

Redundancies in systems of fuzzy/linguistic IF-THEN rules

Antonín Dvořák Martin Štěpnička Lenka Vavříčková

Institute for Research and Applications of Fuzzy Modeling, University of Ostrava
Ostrava, Czech Republic

Email:{antonin.dvorak,martin.stepnicka,lenka.vavrickova}@osu.cz

Abstract

In this paper we will investigate which fuzzy/linguistic rules are redundant in systems of such rules called linguistic descriptions. We present a formal definition of redundancy and show that rules which are seemingly redundant can be in fact indispensable. These results apply for IF-THEN rules which use evaluative linguistic expressions (e.g., *small*, *very big*, etc.) and inference method called *perception-based logical deduction* (PbLD). However, they are also valid for inference systems which use compatible design choices with PbLD.

Keywords: Redundancy, Fuzzy rules, Linguistic description, Perception-based logical deduction

1. Introduction and motivation

In this contribution, we outline our approach to detection of so-called redundancies in systems of fuzzy/linguistic IF-THEN rules (we call such systems *linguistic descriptions*).

Notions of redundancy and inconsistency of fuzzy IF-THEN rules were studied by several authors [1, 2, 3]. But, for our so-called linguistic approach originated by V. Novák [4] these questions were not adequately investigated. We were able to successfully apply our linguistic approach in analysis and forecasting of time series [5, 6], decision making [7] or data mining [8]. There exists also a large scale application of fuzzy control system based on fuzzy/linguistic IF-THEN rules. It is a control of five large aluminium furnaces which is in operation for more than fifteen years without any problems [9]. We found out that this approach has distinct advantages in good interpretability, great robustness, foundations in strong formal logical system etc. However, if we learn our rules from real-world data, as is the case in time series analysis or data mining, then interpretability can be seriously harmed by the presence of redundant rules. Users of these rules can be then buried under a high number of seemingly very similar rules and consequently would not be able to comprehend and modify them. Therefore, theoretically well-founded and efficient algorithm for detection of redundant rules is a necessity.¹

¹Inconsistencies in systems of IF-THEN rules are also un-

The inference mechanism tailored for our fuzzy/linguistic IF-THEN rules is called *perception-based logical deduction* (PbLD). Quite informally, it, for a given input x_0 , picks and fires that rule from the linguistic description, whose antecedent (IF part) has maximal membership degree at x_0 ² If there are several such rules, then it picks that whose antecedent is most specific (for example, *very small* is more specific than *small*, etc.).

A fuzzy IF-THEN rule \mathcal{R}_1 is usually thought to be redundant with respect to rule \mathcal{R}_2 if their consequents are identical and their antecedents are different but not contradictory. For example, let \mathcal{R}_1 and \mathcal{R}_2 be

$\mathcal{R}_1 :=$ IF X is *very small* THEN Y is *big*,

$\mathcal{R}_2 :=$ IF X is *roughly small* THEN Y is *big*.

If the interpretation of linguistic expressions *roughly small* and *very small* is such that every element which is very small is also roughly small in at least the same degree, then rule \mathcal{R}_1 is redundant — nothing changes if it is left out. Initially it seemed to us that elimination of redundancies would consist simply in detection of pairs of rules such as \mathcal{R}_1 and \mathcal{R}_2 and deletion of \mathcal{R}_1 . However, it turned out that it is not that simple. We accepted a natural definition that, informally speaking, a rule is redundant if results of inference mechanism are exactly the same if we leave this rule out. Then, the presence of other rules (with different consequents) can cause that a rule such as \mathcal{R}_1 is all of a sudden not redundant at all. Imagine that third rule is present, namely

$\mathcal{R}_3 :=$ IF X is *small* THEN Y is *small*

If the interpretation of *small* is such that it lies “in-between” *roughly small* and *very small* (see Figure 1), then we cannot leave out rule \mathcal{R}_1 , because for values of X for which it is fired (result of inference should be “big”) suddenly, after \mathcal{R}_1 is left out, \mathcal{R}_3 would be fired (and result of inference would be

pleasant for the interpretability, but they can harm the performance of inference mechanism, too. Their detection and elimination will be a topic of our further research.

²More precisely, the antecedent is a linguistic expression, e.g. *small*. This expression is, after some formal steps, interpreted by some fuzzy set, and the membership degree in this fuzzy set is referred to in the informal expression “antecedent has maximal membership degree at x_0 ”.

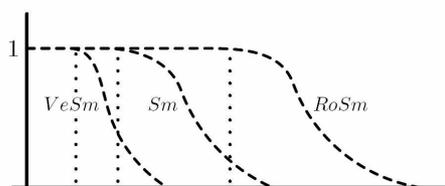


Figure 1: Graphical presentation of fuzzy sets (extensions) that interpret linguistic expressions *very small*, *small* and *roughly small*.

