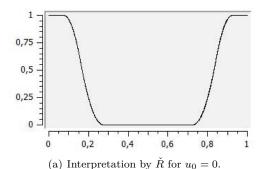


Figure 1: Graphical representations of fuzzy sets that interpret linguistic expressions extremely small, small, more or less small, more or less medium, medium and big.



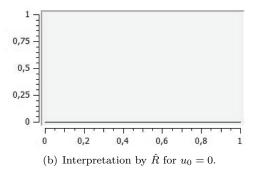


Figure 2: Results of \tilde{R} and \hat{R} interpretations of \mathcal{RB}_1 in case of the inclusive interpretation of extremely small and small.

For example, consider the following fuzzy rule base \mathcal{RB}_1 :

 $\mathcal{R}_1 := \mathsf{IF}\ X \text{ is } extremely \ small\ \mathsf{THEN}\ Y \text{ is } \mathit{big},$

 $\mathcal{R}_2 := \mathsf{IF}\ X \text{ is } small\ \mathsf{THEN}\ Y \text{ is } small,$

 $\mathcal{R}_3 := \mathsf{IF}\ X \text{ is } medium\ \mathsf{THEN}\ Y \text{ is } medium.$

If interpretations of extremely small and small are as in Fig. 1 and the observation is $u_0 = 0$, we expect that the rule \mathcal{R}_1 has to be used, and the result has to correspond to linguistic expression big. However, as can be seen in Fig. 2, results of the conjunctive interpretation of fuzzy IF-THEN rules \hat{R} (2) as well as of the disjunctive interpretation \check{R} (3) are unsatisfactory. The reason is that both rules \mathcal{R}_1 and \mathcal{R}_2 are fired for the observation $u_0 = 0$. Because consequents of these rules are very different (contradictory), the inference mechanism either put both consequents into the conclusion (Fig. 2(a)), or it annihilates both consequents into the constant (empty) fuzzy set (Fig. 2(b)).

Further, we do not view a set of fuzzy IF-THEN rules as a description of an unknown function, but as a set of genuine linguistically expressed logical implications. Based on these considerations, a method called *Perception-based Logical Deduction* (abbr. PbLD) has been developed [13, 14, 15]. *Perception-based*, because to an input we assign only the most fitting linguistic expression(s) from antecedents of a fuzzy rule base, and call them *perceptions*. PbLD then fires the fuzzy rules corresponding to the perceptions. *Logical*, because it understands IF-THEN rules as logical implications and it has

been developed within a formal logical theory.¹

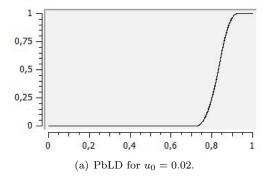
A fuzzy relational model is interpolative (correct) if and only if it models a certain type of continuous behavior [16]. In general, this can be viewed as a feature rather than a disadvantage, but there are situations where even a smooth fuzzy model should be able to provide a discontinuous jump [16]. Consider the following fuzzy rule base \mathcal{RB}_2 :

$$\mathcal{R}_1 := \mathsf{IF}\ X \text{ is } small\ \mathsf{THEN}\ Y \text{ is } -big,$$
 $\mathcal{R}_2 := \mathsf{IF}\ X \text{ is } extremely\ small\ \mathsf{THEN}\ Y \text{ is } +big,$

where X is a distance of a vehicle from a traffic signal when the yellow light appears on it and Y is a desired change of speed of that vehicle. Obviously, the change from braking to speeding up (big negative and positive changes in speed) is something that cannot be provided by a correct (consistent) model using a fuzzy IF-THEN rule base within a fuzzy relational inference. However, for PbLD method it is no problem to provide conclusions based on this rule base in accordance with intuition.

As we mentioned above, PbLD fires the rules whose antecedents fit the observation best. But what does it mean? Consider again the fuzzy rule base \mathcal{RB}_1 . Similarly as in case of the rule base \mathcal{RB}_2 above, there is a substantial difference between consequents corresponding to antecedents extremely small (rule \mathcal{R}_1) and small (rule \mathcal{R}_2). Intuitively, if an observation is extremely small, only \mathcal{R}_1 should be used. Similarly, if an observation is small (but not extremely small) only \mathcal{R}_2 should be used. The idea is that we use the rule which fits the observation best (the membership degree of the observation in fuzzy sets interpreting antecedent linguistic expressions is maximal).² If there are more such rules, then we use the rule whose antecedent is most specific. Consider the interval [0, 1] as the universe for the variable X. If the observation is, say, 0.1, we use the rule \mathcal{R}_2 , because the firing degree of the observation is maximal for it. The result of PbLD inference in this case is depicted in Fig. 3(b) and corresponds to the consequent of the rule \mathcal{R}_2 , that is, small. If the observation moves leftwards to 0.02, then antecedents of both rules \mathcal{R}_1 and \mathcal{R}_2 are to be selected (both firing degrees are 1). However, we use the rule \mathcal{R}_1 only, because the linguistic expression extremely small is more specific than small. We can see the result of PbLD inference in this case in Fig. 3(a). It corresponds to linguistic expression big, that is, to the consequent of the fuzzy rule \mathcal{R}_1 .

Up to now, the study and applications of Perception-based Logical Deduction have been conducted mainly for the case of crisp observations. However, the more general case of fuzzy observations is important from the theoretical as well as



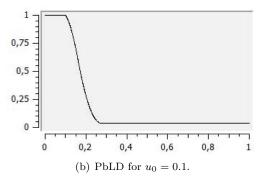


Figure 3: Results of PbLD for fuzzy rule base \mathcal{RB}_1 .

the practical viewpoint.³ Because a crisp observation can be understood as a fuzzy singleton, we intuitively expect that PbLD inference method with crisp observations should be a special case of PbLD with fuzzy observations.

