

Practical inference with systems of gradual implicative rules

Hazaël Jones, Brigitte Charnomordic, Didier Dubois, Serge Guillaume

▶ **To cite this version:**

Hazaël Jones, Brigitte Charnomordic, Didier Dubois, Serge Guillaume. Practical inference with systems of gradual implicative rules. IEEE Transactions on Fuzzy Systems, 2009, 17 (1), pp.61-78. 10.1109/TFUZZ.2008.2007851 . hal-00858551

HAL Id: hal-00858551
https://inria.hal.science/hal-00858551

Submitted on 15 May 2020

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.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

Practical inference with systems of gradual implicative rules

Hazaël Jones, Brigitte Charnomordic, Didier Dubois and Serge Guillaume
IEEE Transactions on Fuzzy Systems, 17(1):61-78, 2009

***Abstract*— A general approach to practical inference with gradual implicative rules and fuzzy inputs is presented. Gradual rules represent constraints restricting outputs of a fuzzy system for each input. They are tailored for interpolative reasoning. Our approach to inference relies on the use of inferential independence. It is based on fuzzy output computation under an interval-valued input. A double decomposition of fuzzy inputs is done in terms of α-cuts and in terms of a partitioning of these cuts according to areas where only a few rules apply. The case of one and two dimensional inputs is considered, as well as higher dimensional cases. An application to a cheese-making process illustrates the approach.**

I. Introduction

Fuzzy logic, as an interface between symbolic and numeric computations, is well-known for its ability to represent the graded nature of some non-Boolean linguistic concepts.

Historically, fuzzy inference systems were devised to perform a reasoning task based upon expert knowledge yielding a continuous numerical ouput, as needed in fuzzy control. Afterwards, many learning methods were added to enhance numerical performance.

Conjunctive rules used in the Mamdani-style fuzzy inference systems [1], represent joint sets of possible input and output values. They cannot be easily interpreted as generalizations of usual Boolean "if-then" statements in propositional logic, since the latter are modelled by material implication [2]. The weak logical behavior of conjunctive rules was pointed out by several authors like Baldwin and Guild [3] and Di Nola et al. [4]. Nevertheless, mainly due to alleged computational difficulties, fuzzy extensions of material implications have been neglected so far, if not simply rejected as proper tools for modeling fuzzy systems. For instance, Mendel [5] dismissed implicative fuzzy rules as being counterintuitive for engineers, and dubbed "engineering implications" the minimum or product operations, that are in fact generalized logical conjunctions.

However, inferring with parallel implicative rules and a precise input is not more computationally difficult than with fuzzy conjunctive rules (it can be done rule by rule). Moreover, it yields normalized fuzzy outputs often more precise than with conjunctive rules. Recently, we outlined several advantages of implicative rules with respect to conjunctive rules [6]. For instance, with conjunctive rules, the more rules in a rule base, the more imprecise its output becomes. This fact is usually hidden by defuzzification. The converse occurs with implicative rules. Their output is all the more precise as more rules are triggered. Furthermore, using conjunctive rules, the fuzzy output width can bias the defuzzified result. In constrast, gradual implicative rules [7] model constraints restricting output values for each input, and have interesting interpolation properties [7], [8]. They are fully compatible with the classical logic view. Among these kinds of rules, the most interesting ones for practical purposes use Goguen implication because of its continuous inference result [2], and Resher-Gaines implication if a non fuzzy (interval) output is needed [7]. Implicative rules are more natural to represent expert knowledge [9] as they model constraints relating input and output values.

In practical applications, fuzzy inputs are useful to account for sensor imprecision and approximate measurements. Furthermore in the case of cascaded fuzzy systems, it makes little sense to defuzzify the output to one system before feeding the next one, since it comes down to neglecting the meta-information concerning the imprecision of results (hence the validity of the eventually defuzzified overall output cannot be assessed).

Note that the recent blossoming of Type 2 fuzzy systems [10], was partly motivated by the need for accounting for higher order uncertainty in fuzzy systems outputs. Since the output of a fuzzy system is usually precise (either due to fuzzification or due to the use of the Takagi-Sugeno approach), this concern may look legitimate. But, arguably, the higher-order uncertainty is already present in the fuzzy output of a Type 1 fuzzy logic system, if rule conclusions are not precise, provided one refrains from defuzzifying it[1]. However, the fuzzy output of Mamdani systems is hard to interpret as often not normalized and with unreasonably wide support. On the contrary, the fuzzy output of consistent implicative fuzzy logic systems is a regular fuzzy interval (provided suitable fuzzy partitions of the input and output space are chosen). It can be summarized by a precise value if needed, and the higher order uncertainty of this value can be measured by some non-specificity index of the fuzzy output.

Hazaël Jones is with INRA and Cemagref, UMR ITAP, UMR ASB, place Pierre Viala 34060 Montpellier Cedex 1 FRANCE (phone: +33 499612070; email: jones@supagro.inra.fr).

Brigitte Charnomordic is with INRA, UMR Analyse des Systèmes et Biométrie 2, place Pierre Viala 34060 Montpellier Cedex 1 FRANCE (phone: +33 499612416; email: bch@supagro.inra.fr).

Didier Dubois is with IRIT, CNRS & Université de Toulouse 118 Route de Narbonne 31062 Toulouse Cedex 09 (phone: +33 561556331; email: dubois@irit.fr).

