



HAL
open science

What are fuzzy rules and how to use them

Didier Dubois, Henri Prade

► **To cite this version:**

Didier Dubois, Henri Prade. What are fuzzy rules and how to use them. *Fuzzy Sets and Systems*, 1996, 84, dedicated to the Memory of Professor Arnold Kaufmann (2), pp.169–185. <10.1016/0165-0114(96)00066-8>. <hal-04039815>

HAL Id: hal-04039815

<https://hal.science/hal-04039815v1>

Submitted on 22 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



ELSEVIER

Fuzzy Sets and Systems 84 (1996) 169–185

FUZZY
sets and systems

What are fuzzy rules and how to use them

Didier Dubois, Henri Prade

*Institut de Recherche en Informatique de Toulouse (I.R.I.T.) – C.N.R.S., Université Paul Sabatier, 118 route de Narbonne,
31062 Toulouse Cedex, France*

Received February 1996

Abstract

Fuzzy rules have been advocated as a key tool for expressing pieces of knowledge in “fuzzy logic”. However, there does not exist a unique kind of fuzzy rules, nor is there only one type of “fuzzy logic”. This diversity has caused many a misunderstanding in the literature of fuzzy control. The paper is a survey of different possible semantics for a fuzzy rule and shows how they can be captured in the framework of fuzzy set and possibility theory. It is pointed out that the interpretation of fuzzy rules dictates the way the fuzzy rules should be combined. The various kinds of fuzzy rules considered in the paper (gradual rules, certainty rules, possibility rules, and others) have different inference behaviors and correspond to various intended uses and applications. The representation of fuzzy unless-rules is briefly investigated on the basis of their intended meaning. The problem of defining and checking the coherence of a block of parallel fuzzy rules is also briefly addressed. This issue has been neglected in the fuzzy control literature although it looks important for validation purposes.

Keywords: Fuzzy rules; Knowledge representation; Gradualness; Certainty; Possibility; Exceptions; Consistency; Approximate reasoning

“Associer le mot flou avec le mot logique est choquant. La logique, au sens vulgaire du mot, est une conception des mécanismes de la pensée qui ne devrait être jamais floue, mais toujours rigoureuse et formelle. [...]

La pensée humaine, superposition d'intuition et de rigueur, c'est-à-dire d'une prise en compte globale ou parallèle (nécessairement floue) et d'une prise en compte logique ou séquentielle (nécessairement formelle), est un mécanisme flou. Les lois de la pensée que nous pouvons faire entrer dans les programmes des ordinateurs sont obligatoirement formelles, les lois de la pensée dans le dialogue homme-homme sont floues.”

Arnold Kaufmann
[23, Chapter 3, p. 191]

1. Introduction

If/then rules offer a convenient format for expressing pieces of knowledge. But it is just a format which can cover different intended semantics and uses. The consequent of a rule may be qualified

with various modalities expressing, for instance, strong ones such as certainty, obligation, or weaker ones such as possibility or feasibility. Rules may have implicit exceptions, or may universally hold. Depending on their interpretation, rules have to be represented and processed in a specific way at the

inference level. Strangely enough, in spite of the acknowledged importance of fuzzy rules and of the great number of works in fuzzy set-based approximate reasoning, there has been very little concern until now about the possible intended meanings of a fuzzy rule, although fuzzy rules seem to play an important role in our thinking mechanisms. This paper is a systematic attempt to analyzing what fuzzy rules are.

Fuzzy if/then rules are rules whose antecedents, consequences or both are fuzzy rather than crisp. The most common and widely used interpretation considers a fuzzy rule “if x is A then y is B ” as a fuzzy point $A \times B$ and a collection of fuzzy rules “if x is A_i then y is B_i ” $i = 1, \dots, n$ as a fuzzy graph providing a rough description of the relation between x and y (e.g., [36, 37]). Then the benefit of having fuzzy antecedents is to provide a basis for an interpolation mechanism. In this view a fuzzy rule is defined by means of a conjunction (for defining a fuzzy Cartesian product $A \times B$) rather than in terms of a multiple-valued logic implication. However, for a long time there has been, especially in an approximate reasoning perspective, a considerable interest for fuzzy rules modelled by multivalued implications under the form $A \rightarrow B$ (see [2, 27, 30] as noticeable examples of early contributions). The variety of implication operators existing in the fuzzy set framework has always been seen as a rich potential for modelling different shades of expert attitude in the inference process (e.g., [24, 25]), although no precise, practical interpretation was provided for the different implication operators.

The proper use of implication-based fuzzy rules is often misunderstood in fuzzy control. For example, in a recent introduction to fuzzy logic by Mendel [26] for a special issue of Proceedings of IEEE on fuzzy logic engineering applications, the author criticizes fuzzy implications that produce output fuzzy sets with an unlimited support, and concludes that it is “a result that violates engineering commonsense”. Although Mendel [26] asks himself if this situation “is a result of a poor choice for $\mu_{A \rightarrow B}$ ”, he apparently does not notice that there indeed exist genuine implications which do not have this behavior, such as Gödel implication, or Goguen implication. Moreover, Mendel [26] interprets the fact that $\forall v, \mu_A(u_0) \rightarrow \mu_B(v) = 1$ as soon as

the input u_0 is such that $\mu_A(u_0) = 0$, as “the rule will be fired”, and concludes that “this does not make much sense from an engineering perspective”. However, anybody familiar with material implication in logic, knows that a formula $p \rightarrow q$ is true when p is false, and that the rule is then not triggered at all. Indeed, there is no way of producing q from $p \rightarrow q$ and $\neg p$. This is an example of a limited, often encountered, view of what a fuzzy rule is, which may work for simple control engineering applications, but which is clearly insufficient in a knowledge engineering perspective.

Several authors [2, 9, 15–18, 28] have pointed out that fuzzy rules of the form “if x is A then y is B ”, modelled by a Cartesian product $A \times B$, have a significantly different meaning from rules modelled in terms of genuine implications $A \rightarrow B$, and should serve different purposes. In the following, such a view is developed further by insisting on the meaning of fuzzy rules. Understanding the semantics of fuzzy rules is a key issue for their intended applications, and for processing them. Moreover fuzzy “unless rules” (of the form “if x is A then y is B unless z is C ”) are also considered in this perspective. “Unless rules” have been already discussed by Driankov and Hellendoorn [7, 8]. They have advocated the interest of these context-dependent rules. They have proposed to model fuzzy “unless rules” by means of a ternary relation whose behavior obeys some inferential constraints. The approach proposed in the paper also briefly relates the semantical view of fuzzy rules previously developed to the possibilistic logic treatment of nonmonotonic reasoning for handling rules with exceptions [3, 11].

Section 2 first discusses different representations of rules in a nonfuzzy setting and then extends these representations to rules with a fuzzy conclusion part. Section 3 introduces the different types of fuzzy rules, derives their representation in the framework of fuzzy sets and possibility theory from their intended semantics, and gives their inference behaviors. Section 4 deals with fuzzy unless rules. Section 5 briefly discusses the coherence of parallel rules. This issue is particularly important for the validation of expert-generated systems of fuzzy rules, and has not been very often considered.

2. From if/then rules to fuzzy rules

To properly understand what a fuzzy rule may mean, it might be useful to start with the meaning of a nonfuzzy rule of the form “if x is A , then y is B ”, and then to investigate what are the different possible meanings of the rule when B becomes fuzzy, while A remains an ordinary set. The case where A , and possibly B , are fuzzy is fully addressed in Section 3. In the whole paper it will be assumed that the fuzzy sets are normalized (i.e., there exists at least one element with a membership grade equal to 1), unless stated otherwise.

2.1. Nonfuzzy rules

Let us consider a rule “if $x \in A$ then $y \in B$ ”, where A and B are ordinary subsets and x and y are variables ranging on the domains U and V respectively. The rule provides a partial and imprecise description of a relationship between the variables x and y . It means in terms of characteristic functions that

$$\text{if } \mu_A(x) = 1 \text{ then } \mu_B(y) = 1. \quad (1)$$

Clearly, the pairs (u, v) of values of the variables (x, y) such that $\mu_A(u) = 1$ and $\mu_B(v) = 0$ are not compatible with (1), while the pairs such that $\mu_A(u) = 1$ and $\mu_B(v) = 1$ agree with (1). However, nothing is said for the pairs (u, v) of values of (x, y) such that $\mu_A(u) = 0$; for them $\mu_B(v)$ remains unconstrained according to (1). Let R be the set of pairs (u, v) of values of (x, y) compatible with (1). It can be checked that its characteristic function is such that

$$\mu_A(u) \wedge \mu_B(v) \leq \mu_R(u, v) \leq n(\mu_A(u)) \vee \mu_B(v) \quad (2)$$

where \wedge denotes the Boolean conjunction operation, (i.e., $a \wedge b = 1$ iff $a = b = 1$ and $a \wedge b = 0$ otherwise), $n(a) = 1$ iff $a = 0$ and $n(a) = 0$ if $a = 1$, and \vee denotes a disjunction operation (i.e., $a \vee b = 0$ iff $a = b = 0$ and $a \vee b = 1$ otherwise). Indeed, R contains at least the pairs such that $\mu_A(u) = 1$ and $\mu_B(v) = 1$, which are such that $\mu_A(u) \wedge \mu_B(v) = 1$. R also contains at most all the pairs such that $\mu_A(u) = 0$ together with the ones such that $\mu_A(u) = 1$ and $\mu_B(v) = 1$, i.e., the pairs such that $n(\mu_A(u)) \vee \mu_B(v) = 1$. (2) can be written equivalently in set theoretic terms under the form

$A \times B \subseteq R \subseteq \bar{A} + B$ where the overbar denotes the complementation and $+$ is the operation dual to the Cartesian product \times : $\bar{A} + \bar{B} = \bar{A} \times \bar{B}$.