“small”). We call \mathcal{R}_1 to be suspicious to be redundant with respect to \mathcal{R}_2 and search mathematical results stating conditions under which it is indeed redundant and under which it is not.

Let us remark that our results apply also to systems of IF-THEN rules and inference mechanisms which use compatible design choices (overlapping interpretations of linguistic expressions like *very small* and *small*, inference mechanisms firing rules based on best fit of inputs and highest specificity, etc.) Concrete shapes of fuzzy sets interpreting evaluative expressions and details of inference mechanisms are not important from this point of view.

2. Theoretical background

In this section we provide notions necessary for the understanding of our results on the detection of redundancies. We need to cover fuzzy/linguistic rules discussed in Section 1 and inference over sets of such rules (our inference method is called *perception-based logical deduction* – PbLD). But first we should look at one of main constituents of IF-THEN rules, namely so-called evaluative linguistic expressions. Then we need to establish an ordering of evaluative expressions which appear in IF-THEN rules according to their specificity. We concentrate on main ideas, details can be found e.g. in [6].

2.1. Evaluative linguistic expressions

One of main constituents of systems of fuzzy/linguistic IF-THEN rules are *evaluative linguistic expressions* [10], in short *evaluative expressions*. They are special expressions of natural language that are used whenever it is important to evaluate a decision situation, to specify the course of development of some process, and in many other specific situations. Typical examples of evaluative linguistic expressions are the expressions *very large*, *extremely expensive*, *more or less hot*, etc. Note that their importance and the potential to model their meaning mathematically have been pointed out by L. A. Zadeh (e.g., in [11, 12, 13]).

A simple form of evaluative expressions keeps the

following structure:

$$\langle \text{linguistic hedge} \rangle \langle \text{atomic evaluative expression} \rangle \quad (1)$$

Atomic evaluative expressions comprise any of the *canonical* adjectives *small*, *medium*, *big*³, abbreviated in the following as *Sm*, *Me*, *Bi*, respectively. It is important to stress that these words are in practice often replaced by other kinds of evaluative words, such as “thin”, “old”, “new”, etc., depending on the context of speech.

Linguistic hedges are specific adverbs that make the meaning of the atomic expression more or less precise. We may distinguish hedges with *narrowing effect*, e.g. *very*, *extremely*, etc. and with *widening effect*, e.g. *roughly*, *more or less*, etc. In general, there can be a finite number of hedges with narrowing effect and a finite number of hedges with widening effect. In practical applications, their number is of course limited to, e.g., five with narrowing effect and five with widening effect. In the following text, we, without loss of generality, use two hedges with narrowing effect (*extremely*, *very*) and two hedges with widening effect (*more or less*, *roughly*). As a special case, the $\langle \text{linguistic hedge} \rangle$ can be empty. This enables us to include atomic evaluative expressions into the class of simple ones and develop a unified theory of their meaning. Note that our hedges are of so-called inclusive type [14], which means that extensions of more specific evaluative expressions are included in less specific ones, see Figure 1. Other (non-inclusive) models are possible as well, but in our opinion they do not capture well the intuitive understanding of hedges.⁴ Evaluative expressions of the form (1) will generally be denoted by script letters \mathcal{A} , \mathcal{B} , etc.

Evaluative expressions are used to evaluate values of some variable X . The resulting expressions are called *evaluative linguistic predications*, and have the form

$$X \text{ is } \mathcal{A}.$$

Examples of evaluative predications are “temperature is very high”, “price is low”, etc.

The model of the meaning of evaluative expressions and predications makes distinction between *intensions* and *extensions* in various *contexts*. The context characterizes a range of possible values. This range can be characterized by a triple of numbers $\langle v_L, v_M, v_R \rangle$, where $v_L, v_M, v_R \in \mathbb{R}$ and $v_L < v_M < v_R$ ⁵. These numbers characterize the minimal, middle, and maximal values, respectively, of the evaluated characteristics (such as “height”) in the specified context of use. Therefore, we

³In many situations, it is advantageous to extend the set of atomic evaluative expressions by the evaluative expression *zero*.

⁴For recent studies on hedges from the point of view of linguistic and pragmatics, see [15].

⁵Let us emphasize that the middle value v_M is not required to be in the exact center of the interval $[v_L, v_R]$.