Serge Guillaume is with Cemagref UMR ITAP 361, rue JF Breton 34196 Montpellier Cedex 5 FRANCE (phone: +33 467046317; email: serge.guillaume@montpellier.cemagref.fr).

[1]The term "defuzzification" to designate the extraction of a precise value from a fuzzy set is a language abuse, as strictly speaking, stripping a fuzzy set from its fuzziness should yield a crisp *set*, not an element thereof.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

Moreover the imprecision produced by a set of implicative rules is rather limited when the rules are informative enough, which enables cascading.

Nevertheless, the practical use of parallel implicative rules with a fuzzy input is difficult, as the inference can no longer be done rule by rule. The aim of this article is to show that under some conditions on input partitions, inference becomes easier due to a double decomposition of the fuzzy input: by α-cut and by partitioning. In the sequel, section II recalls features of conjunctive and implicative rules and compares them according to some expected properties. In section III, we present sufficient conditions to obtain inferential independence, so as to facilitate the calculation of the inference process. Then, in section IV, exact analytical expressions are given for one dimensional systems. In section V we propose a fuzzy input decomposition method based on inferential independence that allows to simplify the inference mechanism, and apply it to the one dimensional case. Section VI adresses the two dimensional case, and section VII outlines an approach to the complex case of higher dimensional systems.

A practical application to the predictive diagnosis of a cheese-making process is outlined in section VIII to illustrate the technique, and some general conclusions are given in section IX.

II. Fuzzy rules: conjunction vs implication

Before examining the semantics of fuzzy rules, let us first recall what is the meaning of a rule in classical logic, i.e. a crisp rule. A crisp rule "If X is A then Z is O" relates two universes of discourse U and W that form the domains of variables X and Z respectively, locally restricting the domains of X and Z to subsets A of U and O of W. Such a rule can be interpreted in two ways according to whether one focuses on its examples or its counterexamples [11]. The examples of the rule precisely form the set of pairs $(u, w) \in A \times O$. Modeled as such, a rule cannot be understood as a constraint since $A \times O$ does not encompass all admissible pairs (u, w) relating U and W. Indeed, the rule does not prevent X from lying outside A. So the rule cannot be understood as the necessity to let $(X, Z) \in A \times O$; it only points out $A \times O$ as one set of explicitly allowed pairs for (X, Z).

On the contrary, the counterexamples of the rules are the set of pairs (u, w) such that $u \in A, w \notin O$. The Cartesian product $A \times O^c$, where O^c is the complement of O, is the set of pairs of values explicitly forbidden by the rule. It means that the set of implicitly allowed pairs of values form the set $(A \times O^c)^c = A^c \cup O = (A^c \times W) \cup (A \times O)$ corresponding to a material implication. This is the usual representation of rules in classical logic. Clearly, to the set $A \times O$ of examples, it adds the set $(A^c \times W)$ of pairs of values uncommitted by the rule. Since a rule refers to both examples and counterexamples, the complete representation of the rule is the pair $(A \times O, A^c \cup O)$ made of explicitly and implicitly permitted values (u, w).

In the case of fuzzy rules A and O are fuzzy sets, and the two fuzzy sets $A \times O$ and $A^c \cup O$ are modelled using fuzzy connectives of conjunction and implication, respectively:

$$\mu_A(u) \wedge \mu_O(w); \quad (1)$$

$$\mu_A(u) \rightarrow \mu_O(w). \quad (2)$$

First we will present commonly used rules: conjunctive rules. Then implicative rules will be described. An interpretation in terms of logic will be given and a comparison will be made according to several properties.

A. Conjunctive Fuzzy Rules

In contrast to logic representations, the most popular representation of fuzzy rules is the Cartesian product of the fuzzy condition and the fuzzy conclusion, following the approach of Mamdani. These rules may have a simple interpretation in terms of guaranteed possibility distributions [2]. For a given variable X, a guaranteed possibility distribution δ_X is associated to statements of the form "$X \in A$ is possible":

$$\forall u \in U, \delta_X(u) \geq \mu_A(u).$$

The statement "$X \in A$ is possible" only means that values in A are possible to some degree. $\delta_X(u) = 1$ indicates that $X = u$ is an actual situation, an observed value. $\delta_X(u) = 0$ indicates no evidence in favor of $X = u$ has been collected yet. It does not forbid situations where the statement is false. δ_X is a lower possibility distribution. Note that this interpretation is at odds with classical logic where asserting a proposition p explicitly forbids situations where p is false.

Conjunctive rules "if X is A then Z is O", can be understood as: "the more X is A, the more possible it is that Z lies in O" [2]. In this approach, the operator "then" is modeled by a conjunction and the rule output is a guaranteed possibility distribution: $\delta_{Z|X} = \mu_A \wedge \mu_O$. The traditional Mamdani conjunction operator is the $\min$.

$\forall u \in U, \forall w \in W, \delta_{Z|X}(u, w)$ can be interpreted as follows: when X is A to some degree, "Z is O" is possible at least to level $\min(\mu_A(u), \mu_O(w))$.

If we consider a precise input u_0 and if $\mu_A(u_0) = \alpha$ with $\alpha \in [0, 1]$, values in O are guaranteed at degree α. So the output O' is given by the truncation of O at level α as shown on figure 1.