One important point has to be addressed in the case of PbLD with fuzzy inputs. The fired fuzzy IF-THEN rule is selected according to the firing degree of the observation. However, for fuzzy inputs, this notion is not defined. We can use a *similarity relation* as in [17], but it can be argued that not all properties usually required for similarity relations are necessary here. In Section 2.4, we propose a *generalized firing degree* to be used for this purpose.

Fuzzy inference systems including PbLD can be studied from the viewpoint of various mathematical properties based on intuitive requirements. For example, interpolativity says that if a fuzzy input (observation) is equal to the antecedent of fuzzy rule \mathcal{R}_i , then the inferred conclusion should be equal to the consequent of \mathcal{R}_i . Other properties include continuity, monotonicity etc. Due to space limitations, we study only interpolativity in this paper (Section 3).

2. PbLD Revisited

2.1. Mathematical Background

We fix the Łukasiewicz algebra

$$\mathcal{L} = \langle [0,1], \wedge, \vee, *, \to 0, 1 \rangle$$

 $^{^1\}mathrm{Logical}$ analysis of PbLD is omitted in this paper due to lack of space.

²We will call this degree a *firing degree* of the observation u_o in fuzzy rule \mathcal{R}_i .

 $[\]overline{\ \ }^3 \text{Fuzzy observations}$ occur, for instance, in hierarchical fuzzy systems.

as the background algebraic structure. Thus, * is the Łukasiewicz t-norm defined as $a*b=(a+b-1)\vee 0$, and \to is its residual Łukasiewicz implication $a\to b=(1-a+b)\wedge 1$, for $a,b\in [0,1]$.

Major components of systems of fuzzy IF-THEN rules are evaluative linguistic expressions [18], evaluative expressions for short. Their importance and the potential for modeling their meaning in a mathematical manner were discussed, e.g., in [1].

A simple form of an evaluative expression has the following structure:

 $\langle \text{linguistic hedge} \rangle \langle \text{atomic evaluative expression} \rangle$.

An atomic evaluative expression is one of the canonical adjectives: small, medium and big, which we abbreviate in the following as Sm, Me and Bi, respectively.⁴ In practice, these words are often replaced by other evaluative words, such as "thin", "old", or "new", depending on the context of speech.

Linguistic hedges are specific adverbs that make the interpretations of atomic expressions more or less precise. We may distinguish between hedges with a narrowing effect and with a widening effect, (a special case is the empty hedge). Of course, the number of hedges is limited in practical applications. In the following, without loss of generality, we use the hedges introduced in Table 1, which were successfully employed in real applications [19]. They are also implemented in the LFLC software package [20]. Note that our hedges are of the socalled inclusive type [21], which means that the interpretations (fuzzy sets) of more specific evaluative expressions are included in those of less specific ones, as shown in Fig. 1. Note also that we always suppose the interpretations of evaluative expressions to be *normal* fuzzy sets.

| Narrowing effect | Widening effect |
|--------------------|--------------------|
| very (Ve) | more or less (ML) |
| significantly (Si) | roughly (Ro) |
| extremely (Ex) | quite roughly (QR) |

Table 1: Linguistic hedges and their abbreviations.

The theory of evaluative expressions was extended in [15] with the following partition axiom adding the assumption that fuzzy sets which model evaluative expressions cannot overlap in degree 1 if these expressions are not of the same atomic type. Indeed, no element u in any universe is naturally assumed to belong in degree 1 to a fuzzy set of small objects, as well as of medium or big objects, regardless of the influence of the widening or narrowing effect of applied linguistic hedges.

Axiom 1 Let A_1, A_2 be evaluative expressions with different atomic expressions that are modeled by fuzzy sets A_1, A_2 on U. Then, for all $u \in U$ it holds that

$$A_1(u) + A_2(u) < 2.$$

Whenever we use the above recalled theory of evaluative linguistic expressions with linguistic hedges of an inclusive type, the use of a single fuzzy relation – either \hat{R} or \check{R} – is not appropriate anymore, as we discussed in Section 1.2, see also [17]. In order to distinguish the situation from the, say, "standard" fuzzy rule base, which deal with fuzzy partitions and may freely use a single fuzzy relation as an appropriate model, the set of rules (1) will be called *linguistic description* and denoted by $LD = \{\mathcal{R}_1, \dots, \mathcal{R}_n\}^5$ The following conventions will be kept throughout the rest of this paper: We will denote by N_n the set $\{1, \ldots, n\}$ of natural numbers. If LD is given, then A_i and B_i , $i \in N_n$, will denote antecedent and consequent evaluative expressions from the *i*-th fuzzy IF-THEN rule \mathcal{R}_i , respectively. Further, A_i and B_i will denote interpretations (fuzzy sets) of A_i and B_i , respectively, $A_i \in \mathcal{F}(U)$ and $B_i \in \mathcal{F}(V)$, where U and V are closed real intervals.