Thus, (2) expresses that any representation of the rule “if $x \in A$ then $y \in B$ ” is lower bounded by the representation of the conjunction “ $x \in A$ and $y \in B$ ” and upper bounded by the representation of the material implication $x \in A \Rightarrow y \in B$, i.e., $x \notin A$ or $y \in B$. This view is the starting point of the modelling of a rule “if $x \in A$ then $y \in B$ ” in terms of “conditional objects” (or “conditional events”), namely, “ $y \in B$ given that $x \in A$ ”, or $B|A$ for short. See [22] for a study of the algebraic aspects of conditional objects and [21] for their logical study in the scope of nonmonotonic reasoning applications. Although Bouchon [4] has studied fuzzy rules in the scope of conditioning, and Weber [32] has already studied extensions of conditional objects when A and B are fuzzy, we shall not investigate fuzzy rules in that perspective here. Rather we are going to deal separately with the two inequalities in (2) and to view each of them as the partial specification of a possibility distribution $\pi_{y|x}$ pertaining to the value of y given the value of x . Namely

– the inequality

$$\pi_{y|x}(v, u) \geq \mu_A(u) \wedge \mu_B(v), \quad (3)$$

means that all the values v in B are allowed for y when x takes a value in A , i.e., $\pi_{y|x}(v, u) = 1$, $\forall v \in B, \forall u \in A$; for pairs (u, v) such that $u \in A$ and $v \notin B$, as well as for pairs (u, v) such that $u \notin A$ nothing is said by (3) since for them we only have the trivial constraint $\pi_{y|x}(v, u) \geq 0$. Thus, the inequality (3) corresponds to the semantics of a rule expressing that “if $x \in A$, all the values in B are possible (or if we prefer admissible, feasible, allowed) for y ”;

– the inequality

$$\pi_{y|x}(v, u) \leq n(\mu_A(u)) \vee \mu_B(v), \quad (4)$$

expresses that the pairs (u, v) of values of (x, y) such that u belongs to A and v does not belong to B are forbidden ($\forall u \in A, \forall v \notin B, \pi_{y|x}(v, u) \leq 0$); then the value of y is necessarily among the elements of B , when x takes a value u outside B are impossible. When x takes a value u outside A , the possible values for y are unrestricted, since

then $\pi_{y|x}(v, u) \leq 1$. Thus, the inequality (4) means that “if $x \in A$, it is *certain* that the value of y is in B ”.

The two semantics, respectively in terms of possibility (or feasibility) and in terms of certainty, of the two models of an if/then rule appear to be clearly distinct by the way conclusions are combined when dealing with two parallel rules. Indeed, the two rules “if $x \in A_1$ then $y \in B_1$ ” and “if $x \in A_2$ then $y \in B_2$ ” induce the following constraints on the possibility distribution $\pi_{y|x}$ describing the relationship between x and y :

– in the conjunction-based model

$$\pi_{y|x}(v, u) \geq \mu_{A_1}(u) \wedge \mu_{B_1}(v),$$

$$\pi_{y|x}(v, u) \geq \mu_{A_2}(u) \wedge \mu_{B_2}(v),$$

and then

$$\pi_{y|x}(v, u) \geq (\mu_{A_1}(u) \wedge \mu_{B_1}(v)) \vee (\mu_{A_2}(u) \wedge \mu_{B_2}(v)).$$

This expresses that interpreting the rules as “if $x \in A_1$, all the values in B_1 are possible for y ” and “if $x \in A_2$, all the values in B_2 are possible for y ”, we indeed conclude that “if $x \in A_1 \cap A_2$ then all the values in B_1 or B_2 , i.e., in $B_1 \cup B_2$, are possible for y ”;

– in the implication-based model

$$\pi_{y|x}(v, u) \leq n(\mu_{A_1}(u)) \vee \mu_{B_1}(v),$$

$$\pi_{y|x}(v, u) \leq n(\mu_{A_2}(u)) \vee \mu_{B_2}(v),$$

and then

$$\pi_{y|x}(v, u) \leq [n(\mu_{A_1}(u)) \vee \mu_{B_1}(v)] \wedge [n(\mu_{A_2}(u)) \vee \mu_{B_2}(v)].$$

This expresses that interpreting the rules as “if $x \in A_1$, it is certain that the value of y is in B_1 ” and “if $x \in A_2$, it is certain that the value of y is in B_2 ”, we indeed conclude that “if $x \in A_1 \cap A_2$ then it is certain that the value of y is in B_1 and in B_2 , i.e., in $B_1 \cap B_2$ ”.

Thus depending on the interpretation of the rules, conclusions of parallel rules will have to be combined disjunctively in the conjunction-based model (whose semantics is in terms of possibility/feasibility), and conjunctively in the implication-based model (whose semantics is in terms of necessity/certainty).

Informally, the conjunction-based model of a rule comes down to viewing the rule as a piece of data (the rule is then considered as a pair (A, B)). The disjunctive combination of rules is then a data accumulation procedure. On the contrary, the implication-based model of a rule comes down to viewing the rule as a constraint that restricts a set of possible solutions. The conjunctive combination is then a refinement, a solution elimination procedure aiming at improving precision (see also [28]).

Note. The two semantics cannot always be enforced simultaneously in the sense that the two sets of constraints $A_1 \times B_1 \subseteq R_1 \subseteq \bar{A}_1 + B_1$ and $A_2 \times B_2 \subseteq R_2 \subseteq \bar{A}_2 + B_2$ are sometimes incompatible with a unique relation $\mu_{R_1}(u, v) = \mu_{R_2}(u, v) = \pi_{y|x}(v, u)$, $\forall u, \forall v$, since $(A_1 \times B_1) \cup (A_2 \times B_2)$ is, generally not included into $(\bar{A}_1 + B_1) \cap (\bar{A}_2 + B_2)$.

2.2. Rules with fuzzy conclusion parts

If/then rules with a nonfuzzy condition part and a fuzzy conclusion part express the existence of more or less possible conclusions when the condition is satisfied. A particular case of such a situation is an uncertain rule. Following an idea, already presented in Dubois and Prade [14], an uncertain rule can be obtained from a nonfuzzy, completely certain rule, under a default information assumption pertaining to a subpart of its condition part. Namely, let us consider the rule

“if $x \in A$ and $z \in C$ then $y \in B$ ”,

where A, B and C are ordinary subsets of U, V and W respectively. Let us assume that we have some default information about the value of z with respect to C , of the type “it is almost sure that $z \in C$ ”. Such a piece of information can be modelled by a simple possibility distribution of the form

$$\pi_z(w) = \begin{cases} 1 & \text{if } w \in C \\ \lambda < 1 & \text{if } w \notin C \end{cases} = \max(\mu_C(w), \lambda). \quad (5)$$

This special kind of fuzzy set with unlimited support expresses that the degree of certainty of “ $z \in C$ ” is then equal to $N(C) = \inf_{w \notin C} 1 - \pi_z(w) = 1 - \lambda$. The smaller λ , the less possible $z \notin C$ and the more certain the default information “ $z \in C$ ”. Using the

implication-based model of the rule, we have

$$\pi_{y|x,z}(v, u, w) \leq n(\mu_A(u)) \vee n(\mu_C(w)) \vee \mu_B(v). \quad (6)$$

Combining conjunctively π_z and $\pi_{y|x,z}$ and projecting on $V \times U$, according to Zadeh's [35] possibility theory-based approach to approximate reasoning, we easily get

$$\begin{aligned} \sup_w \min(\pi_z(w), \pi_{y|x,z}(v, u, w)) &= \pi_{y|x}(v, u) \\ &\leq \max(n(\mu_A(u)), \mu_B(v), \lambda). \end{aligned} \quad (7)$$

Thus the possibility distribution which is obtained expresses that the possible values for y , when the value of x is in A , are in B with a possibility degree at most equal to 1, and outside B with a possibility at most λ . Let B^* be the fuzzy set defined by $\mu_{B^*}(v) = \max(\mu_B(v), \lambda)$, then we recognize that (7) is an implication-based model of the rule "if $x \in A$ then y is restricted by B^* ", where the underlying implication is defined by

$$\begin{aligned} a \rightarrow b^* &= \max(n(a), b^*) = \max(n(a), b, \lambda) \\ &= \begin{cases} \lambda & \text{if } a = 1 \text{ and } b = 0 \\ 1 & \text{otherwise.} \end{cases} \end{aligned}$$

The fuzzy rule so obtained is actually a crisp rule with an uncertain conclusion of the form "if $x \in A$ then y is B is $(1 - \lambda)$ -certain".¹

Thus, a special kind of rule with a fuzzy conclusion part having unlimited support has been obtained here as an approximation of a rule without any uncertainty, taking into account default information. Limited support fuzzy sets in the conclusion part of the rule may be obtained as well. Let us now consider the two rules

"if $x \in A$ and $z \in C$ then $y \in B$ ",

"if $x \in A$ and $z \notin C$ then $y \in B'$ ".