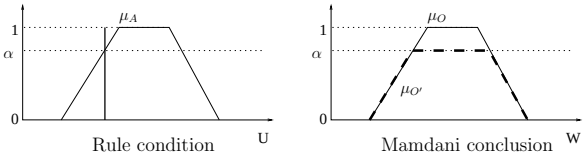


Fig. 1. Inference with Mamdani rules

In a knowledge base $K = \{A_i \times O_i, i = 1, ..., n\}$ of n parallel fuzzy rules (having the same input space U and the same output space W), rule aggregation is disjunctive. As a rule suggests outputs with a guaranteed possibility degree, when two or more rules are fired, all the corresponding

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

outputs are guaranteed, each one at least to level $\delta^i_{Z|X}, \forall i$. The final possibility distribution will then be:

$$\delta_K \geq \max_{i=1,\ldots,n} \delta^i_{Z|X} \tag{3}$$

The maximum represents a lower bound of possibility degrees. Clearly, $\delta_{Z|X}(u,w) = 0$ means that if $X = u$, no rule can guarantee that w is a possible value for Z. Ignorance is then represented by a null distribution: $\delta_{Z|X}(u,w) = 0, \forall w$.

B. Implicative Fuzzy Rules

The interpretation of implicative rules is based on a straightforward application of Zadeh's theory of approximate reasoning [12]. According to Zadeh, each piece of knowledge can be considered as a fuzzy restriction on a set of possible worlds. It extends the conventions of classical logic.

The statement " X is A_i " can be depicted as:

$$\forall u \in U, \pi_X(u) \leq \mu_{A_i}(u) \tag{4}$$

where $\pi_X(u)$ is a (potential) possibility distribution. " X is A_i " now means: " X must be in A_i ", it represents a constraint, i.e., negative information in the sense that it points out forbidden values.

In view of the above discussions, the two possibility distributions $\delta_{Z|X}$ and $\pi_{Z|X}$ have very different semantics: degrees of possibility expressed by $\pi_{Z|X}$ are potential : $\pi_{Z|X}(u,w) = 1$ means that nothing forbids (u,w) from further consideration, while $\pi_{Z|X}(u,w) = 0$ means that (u,w) is forbidden by the rule.

The difference of nature between conjunctive and implicative rules has impact when combining several rules together: while several conjunctive rules are combined disjunctively (as they point to more examples than a single rule), implicative rules are combined conjunctively, because several constraints lead to a more restricted feasible set of allowed situations than a single constraint:

$$\pi_{Z|X}(u,w) = \min_{i=1,\ldots,n} \pi^i_{Z|X} = \min_{i=1,\ldots,n} (\mu_{A_i}(u) \rightarrow \mu_{O_i}(w)) \tag{5}$$

Rule aggregation is conjunctive because the possibility in the sense of (4) is not guaranteed: a value estimated as possible by a rule can be forbidden by other rules.

There are different kinds of implicative rules: certainty rules and gradual rules [2]. In this article, we only focus on gradual rules. The behavior of gradual implicative rules, "the more X is A, then the more Z is O", depends on the selected implication. We consider in this paper the following residuated implications :

- Resher-Gaines: $a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b \\ 0 & \text{otherwise} \end{cases}$
- Gödel: $a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b \\ b & \text{otherwise} \end{cases}$
- Goguen: $a \rightarrow b = \begin{cases} \min(1, b/a) & \text{if } a \neq 0 \\ 1 & \text{otherwise} \end{cases}$

Figure 2 clearly shows that under a precise input u_0 the resulting output affects the shape of the conclusion part while maintaining the output values within the support of the rule conclusion. In all cases the core of the output gets larger as the input membership value decreases, thus relaxing the constraint expressed in the rule conclusion at level 1. In the case of Goguen implication, the output membership function remains continuous, if $\mu_A(u_0) > 0$, while Gödel implication almost always results in a discontinuous output. In the case of Resher-Gaines implication, the output coincides with the core of the output obtained by all other residuated implications, a crisp interval in practice, that gets wider as the input membership value decreases [7], [13]. In particular, if $\mu_A(u_0) = 1$ and the core of the output is a singleton, the output is precise.

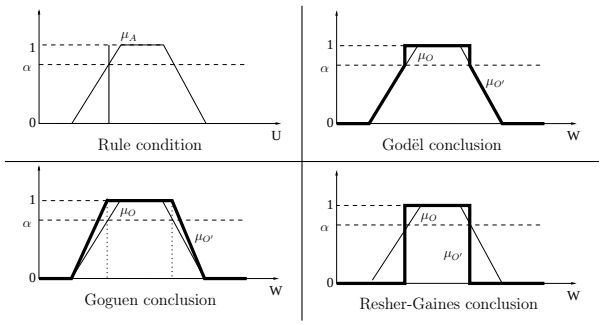


Fig. 2. Inference with one gradual rule and a precise input

Modus Ponens in classical logic is: $A \wedge (A \rightarrow O) \models O$ where $\models$ represents the logical inference. In fuzzy logic, modus ponens can be non-trivially extended to Generalised Modus Ponens (GMP) [14] $A' \wedge (A \rightarrow O) \models O'$. In the presence of an approximate fact A' and the implication $A \rightarrow O$, we are able to calculate O' defined by:

$$\mu_{O'}(v) = \sup_{u \in U} \big(\mu_{A'}(u) \top (\mu_A(u) \rightarrow \mu_O(w))\big) \tag{6}$$

The output O' constrains the value of the output variable. When an operator $\rightarrow$ (implication) is obtained from $\top$ (conjunction) by residuation, the GMP $A' \wedge (A \rightarrow O) \models O'$ is recovered for fuzzy rules [15]. Note that for pure (Resher-Gaines) gradual rules, modus ponens is strengthened: from $A' \subset A$ and $A \rightarrow O$, a conclusion more precise than O can be obtained.