It should be noted that usually the $Topic_{LD}$ contains intensions of evaluative predications which are compound by a conjunction of more than one evaluative predication. In other words, we usually meet the following situation

$$(X \text{ is } \mathcal{A}_i) = (X_1 \text{ is } \mathcal{A}_{i_1}) \text{ AND } \cdots \text{ AND } (X_K \text{ is } \mathcal{A}_{i_K}),$$

$$(X \text{ is } \mathcal{A}_j) = (X_1 \text{ is } \mathcal{A}_{j_1}) \text{ AND } \cdots \text{ AND } (X_K \text{ is } \mathcal{A}_{j_K}).$$

In this case, the ordering \leq_{LE} is preserved with respect to the components:

$$\mathcal{A}_i \leq_{LE} \mathcal{A}_j \quad \text{if} \quad \mathcal{A}_{i_k} \leq_{LE} \mathcal{A}_{j_k} \quad \text{for all } k = 1, \dots, K \quad (8)$$

and the extension of the compound linguistic predication is given as follows

$$\begin{aligned} & (\text{Int}(X \text{ is } \mathcal{A}_i)(w_1, \dots, w_K))(u_1, \dots, u_K) = \\ & = \bigwedge_{k=1}^K (\text{Int}(X_k \text{ is } \mathcal{A}_i)(w_k))(u_k). \end{aligned} \quad (9)$$

Then, the final ordering $\leq_{(u_0, w)}$ is analogous to the one-dimensional one, first we determine a_i and a_j according to (6-7) with having in mind the extension of the compound predication (9) and in case of equal values, we determine the ordering \leq_{LE} of the expressions appearing in the components of the compound one as given by (8).

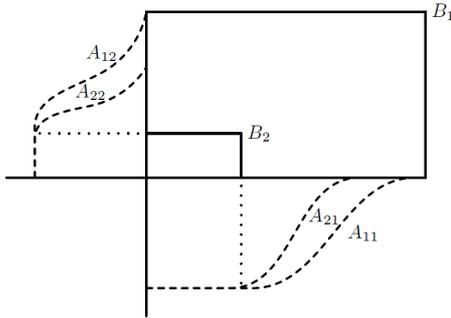


Figure 2: Visualization of fuzzy rules $\mathcal{R}_1, \mathcal{R}_2$. Displayed rectangles symbolically denote areas covered by antecedents of given fuzzy rules with their respective consequents $\mathcal{B}_1, \mathcal{B}_2$.

On Figure 2, we provide readers with a visualization of two fuzzy rules with two input variables. This double-input example is fully sufficient for intuitive understanding and thus, will be used for an illustration of distinct situations in the rest of the paper. Note that the rectangles denote areas covered by antecedents of given fuzzy rules. Each rectangle is also denoted by a respective consequent (B_1 or B_2). Thus, for the sake of brevity and simplification of the notation, we will use only the rectangles to display the areas covered by antecedents jointly with the labels denoting the respective consequents, as on Figure 3. We find such a form of illustration of fuzzy rules to be easy-to-read and fully informative.

2.4. Perception-based logical deduction

This is a special inference method aimed at the derivation of results based on fuzzy/linguistic IF-THEN rules. A perception is understood as an evaluative expression assigned to the given input value in the given context. The choice of perception is not arbitrary. It depends on the topic of the specified linguistic description, too. In other words, perception is chosen among evaluative expressions which occur in antecedents of IF-THEN rules. For details, see [4, 6].

Based on the ordering of linguistic predications we define a special function of *local perception*

$$LPerc^{LD} : w \times W^K \longrightarrow \mathcal{P}(Topic_{LD})$$

assigning to each value $u_0 = [u_1, \dots, u_K] \in w$ for $w = [w_1, \dots, w_K] \in W^K$ a subset of intensions

$$LPerc^{LD}(u_0, w) = \{ \text{Int}(X \text{ is } \mathcal{A}_i) \mid \forall \text{Int}(X \text{ is } \mathcal{A}_j) \in Topic_{LD} : ((X \text{ is } \mathcal{A}_j) \leq_{(u_0, w)} (X \text{ is } \mathcal{A}_i)) \Rightarrow ((X \text{ is } \mathcal{A}_j) = (X \text{ is } \mathcal{A}_i)) \} \quad (10)$$

Let LD be a linguistic description (4). Let us consider a context $w \in W$ for the variable X and a context $w' \in W$ for Y . Let an observation $X = u_0$ in the context w be given, where $u_0 \in w$. Then, the following *rule of perception-based logical deduction* (r_{PbLD}) is valid:

$$r_{PbLD} : \frac{LPerc^{LD}(u_0, w), LD}{C} \quad (11)$$

where C is the conclusion which corresponds to the observation in a way described below. For purposes of this paper, it will be sufficient to explain this deduction rule informally. Details can be found e.g. in [6]. Inputs to this inference rule are linguistic description LD and local perception $LPerc^{LD}(u_0, w)$ from (10). This local perception is formed by a set of evaluative expressions from antecedents of IF-THEN rules (i.e., from the topic) of the given linguistic description. Formula (10) chooses these antecedents, which best fit the given numerical input u_0 and are most specific according to the ordering $\leq_{(u_0, w)}$.⁸