Then $\pi_{y|x,z}$ is upper bounded by (6) and by $n(\mu_A) \vee \mu_C \vee \mu_{B'}$. Combining π_z (defined by (5)) and $\pi_{y|x,z}$ and projecting the result, it can be checked that we finally obtain

$$\pi_{y|x}(v, u) \leq \max[n(\mu_A(u)), \max(\mu_B(v), \min(\lambda, \mu_{B'}(v)))] \quad (8)$$

Again we obtain the implication-based model of a rule of the form "if $x \in A$ then y is restricted by B^* " with $\mu_{B^*}(v) = \max(\mu_B(v), \min(\lambda, \mu_{B'}(v)))$, i.e., B^* is a fuzzy set with support $B' \cup B$ and core B . More generally, starting with a collection of rules

"if $x \in A$ and $z \in C_i$ then $y \in B_i$ " for $i = 1, \dots, n$ (9)

where the C_i 's make a partition of W , and using a default information of the form

$$\pi_z(w) = \lambda_i \text{ if } z \in C_i \text{ and } \lambda_1 = 1 > \dots > \lambda_n > 0, \quad (10)$$

by combination and projection of (10) and of the upper bounds of $\pi_{y|x,z}$ induced by the rules (9), it can be shown that we get

$$\pi_{y|x}(v, u) \leq \max(n(\mu_A(u)), \mu_{B^*}(v)),$$

with

$$\mu_{B^*}(v) = \max_{i=1..n} \min(\lambda_i, \mu_{B_i}(v)). \quad (11)$$

Moreover B^* is a fuzzy set whose λ_i -level cut is B_i for $i = 1, \dots, n$, if $B_1 \subset B_2 \subset \dots \subset B_n$.

Using the conjunction-based model of the rules, we can also obtain a rule with a fuzzy conclusion by an approximating process. Indeed the conjunction-based model of rules of the form (9) is

$$\pi_{y|x,z}(v, u, w) \geq \bigvee_{i=1..n} \mu_A(u) \wedge \mu_{C_i}(w) \wedge \mu_{B_i}(v), \quad (12)$$

which gives with the default information (10), by combination and projection

$$\pi_{y|x}(v, u) \geq \min(\mu_A(u), \mu_{B^*}(v)), \quad (13)$$

with $\mu_{B^*}(v) = \max_{i=1..n} \min(\lambda_i, \mu_{B_i}(v))$ again.

To conclude, a rule with a fuzzy conclusion part can be seen as approximating, but also summarizing, a collection of nonfuzzy rules, taking into account a default information. This view can be extended to continuous membership functions when V and W are subsets of the real line.

¹ This kind of fuzzy rules is the one encoded by weighted clauses in possibilistic logic [11], where the fuzzy set of models of the clause $(\neg p \vee q, \alpha)$ representing the piece of information "it is α -certain that if p is true then q is true", is given by $\mu_{M(\neg p \vee q, \alpha)} = \max(\mu_{M(\neg p \vee q)}, 1 - \alpha)$ where $M(\neg p \vee q)$ is the classical set of models of $\neg p \vee q$. The term "model" is understood here in the sense of logic, i.e., the set of interpretations which makes a formula true.

3. Representation of different kinds of fuzzy rules

3.1. Making the condition part of a rule fuzzy

We just saw that the conclusion-part of a rule “if $x \in A$ then y is B^* ” can be made fuzzy in order to express the existence of more or less possible values of y when the condition $x \in A$ holds for true. The reason for having the condition part, say, x is A , fuzzy is of a different nature.

Let us consider a rule of the form “if x is A then y is B^* ” where A and/or B might be fuzzy. When A is a fuzzy set, we distinguish between typical values in the core of A (i.e., the u 's such that $\mu_A(u) = 1$) and less typical values which are such that $1 > \mu_A(u) > 0$. Then the idea is to say that if the value of x is no longer in the core of A , but still close to the values u such that $\mu_A(u) = 1$, the possible values of y are in some subset not too much different from B . The ways B can be modified in order to represent the set of possible values of y when the value of x is no longer in the core of A but still close to it, will depend on the representation of the rule that is used.

If we use an *implication-based* model, B is viewed as an upper bound of the possible values of y when x is in A , i.e., B restricts the possible values of y . Then it seems natural, when the input condition is no longer fully satisfied, to modify B into a *less restrictive* fuzzy set (where the possibility degrees attached to values can only increase). This can be done in two different ways (e.g., [5, 6, 16]).

- One way is to attach some uncertainty to B by expressing that the possibility degree of values outside of the support of B is no longer strictly zero, but progressively increases as the value of x moves away from the core of A . This can be done whether B is fuzzy or not. If the support of B is the whole domain V , the possibility degrees of the less possible values for y will simply increase when x moves away from the core of A . This corresponds to a rule of the type “the more x is A , the more certain y is B^* ” (in other words the less possible the values outside B). An example of such a rule is “the younger a person is, the more certain the person is single”.
- The other way of using a genuine implication consists in enlarging the core of B in such a way

that when the value of x is in the α -cut of A ($x \in \{u, \mu_A(u) \geq \alpha\}$), all the values v in the α -cut of B become fully possible (i.e., possible at the degree 1). Thus, when the value of x is somewhat close to the core of A , the value of y remains somewhat close to the core of B . Clearly, this second way requires that B be a fuzzy set too. Then there still exist two slightly different manners to implement it depending on the interpretation of the rule when x is in the core of A : we may either consider that the possible values of y are restricted by the core of B , or by B itself (unmodified in this case). With the first manner, we model a purely gradual rule of the form “the more x is A , the more y is B^* ”, while with the second one some fuzziness on the conclusion is acknowledged (since we associate a fuzzy set of values to y).

If we use a *conjunction-based* model, B is viewed as a lower bound of the possible values of y when x is in A , i.e., B is viewed as a set of values which are guaranteed to be possible for y . Then, when the value of y moves away from the core of A , we may only think of reducing the minimal level of guaranteed possibility or the range of values whose possibility is assessed. Again there are two cases [18]:

- In the first case, the lower bound on the level of possibility of the values in B will be diminished down to the degree of membership of x to A . This can be done whether B is fuzzy or not. This corresponds to a rule of the type “the more x is A , the more possible y is B^* ”. An example of such a rule is “the older a person is, the more possible the person has/had been married”.
- In the second case, we can consider that when the value of x is in the core of A , all the values in the support of B are possible without restriction, while, as x gets away from the core of A , the subset of values with some guaranteed possibility for y becomes smaller and smaller inside B . Clearly, in this second interpretation B should be a fuzzy set. Again, there are two subcases here: when x is in the core of A , we can consider that either, only the values in the support of B (i.e., $\{v, \mu_B(v) > 0\}$) have a degree of possibility guaranteed to be equal to 1, or the lower bound on degrees of possibility for y is described by the

fuzzy set B itself. In the first subcase, we capture a rule of the type “the more x is A , the larger the set of possible values for y around the core of B ”. An example of such a rule is “the more experienced a cook, the larger the set of dishes he can prepare”. In the second subcase some fuzziness is also acknowledged.

In the following, it is shown that these different approaches we may have in mind for a fuzzy rule can be indeed captured by different extensions of the framework described in Section 2.

3.2. Modelling fuzzy rules

We now restate the main results of Dubois and Prade, [18] about the representation of the different kinds of fuzzy rules. In this reference, the reader can find detailed justifications, in terms of possibility and certainty qualifications and in terms of graduality, of the representations proposed for the different rules.

3.2.1. Certainty rules

A first kind of fuzzy rule corresponds to statements of the form “the more x is A , the more certain y lies in B ”. Interpreting the rule as “ $\forall u$, if $x = u$, it is at least $\mu_A(u)$ -certain that y lies in B ”, the degree $1 - \mu_A(u)$ assesses the possibility that y is outside of B when $x = u$, since the more x is A , the less possible y lies outside B , and the more certain y lies in B (indeed the certainty of an event corresponds to the impossibility of the contrary event). By application of (7) we get the following constraint for the conditional possibility distribution modelling the rule

$$\forall u \in U, \forall v \in V, \pi_{y|x}(v, u) \leq \max(1 - \mu_A(u), \mu_B(v)). \tag{14}$$

We recognize Kleene–Dienes implication ($a \rightarrow b = \max(1 - a, b)$) for combining A and B in (14).

In the particular case where A is an ordinary subset, (14) yields

$$\forall u \in A, \pi_{y|x}(v, u) \leq \mu_B(v) \tag{15}$$

$\forall u \notin A, \pi_{y|x}(v, u)$ is completely unspecified.

This corresponds to the implication-based modelling of a fuzzy rule with a nonfuzzy condition part

described in Section 2.2. Note that B may be any kind of fuzzy set in (14), hence in (15) as well. Particularly, B may itself include some uncertainty; for instance, the membership function of B may be of the form $\mu_B = \max(\mu_{\hat{B}}, \lambda)$ in order to express that when x is A , \hat{B} is the range (maybe fuzzy but with a limited support) of y with a certainty $1 - \lambda$ (any value outside the support of \hat{B} remains a possible value for y with a degree equal to λ). Default rules are conveniently modelled by (14). For instance the rule “if x is a bird then it is rather certain that it flies” is modelled with A not fuzzy and B of the form $\mu_B = \max(\mu_{\hat{B}}, \lambda)$, while the rule “the more typical a bird is, the more certain it flies” is modelled by means of a fuzzy A and a non-fuzzy B .

As can be seen from (14), B is modified by applying a function $\mu_\tau(t) = \max(t, 1 - \alpha)$, $\forall t \in [0, 1]$ to μ_B with $\alpha = \mu_A(u)$, (14) can be rewritten $\pi_{y|x} \leq \mu_\tau \circ \mu_B$; τ is pictured in Fig. 1(b) and can be seen as a fuzzy truth-value in the sense of Zadeh [34]. This fuzzy truth-value approach to fuzzy rules is originally due to Baldwin [1].