C. Rule Behavior Comparison

In line with their different meanings, conjunctive and implicative rules do not behave similarly. In the presence of fuzzy inputs or cascading systems of fuzzy rules, conjunctive rules have some unnatural behavior.

1) Interpolation between rules: The interpolation mechanism used for Mamdani rules is described in depth in [16]. Let us consider input/output partitions such as $core(A_i) = \{a_i\}$ and $supp(A_i) = [a_{i-1}, a_{i+1}]$, with $a_{i-1} < a_i < a_{i+1}$

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

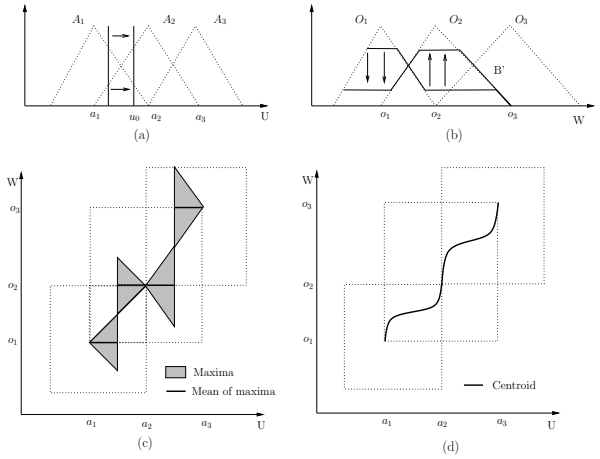


Fig. 3. Interpolation with Mamdani rules

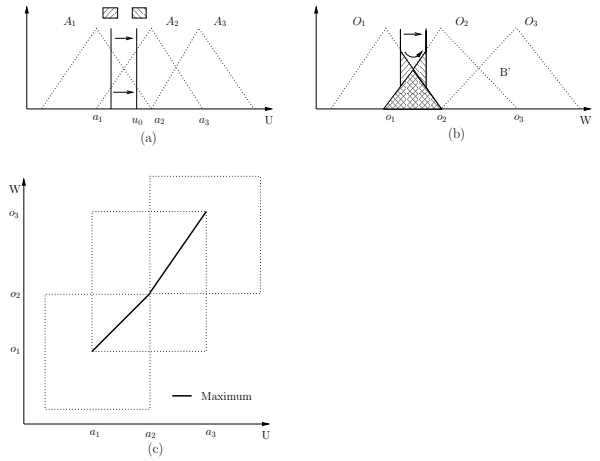


Fig. 4. Interpolation with gradual implicative rules (Godël implication)

a) Conjunctive possibility rules: Figure 3 shows the output possibility distribution inferred by three Mamdani rules, $A_i \wedge O_i$ $(i = 1, 2, 3)$, when input u_0 moves from a_1 to a_2 (see subfigure a): only truncation levels of O_1 and O_2 are affected (see subfigure b). A defuzzification step is always needed. Subfigures (c) and (d) respectively show results using mean of maxima and centroid defuzzifications. Only the centroid defuzzification leads to a continuous function, which is generally monotonic. However, contrary to what could be expected, this function is not linear. In fact it has been shown that in some configurations, a set of fuzzy rules qualitatively expressing a monotonic behavior may fail to produce a monotonic control law ([17] and [18]).

b) Gradual implicative rules: Figure 4 illustrates the case of three gradual rules $A_i \rightarrow O_i (i = 1, 2, 3)$. Due to the fuzzy partition structure, the maximum is unique (b) and defuzzification is not necessary in that case. Subfigure (c) shows the linear evolution of this unique maximum. This subfigure holds for all residuated implications, as they yield the same core.

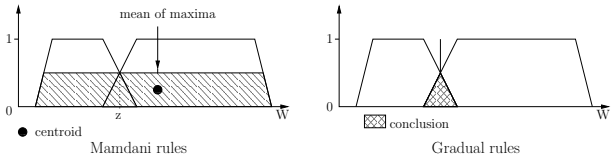


Fig. 5. Fuzzy set width influence

2) Influence of the specificity of the rules: Let us consider two rules triggered at the same level.

a) Conjunctive possibility rules: When two trapezoidal output fuzzy sets have equal widths, the inferred value (mean of maxima or centroid) is equal to z such that $\mu_{O_1}(z) = \mu_{O_2}(z)$. This result is the one expected. Nevertheless, if one output set is wider than the other, the defuzzified value moves towards the wider one, which is counter-intuitive, as shown in the left part of figure 5.

b) Gradual implicative rules: This behavior is impossible with gradual implicative rules because rules are aggregated in a conjunctive way. In fact the result of triggering two gradual rules is more precise than the result of triggering a single rule. This is totally the opposite situation for conjunctive rules, even with precise inputs. So there is a natural expectation of limited imprecision of results when triggering fuzzy implicative rules with fuzzy inputs, including the case when such fuzzy inputs result from a previous inference step.