A specificity ordering relation on the set of evaluative expressions is defined in order to allow us to state the relationships (inclusions) among evaluative expressions (or their models). First, let us define the ordering \leq_H on the set of hedges that can be defined on the hedges from Table 1 as follows:

$$\operatorname{Ex} \leq_H \operatorname{Si} \leq_H \operatorname{Ve} \leq_H \langle \operatorname{empty} \rangle \leq_H \operatorname{ML} \leq_H \operatorname{Ro} \leq_H \operatorname{QR}.$$

Based on \leq_H , we define the ordering \leq_{LE} of evaluative expressions. Let \mathcal{A}_1 , \mathcal{A}_2 be two evaluative expressions such that $\mathcal{A}_1 := \langle \text{hedge} \rangle_1 \mathcal{A}$ and $\mathcal{A}_2 := \langle \text{hedge} \rangle_2 \mathcal{A}$, where \mathcal{A} is an atomic expression. Then, we write

$$\mathcal{A}_1 \leq_{\mathrm{LE}} \mathcal{A}_2$$

if $\langle \text{hedge} \rangle_1 \leq_H \langle \text{hedge} \rangle_2$. In other words, evaluative expressions of the same type (with identical atomic expressions) are ordered according to their specificity, which is given by the hedges that appear in these expressions. Evaluative expressions with different atomic expressions cannot be ordered by \leq_{LE} .

Further, we adopt the extension of the theory of evaluative expressions by the following *inclusion axiom*.

Axiom 2 Let A_1, A_2 be two nonequal evaluative expressions ordered as $A_1 \leq_{\text{LE}} A_2$ and modeled by fuzzy sets A_1, A_2 , respectively. Then,

$$A_1 \subseteq A_2$$
 and $\operatorname{Ker}(A_1) \subset \operatorname{Ker}(A_2)$,

where Ker(A) denotes the kernel of a fuzzy set A.

⁴If necessary, it is possible to extend the range of atomic evaluative expressions using a finite number of additional expressions, e.g., *lower medium* or *upper medium*, such that Axioms 1 and 2 below hold.

 $^{^5}LD$ is viewed as a set, hence we omit multiple occurrences and each rule may be contained in LD only once.

Finally, we define the ordering of antecedent fuzzy sets (for a given linguistic description) with respect to an observation.

Definition 1 Let LD be a linguistic description and $u_0 \in U$. We write

$$A_i \leq_{u_0} A_i$$

either if
$$A_i(u_0) > A_j(u_0)$$
,
or if $A_i(u_0) = A_j(u_0)$ and $A_i \leq_{\text{LE}} A_j$.

It should be noted that antecedents of a linguistic description usually contain more than one variable. In this case, antecedent variables are compounded by conjunction and the ordering $\leq_{\rm LE}$ of compound evaluative expressions is extended in a straightforward way.

2.2. "Original" PbLD

A perception is understood to be a set of indices of fuzzy IF-THEN rules (for a given LD) that is assigned to an input. Intuitively, each index from a perception corresponds to the antecedent evaluative expression that fits best the given input.

Definition 2 Let LD be a linguistic description. The *local perception* function is a mapping P^{LD} : $U \to \mathcal{P}(N_n)$ that assigns to each value $u_0 \in U$ a set of indices

$$P^{LD}(u_0) = \{ i \in N_n \mid A_i(u_0) > 0 \& \& (\forall j \in N_n) ((A_j \le_{u_0} A_i) \Rightarrow (\mathcal{A}_j = \mathcal{A}_i)) \}.$$
(8)

The local perception function has a key role in the PbLD inference expressed as the $rule\ of\ PbLD$.

Definition 3 Let LD be a linguistic description. Let us be provided with an observation $u_0 \in U$. Then, the *rule of PbLD* (r_{PbLD}) is given as follows:

$$r_{PbLD}: \frac{P^{LD}(u_0), LD}{C},$$

where

$$C = \bigcap \{ C_j \mid j \in P^{LD}(u_0) \& \\ \& (\forall v \in V) (C_j(v) = A_j(u_0) \to B_j(v)) \},$$

where \cap is the Gödel intersection.

Informally, the inputs for this inference rule are the linguistic description LD and local perception $P^{LD}(u_0)$. This local perception is formed by a set of indices corresponding to antecedent fuzzy sets that are selected according to (8). Formula (8) selects indices of these antecedents that are the best fit for the given numerical input u_0 . Then, the individual conclusions C_j are computed as $A_j(u_0) \to B_j(v)$ for all $v \in V$, that is, for each $j \in P^{LD}(u_0)$, we take the j-th IF-THEN rule and compute the conclusion C_j , forgetting for the moment remaining IF-THEN rules from LD. Finally, we aggregate fuzzy sets C_j by intersection.

2.3. "Balancing" PbLD

As we explained in Section 1.2, the motivation for the introduction of PbLD was to design an inference method that comply with understanding of fuzzy IF-THEN rules as linguistically expressed logical implications. These fuzzy IF-THEN rules contain evaluative expressions in their antecedents and consequents. We argue that interpretations corresponding to these expressions must be inclusive, that is, if A_1 and A_2 are interpretations of extremely small and small, respectively, then A_1 should be a fuzzy subset of A_2 .

However, arguments in favor of use of the perception-based approach apply mainly to the case when antecedents of fired rules are in an inclusion, i.e., they have the same atomic expression. Nevertheless, consider the case when two rules have nonzero firing degrees and their antecedents have different atomic expressions. According to the original PbLD from the previous subsection, we should use only the rule which is fired to a higher degree. But, the information coming from the second rule is lost in this case. It seems reasonable to use both rules with their respective firing degrees and then aggregate the conclusions conjunctively (mimicking the standard implicative approach represented by \hat{R}).

Therefore, we come up with a new "balancing" PbLD that employs this idea and applies the perceptions only on the chains of antecedents with the same atomic expressions. The only necessary modification is the one of \leq_{u_0} (Definition 1).