3.2.2. Gradual rules

This second kind of implication-based fuzzy rule has been discussed in detail in [19]. Gradual rules correspond to statements of the form the more x is A , the more y is B . Statements involving “the less” in place of “the more” are easily obtained by changing A or B into their complements \bar{A} and \bar{B} due to the equivalence between “the more x is A ” and “the less x is \bar{A} ” (with $\mu_{\bar{A}} = 1 - \mu_A$). More precisely, the intended meaning of a gradual rule can be understood in the following way: “the greater the degree of membership of the value of x to the fuzzy set A and the more the value of y is considered to be in relation (in the sense of the rule) with the value of x , the greater the degree of membership to B should be for this value of y ”, i.e.

$$\forall u \in U, \min(\mu_A(u), \pi_{y|x}(v, u)) \leq \mu_B(v) \tag{16}$$

or, using the equivalence $\min(a, t) \leq b \Leftrightarrow t \leq a \rightarrow b$ where \rightarrow denotes Gödel’s implication,

$$\forall u \in U, \pi_{y|x}(v, u) \leq \mu_A(u) \rightarrow \mu_B(v) = \begin{cases} 1 & \text{if } \mu_A(u) \leq \mu_B(v), \\ \mu_B(v) & \text{if } \mu_A(u) > \mu_B(v). \end{cases} \tag{17}$$

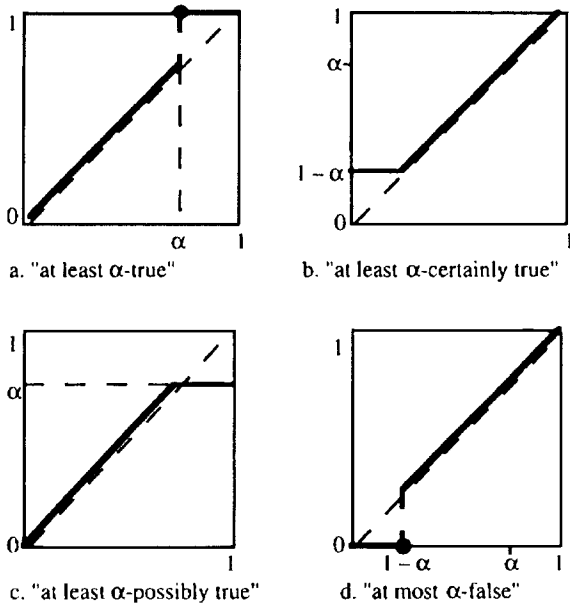


Fig. 1. Four basic types of fuzzy truth-values

Note that if A is not fuzzy, we recover $\pi_{y|x}(v, u) \leq \mu_B(v)$ if $u \in A$ and $\pi_{y|x}(v, u) \leq 1$ if $u \notin A$, i.e., the implication-based model of a rule with a fuzzy conclusion part. Eq. (17) can be equivalently written

$$\forall u \in U, \pi_{y|x}(v, u) \leq \max(\mu_{[\mu_A(u), 1]}(\mu_B(v)), \mu_B(v)) = \mu_{[\mu_A(u), 1] \cup T}(\mu_B(v)) \quad (18)$$

where $\mu_{[\mu_A(u), 1]}$ is the characteristic function of the interval $[\mu_A(u), 1]$ and where T is the fuzzy set on $[0, 1]$, defined by $\forall t \in [0, 1], \mu_T(t) = t$, which models the fuzzy truth-value “true” in fuzzy logic [34]. If we remember that “ x is A is τ -true” is represented by the possibility distribution [34]

$$\forall u \in U, \pi_x(u) = \mu_\tau(\mu_A(u)), \quad (19)$$

where τ is a fuzzy truth-value modelled by the fuzzy set τ of $[0, 1]$, we can interpret the meaning of gradual rules in the following way using (18) and (19): “ $\forall u \in U$, if $x = u$ then it is at least $\mu_A(u)$ -true that y is B ”. The membership function of the fuzzy truth-value “at least α -true” is pictured in Fig. 1(a). As it can be seen, it is not a crisp “at least α -true”

(which would correspond to the ordinary subset $[x, 1]$), but a fuzzy one in agreement with the truth-qualification in the sense of (19), indeed “at least 1-true” corresponds to the fuzzy truth-value T .

If we are only looking for the crisp possibility distributions $\pi_{y|x}$ (i.e., the $\{0, 1\}$ -valued ones) which satisfy (16), because we assume that there is a crisp relation between y and x which underlies the rule “the more x is A , the more y is B ”, then we obtain the constraint

$$\forall u \in U, \pi_{y|x}(v, u) \leq \begin{cases} 1 & \text{if } \mu_A(u) \leq \mu_B(v) \\ 0 & \text{if } \mu_A(u) > \mu_B(v) \end{cases} = \mu_{\{\mu_A(u), 1\}}(\mu_B(v)), \quad (20)$$

which expresses that the rule is now viewed as meaning: “ $\forall u \in U$, if $x = u$ then it is at least $\mu_A(u)$ -true that y is B ”, where the truth-qualification is understood in a crisp sense. In terms of α -level cuts it is equivalent to $\forall \alpha, x \in A_\alpha \Rightarrow y \in B_\alpha$ which clearly expresses that the greater the membership grade of x , i.e., the closer the value of x to typical values in A , the greater the membership grade of y and the closer the value of y to typical values in B . The implication $a \rightarrow b = 1$ if $a \leq b$ and $a \rightarrow b = 0$ if $a > b$ is called Rescher-Gaines implication in the fuzzy set literature. Then the fuzzy rule summarizes the collection of ordinary rules “if $x \in A_\alpha$ then $y \in B_\alpha$,” for $\alpha \in (0, 1]$. Indeed, it can be checked that $\pi_{y|x}$ is then upper bounded by the membership function of the crisp relation $R = \bigcap_{\alpha \in (0, 1]} \bar{A}_\alpha \cup B_\alpha$.

Observe a striking difference between (18) and (20): in (18), B is modified into a fuzzy set larger than B ($\mu_A(u) \rightarrow \mu_B(v) \geq \mu_B(v), \forall v$), while in (20) the conclusion of the rule when $x = u$ is a crisp subset which is not larger than B (nor included in B); it just focuses on the values of B with high membership grades.

3.2.3. Possibility rules

A first kind of conjunction-based fuzzy rule corresponds to statements of the form “the more x is A , the more possible B is a range for y ”. If we interpret this rule as “ $\forall u$, if $x = u$, it is at least $\mu_A(u)$ -possible that B is a range for y ”, a straightforward extension of (13), yields the following constraint on the conditional possibility distribution $\pi_{y|x}(\cdot, u)$ representing

the rule when $x = u$

$$\forall u \in U, \forall v \in V, \min(\mu_A(u), \mu_B(v)) \leq \pi_{y|x}(v, u) \quad (21)$$

Then the degree of possibility of the values in B is lower bounded by $\mu_A(u)$, and the values in B are considered as possible values for y at least at the degree $\mu_A(u)$. As expressed by (21), B is modified into a fuzzy set with membership function $\mu_\tau \circ \mu_B$ where $\mu_\tau(t) = \min(t, \alpha)$, $\forall t \in [0, 1]$ and $\alpha = \mu_A(u)$; τ can be seen as a special kind of fuzzy truth value and is pictured in Fig. 1(c). When A is not fuzzy, we recover $\mu_B(v) \leq \pi_{y|x}(v, u)$ if $u \in A$ and $0 \leq \pi_{y|x}(v, u)$ if $u \notin A$, i.e., the conjunction-based representation of a rule with a fuzzy conclusion part discussed in Section 2.2.

3.2.4. Anti-gradual rules

This is a second kind of conjunction-based fuzzy rules. The inequality (16) looks like (21) when exchanging $\mu_B(v)$ and $\pi_{y|x}(v, u)$, while (17), which is equivalent to (16), is analogous to (14) in the sense that in both cases $\pi_{y|x}$ is bounded from above by a multiple-valued logic implication function. It leads formally to considering the inequality constraint obtained from (14) by exchanging $\mu_B(v)$ and $\pi_{y|x}(v, u)$, i.e.,

$$\forall u \in U, \forall v \in V, \max(\pi_{y|x}(v, u), 1 - \mu_A(u)) \geq \mu_B(v). \quad (22)$$

This corresponds to a new kind of fuzzy rules, the intended meaning of which we now investigate. Eq. (22) is perhaps more easily understood by taking the complement to 1 of each side of the inequality, i.e.,

$$\forall u \in U, \forall v \in V,$$

$$1 - \mu_B(v) \geq \min(\mu_A(u), 1 - \pi_{y|x}(v, u)),$$

which can be interpreted as “the more x is A and the less y is related to x , the less y is B ”. Using the equivalences $\min(a, 1 - t) \leq 1 - b \Leftrightarrow 1 - t \leq a \rightarrow (1 - b) \Leftrightarrow t \geq 1 - (a \rightarrow (1 - b))$, where \rightarrow is Gödel’s implication, we can still write (22) in the form

$$\forall u \in U, \forall v \in V,$$

$$\pi_{y|x}(v, u) \geq \begin{cases} 0 & \text{if } \mu_A(u) + \mu_B(v) \leq 1 \\ \mu_B(v) & \text{if } \mu_A(u) + \mu_B(v) > 1 \end{cases}$$

$$\begin{aligned} &= \min(\mu_{(1 - \mu_A(u), 1]}(\mu_B(v)), \mu_B(v)) \\ &= \mu_{(1 - \mu_A(u), 1] \cap T}(\mu_B(v)). \end{aligned} \quad (23)$$