3) Rule accumulation: Adding a conjunctive rule enlarges the output possibility distribution. Then a rule system is never inconsistent even if the rule base includes conflicting rules from a knowledge representation point of view. When many rules are added to the rule base, the output possibility distribution approaches the membership function of the whole referential. That behavior, often hidden by defuzzification, is not intuitive because we might think that adding new rules (hence new information) to the knowledge base would lead to a more accurate system. If conjunctive fuzzy systems have to be cascaded, it is clear that using the fuzzy output of the first system as a fuzzy input for the second one may lead to unreasonably imprecise responses.

Implicative rules formulate constraints on possible input/output mappings. The more rules there are in a rule base, the more precise the output fuzzy set becomes, at the risk of reaching inconsistency. Inconsistency arises when for a given input $u \in U, \pi_{Z/X}(u, w) < 1, \forall w$. This feature is interesting because it allows to check logical consistency of the rule base [19].

4) Inference Mechanism: With conjunctive rules, the output O' is equal to:

$$O' = A' o (\bigcup_{i=1}^{n} A_i \wedge O_i) = \bigcup_{i=1}^{n} (A' o (A_i \wedge O_i)) \qquad (7)$$

because of the commutativity of the $\sup - \min$ composition (denoted o) and the $\bigcup$ operator, the maximum for Mam-

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

dani systems. This method, named FITA[2], corresponds to the right-hand side of equation 7. The inference mechanism is easy to implement because the inference can be performed rule by rule.

With implicative rules, the output O' is given by:

$$O' = A' o \bigcap_{i=1}^{n} (A_i \rightarrow O_i) \tag{8}$$

where $\bigcap$ is the minimum operator. When A' is a precise input, operators o and $\bigcap$ commute, the output can then be written:

$$O' = \bigcap_{i=1}^{n} ((A' o A_i) \rightarrow O_i)) \tag{9}$$

This formalisation corresponds to the FITA method for computing inference results.

However, when the input A' is imprecise or fuzzy, the commutativity between $\sup - \min$ composition and the $\bigcap$ operator is no longer possible [14]. Only the expression (8) which is a FATI[3] inference is correct. For an approximate fact, the following inclusion is true :

$$A' o \left(\bigcap_{i=1}^{n} A_i \rightarrow O_i \right) \subseteq \bigcap_{i=1}^{n} \left((A' \ o \ A_i \ ) \rightarrow \ O_i \right)$$

The FITA method only gives an upper approximation of the result.

Currently there are almost no practical methods for computing inference with implicative fuzzy rules. One method had been developed in [20] for Gödel implication when the fuzzy sets in condition parts are one-dimensional and have overlapping cores. Another technique proposed by Ughetto and al. [21] is devoted to Resher-Gaines implications with one-dimensional inputs; it presupposes an explicit calculation of the (crisp) relation defined by a set of gradual rules, in the form of two piecewise linear functions.

D. Other fuzzy interpolation techniques

In this subsection we briefly discuss the difference between gradual rule-based inference and other fuzzy interpolation methods. Many fuzzy systems rely on the proposal previously made by Takagi and Sugeno [22] to simplify Mamdani-like systems, turning the fuzzy conclusions of rules into precise ones. Then, using the centroid defuzzification, the fuzzy system computes a standard interpolation between precise conclusions, weighted by the degrees of activation of rules, due to a precise input. There is a precise connection between Takagi-Sugeno systems and gradual rules. In the one dimensional case, if strong partitions are used for inputs and output, Takagi-Sugeno inference coincides with gradual rule inference, both of which generalize linear interpolation [23]. In particular, a precise input yields a precise output. In the multidimensional case, this equivalence no longer holds because the output of a gradual rule system under a precise multidimensional input is generally an interval [13]. Nevertheless it is possible to devise a gradual rule system so that the output interval contains the precise output of some prescribed T-S system. In fact, even if gradual rule systems have interpolation capabilities built in the logic, their scope is to reflect the imprecision pervading the input and the rules in their output results, while T-S systems aim at modeling a generalized form of precise interpolation by means of rules having fuzzy conditions.

The nD Takagi-Sugeno method essentially comes down to extending an interpolation equation from 2D to nD. On the contrary the nD gradual rule approach extends the 2D generalized modus ponens, hence the result will be imprecise (an interval) even if the input is precise. So, contrary to T-S approach, gradual rules not only interpolate but also propagate imprecision (present because of the granular nature of a fuzzy rule-based system) as well, hence handle uncertainty without resorting to type 2 fuzzy sets.

Other interpolation methods exist for fuzzy systems having rules whose condition parts fail to cover the input domain, starting with works by Koczy and Hirota [24]. Usually, such methods start with a given classical numerical interpolation scheme, and extend it to fuzzy data expressed by scarce fuzzy rules. Reasoning alpha-cut-wise often leads to difficulty because the obtained output intervals for each membership levels may fail to be nested. Jenei and colleagues [25], [26] provide an extensive analysis of fuzzy interpolation techniques with a set of requirements fuzzy interpolation should satisfy. The latter family of techniques is driven by the necessity to produce an output result despite the scarcity of information, while the gradual rule approach is tailored not to produce an output result when a logical inconsistency is detected [19], a conflict resulting from handling too much information.

III. Inferential Independence

To design a practical algorithm for implicative inference, we use the interesting property of inferential independence [27], leading to well-conditioned systems. Section III-A recalls the main results available in the literature, that will be used in section III-B.