Once one or more antecedents $\text{Int}(X \text{ is } \mathcal{A}_i) \in Topic_{LD}$, $i = 1, \dots, L$ are chosen according to (10), we compute for any of them conclusions C_{i_i} by a procedure which for our purposes can be imagined as a standard fuzzy relational inference mechanism by means of an appropriate fuzzy relational composition.⁹ It means that for $\text{Int}(X \text{ is } \mathcal{A}_i)$ we take i -th IF-THEN rule from LD and compute the conclusion, for the time being forgetting other IF-THEN

⁸Note that usually the set $LPerc^{LD}(u_0, w)$ has only one member, only occasionally there are more antecedents which fit the given input equally and cannot be distinguished using $\leq_{(u_0, w)}$.

⁹Let us note that, besides the well-known CRI inference, there is also a Bandler-Kohout subproduct that is an equally appropriate fuzzy relational composition for being used as an inference mechanism, see [18]. For an exhaustive investigation of fuzzy relational compositions, we refer to [19].

rules from LD . Then, if there were $L > 1$ members in $LPerc^{LD}(u_0, w)$, we aggregate corresponding L conclusions C_{i_i} to one final conclusion C by the intersection operation.

The idea of assigning local perceptions is not restricted only to the topic of a given linguistic description. If we generalize it slightly, we can learn the linguistic description on the basis of the given data. More details about this method can be found in [20]. Let us remark that we have successfully implemented this method in the software system LFLC (see [21]).

3. Redundancy

Linguistic descriptions may suffer from several problems, such as redundancy or inconsistency. The danger of an existence of such problems is even strengthened in cases when the discussed fuzzy rule base is automatically generated from given data. For instance the redundancy, that is the existence of redundant fuzzy rules in a given linguistic description, is usually caused by redundant measurements that are used to generate a linguistic description, see the properties of the learning procedure in [20]. Such situations are not typical only for fuzzy control but also for further applications, such as analysis and forecast of time series with the help of fuzzy rules, see [6, 5].

Generally, redundancy in fuzzy rules is observed as an existence of fuzzy rules with distinct but not contradictory antecedents (an antecedent is fully overlapped by an antecedent of another rule) and identical consequents. However, we will show that the situation is not that straightforward and that the redundancy phenomenon requires further formal investigation.

Remark 1 *Let us fix the notation for the rest of the paper. Let us consider a linguistic description LD and let us be given an observation u_0 in a given context $w \in W$. Let C be the conclusion derived from u_0 based on LD using the rule of perception based logical deduction given by (11). Then this fact will be denoted by*

$$r_{PbLD}(LPerc^{LD}(u_0, w)) : C. \quad (12)$$

Furthermore, suppose that fuzzy sets C, D on the same universe are equivalent. Then this fact will be denoted by $C \equiv D$.

Definition 1 *Let $LD = \{\mathcal{R}_1, \dots, \mathcal{R}_m\}$ be a linguistic description (4). Rule \mathcal{R}_i is redundant in LD if $D_1 \equiv D_2$ for each value $u_0 \in w, w \in W$, where*

$$\begin{aligned} r_{PbLD}(LPerc^{LD}(u_0, w)) &: D_1, \\ r_{PbLD}(LPerc^{LD'}(u_0, w)) &: D_2 \end{aligned}$$

and where $LD' = LD \setminus \{\mathcal{R}_i\}$.

As we have mentioned, redundancy is observed as an existence of fuzzy rules with distinct overlapping antecedents and identical consequents. But as we will show below, sometimes such an intuitively redundant fuzzy rule does not have to be always redundant with respect to a formal definition of the redundancy. Therefore, such a rule will be called *suspicious from a redundancy* and a further analysis of its potential redundancy turns out to be necessary.