Definition 4 Let LD be a linguistic description and let A_i, A_j be such that $A_i := \langle \text{hedge} \rangle_i A$ and $A_j := \langle \text{hedge} \rangle_j A$ where A is an atomic expression. Let $u_0 \in U$. We write

$$A_i \leq_{u_0} A_i$$

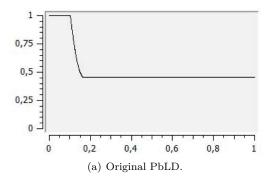
either if
$$A_i(u_0) > A_j(u_0)$$
,
or if $A_i(u_0) = A_j(u_0)$, and $A_i \leq_{\text{LE}} A_j$.

The definition of the local perception as well as of the rule r_{PbLD} remain unmodified. However, the behavior of the inference mechanism adopts the above discussed balancing character. Let us demonstrate the difference in behavior of the original and the balancing PbLD on an example. Consider the following linguistic description \mathcal{RB}_3 :

$$\mathcal{R}_1 := \mathsf{IF}\ X \text{ is } \mathit{Ex}\ \mathit{Sm}\ \mathsf{THEN}\ \mathit{Y} \text{ is } \mathit{Bi},$$
 $\mathcal{R}_2 := \mathsf{IF}\ \mathit{X} \text{ is } \mathit{Ro}\ \mathit{Sm}\ \mathsf{THEN}\ \mathit{Y} \text{ is } \mathit{Ro}\ \mathit{Sm},$ $\mathcal{R}_3 := \mathsf{IF}\ \mathit{X} \text{ is } \mathit{Ro}\ \mathit{Me}\ \mathsf{THEN}\ \mathit{Y} \text{ is } \mathit{Ro}\ \mathit{Me}.$

In Figs. 4 and 5, differences between the original and the balancing PbLD can be seen in cases when two rules whose antecedents have different atomic

 $^{^6{\}rm Both}$ rules are used only in the rare case when both firing degrees are equal.



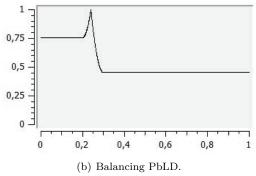


Figure 4: Results of the original and the balancing PbLD for linguistic description \mathcal{RB}_3 and $u_0 = 0.240$.

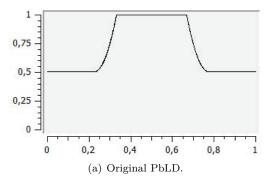
expressions (rules \mathcal{R}_2 and \mathcal{R}_3) have non-zero firing degrees. Results of the original PbLD (Figs. 4(a) and 5(a)) contain the conclusion of the rule with the highest firing degree only. For example, in Fig. 4(a), the rule \mathcal{R}_2 fires and the conclusion corresponds to its consequent roughly small. This conclusion is modified ("shifted up"), because the firing degree of the observation ($u_0 = 0.240$ in this case) is less than 1.

For the balancing PbLD (Figs. 4(b) and 5(b)), both rules \mathcal{R}_2 and \mathcal{R}_3 are fired and the inference uses information from both rules in the construction of its conclusion. For example, in Fig. 5(b), the firing degree of the observation $u_0 = 0.257$ in the rule \mathcal{R}_3 is higher than that in \mathcal{R}_2 . Hence, the result "has more from" the consequent of \mathcal{R}_3 (roughly medium) in it than from \mathcal{R}_2 (roughly small).

2.4. PbLD with Fuzzy Inputs

Now we propose the generalization of both PbLD variants to the case of fuzzy inputs in such a way that the inference with crisp inputs would be a special case of the one with the fuzzy inputs, as it holds for fuzzy relational inference.

Indeed, if we consider crisp inputs and the fuzzy relational inference, the inference mechanism itself is not important. Irrespective of using the CRI or the BK-subproduct, the inference reduces to a substitution of the input $u_0 \in U$ into the fuzzy relation R that models the fuzzy rule base, i.e. $(A_0@R)(v) = R(u_0, v)$ for $A_0 = \chi_{u_0}$, where χ denotes the characteristic function of a set and $\emptyset \in \{\circ, \lhd\}$.



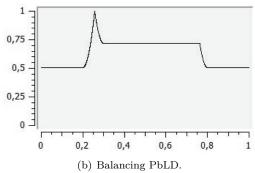


Figure 5: Results of the original and the balancing PbLD for linguistic description \mathcal{RB}_3 and $u_0 = 0.257$.

Let the orderings \leq_{H} and \leq_{LE} be defined as above. Let the *generalized firing degree* of an observation $A_0 \in \mathcal{F}(U)$ with respect to a fuzzy set A_i for a given LD be defined as follows

$$A_i(A_0) = A_0 @ A_i, \quad @ \in \{ \circ, \lhd \}$$
 (9)

that is, depending on the image:

$$A_i(A_0) = A_0 \circ A_i = \bigvee_{u \in U} (A_0(u) * A_i(u)), \text{ or }$$

$$A_i(A_0) = A_0 \lhd A_i = \bigwedge_{u \in U} (A_0(u) \to A_i(u)).$$

Then, for the original PbLD, we may introduce the following ordering of antecedent fuzzy sets with respect to a fuzzy observation.

Definition 5 Let LD be a linguistic description. Let $A_0 \in \mathcal{F}(U)$. We write

$$A_i \leq_{A_0} A_i$$

either if
$$A_i(A_0) > A_j(A_0)$$
,
or if $A_i(A_0) = A_j(A_0)$ and $A_i \leq_{\text{LE}} A_j$.

Analogously, the ordering \leq_{A_0} may be introduced for the balancing PbLD.