Unsurprisingly, the lower bound of $\pi_{y|x}(v, u)$ which is obtained, is a multiple-valued logic conjunction function of $\mu_A(u)$ and $\mu_B(v)$ (indeed $f(a, b) = 0$ if $a + b \leq 1$ and $f(a, b) = b$ otherwise, is such that $f(0, 0) = f(0, 1) = f(1, 0) = 0$ and $f(1, 1) = 1$). Moreover (23), when A is not fuzzy, reduces to the conjunction-based model of a rule with a fuzzy conclusion part. From (23) we see that this type of rules can be interpreted in the following way: the values whose degree of membership to B is not more than $1 - \mu_A(u)$ do not belong to the guaranteed possible values of y according to the rule. The rule could be understood, in terms of fuzzy truth values in the sense of (19), as “if $x = u$ then it is at least $(1 - \mu_A(u))$ -true that y is B ” (however in a way different from the fuzzy truth value pictured in Fig. 1(a)), or if we prefer “if $x = u$ then it is at most $\mu_A(u)$ -false that y is B ”. The membership function of the corresponding fuzzy truth-value which modifies B in (23) is given in Fig. 1(d). As it can be observed by comparing Figs. 1(a) and 1(d), they correspond to two points of view in (fuzzy) truth-qualification of level α , one leaving completely possible the values corresponding to degrees of truth greater than α (enlarging the core), the other assigning no possibility to the values corresponding to the degrees of truth less than or equal to $1 - \alpha$ (reducing the support). This agrees with the last kind of rules described in Section 3.1. Indeed (23) expresses that the values which sufficiently belong to B are guaranteed as possible for y (in relation to their level of belonging to B).

Let us now consider the crisp solution to (22). Taking x in the α -cut of A , i.e., $x \in \{u, \mu_A(u) \geq \alpha\}$, it expresses that the values v such that $\mu_B(v) \leq 1 - \alpha$ have no guaranteed possibility ($\pi_{y|x}(v, u) \geq 0$) while other values are fully possible. Thus if $x \in A_\alpha$ it corresponds to saying that the pairs (x, y) with guaranteed possibility 1 are of the form

$$(x, y) \in A_\alpha \times B_{\frac{1-\alpha}{1-x}},$$

where $B_{\frac{1-\alpha}{1-x}}$ denotes the strict $(1 - \alpha)$ -cut of B , i.e., $\{v, \mu_B(v) > 1 - \alpha\}$. Thus the larger $\mu_A(x)$, the larger the subset of values for y guaranteed as possible.

More generally, with (23), the larger $\mu_A(x)$ the larger the subset of values for y , which are guaranteed to be somewhat possible. Thus, these rules may be called “anti-gradual rules”, since gradual rules are such that “the larger $\mu_A(x)$, the smaller the core of the subset of possible values for y ”.

Let **comp** and **ant** be two transformations on the truth-scale reflecting the ideas of complementation and antonymy respectively, and defined by **comp**($f(t)$) = $1 - f(t)$ and **ant**($f(t)$) = $f(1 - t)$, $\forall t \in [0, 1]$ and f ranging in $[0, 1]$. Note that **comp** ◦ **ant** = **ant** ◦ **comp**, and that **comp** and **ant** are involutive.

Suppose f is the membership function of “at least α -true” as in Fig. 1(a) then **comp**(at least α -true) = at most α -true, and **ant**(at least α -true) = at least α -false = at most $(1 - \alpha)$ -true.² If f is the membership function of “at least α -certainly true” then **comp**(at least α -certainly true) = at least α -possibly false and **ant**(at least α -certainly true) = at least α -certainly false. As a consequence, it can be seen on Fig. 1 and formally checked that the four fuzzy truth-values we have introduced satisfy the two duality relations

at least α -certainly true = **comp** ◦ **ant** (at least α -possibly true), (24)

at least α -true = **comp** ◦ **ant** (at most α -false). (25)

Note that when there are only two degrees of truth, 0 (false) and 1 (true), “at least α -certainly true” corresponds to the possibility distribution $\pi(\text{true}) = 1$, $\pi(\text{false}) \leq 1 - \alpha$ and “at least α -possibly true” to $\pi(\text{false}) \geq 0$ and $\pi(\text{true}) \geq \alpha$, while the two other (fuzzy) truth-values make no sense. Dually when there are only two degrees of possibility, 0 (complete impossibility) and 1 (complete possibility), then the representations of “at least α -true” (in the sense of Fig. 1(a)) and of “at most α -false” (in the sense of Fig. 1(d)) respectively coincide with the ordinary subsets of truth degrees $[\alpha, 1]$ and $(1 - \alpha, 1]$.

² Mind that the truth-value of Fig. 1(d) can be also read “at least $(1 - \alpha)$ -true” with an understanding of graded fuzzy truth different from the one in Fig. 1(a). In Fig. 1(a), the fuzzy truth value $T (\forall t, \mu_T(t) = t)$ is modified by increasing the high membership degrees to 1, while in Fig. 1(d) the low membership degrees are decreased to 0.

3.3. Inference with a fuzzy rule and a precise input

The four fuzzy truth-values pictured in Fig. 1 (with $\alpha = \mu_A(u)$) can be viewed as representing modifiers m (in the sense of Zadeh [33]) which modify the fuzzy set B into a fuzzy set B' such that $\mu_{B'} = m(\mu_B)$, in order to specify the subset of interest for y in the various rules for a precise input $x = u$; see Bouchon-Meunier [5] on the role of modifiers in approximate reasoning. Summarizing, in the case of

- *certainty rules*, the possibility distribution $\pi_{y|x}(\cdot, u)$ is bounded from above by $m(\mu_B)$ with $m(t) = \max(t, 1 - \mu_A(u))$, i.e., B is drowned in a level of indetermination $1 - \alpha$;
- *gradual rules*, the possibility distribution $\pi_{y|x}(\cdot, u)$ is bounded from above by $m(\mu_B)$ with $m(t) = \mu_A(u) \rightarrow t$ (where \rightarrow denotes Gödel’s implication), i.e., the core of B is enlarged;
- *possibility rules*, the possibility distribution $\pi_{y|x}(\cdot, u)$ is bounded from below by $m(\mu_B)$ with $m(t) = \min(\mu_A(u), t)$, i.e., B is truncated up to the height $\alpha = \mu_A(u)$;
- *anti-gradual rules*, the possibility distribution $\pi_{y|x}(\cdot, u)$ is bounded from below by $m(\mu_B)$ with $m(t) = 0$ if $\mu_A(u) + t \leq 1$ and $m(t) = t$ otherwise, i.e., the support of B is diminished, reduced.

Fig. 2 gives the output B' , corresponding to a precise input $x = u$, for the different meanings of a rule. It includes the above four types of rules, together with the “crisp” counterparts of gradual and anti-gradual rules (Fig. 2(c) and 2(g)). For these two types of rules, B' is an ordinary subset, and thus B' does not encode any uncertainty. These two types of rules take advantage of the gradualness of the membership functions of A and B in order to (partially) describe the relationship between x and y .

All the models of fuzzy rules which have been discussed until now in this paper only require purely ordinal scales for the assessment of the degrees of membership, possibility or certainty. Indeed, all the connective operators which have been used are based on min operation and on an order-reversing operation (i.e., the complementation to 1 when using the interval $[0, 1]$); more generally, negation is defined by reversing the scale, e.g., if the scale is finite: $e_0 < \dots < e_i < \dots < e_m$, the negation

operation n is defined by $n(e_i) = e_{m-i}$. See [13] for the definition of the connective operators different from min and the negation. If we have a scale with a “richer” structure, we may thus use the product

instead of min, it leads to use Goguen’s implication ($a \rightarrow b = 1$ if $a = 0$, $a \rightarrow b = \min(1, b/a)$ if $a \neq 0$; see Fig. 2(e)) in place of Gödel’s for instance; see [13] for the other operators related to product. If $\max(0,$