Definition 2 *Let LD be a linguistic description (4), let $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$. Rule \mathcal{R}_i is suspicious from redundancy with respect to \mathcal{R}_j if $C_1 \equiv C_2$ for each value $u_0 \in w, w \in W$, where*

$$r_{PbLD}(LPerc^{\{\mathcal{R}_i, \mathcal{R}_j\}}(u_0, w)) : C_1$$

and

$$r_{PbLD}(LPerc^{\{\mathcal{R}_j\}}(u_0, w)) : C_2.$$

Theorem 1 *Let LD be a linguistic description (4), let $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$. Rule \mathcal{R}_i is suspicious from redundancy with respect to \mathcal{R}_j if and only if $\mathcal{A}_i \leq_{LE} \mathcal{A}_j$ and $B_i \equiv B_j$.*

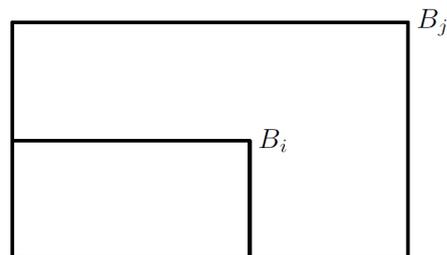


Figure 3: Visualization of fuzzy rule \mathcal{R}_i that is suspicious from redundancy w.r.t. fuzzy rule \mathcal{R}_j . Displayed rectangles symbolically denote areas where the respective fuzzy rules fire. Both rectangles are black and solid to symbolize that $B_i \equiv B_j$.

Theorem 1 claims that a fuzzy rule with an antecedent overlapped by an antecedent of another rule with the identical consequent is suspicious from the redundancy w.r.t. this rule. Furthermore, there are no other fuzzy rules that could be suspicious from redundancy w.r.t. another fuzzy rule besides those that meet the above mentioned situation. In other words, Theorem 1 specifies fuzzy rules that makes sense to investigate. The situation is displayed on Figure 3 where one can see that nothing changes if fuzzy rule \mathcal{R}_i is removed, unless other rules are involved.

The involvement of other rules may be demonstrated easily. Let us consider a linguistic description LD with $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$ where fuzzy rule \mathcal{R}_i is suspicious from the redundancy with respect to fuzzy rule \mathcal{R}_j . Then there can exist a fuzzy rule

$\mathcal{R}_k \in LD$ with a consequent \mathcal{B}_k distinct from consequent \mathcal{B}_i that “cancels” the redundancy from which \mathcal{R}_i was suspicious. For a visualization of such a situation we refer to Figure 4.

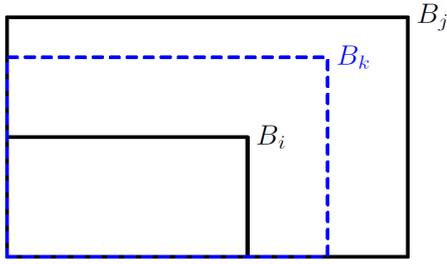


Figure 4: Visualization of a situation when fuzzy rule \mathcal{R}_k “cancels” the potential redundancy of fuzzy rule \mathcal{R}_i w.r.t. \mathcal{R}_j . Area where \mathcal{R}_k fires, is displayed by blue dashed line in order to symbolize that $\mathcal{B}_k \neq \mathcal{B}_i(\mathcal{B}_j)$.

Naturally, one could release a hypothesis stating the situation when fuzzy rule \mathcal{R}_i , that is suspicious from redundancy w.r.t. \mathcal{R}_j , is not redundant. Formally, such a hypothesis could be formulated as follows.

Hypothesis 1 Let $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$ and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . If there exists a rule $\mathcal{R}_k \in LD$ such that

- (i) $\mathcal{B}_k \neq \mathcal{B}_i$,
- (ii) $\mathcal{A}_k \leq_{LE} \mathcal{A}_j$,
- (iii) $\mathcal{A}_k \not\leq_{LE} \mathcal{A}_i$,

then \mathcal{R}_i is NOT redundant in LD .

However, after a careful investigation one may find a counterexample that declines Hypothesis 1. It is sufficient to consider another fuzzy rule $\mathcal{R}_p \in LD$ with the consequent $\mathcal{B}_p \equiv \mathcal{B}_i$ and with an appropriate antecedent, see Figure 5. In this case, fuzzy rule $\mathcal{R}_k \in LD$ cancels the suspicion of the fuzzy rule \mathcal{R}_i , but fuzzy rule $\mathcal{R}_p \in LD$ cancels the cancellation provided by $\mathcal{R}_k \in LD$. Thus, Hypothesis 1 does not hold.

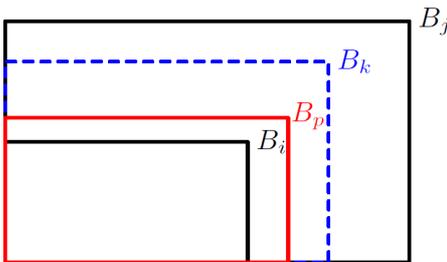


Figure 5: Scheme of fuzzy rules that contradicts Hypothesis 1 because \mathcal{R}_p with $\mathcal{B}_p \equiv \mathcal{B}_i(\mathcal{B}_j)$ cancels the cancellation of \mathcal{R}_k .