Definition 6 Let LD be a linguistic description and let A_i, A_j be such that $A_i := \langle \text{hedge} \rangle_i A$ and $A_j := \langle \text{hedge} \rangle_j A$, where A is an atomic expression. Let $A_0 \in \mathcal{F}(U)$. We write

$$A_i \leq_{A_0} A_i$$

either if
$$A_i(A_0) > A_j(A_0)$$
,
or if $A_i(A_0) = A_j(A_0)$ and $A_i \leq_{\text{LE}} A_j$.

The local perception will then be given as follows.

Definition 7 Let LD be a linguistic description. The local perception is a mapping $P^{LD}: \mathcal{F}(U) \to \mathcal{P}(N_n)$ that assigns to each fuzzy set $A_0 \in \mathcal{F}(U)$ the set

$$P^{LD}(A_0) = \{ i \in N_n \mid A_i(A_0) > 0 \& \\ \& (\forall j \in N_n) ((A_j \leq_{A_0} A_i) \Rightarrow (\mathcal{A}_j = \mathcal{A}_i)) \}.$$
 (10)

Definition 8 Let LD be a linguistic description. Let us be provided with an observation $A_0 \in \mathcal{F}(U)$. Then, the *rule of PbLD* (r_{PbLD}) is given as follows:

$$r_{PbLD}: \frac{P^{LD}(A_0), LD}{C}, \tag{11}$$

where

$$C = A_0 @ \hat{R}_{A_0}$$
 (12)

and, for all $u \in V$ and all $v \in V$,

$$\hat{R}_{A_0}(u,v) = \bigwedge_{i \in P^{LD}(A_0)} (A_i(u) \to B_i(v)).$$

Note that we suppose the use of the same image operation $@ \in \{\circ, \lhd\}$ in (12) as in (9). One can easily check that, in case of a crisp (singleton) input, the above definitions are equivalent to Definitions 1–4. Indeed, let us be given an input $A_0 = \chi_{u_0}$ for some $u_0 \in U$. Then $A_i(A_0) = A_0@A_i = A_i(u_0)$ for $@ \in \{\circ, \lhd\}$. Thus, \leq_{u_0} and \leq_{A_0} coincide (for both PbLD variants). Furthermore, in the case of $@ = \circ$:

$$C(v) = \bigvee_{u \in U} \left(A_0(u) * \bigwedge_{i \in P^{LD}(A_0)} (A_i(u) \to B_i(v)) \right)$$
$$= \bigwedge_{i \in P^{LD}(A_0)} (A_i(u_0) \to B_i(v))$$

and, similarly, for the case of $@= \lhd$:

$$C(v) = \bigwedge_{u \in U} \left(A_0(u) \to \bigwedge_{i \in P^{LD}(A_0)} (A_i(u) \to B_i(v)) \right)$$
$$= \bigwedge_{i \in P^{LD}(A_0)} (A_i(u_0) \to B_i(v)).$$

Hence, one can see that irrespective of the chosen image operator @, the inferred output coincides with the one obtained from the crisp input u_0 when we use the r_{PbLD} from Definition 3.

3. Fuzzy Interpolation

Each inference mechanism should possess a fundamental property — preservation of modus ponens. If an input fuzzy set $A_0 \in \mathcal{F}(U)$ is equal to one of the antecedents, say to the A_i , then the inferred

output $B_0 \in \mathcal{F}(V)$ should be equal to the respective consequent B_i . This requirement leads to the following system of fuzzy relational equations

$$A_i@R = B_i, \qquad i = 1, \dots, n$$
 (13)

where $R \in \mathcal{F}(U \times V)$ is unknown. A fuzzy relation R which satisfies (13) is called a *solution* of the system and we say that R interpolates (A_i, B_i) . In such case, R can be seen as a correct model of the given fuzzy rule base in the given fuzzy inference system.

Obviously, not all systems (13) are solvable, i.e., not for all sets of pairs $(A_i, B_i)_{i=1}^n$ there exists a fuzzy relation that interpolates them. The question of solvability of such systems was addressed by many researchers. In this section, we recall only the most fundamental results [22, 23, 24].

Theorem 1 System (13) with $@=\circ (@=\lhd)$ is solvable if and only if \hat{R} (\check{R}) is a solution of this system. In case of solvability, \hat{R} (\check{R}) is the greatest (the least) solution of (13) with $@=\circ (@=\lhd)$.

Theorem 1 actually states that the implicative model \hat{R} (Mamdani-Assilian model \check{R}) should be the first choice whenever dealing with the inference modelled by \circ (\lhd). If there exist some reasons (e.g., preservation of robustness, low computational complexity [9] or the existence of hierarchical inference that is identical to the non-hierarchical one [9, 25]), why the combination of Mamdani-Assilian model \hat{R} and the CRI inference \circ (or the combinations of the implicative model \hat{R} and the BK-subproduct inference \lhd) should be preferred, one should first of all check whether the interpolativity is preserved also in this case. The answer to this question is provided by the following theorems that collect results from [23, 26].

Theorem 2 Let all A_i , $i \in N_n$, be normal. Then \check{R} (\hat{R}) is a solution of (13) with $@=\circ (@=\lhd)$ if and only if the condition

$$\bigvee_{u \in U} (A_i(u) * A_j(u)) \le \bigwedge_{v \in V} (B_i(v) \leftrightarrow B_j(v)) \quad (14)$$

holds for any $i, j \in N_n$.

Due to the monotonicity of images, namely:

$$A@R_1 \subseteq A@R_2, @\in \{\circ, \lhd\}$$

for any $R_1, R_2 \in \mathcal{F}(U \times V)$ such that $R_1 \subseteq R_2$, we can state the following corollary.