A. Definitions and results

A rule system $\{A_j \rightarrow O_j, j = 1, \ldots, n\}$ is well-conditioned if it produces the output fact O_i when fed with the input fact A_i, for any $i = 1, \ldots, n$:

$$\forall i, A_i o \bigcap_{j} (A_j \rightarrow O_j) = O_i$$

More often than not, this condition is not satisfied, and the output is more precise:

$$A_i o \bigcap_{j} (A_j \rightarrow O_j) = O_i' \subset O_i$$

[2]FITA means "First Infer Then Aggregate"
[3]FATI means "First Aggregate Then Infer"

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

According to Morsi[28], if we substitute each rule conclusion with the inferred output O'_i, the system $A_j \rightarrow O'_j$ is well-conditioned:

$$A_i o \bigcap_j (A_j \rightarrow O'_j) = O'_i$$

Morsi's proof uses residuated implication properties [29] verified by Gödel and Goguen operators and the relation: $\bigcap_j (A_j \rightarrow O_j) = \bigcap_j (A_j \rightarrow O'_j)$ proved in [28]. In a well-conditioned system, rules are said to be *inferentially independent*.

This way of doing requires an inference step. Alternatively, the inferential independence property can be guaranteed by a proper design of fuzzy input partitions.

B. Sufficient conditions for a well-conditioned system

In the sequel, we look for a form of fuzzy input partition ensuring a well-conditioned system. Two cases are to be considered: residuated implications (Gödel and Goguen) and Resher-Gaines implication. The following result does not work for Resher-Gaines but it is true for all residuated implications obtained from a continuous t-norm.

Theorem: A system of fuzzy implicative fuzzy rules $\{A_i \rightarrow O_i, i = 1, \ldots, n\}$, modeled by residuated implications is well-conditioned as soon as

$\forall\ i = 1, \ldots, n\ \exists\ x \in core(A_i)$, $\mu_{A_j}(x) = 0$, $\forall\ j \neq i$.

Proof: Let $\top$ be a continuous triangular norm on $[0, 1]$, and $\rightarrow$ be the corresponding residuated implication $a \rightarrow b = \sup\{c, a \top c \leq b\}$. The $\max - \min$ composition is generalized into a $\max - \top$ composition. From equation (6), and because of the conjunctive aggregation of implicative rules, we require: $\forall z \in W$,

$$\sup_{x \in U} \mu_{A_i}(x) \top \min_{j \in N} \Big(\mu_{A_j}(x) \rightarrow \mu_{O_j}(z)\Big) = \mu_{O_i}(z)$$

We can shift $\mu_{A_i}(x)$ and t-norm $\top$ inside of $\min$. We are looking for sufficient conditions for the equality: $\forall z$,

$$\sup_{x \in U} \min_{j \in N} \Big(\mu_{A_i}(x) \top (\mu_{A_j}(x) \rightarrow \mu_{O_j}(z))\Big) = \mu_{O_i}(z)$$

to hold. This sufficient condition is equivalent to:

$\forall z, \exists x \in U$,

$$\min_{j \in N} \Big(\mu_{A_i}(x) \top (\mu_{A_j}(x) \rightarrow \mu_{O_j}(z))\Big) = \mu_{O_i}(z)$$

Then, the following conditions are sufficient to ensure this equality:

$\forall z, \exists x \in U$,

$$\forall j \neq i, \mu_{A_i}(x) \top (\mu_{A_j}(x) \rightarrow \mu_{O_j}(z)) \geq \mu_{O_i}(z) \quad (10)$$

and

$$\mu_{A_i}(x) \top (\mu_{A_i}(x) \rightarrow \mu_{O_i}(z)) = \mu_{O_i}(z) \quad (11)$$

Choosing $x \in core(A_i)$, equation (11) obviously holds since $1 \rightarrow \mu_{O_i}(z) = \mu_{O_i}(z)$ for residuated (hence Gödel and Goguen) implications.

Now, we must deal with equation (10). If we consider x in the core of A_i, then $\mu_{A_i}(x) = 1$. A sufficient condition is then:

$$\forall z, \exists x \in core(A_i), \forall j \neq i, \mu_{A_j}(x) \rightarrow \mu_{O_j}(z) \geq \mu_{O_i}(z) \quad (12)$$

There are two cases:

- $\mu_{A_j}(x) > \mu_{O_j}(z)$: then equation (12) is not usually true. If this strict inequality holds $\forall x \in core(A_i)$, the system is not well-conditioned.
- $\mu_{A_j}(x) \leq \mu_{O_j}(z)$: then equation (12) is always true.

Fuzzy systems will ever respect the latter inequality condition $\mu_{A_j}(x) \leq \mu_{O_j}(z)$, if the following property holds: at least one value in a fuzzy set core does not belong to the support of other input fuzzy sets. i.e. as we can see on figure 6, $\exists\ x \in core(A_i), \mu_{A_j}(x) = 0, \forall\ j \neq i$. Q.E.D.

This proof holds for a n-dimensional-input system as well (interpreting x as a vector of coordinates).

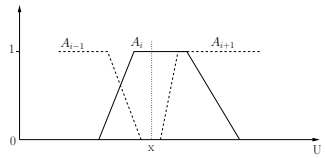


Fig. 6. A fuzzy partition allowing inferential independence

For strong input fuzzy partitions (see figure 8) the following stronger property holds: $\forall\ j \neq i$, $\forall\ x \in core(A_i)$, $\mu_{A_j}(x) = 0$. Hence the system is always well-conditioned in this case. An interesting property useful for inference is that for strong fuzzy partitions, with $x \in core(A_i)$, the system output is O_i for Gödel and Goguen implications.