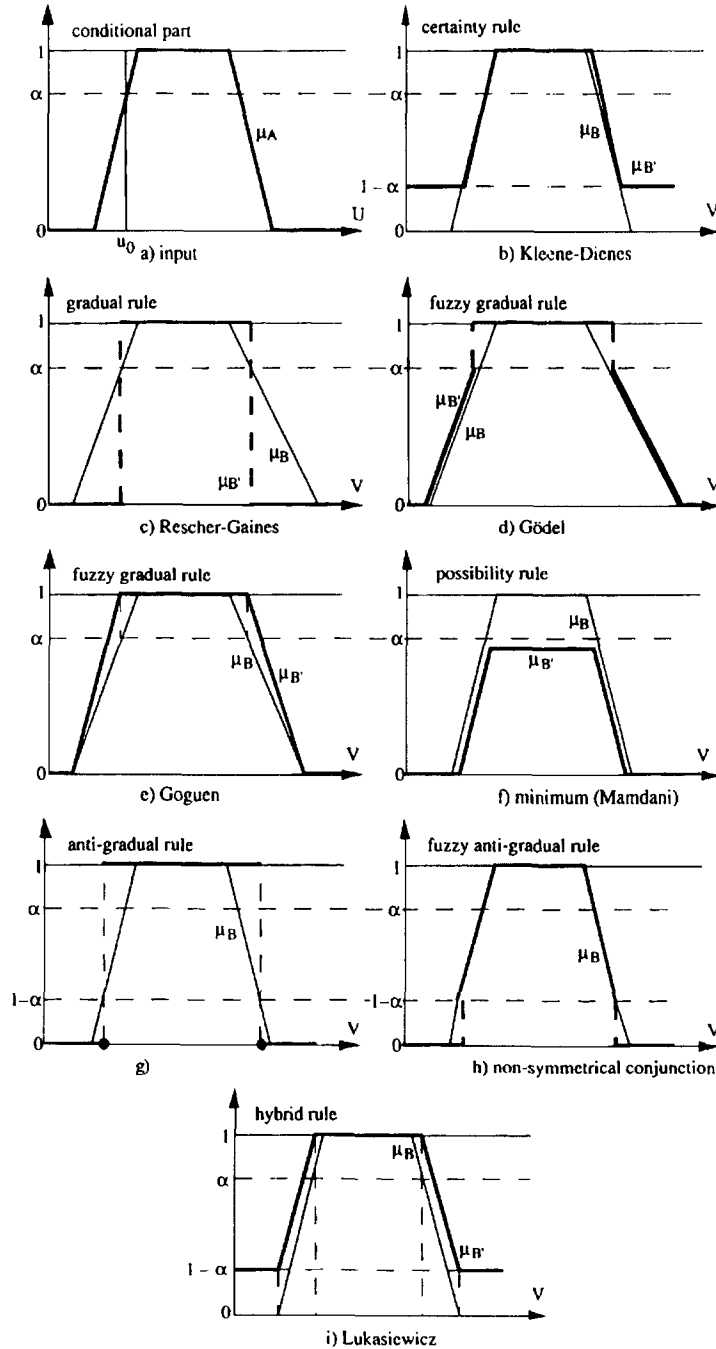


Fig. 2.

$a + b - 1$) is used as the basic conjunction, then there only exists one type of conjunction and one type of implication (see [13]). The associated implication is Lukasiewicz implication which combines the effects of Gödel-like and Kleene–Dienes-like implications, i.e., the simultaneous enlargement of the core and the introduction of a level of indetermination; see Fig. 2(i).

For the sake of brevity, we do not deal here with the effects of fuzzy inputs in the different cases; see, e.g., [16] on this question.

4. Fuzzy “if... then... unless...” rules

Fuzzy “if... then... unless...” rules of the form “if x is A then y is B unless z is C ” are often encountered in the expression of expert knowledge, since the unless part gives some proviso for exceptions.

Driankov and Hellendoorn [8] (see also [7]) have proposed an approach to the modelling of fuzzy “if... then... unless...” rules. These authors propose to model the rule “if x is A then y is B unless z is C ” as the conjunctive combination of the two rules “if x is A and z is not C then y is B ” and “if x is A and z is C then y is not B ”, each of the rules being represented by means of Gödel implication.

In the light of the different semantics that we can imagine for fuzzy “if... then...” rules, and which have been discussed in Sections 2 and 3, it is clear that there exist many types of fuzzy “if... then... unless...” rules. For each type of fuzzy “if... then...” rule we described, we can imagine the associated fuzzy “if... then... unless...” rule, and what is the intended meaning of having C fuzzy. In the following we briefly discuss only a few types of such rules. Also the meaning of the unless part of the rule has to be made precise. Apart from “if x is A and z is not C then y is B ”, does the rule mean that “if x is A and z is C then y is not B ”, as suggested by [8] or only that “if x is A and z is C then y is unknown”? Moreover, do we expect to conclude that “ y is B when x is A if nothing is known on z ”?

Default reasoning, based on certainty rules, offers good examples of fuzzy “if... then... unless...” rules. Let us, for instance, consider the rule

“usually, birds fly unless they are very big”. Such a rule means that the less true a bird is very big, the more certain it flies. This understanding of the rule leads to a model of the form

$$\pi_{y|x,y}(v, u, w) \leq \max(\mu_B(v), \mu_C(w), 1 - \mu_A(u)), \quad (26)$$

for the rule “if x is A then y is B unless z is C ” (with $A = \text{bird}$, $B = \text{flying}$, $C = \text{very big}$). In our example, A is not fuzzy, but B may include some uncertainty pervading the conclusion (to take into account the “usually” in the above rule). As can be seen in (26), the level of uncertainty pervading the conclusion of the rule when $\mu_A(u) = 1$, is the maximum of the one, λ , associated with B (if B is of the form $\mu_B = \max(\mu_{B^*}, \lambda)$) and of the degree $\mu_C(z)$ which estimates to what extent the bird is very big in our example. When z is precisely known, $z = w$, and $\mu_C(w) = 0$, (26) reduces to the standard certainty rule as expected. When $\mu_C(w) = 1$, nothing can be concluded about y ; this is also the case if z is unknown (except if we add a default information of the type “generally z is not C ”, as already discussed in Section 2.2). If we want to conclude also that y is not B when z is C , we add the other constraint

$$\pi_{y|x,z}(v, u, w) \leq \max(1 - \mu_{B^*}(v), 1 - \mu_C(w), 1 - \mu_A(u)). \quad (27)$$

Note that B^* is used in (27) instead of B in order to preserve the normalization assuming that the support of B^* is strictly included in V (if $\mu_B = \max(\mu_{B^*}, \lambda)$). By combining (26) and (27), and assuming that B^* is not fuzzy and $\lambda = 0$ for simplicity, we get

$$\begin{aligned} \pi_{y|x,z}(v, u, w) \\ \leq \max(1 - \mu_A(u), \min[\max(\mu_{B^*}(v), \mu_C(w)), \\ \max(1 - \mu_{B^*}(v), 1 - \mu_C(w))]). \end{aligned} \quad (28)$$

As it can be seen, when $\mu_A(u) = 1$, we obtain a *non-normalized* upper bound for $\pi_{y|x,z}$ if $z = w$ is such that $0 < \mu_C(w) < 1$. This is due to the uncertainty pervading the situation. In our example, we cannot say clearly if the considered bird is very big or not; it is somewhat in between. If $\mu_A(u) = 1$ and $\mu_C(w) = 1 - \mu_C(w) = 1/2$, (28) yields $\pi_{y|x}(v, u) \leq 1/2, \forall v$, which expresses a conflicting situation. When $\mu_C(w) < 1/2$, B^* emerges with a somewhat

greater level of possibility, while if $\mu_C(w) > 1/2$, it is its complement which is emerging (this could be emphasized by normalizing the result in each case). The denormalization of the result should not be surprising. It expresses a partial inconsistency induced by the uncertainty about the specific subclass to which the situation belongs. This type of problem is encountered in nonmonotonic reasoning where the existence in the knowledge base of conditional rules referring to more or less specific contexts and having conflicting conclusions, creates a partial inconsistency of the possibilistic logic representation of the knowledge base [11]. Then a special inference machinery is needed to get the expected conclusions; see [3]. This machinery needs to be extended in order to allow for possible uncertainty about the context to which the current situation belongs, when this situation is described in terms of vague predicates. Note that we have a similar denormalization of the result if (27) is combined with $\pi_{y|x} \leq \max(\mu_B, 1 - \mu_A)$, representing the rule “usually birds flies” in our example, instead of being combined with (26), as done in (28).

Clearly, fuzzy unless rules are not only connected with certainty rules. We can for instance, imagine a rule of the type “the more x is A , the more y is B , unless z is C ”. Such a rule can be easily understood and represented when C is an ordinary subset: the unless part restricts the rule to the situation where $z \notin C$; then if $z \in C$, the relation between x and y may be completely different. When C is fuzzy, our understanding of the rule may be, for instance, “the more z is C , the more certain it is that the more x is A , the more y is B ”, which can be clearly modelled by the constraint

$$\pi_{y|x,z}(v, u, w) \leq \max(\mu_{A \rightarrow B}(u, v), 1 - \mu_C(w)), \quad (29)$$

where $A \rightarrow B$ is modelled by Gödel or by Rescher–Gaines implication. This expresses that the smaller $\mu_C(z)$, the less certain we are that the gradual rule holds.

Possibility rules may also have an unless part, like in the rule “the more is A , the more possible y is B unless z is C ”. If C is an ordinary subset the semantics is clear. If C is fuzzy, several interpretations can be thought of. For instance, it might mean that “the more x is A and the less z is C , the more

possible y is B ”, which leads to the representation

$$\pi_{y|x,z}(v, u, w) \geq \min(\mu_A(u), 1 - \mu_C(w), \mu_B(v)). \quad (30)$$

A complete investigation of all the types of fuzzy unless rules is beyond the purpose of this paper. We only emphasize that their proper modelling requires the proper modelling of simpler if/then rules and the precise understanding of their intended meaning.

5. Coherence of parallel fuzzy rules

Fuzzy rules are often used in parallel (by parallel we mean that they relate the same variables, i.e., the same variable(s) x_1, \dots, x_n appear(s) in all the condition parts of the rules, while all the conclusions pertain to the same variable, say y). Then an important issue is then to make sure that a set of fuzzy rules is coherent. By coherent, it is meant that for any input the result of the system of rules is a normalized fuzzy set, i.e., at least one value remains completely possible for the variable in the conclusion part whatever the input is. Since conjunction-based fuzzy rules represented by lower bounds of a possibility distribution are combined disjunctively nothing in the constraints representing the rules can forbid the possibility distribution to be normalized for any input. Thus, an incoherence in the system of rules cannot take place with conjunction-based representations³. By contrast, implication-based rules can be incoherent, since after combination by means of the min operation, the resulting upper bound of the possibility distribution representing the collection of rules can be no longer normalized for some input value. In the following, we briefly state some results presented in [20] on the coherence of parallel implication-based rules. More details are given in [31].