Corollary 1 Let $R_1, R_2 \in \mathcal{F}(U \times V)$ be two solutions of system (13) with $@=\circ (@=\lhd)$. Then any $R \in \mathcal{F}(U \times V)$, such that $R_1 \subseteq R \subseteq R_2$, is a solution of this system, too.

3.1. Interpolativity in Case of the Original Perception-based Logical Deduction

The interpolativity of the PbLD turns into an investigation of the following problem:

Given LD and the antecedent $A_i \in \mathcal{F}(U)$ of the i-th rule, i = 1, ..., n, it should hold that after the application of r_{PbLD} :

$$r_{PbLD}: \frac{P^{LD}(A_i), LD}{C},$$

the conclusion C is equal to the consequent $B_i \in \mathcal{F}(V)$ of this *i*-th rule.⁷

In the rest of the paper, we suppose that all linguistic descriptions LD fulfill the following condition:

$$(\forall i \in N_n)(\forall j \in N_n)((i \neq j) \Rightarrow (\mathcal{A}_i \neq \mathcal{A}_j)) \quad (15)$$

Further, let LD be given. Let A_i , $i \in N_n$, be one of the antecedent fuzzy sets of LD. Then we will denote by C_i the result of r_{PbLD} (11) for $A_0 = A_i$.

Lemma 1 Let LD be a linguistic description, let $i \in N_n$. If $@= \circ$ then $P^{LD}(A_i) = \{i\}$ if and only if for all $j \in N_n$ such that $j \neq i$,

- a) either A_i and A_j have different atomic evaluative expressions (e.g., very small and more or less medium);
- b) or $A_i \leq_{\text{LE}} A_j$.

Sketch of the proof: According to (10),

$$P^{LD}(A_i) = \{ j \in N_n \mid A_j(A_i) > 0 \& (\forall k \in N_n) \\ ((A_k \leq_{A_i} A_j) \Rightarrow (A_k = A_j)) \}$$

and $A_i(A_i) = 1$ for $@ = \circ$.

For any $j \in N_n$ such that a) holds, due to the partition axiom, we get $A_j(A_i) < 1 = A_i(A_i)$ from which we get $j \notin P^{LD}(A_i)$.

For any $j \in N_n$ such that b) holds, $A_j(A_i) = 1$ but $A_j \not\leq_{A_i} A_i$ from which we get $j \notin P^{LD}(A_i)$.

If a) or b) holds for any $j \neq i$ then $P^{LD}(A_i) = \{i\}$. The opposite implication is obvious.

Corollary 2 If $@ = \circ$ and a) or b) from Lemma 1 holds for all $i \in N_n$, then the interpolativity is preserved.

Although Corollary 2 states a positive result in a sense of sufficient conditions for the preservation of the interpolativity, one has to note the price we pay for it, particularly, how much restrictive the assumptions are, which may be seen also from the following Lemma and mainly from its Corollary. **Lemma 2** If $@= \circ$ and there are $i, j \in N_n$, such that $i \neq j$ and $\mathcal{A}_j \leq_{\mathrm{LE}} \mathcal{A}_i$ (cf. conditions a) and b) of Lemma 1) and, at the same time, the corresponding \mathcal{B}_i and \mathcal{B}_j are not equal, then the interpolativity is NOT preserved.

Sketch of the proof: Let us fix some arbitrary i. The proof splits into two possibilities. First, assume that A_j is the least antecedent fuzzy set with respect to \leq_{A_i} . Then $P^{LD}(A_i) = \{j\}$ and

$$C_i(v) = \bigvee_{u \in U} A_i(u) * (A_j(u) \to B_j(v))$$

and we need to check whether $C_i = B_i$, i.e., that $C_i(v) \leq B_i(v)$ and $C_i(v) \geq B_i(v)$ holds for all $v \in V$. The first inequality holds if and only if

$$\bigvee_{u \in U} (A_i(u) * (A_j(u) \to B_j(v))) \le B_i(v),$$

which by adjunction turns into

$$A_i(u) \to B_i(v) \le A_i(u) \to B_i(v), \quad \forall u \in U.$$
 (16)

Inequality (16) has to be preserved for all u and thus, it has to be preserved also for such $u \in U$ for which $A_i(u) = 1$ and $A_j(u) < 1$. However, for such u, the following inequality would have to be preserved

$$A_i(u) \to B_i(v) \le B_i(v)$$

which is not possible in the underlying Łukasiewicz algebra where the inequality turns into

$$1 - A_j(u) + B_j(v) \le B_i(v)$$

and the left-hand side of the equation is strictly greater than 0 for arbitrary $v \in V$ which is not the case of the right-hand side.

The second possibility is that there exists $k \in N_n$ such that A_k it is the least antecedent with respect to \leq_{A_i} and thus, $A_k \leq_{A_i} A_j$. Then either $B_k \neq B_i$ and the proof goes as above for the first case, or $B_k = B_i \neq B_j$. Then, the interpolativity would be preserved for i, however, for j, we would obtain $P^{LD}(A_j) = \{k\}$ and, using the same technique, we would get $C_j(v) \not\leq B_j(v)$ for some $v \in V$ and thus, $C_j \neq B_j$.

Corollary 3 Let $@=\circ$. Then the interpolativity is preserved if and only if for all $i, j \in N_n$ such that $A_j \leq_{\text{LE}} A_i$, the corresponding consequents $\mathcal{B}_i, \mathcal{B}_j$ are equal.