For the Resher-Gaines implication, equation (11) holds if, $\forall z$, we choose x such that $\mu_{A_i}(x) = \mu_{O_i}(z)$. Then equation (10) will hold if and only if $\mu_{A_j}(x) \leq \mu_{O_j}(z)$. Assume strong input and output partitions. Then, in the one-dimensional case, only adjacent rules $A_i \rightarrow O_i$, $A_{i-1} \rightarrow O_{i-1}$ and $A_{i+1} \rightarrow O_{i+1}$ are triggered. Then for $j \neq \{i, i+1\}$ equation (10) trivially holds. For $j = i+1$, this equation reads :

$$\mu_{A_i}(x) \top ((\mu_{A_{i+1}}(x) \rightarrow (\mu_{O_{i+1}}(z))) \geq \mu_{O_i}(z)$$

Because of the strong partition assumption, the equation is equivalent to:

$$\mu_{A_i}(x) \top ((1 - \mu_{A_i}(x) \rightarrow (1 - \mu_{O_i}(z))) \geq \mu_{O_i}(z)$$

which holds if $\mu_{A_i}(x) = \mu_{O_i}(z)$ for Resher-Gaines implication. This behavior is also true for $j = i - 1$.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

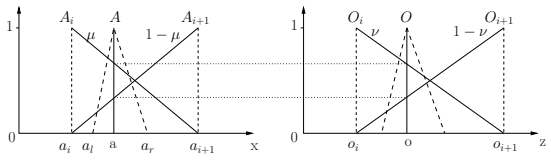


Fig. 7. Notation (Resher-Gaines implication)

In the one dimensional case, exact analytical expressions can be calculated for the inference result. We give them for all different implication types.

IV. Analytical expressions for inference with a single fuzzy input

In the one dimensional case it is possible to provide analytical expressions of the inference result. Let us consider a fuzzy input A as a unimodal fuzzy interval whose support lies in the interval between the cores of two subsequent rule conditions (which is the most complex case). Therefore, the μ function is invertible.

The output is given by:

$$\mu_O(z) = \sup_{x\in U} \mu_A(x) \top \min\Big(\mu_{A_i}(x) \to \mu_{O_i}(z), \mu_{A_{i+1}}(x) \to \mu_{O_{i+1}}(z)\Big)$$

As we deal with strong partitions this is also, letting $\mu = \mu_{A_i}$, and $\nu = \mu_{O_i}$ for short:

$$\mu_O(z) = \sup_{x\in U} \mu_A(x) \top \min\Big(\mu(x) \to \nu(z), (1-\mu(x)) \to (1-\nu(z))\Big) \tag{13}$$

Let us study the different implications. Figure 7 illustrates some notations. Let's denote $e = \min\Big(\mu(x) \to \nu(z), (1-\mu(x)) \to (1-\nu(z))\Big)$.

1) Resher-Gaines implication: Let us consider the case of Resher-Gaines implication. It is clear that $e > 0$ if and only if $\mu(x) = \nu(z)$, and then $e = 1$. Hence x is equal to $\mu^{-1}(\nu(z))$, the inference result O_{RG} is such as:

$$\mu_{O_{RG}}(z) = \mu_A(\mu^{-1}(\nu(z))) \tag{14}$$

In other words, if the input partition is the same as the output partition, the computed output is the same as the fuzzy input.

2) Gödel implication: Now in the case of Gödel implication, $\top = min$, and the inference process reads, distinguishing 3 cases:

- $\mu(x) = \nu(z)$ then $e = 1$
- $\mu(x) > \nu(z)$ then $e = \nu(z)$
- $\mu(x) < \nu(z)$ then $e = 1 - \nu(z)$

From equation 13, we can deduce:

$$\begin{aligned}\mu_{O_{God}}(z) &= \max\Bigg(\mu_A(\mu^{-1}(\nu(z))), \sup_{\mu(x)<\nu(z)} \min\Big(\mu_A(x), \\ &\quad 1-\nu(z)\Big), \sup_{\mu(x)>\nu(z)} \min\Big(\mu_A(x), \nu(z)\Big)\Bigg) \\ &= \max\Bigg(\mu_A(\mu^{-1}(\nu(z))), \min\Big(1-\nu(z), \\ &\quad \sup_{\mu(x)<\nu(z)} \mu_A(x)\Big), \min\Big(\nu(z), \sup_{\mu(x)>\nu(z)} \mu_A(x)\Big)\Bigg)\end{aligned}$$

Note that, for a given z, $\{x|\mu(x) > \nu(z)\}$ is of the form $[a_i, \mu^{-1}(\nu(z))[$, and so, the possibility degree $\sup_{\mu(x)>\nu(z)} \mu_A(x))$ is 1 if $\mu^{-1}(\nu(z)) > a$ (the core of A), and $\mu_A(\mu^{-1}(\nu(z)))$ otherwise. In other words, it is the membership degree of $\mu^{-1}(\nu(z))$ to the fuzzy interval $[a_i, A]$.