Nevertheless, Hypothesis 1 may be rewritten into a valid theorem if we consider a linguistic description that has only three rules.

Theorem 2 Let $LD = \{\mathcal{R}_i, \mathcal{R}_j, \mathcal{R}_k\}$ and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . If for \mathcal{R}_k the following holds

- (i) $\mathcal{B}_k \neq \mathcal{B}_i$,
- (ii) $\mathcal{A}_k \leq_{LE} \mathcal{A}_j$,
- (iii) $\mathcal{A}_k \not\leq_{LE} \mathcal{A}_i$,

then \mathcal{R}_i is NOT redundant in LD .

There exists another hypothesis which seems to be naturally valid. Let us be given a linguistic description LD , let $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$ and let rule \mathcal{R}_i be suspicious from the redundancy with respect to a rule \mathcal{R}_j . There can exist a rule $\mathcal{R}_k \in LD$ with $\mathcal{B}_k \neq \mathcal{B}_i$ that “cancels” the redundancy from which \mathcal{R}_i was suspicious, see Figure 6.

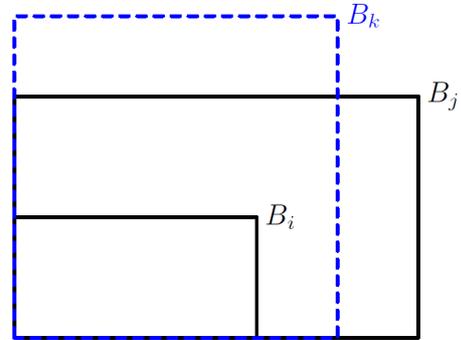


Figure 6: Visualization of another situation when fuzzy rule \mathcal{R}_k “cancels” the potential redundancy of fuzzy rule \mathcal{R}_i w.r.t. \mathcal{R}_j . Area where \mathcal{R}_k fires, is displayed by blue dashed line in order to symbolize that $\mathcal{B}_k \neq \mathcal{B}_i(\mathcal{B}_j)$.

Hypothesis 2 Let $\{\mathcal{R}_i, \mathcal{R}_j\} \in LD$ and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . If there exists a rule $\mathcal{R}_k \in LD$ such that

- (iv) $\mathcal{B}_k \neq \mathcal{B}_i$,
- (v) $\mathcal{A}_k \parallel_{LE} \mathcal{A}_j$, but $\mathcal{A}_k, \mathcal{A}_j$ have the same atomic expression,
- (vi) $\mathcal{A}_i \leq_{LE} \mathcal{A}_k$,

then \mathcal{R}_i is NOT redundant in LD .

In both Hypothesis 1 and Hypothesis 2 it is necessary to assume the consequents of canceling rules \mathcal{B}_k to be different than the consequent \mathcal{B}_i that appears in fuzzy rules \mathcal{R}_i as well as \mathcal{R}_j . This is specified by conditions (i) and (iv), respectively. The difference between Hypothesis 1 and Hypothesis 2 consists in the “placement” of the canceling fuzzy rule \mathcal{R}_k in a sense of determination the area (with respect to \mathcal{R}_i and \mathcal{R}_j) where this rule fires. In the case of Hypothesis 1, the canceling rule \mathcal{R}_k fires in an area

that is fully embedded into the area where \mathcal{R}_j fires, which is given by condition (ii). More specifically, the antecedent of \mathcal{R}_j fully overlaps the antecedent of \mathcal{R}_k , see Figure 4. Moreover, the antecedent of \mathcal{R}_k must not be fully overlapped by the antecedent of \mathcal{R}_i .

In the case of Hypothesis 2, the situation is different. The antecedent of \mathcal{R}_k neither overlaps the one of \mathcal{R}_j nor is overlapped by the one of \mathcal{R}_j , see Figure 6. This is given by condition (v). On the other hand, the antecedent \mathcal{R}_k must necessarily overlap the antecedent of \mathcal{R}_i , which is given by condition (vi). This is a significant difference to the condition (iii) from Hypothesis 1 which allowed also incomparable position instead of the full overlap.

Analogously to the case of Hypothesis 1, neither Hypothesis 2 hold, see the counterexample on Figure 7. There can exist a rule $\mathcal{R}_p \in LD$ with the consequent $\mathcal{B}_p \equiv \mathcal{B}_i$ and with an appropriate antecedent that cancels the cancellation of the suspicion of the rule \mathcal{R}_i provided by rule \mathcal{R}_k .

However, Hypothesis 2 may be rewritten into a valid theorem where we again consider a linguistic description that consist of just three rules.