One may see that \circ is not very appropriate here. The reason is that the local perception chooses narrower antecedents, not the most similar ones. For example, if we had antecedents $very \ small$ and small in our LD, and the observation was the wider one (small in our case), the local perception would fire only the rule with the antecedent $very \ small$ which is very unintuitive and makes the interpolativity problematic. Now, let us focus on $@= \lhd$.

⁷Note the difference compared to (13), where a single fuzzy relation has to be a solution of the whole system, while here, not all rules are fired and each equation is solved separately.

Lemma 3 Let $i \in N_n$. If $@= \lhd$, then $P^{LD}(A_i) = \{i\}$ for all $i \in N_n$.

Sketch of the proof: Clearly $A_i(A_i) = 1$. Consider $j \in N_n$ such that $A_i = \langle \text{hedge} \rangle_i \mathcal{A}'$, $A_j = \langle \text{hedge} \rangle_j \mathcal{A}''$, where \mathcal{A}' and \mathcal{A}'' are atomic evaluative expressions. Then, there exist two possibilities:

First, let either $\mathcal{A}' \neq \mathcal{A}''$ or let $\mathcal{A}' = \mathcal{A}''$ and $\mathcal{A}_i \leq_{\text{LE}} \mathcal{A}_i$. Then

$$A_j(A_i) = \bigwedge_{u \in U} (A_i(u) \to A_j(u)) < 1.$$

Second, let $\mathcal{A}' = \mathcal{A}''$ and $\mathcal{A}_i \leq_{\mathrm{LE}} \mathcal{A}_i$. Then

$$A_j(A_i) = \bigwedge_{u \in U} (A_i(u) \to A_j(u)) = 1$$

but, due to the order $A_i \leq_{\text{LE}} A_j$, $j \notin P^{LD}(A_i)$. Therefore, $P^{LD}(A_i) = \{i\}$ in both cases.

Corollary 4 If $@= \lhd$ then the interpolativity is preserved.

Sketch of the proof: Due to Lemma 3 we have $P^{LD}(A_i) = \{i\}$ for any $i \in N_n$. Then the proof continues as follows:

$$C_i(v) = \bigwedge_{u \in U} (A_i(u) \to (A_i(u) \to B_i(v)))$$

= 1 \to B_i(v) = B_i(v).

3.2. Interpolativity in Case of the Balancing Perception-based Logical Deduction

In Section 3.1, we have shown that the case of $@= \circ$ does lead to restrictive assumptions, if we want to keep the interpolation property. Moreover, we have also shown somewhat unintuitive behavior of the inference mechanism in this setting. Of course, this undesirable behavior is preserved also for the balancing variant of the PbLD. Therefore, in this section, we will focus only on the case of $@= \lhd$.

Lemma 4 If @= < then the interpolativity can be guaranteed if and only if

$$\bigvee_{u \in U} (A_j(u) * A_i(u)) \le \bigwedge_{v \in V} (B_i(v) \to B_j(v)) \quad (17)$$

for all $i \in N_n$ and for all $j \in P^{LD}(A_i)$.

Sketch of the proof: Let us fix any $i \in N_n$. The corresponding fuzzy relational equation:

$$A_i \lhd R = B_i \tag{18}$$

is always solvable as (14) in the case of a single equation turns into

$$\bigvee_{u \in U} (A_i(u) * A_i(u)) \le \bigwedge_{v \in V} (B_i(v) \leftrightarrow B_i(v))$$

which always holds. Thus, according to Theorem 2, $\hat{R}_i(u,v) = A_i(u) \to B_i(v)$ is a solution of the equation (18) and, according to Theorem 1, the least solution of this equation is $\check{R}_i(u,v) = A_i(u) * B_i(v)$.

In order to prove that \hat{R}_{A_i} defined as

$$\hat{R}_{A_i}(u,v) = \bigwedge_{j \in P^{LD}(A_i)} (A_j(u) \to B_j(v))$$

is a solution of the equation (18), it is sufficient to prove that it lies between \check{R}_i and \hat{R}_i (c.f. Corollary 1).

Since $i \in P^{LD}(A_i)$, then obviously $\hat{R}_{A_i} \subseteq \hat{R}_i$ holds. As \check{R}_i is the least solution, \hat{R}_{A_i} will become a solution if and only if the second inclusion $\check{R}_i \subseteq \hat{R}_{A_i}$ is preserved. The inclusion can be expanded as follows

$$A_i(u) * B_i(v) \le \bigwedge_{j \in P^{LD}(A_i)} (A_j(u) \to B_j(v)) \quad \forall u \forall v$$

and will be preserved if and only if for all $j \in P^{LD}(A_i)$ the following inequality will be preserved

$$A_i(u) * B_i(v) \le A_i(u) \to B_i(v) \quad \forall u \forall v$$

which, using adjunction, associativity of * and again adjunction, turns to be equivalent to the preservation of inequality (17).

4. Concluding Remarks

The results should be seen as follows. The combination of the implicative rules and the BK-subproduct, which may be very promising from practical aspects [9] but is less preferable form the interpolativity point of view, may be successfully employed, if one uses a specific inference mechanism (PbLD) and specific model of meaning of evaluative expressions (the theory of evaluative linguistic expressions). The use of this model of evaluative expressions would be impossible in the standard fuzzy relational approach due to the full inclusion of the corresponding fuzzy sets for evaluative expressions with the same atomic expression.

The balancing PbLD represents a sort of compromise between the original PbLD and the standard fuzzy relational model \hat{R} , where the perception is applied only on the subsets of antecedents (rules) with the same type of expression. Then, all the fired rules are used together as in the traditional way when using \hat{R} . Thus, the interpolativity is not that easily guaranteed as in the case of the original PbLD. However, it still may be preserved even in the case of BK-subproduct, which is not very often possible when using \hat{R} . Note that the inequality (17) is less restrictive than the well-known inequality (14) valid for standard fuzzy relational inference schemes. The reason is that less rules are fired and moreover, mainly, for each input, we "build" a different fuzzy relation and thus, we do not need to build a single one that would comply with all inputs/antecedents.