Consider first a set of two crisp rules “if x is A_i then y is B_i ”, $i = 1, 2$. If $x \notin A_i$ for some i then rule i enforces nothing on the value of y . Hence it is always possible to assign a value to y . The only

³ A form of incoherence might only appear if after defuzzification of a nonconvex result (obtained by disjunctive combination), the precise conclusion, thus obtained, does not belong to the support of the non-convex fuzzy set which is defuzzified.

situation that may lead to an incoherence is when $x \in A_1 \cap A_2$. Then due to the rules, $y \in B_1 \cap B_2$, and this is impossible if and only if $B_1 \cap B_2 = \emptyset$. Consequently two crisp rules are coherent if and only if

$$A_1 \cap A_2 \neq \emptyset \text{ implies } B_1 \cap B_2 \neq \emptyset. \quad (31)$$

This is clearly equivalent to

$$\forall u, \exists v, \min(\max(1 - \mu_{A_1}(u), \mu_{B_1}(v)), \max(1 - \mu_{A_2}(u), \mu_{B_2}(v))) = 1. \quad (32)$$

When n rules are considered, the extension of (32) is straightforward, while (31) must be written for all subsets of the n rules whose condition parts overlap.

Let us consider the case of certainty rules. The coherence of two certainty rules is defined by (32) where the membership functions are valued on $[0, 1]$. Indeed (32) expresses that the projection on U of the combination of the two rules does not induce any restriction on the variable x , since the degree of possibility of any possible value $u \in U$ of x remains equal to 1. It is easy to see that (32) holds for any pair of rules and any $u \notin \text{support}(A_1) \cap \text{support}(A_2)$ for a suitable choice of v in $\text{core}(B_1)$ or $\text{core}(B_2)$. When $u \in \text{support}(A_1) \cap \text{support}(A_2)$ then the only way to satisfy (32) is to select v such that $\mu_{B_1}(v) = \mu_{B_2}(v) = 1$, i.e., $v \in \text{core}(B_1) \cap \text{core}(B_2)$. It means that the characteristic condition for the coherence of two crisp rules (31) is easily adapted to certainty rules as follows: two certainty rules are coherent if and only if

$$\begin{aligned} \text{support}(A_1) \cap \text{support}(A_2) \neq \emptyset \text{ implies} \\ \text{core}(B_1) \cap \text{core}(B_2) \neq \emptyset. \end{aligned} \quad (33)$$

In other words, we are back to a coherence problem in classical logic. This result generalizes to subsets of rules with coherent condition parts [20].

Coherence condition (32) writes for a knowledge base made of two *gradual* rules

$$\forall u, \exists v, \min(\mu_{A_1}(u) \rightarrow \mu_{B_1}(v), \mu_{A_2}(v) \rightarrow \mu_{B_2}(v)) = 1, \quad (34)$$

with $\mu_{A_i}(u) \rightarrow \mu_{B_i}(v) = 1$ if $\mu_{A_i}(u) \leq \mu_{B_i}(v)$ and $\mu_{A_i}(u) \rightarrow \mu_{B_i}(v) = \mu_{B_i}(v)$ otherwise. Thus (34) is

equivalent to

$$\begin{aligned} \forall u \in \text{support}(A_1) \cap \text{support}(A_2), \exists v, \\ \mu_{A_1}(u) \leq \mu_{B_1}(v) \text{ and } \mu_{A_2}(u) \leq \mu_{B_2}(v). \end{aligned} \quad (35)$$

This is still equivalent to

$$\forall \alpha \in (0, 1], (A_1)_\alpha \cap (A_2)_\alpha \neq \emptyset \Rightarrow (B_1)_\alpha \cap (B_2)_\alpha \neq \emptyset, \quad (36)$$

where A_α, B_α denote the α -cuts of the fuzzy sets A and B ($A_\alpha = \{u, \mu_A(u) \geq \alpha\}$). These coherence conditions have been already pointed out in [12].

An easy-to-establish consequence of (35) and (36) is

$$\text{hgt}(A_1 \cap A_2) \leq \text{hgt}(B_1 \cap B_2). \quad (37)$$

This is a necessary condition for coherence. However it is not sufficient as shown on Fig. 3.

In [20], it is shown that in the case of triangular membership functions (37) becomes a necessary and sufficient condition, when each fuzzy set only overlaps with its immediate neighbours and the slopes of the decreasing and increasing parts of the A_i 's are proportional to the ones of the corresponding parts of the B_i 's. Namely let a_i denote the peak of A_i and $[\underline{a}_i, \bar{a}_i]$ its support. If b_i, \underline{b}_i and \bar{b}_i are defined likewise then the condition on slopes is

$$\frac{a_1 - \bar{a}_1}{b_1 - \bar{b}_1} = \frac{a_2 - \underline{a}_2}{b_2 - \underline{b}_2}$$

This encompasses the case where all the A_i 's are symmetrical and identical up to a translation and the same condition holds for the B_i 's. A well-known particular case of coherence is encountered in practice in fuzzy control when triangular membership functions such that $\forall i = 1, n - 1, \forall u, \mu_{A_i}(u) + \mu_{A_{i+1}}(u) = 1$ (which entails $\text{hgt}(A_i \cap A_{i+1}) = 1/2$) are used on the domain of U and the B_i 's make a similar fuzzy partition. Practical methods for testing the coherence of a set of gradual rules are under study [31]; especially some results are given to cut down the combinatorial complexity in the case of rules whose conditions are defined on real intervals.

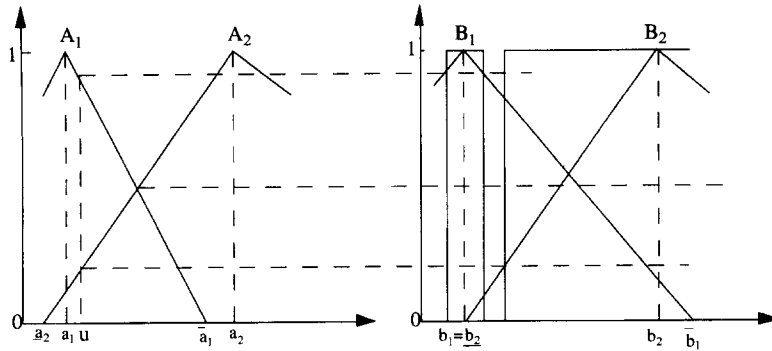


Fig. 3. Condition (33) does not hold: $\{v | \mu_{B_1}(v) \geq \mu_{A_1}(u)\} \cap \{v | \mu_{B_2}(v) \geq \mu_{A_2}(u)\} = \emptyset$

6. Concluding remarks

The representation of the intended meaning of fuzzy rules has received little attention until now in spite of the enormous quantity of existing literature on approximate reasoning and fuzzy controllers. The paper has recalled different kinds of fuzzy rules based on very simple semantical considerations. Four types of fuzzy rules emerge corresponding to very standard alterations of a possibility distribution: enlarging its core, shrinking its support, truncating its height or, drowning it in a uniform level of uncertainty. We can also distinguish between fuzzy rules where some gradual uncertainty is expressed and pure gradual rules where fuzzy sets are used in the rules only for describing a local relation between variables by means of gradual properties (as in rules of the type “the more x is A , the more y is B ”); then in this latter case a precise input yields a nonfuzzy conclusion.

The proposed semantics systematically exploits the gradedness of properties and/or modalities expressing certainty or mere possibility. Implication-based and conjunction-based representations have been contrasted from the points of view of their meanings, of the underlying combination mode and of coherence analysis. An open question is to decide what kind of fuzzy rule is best suited to fuzzy control and related tasks. The almost systematic use of conjunction-based representations, since Mamdani’s early works, including Sugeno’s [29] fuzzy rules with precise conclusions, is surprising from the point of view of logic: conjunction-

based rules do not fit with the usual meaning of rules in expert systems. One possible explanation may be that systems engineers and knowledge-engineers do not have the same background. The former are usually data-driven, are used to identification methods where models derive from data. They also view fuzzy rules as data, and their reasoning methodology is based on accumulating data. Incoherence issues are not considered because incoherence is unavoidable in a set of data. On the contrary knowledge engineers are knowledge-driven, and a fuzzy rule is a piece of knowledge that constrains a set of possible situations. The more fuzzy rules, the more information, and the more precise one can get. This methodology that proceeds by successive refinements clearly does not work if a set of rules is not coherent. A lot remains to be done in order to reconcile the data-driven and the knowledge-driven approach to information technology. Fuzzy rules clearly stand at the cross-road of the two traditions.

A proper representation of the semantic contents of expert rules is clearly a basic issue in approximate reasoning and looks important for a proper use of fuzzy rules in applications. Also, a proper representation of the available pieces of knowledge is a requirement for analyzing a knowledge base made of fuzzy rules, and for validating it. Gradual rules, possibility rules, and certainty rules are very different in their intended use and should not be confused. Gradual rules are oriented towards the interpolation and the extrapolation of conclusions since the fuzzy sets in their condition parts then

encode an idea of similarity, viewing such a rule as expressing that “the closer x is to the core of A , the closer y is to the core of B ”. Pure gradual rules provide a partial but nonfuzzy description of the relation between variables x and y (since $A \rightarrow B$ is a crisp subset of $U \times V$, although A and B are fuzzy for expressing gradedness). Gradual rules can approximate functions; see [10]. Anti-gradual rules express that the set of *feasible* values for y is all the larger as the condition part of the rule is better satisfied (this does not exclude that other values might be feasible as well if it is asserted by other rules); anti-gradual rules contrast with gradual rules which restrict the set of possible values for y all the more as $\mu_A(x)$ is greater. Possibility rules also provide weighted conclusions which can be accumulated disjunctively. This agrees with forms of reasoning like analogical reasoning or abductive reasoning where distinct pieces of acceptable, plausible conclusions are gathered, such as possible actions in a current situation (which resembles situations described in the rules), or suggested explanations for a given set of manifestations.