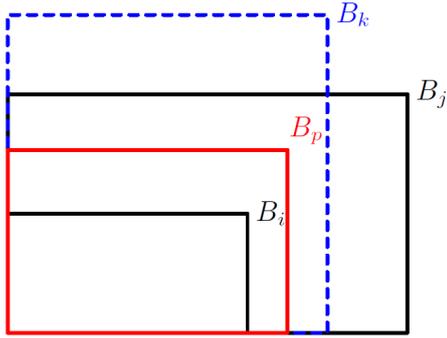


Figure 7: Scheme of fuzzy rules that contradicts Hypothesis 2 because \mathcal{R}_p with $\mathcal{B}_p \equiv \mathcal{B}_i(\mathcal{B}_j)$ cancels the cancellation of \mathcal{R}_k .

Theorem 3 Let $LD = \{\mathcal{R}_i, \mathcal{R}_j, \mathcal{R}_k\}$ and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . If for \mathcal{R}_k the following holds

- (iv) $\mathcal{B}_k \not\equiv \mathcal{B}_i$,
- (v) $\mathcal{A}_k \parallel_{LE} \mathcal{A}_j$, but $\mathcal{A}_k, \mathcal{A}_j$ have the same atomic expression,
- (vi) $\mathcal{A}_i \leq_{LE} \mathcal{A}_k$,

then \mathcal{R}_i is NOT redundant in LD .

Theorem 2 and Theorem 3 were formulated for a linguistic description that consist of only three rules, which makes their importance from a practical point of view rather low. Nevertheless, their existence is justified by the following theorem that stems from them. This theorem already provides us with a general result for arbitrary number of fuzzy rules.

Theorem 4 Let $LD = \{\mathcal{R}_1, \dots, \mathcal{R}_m\}$ be a linguistic description (4) and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . If there exist no rule $\mathcal{R}_k \in LD$ such that either (i)-(iii) or (iv)-(vi) holds, then \mathcal{R}_i is redundant in LD .

Immediately, we can state the following Corollary of Theorem 4.

Corollary 1 Let $LD = \{\mathcal{R}_i, \mathcal{R}_j, \mathcal{R}_k\}$ and let \mathcal{R}_i be suspicious from the redundancy with respect to \mathcal{R}_j . Rule \mathcal{R}_i is redundant if and only if neither (i)-(iii) nor (iv)-(vi) holds for \mathcal{R}_k .

The main contribution of Theorem 4 is that it provides a full classification of “canceling fuzzy rules”. It means, that due to this theorem we know that only fuzzy rules specified either by (i)-(iii) or by (iv)-(vi) may cancel the suspicion from redundancy and no other fuzzy rules may be responsible for that.

4. Conclusions and future work

We have introduced our approach to detection of so-called redundancies in systems of fuzzy/linguistic IF-THEN rules (linguistic descriptions). As we have shown, intuitively redundant rules are not always redundant from a formal point of view and thus, a deeper and formally correct approach had to be introduced.

Our approach is based on detecting the rules that are suspicious from redundancy and their further investigation. Full classification of the rules that are suspicious from redundancy has been provided. Intuitively valid hypotheses determining situations, when a rule suspicious from redundancy is not redundant were disproved. However, both hypotheses led us to a full classification of “canceling rules”, i.e., to a full classification of rules that may cancel the suspicion from redundancy.

After this preliminary and introductive study with promising results, we plan to focus our future work on further investigation of situations that occur when canceling rules exist. These situations may lead either to a true confirmation of the cancellation of the suspicion, or to a determination of the fact that a suspicious rule is indeed redundant. The final goal is a body of such theoretical results that allow us to release a deterministic algorithm that can automatically check (and remove) redundancies in any linguistic description consisting of a finite number of fuzzy/linguistic IF-THEN rules.

When a linguistic description is higher-dimensional (i.e., when it consists of IF-THEN rules with more than one antecedent variables, see (9) and discussion above it) then it may seem that there usually will not be many IF-THEN rules whose antecedents can be ordered using the ordering $\leq_{(u_0, w)}$ from Section 2.3. Hence, it seems that our method is not too useful then. However,

if these IF-THEN rules are generated automatically from data by some learning procedure, then redundant rules can occur quite often, no matter what the dimension of the linguistic description is. And their detection and removal can be really useful from the point of view of performance and interpretability. On the other hand, if there are no rules suspicious from redundancy found, then we can say that there are no redundancies in the strict formal sense of Definition 1. Generally speaking, this formal understanding of redundancy, which stresses the fact that original and new linguistic descriptions are equivalent from the point of view of their behavior, is significantly different in comparison with other approaches aiming mainly at a simplification of linguistic descriptions [2, 3] that use various techniques, e.g. rules merging. Of course, their use may be also beneficial. However, there is no guarantee that the output of simplified linguistic descriptions is equivalent with the output of the original one. Finally, in case of learning high-dimensional fuzzy/linguistic IF-THEN rules, we can restrict the number of linguistic hedges (Section 2.1) which possibly reduces the number of rules (because similar data records produce identical rules, which are pruned out).