Thus, one does not get a single "winner" among the models of implicative fuzzy rules and inference mechanisms. However, a wider choice is given at disposal with theoretical results providing users with information saying when a possible setting will (or will not) work correctly.

Acknowledgements

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project (CZ.1.05/1.1.00/02.0070).

References

- L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I-III. *Inform. Sciences* 8 (1975) 199–250, 8 (1975) 301–357, 9 (1975) 43–80, 1975.
- [2] D. Dubois and H. Prade. What are fuzzy rules and how to use them. Fuzzy Sets Syst., 84:169– 185, 1996.
- [3] E. H. Mamdani and S. Assilian. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man Mach. Stud.*, 7:1–13, 1975.
- [4] J. M. Alonso, C. Castiello, and C. Mencar. Interpretability of fuzzy systems: Current research trends and prospects. In W. Pedrycz J. Kacprzyk, editor, *Handbook of Computational Intelligence*, chapter 14, pages 181–199. Springer, Berlin, 2015.
- [5] H. Jones, B. Charnomordic, D. Dubois, and S. Guillaume. Practical inference with systems of gradual implicative rules. *IEEE T. Fuzzy* Syst., 17(1):61–78, 2009.
- [6] M. Daňková. On approximate reasoning with graded rules. Fuzzy Sets Syst., 158:652–673, 2007.
- [7] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE T. Syst. Man Cyb.*, 3(1):28–44, 1973.
- [8] W. Pedrycz. Applications of fuzzy relational equations for methods of reasoning in presence of fuzzy data. *Fuzzy Sets Syst.*, 16:163–175, 1985.
- [9] M. Štěpnička and B. Jayaram. On the suitability of the Bandler-Kohout subproduct as an inference mechanism. *IEEE T. Fuzzy Syst.*, 18(2):285–298, 2010.
- [10] M. Štěpnička, U. Bodenhofer, M. Daňková, and V. Novák. Continuity issues of the implicational interpretation of fuzzy rules. Fuzzy Sets Syst., 161:1959–1972, 2010.
- [11] D. Dubois, H. Prade, and L. Ughetto. Checking the coherence and redundancy of fuzzy knowl-

- edge bases. *IEEE T. Fuzzy Syst.*, 5(6):398–417, August 1997.
- [12] U. Bodenhofer, M. Daňková, M. Štěpnička, and V. Novák. A plea for the usefulness of the deductive interpretation of fuzzy rules in engineering applications. In *Proc. 16th IEEE* Int. Conf. on Fuzzy Systems, pages 1567–1572, London, July 2007.
- [13] V. Novák. Perception-based logical deduction. In B. Reusch, editor, Computational Intelligence, Theory and Applications, pages 237–250. Springer, Berlin, 2005.
- [14] V. Novák and I. Perfilieva. On the semantics of perception-based fuzzy logic deduction. *Int.* J. Intell. Syst., 19:1007–1031, 2004.
- [15] A. Dvořák, M. Štěpnička, and L. Štěpničková. On redundancies in systems of fuzzy/linguistic IF-THEN rules under perception-based logical deduction inference. Fuzzy Sets Syst., in press.
- [16] I. Perfilieva and V. Novák. System of fuzzy relation equations as a continuous model of ifthen rules. *Inform. Sciences*, 177:3218–3227, 2007.
- [17] A. Dvořák. On preselection of rules in fuzzy logic deduction. *Int. J. Uncertain. Fuzz.*, 8(5):563–572, 2000.
- [18] V. Novák. A comprehensive theory of trichotomous evaluative linguistic expressions. Fuzzy Sets Syst., 159(22):2939–2969, 2008.
- [19] M. Štěpnička, A. Dvořák, V. Pavliska, and L. Vavříčková. A linguistic approach to time series modeling with the help of F-transform. Fuzzy Sets Syst., 180:164–184, 2011.
- [20] A. Dvořák, H. Habiballa, V. Novák, and V. Pavliska. The software package LFLC 2000 - its specificity, recent and perspective applications. Comput. Ind., 51:269–280, 2003.
- [21] M. De Cock and E.E. Kerre. Fuzzy modifiers based on fuzzy relations. *Inform. Sciences*, 160:173–199, 2004.
- [22] B. De Baets. Analytical solution methods for fuzzy relational equations. In D. Dubois and H. Prade, editors, *The Handbook of Fuzzy Set Series Vol. 1*, pages 291–340. Academic Kluwer Publ., Boston, 2000.
- [23] F. Klawonn. Fuzzy points, fuzzy relations and fuzzy functions. In V. Novák and I. Perfilieva, editors, *Discovering the World with Fuzzy Logic*, pages 431–453. Springer, Berlin, 2000.
- [24] I. Perfilieva and L. Nosková. System of fuzzy relation equations with inf-→ composition: Complete set of solutions. Fuzzy Sets Syst., 159:2256–2271, 2008.
- [25] B. Jayaram. On the law of importation $(x \land y) \rightarrow z \equiv (x \rightarrow (y \rightarrow z))$ in fuzzy logic. *IEEE T. Fuzzy Syst.*, 16:130–144, 2008.
- [26] L. Nosková. Systems of fuzzy relation equation with inf- \rightarrow composition: solvability and solutions. J. Electr. Eng., 12(s):69–72, 2005.