Certainty rules are more in the spirit of deductive reasoning since their conclusions are combined conjunctively, even if they are pervaded with uncertainty. Certainty rules with fuzzy condition parts enable us to relate typicality and certainty, by interpreting “the more x is A ” as “the more typical x is”, as in the example “the more typical a bird, the more certain it flies”. In such a case, typicality can be appreciated in terms of the weight of the bird, the length of the wings and so on. So, such a rule requires a rather precise information about the considered bird in order to get a useful conclusion. This differs from the situation dealt with by generic rules having exceptions as in the classical example “birds fly”, “penguins are birds” and “penguins do not fly”; there, the “fuzzy” class of birds which fly is not explicitly defined. However, some exceptions which do not belong to this class are encountered in other rules of the knowledge base (here, the penguins). See [3] for a treatment of rules having exceptions in the framework of possibility theory: the rule “generally, if p then q ” is then understood as “the possibility of having p and q true is strictly greater than the possibility of having p and not q true”. In this latter problem, rules are not fuzzy by

themselves and the fuzzy set of interpretations compatible with a set of default rules is defined on a referential which is not ordered in general. Yet, a set of rules with exceptions leads to considering that some situations are more plausible than others, i.e., induce a nonbinary possibility distribution. Fuzziness is thus attached to the whole knowledge base. On the contrary, fuzzy rules with fuzzy condition parts are usually defined on referentials which are totally ordered, such as numerical universes (even if it is not compulsory) and fuzzy sets are local, attached to predicates. Fuzzy rules thus play an important role at the interface between reasoning with symbolic categories and the handling of numerical data.

References

- [1] J.F. Baldwin, A new approach to approximate reasoning using a fuzzy logic, *Fuzzy Sets and Systems* **2** (1979) 309–325.
- [2] J.F. Baldwin and N.C.F. Guild, Modelling controllers using fuzzy relations, *Kybernetes*, **9** (1980) 223–229.
- [3] S. Benferhat, D. Dubois and H. Prade, Representing default rules in possibilistic logic, in: B. Nebel, C. Rich, W. Swartout, Eds., *Proc. 3rd Internat. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)* Cambridge, MA, (1992) 673–684.
- [4] B. Bouchon, Fuzzy inferences and conditional possibility distributions, *Fuzzy Sets and Systems*, **23** (1987) 33–41.
- [5] B. Bouchon-Meunier, *La Logique Floue*. (Presses Universitaires de France, Paris, 1993).
- [6] B. Bouchon-Meunier and S. Després, Acquisition numérique/symbolique de connaissances graduelles, *Actes des 3èmes Journées Nationales du PRC-GDR Intelligence Artificielle, Paris* (5–7 mars 1990) 127–138.
- [7] P. Doherty, D. Driankov and H. Hellendoorn, (1992) Fuzzy if-then-unless rules and their implementation, Research Report LiTH-IDA-R92–21, Dept. of Computer & Inform. Science, Linköping Univ., 581 83 Linköping, Sweden.
- [8] D. Driankov and H. Hellendoorn, (1992) Fuzzy logic with unless-rules. Tech. Report IDA-RKL-92-TR50, Dept. of Computer & Inform. Science, Linköping Univ., 581 83 Linköping, Sweden.
- [9] D. Driankov, H. Hellendoorn and M. Reinfrank, *An Introduction to Fuzzy Control*. (Springer, Berlin, 1993).
- [10] D. Dubois, M. Grabisch and H. Prade, Gradual rules and the approximation of control laws, in: H.T. Nguyen, M. Sugeno, R. Tong, R. Yager, Eds., *Theoretical Aspects of Fuzzy Control* (Wiley, New York, (1995) 147–181.
- [11] D. Dubois, J. Lang and H. Prade, Automated reasoning using possibilistic logic: Semantics, belief revision, and variable certainty weights, *IEEE Trans. Data Knowledge Eng.*, **6** (1994) 64–71.

- [12] D. Dubois, R. Martin-Clouaire and H. Prade, Practical computing in fuzzy logic, in: M.M. Gupta, T. Yamakawa, Eds., *Fuzzy computing* (North-Holland, Amsterdam, 1988) 11–34.
- [13] D. Dubois and H. Prade, A theorem on implication functions defined from triangular norms, *Stochastica* **8**(3) (1984) 267–279.
- [14] D. Dubois and H. Prade, On the combination of uncertain or imprecise pieces of information in rule-based systems – A discussion in the framework of possibility theory, *Internat. J. of Approx. Reason.* **2**(1) (1988) 65–87.
- [15] D. Dubois and H. Prade, A typology of fuzzy “if... then...” rules. *Proc. 3rd Internat Fuzzy Systems Assoc. (IFSA) Congress* (Seattle, WA 6–11 August 1989) 782–785.
- [16] D. Dubois and H. Prade, Fuzzy sets in approximate reasoning – Part 1: Inference with possibility distributions, *Fuzzy Sets and Systems*, **40** (1991) 143–202.
- [17] D. Dubois and H. Prade, (1991b) Basic issues on fuzzy rules and their application to fuzzy control, in: D. Driankov, P.W. Eklund and A.L. Ralescu, Eds., *Fuzzy Logic and Fuzzy Control, Proc. IJCAI’91 Workshops on Fuzzy Logic and Fuzzy Control*, August 1991) Lecture Notes in Artificial Intelligence, Vol. 833, (Springer, Berlin, 1994) 3–14.
- [18] D. Dubois and H. Prade, Fuzzy rules in knowledge-based systems – Modelling gradedness, uncertainty and preference, in: R.R. Yager and L.A. Zadeh, Eds., *An introduction to Fuzzy Logic Applications in Intelligent Systems* (Kluwer, Dordrecht, (1992) 45–68.
- [19] D. Dubois and H. Prade, Gradual inference rules in approximate reasoning, *Inform. Sci.* **61** (1992) 103–122.
- [20] D. Dubois and H. Prade, On the validation of fuzzy knowledge bases. in: S.G. Tzafestas and A.N. Venetsanopoulos, Rds., *Fuzzy Reasoning in Information, Decision and Control Systems* (Kluwer, Dordrecht, (1994) 31–49.
- [21] D. Dubois and H. Prade, Conditional objects as non-monotonic consequence relationships, *IEEE Trans. Systems, Man Cybernet.* **24**(12), (1994) 1724–1739.
- [22] I.R. Goodman, H.T. Nguyen and E.A. Walker, *Conditional Inference and Logic for Intelligent Systems – A Theory of Measure-Free Conditioning* (North-Holland, Amsterdam, 1991).
- [23] A. Kaufmann, *Introduction à la Théorie des Sous-Ensembles Flous – Vol. 1: Eléments Théoriques de Base* (Masson, Paris, 1973); English translation (Academic Press, New York, 1975).
- [24] A. Kaufmann, *Nouvelles Logiques pour l’Intelligence Artificielle* (Hermès, Paris, 1987).
- [25] A. Kaufmann, *Le Paramétrage des Moteurs d’Inference* (Hermès, Paris, 1988).
- [26] J.M. Mendel, Fuzzy logic systems for engineering: A tutorial, *Proc. IEEE* **83**(3), (1995) 345–377.
- [27] M. Mizumoto, H.J. Zimmermann, Comparison of fuzzy reasoning methods, *Fuzzy Sets and Systems*, **8**, (1982) 253–283.
- [28] E.H. Ruspini, On the semantics of fuzzy logic, *Internat. J. of Approx. Reason.* **5** (1991) 45–88.
- [29] M. Sugeno, An introductory survey of fuzzy control, *Inform. Sci.*, **36** (1985) 59–83.
- [30] E. Trillas and L. Valverde, On mode and implication in approximate reasoning, in: M.M. Gupta, A. Kandel, W. Bandler and J.B. Kiszka, Eds., *Approximate Reasoning in Expert Systems* (North-Holland, Amsterdam, 1985) 157–166.
- [31] L. Ughetto, D. Dubois and H. Prade, La cohérence des systèmes de règles graduées. Actes des 5èmes Journées Francophones “Logique Floue et Applications” (LFA’95), Paris, Nov. 27–28, Cepadue, Toulouse, 103–110.
- [32] S. Weber, Decomposable measures for conditional objects. (abstracts), in: E.P. Klement, S. Weber, Eds., *Proc. 10th Internat. Seminar on Fuzzy Set Theory – Measures and Integrals – Alternative Concepts* Linz, Austria (12–16 September, 1988) (Universitätsdirektion, J. Kepler Universität, 4040 Linz, Austria, 1988) 53–61.
- [33] L.A. Zadeh, A fuzzy set-theoretic interpretation of linguistic hedges, *J. Cybernetics*, **2**(3), (1972) 4–34.
- [34] L.A. Zadeh, PRUF – A meaning representation language for natural languages, *Internat. J. Man-Machine Studies*, **10** (1978) 395–460.
- [35] L.A. Zadeh, A theory of approximate reasoning, in: J.E. Hayes, D. Mitchie and L.I. Mikulich, Eds., *Machine Intelligence, Vol. 9* (Wiley, New York, 1979) 149–194.
- [36] L.A. Zadeh, The calculus of fuzzy if/then rules, *AI Expert* **7**(3) (1992) 23–27.
- [37] L.A. Zadeh, Soft computing and fuzzy logic, *IEEE Software* (November 1994) 48–56.