Acknowledgement

We gratefully acknowledge support of the project DAR 1M0572 of the MŠMT ČR.

References

- [1] Babuška R. and M. Setnes. Data-driven construction of transparent fuzzy models. In H.B. Verbruggen, H.-J. Zimmermann, and R. Babuška, editors, *Fuzzy Algorithms for Control*, pages 83–106. Academic Kluwer Publ., Boston, 1999.
- [2] M. Setnes. *Fuzzy rule base simplification using similarity measures*. PhD thesis, MSc Thesis. Delft University of Technology, Delft, Netherlands, 1995.
- [3] M. Setnes, V. Lacroze, and A. Titli. Complexity reduction methods for fuzzy systems. In H.B. Verbruggen, H.-J. Zimmermann, and R. Babuška, editors, *Fuzzy Algorithms for Control*, pages 185–218. Academic Kluwer Publ., Boston, 1999.
- [4] V. Novák. Perception-based logical deduction. In B. Reusch, editor, *Computational Intelligence, Theory and Applications*, Advances in Soft Computing, pages 237–250, Berlin, 2005. Springer.
- [5] V. Novák, M. Štěpnička, A. Dvořák, I. Perfilieva, V. Pavliska, and L. Vavříčková. Analysis of seasonal time series using fuzzy approach. *International Journal of General Systems*, 39:305–328, 2010.
- [6] M. Štěpnička, A. Dvořák, V. Pavliska, and L. Vavříčková. Linguistic approach to time series modeling with the help of F-transform. *Fuzzy Sets and Systems*, to appear.
- [7] V. Novák, I. Perfilieva, and N.G. Yarushkina. A general methodology for managerial decision making using intelligent techniques. In *Recent Advances in Decision Making*, pages 4–22. Springer, 2010.
- [8] V. Novák, I. Perfilieva, A. Dvořák, Q. Chen, Q. Wei, and P. Yan. Mining pure linguistic associations from numerical data. *International Journal of Approximate Reasoning*, 48:4–22, 2010.
- [9] V. Novák and J. Kovář. Linguistic if-then rules in large scale application of fuzzy control. In Da Ruan and E.E. Kerre, editors, *Fuzzy IF-THEN Rules in Computational Intelligence. Theory and Applications*, pages 223–241. Academic Kluwer Publ., Boston, 2000.
- [10] V. Novák. A comprehensive theory of trichotomous evaluative linguistic expressions. *Fuzzy Sets and Systems*, 159(22):2939–2969, 2008.
- [11] L. A. Zadeh. Precisiated natural language. *AI Magazine*, 25:74–91, 2004.
- [12] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst. Man Cybern.*, 3(1):28–44, 1973.
- [13] L. A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning I–III. *Inform. Sci.* 8 (1975) 199–250, 8 (1975) 301–357, 9 (1975) 43–80, 1975.
- [14] M. De Cock and E.E. Kerre. Fuzzy modifiers based on fuzzy relations. *Information Sciences*, 160:173–199, 2004.
- [15] G. Kaltenböck, W. Mihatsch, and S. Schneider, editors. *New Approaches to Hedging*. Emerald, Bingley, 2010.
- [16] E. Hajičová, B. Partee, and P. Sgall. *Topic-focus Articulation, Tripartite Structures, and Semantic Content*. Kluwer, Dordrecht, 1998.
- [17] P. Sgall, E. Hajičová, J. Panevová, and J. Mey. *The meaning of the sentence in its semantic and pragmatic aspects*. Kluwer, Boston, 1986.
- [18] M. Štěpnička and B. Jayaram. On the suitability of the Bandler-Kohout subproduct as an inference mechanism. *IEEE Transactions on Fuzzy Systems*, 18(2):285–298, 2010.
- [19] L. Běhounek and M. Daňková. Relational compositions in fuzzy class theory. *Fuzzy Sets and Systems*, 160(8):1005 – 1036, 2009.
- [20] R. Bělohlávek and V. Novák. Learning rule base of the linguistic expert systems. *Soft Computing*, 7:79–88, 2002.
- [21] A. Dvořák, H. Habiballa, V. Novák, and V. Pavliska. The software package LFLC 2000 - its specificity, recent and perspective applications. *Computers in Industry*, 51:269–280, 2003.