Similarly, $\sup_{\mu(x)<\nu(z)} \mu_A(x)$ is 1 if $\mu^{-1}(\nu(z)) < a$, and $\mu_A(\mu^{-1}(\nu(z)))$ otherwise. In other words, it is the degree of membership of $\mu^{-1}(\nu(z))$ to the fuzzy interval $[A, a_{i+1}]$ [30]. Hence

$$\mu_{O_{God}}(z) = \max\Bigg(\mu_A(\mu^{-1}(\nu(z))), \min\Big(1-\nu(z), \mu_{[A,a_{i+1}]}(\mu^{-1}(\nu(z)))\Big), \min\Big(\nu(z), \mu_{[a_i,A]}(\mu^{-1}(\nu(z)))\Big)\Bigg)$$

So, the inference result O_{God} has the same core $o = \nu^{-1}(\mu(a))$ as O_{RG} and is such that:

$$\mu_{O_{God}}(z) = \begin{cases} \max(\mu_{O_{RG}}(z), 1-\nu(z)) & \text{if } z < o,\ o = core(O) \\ \max(\mu_{O_{RG}}(z), \nu(z)) & \text{if } z > o,\ o = core(O) \end{cases}$$

3) Goguen implication: For Goguen implication, $e = \min\Big(\min(1, \frac{\nu(z)}{\mu(x)}), \min(1, \frac{1-\nu(z)}{1-\mu(x)})\Big)$ and $\top = *$. We know that $\mu_A(x) = 0, \forall x \notin]a_l, a_r[$; we will then only consider the interval $]a_l, a_r[$. For a given z_0, we denote $x_0 = \mu^{-1}(\nu(z_0))$. Then we have 3 cases:

- $\mu(x) = \nu(z) \Leftrightarrow x = x_0$ then $e = 1$
- $\mu(x) > \nu(z) \Leftrightarrow x \in]a_l, x_0[$ then $e = \min\Big(\frac{\nu(z)}{\mu(x)}, 1\Big) = \frac{\nu(z)}{\mu(x)}$
- $\mu(x) < \nu(z) \Leftrightarrow x \in]x_0, a_r[$ then $e = \min\Big(1, \frac{1-\nu(z)}{1-\mu(x)}\Big) = \frac{1-\nu(z)}{1-\mu(x)}$

The result for Goguen is now given by:

$$\begin{aligned}\mu_{O_{Gog}}(z) &= \max\Big(\mu_A(x_0), \sup_{x\in]a_l,x_0[} \mu_A(x) * \frac{\nu(z)}{\mu(x)}, \\ &\quad \sup_{x\in]x_0,a_r[} \mu_A(x) * \frac{1-\nu(z)}{1-\mu(x)}\Big)\end{aligned}$$

Then, there are two cases:

- $z < o$:
  First, let study $\sup_{x\in]a_l,x_0[} \mu_A(x) * \frac{\nu(z)}{\mu(x)}$. On $]a_l, x_0[$, $\mu_A(x)$ is increasing and $\mu(x)$ is decreasing. For $x = x_0, \mu(x) = \nu(z)$ so $\sup_{x\in]a_l,x_0[} \mu_A(x) * \frac{\nu(z)}{\mu(x)} = \mu_A(x_0)$. Then, we study the interval $]x_0, a_r[$. This study is more complex and gives the result $\sup_{x\in]x_0,a_r[} \mu_A(x) * \frac{1-\nu(z)}{1-\mu(x)} = \frac{1-\nu(z)}{1-\mu(a)}$. We do not give details here, but a

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

geometrical demonstration proves that $\forall z < o, \frac{1-\nu(z)}{1-\mu(a)}$ is always greater than $\mu_A(x_0)$.
So the final result is equal to:

$$\mu_{O_{Gog}}(z) = \max\left(\mu_A(x_0), \mu_A(x_0), \frac{1-\nu(z)}{1-\mu(a)}\right) = \frac{1-\nu(z)}{1-\mu(a)}$$

- $z > o$:
  The study of $\sup_{x\in]a_l,x_0[} \mu_A(x) * \frac{\nu(z)}{\mu(x)}$ gives us $\frac{\nu(z)}{\mu(a)}$. In the interval $]x_0, a_r[$, $\sup_{]x_0,a_r[} \mu_A(x) * \frac{1-\nu(z)}{1-\mu(x)} = \mu_A(x_0)$. As previously one can show that $\frac{\nu(z)}{\mu(x)} > \mu_A(x_0)$.
  As a consequence, the result of the inference for $z > o$ is:

$$\mu_{O_{Gog}}(z) = \max\left(\mu_A(x_0), \frac{\nu(z)}{\mu(a)}, \mu_A(x_0)\right) = \frac{\nu(z)}{\mu(a)}$$

So, the result O_{gog} of the inference has the same core $o = \nu^{-1}(\mu(a))$ as O_{RG} and is such that:

$$\mu_{O_{Gog}}(z) = \begin{cases} \frac{1-\nu(z)}{1-\mu(a)} & \text{if } z < o \\ \frac{\nu(z)}{\mu(a)} & \text{if } z > o \end{cases}$$

Let us stress that all analytical expressions given here are only valid for a fuzzy input lying in the overlapping area between two fuzzy sets of the input partition. The case of a fuzzy input lying within a fuzzy set core is obvious.

V. 1D INFERENCE ALGORITHM

We now use strong input fuzzy partitions and the inferential independence property to design a practical inference process by input decompositions. These decompositions are instrumental due to the following property of a fuzzy relation R:

$$(A \cup A')oR = (AoR) \cup (A'oR) \tag{15}$$

where o is a sup-t-norm composition and $\cup$ is the maximum operation.

We first consider one-dimensional inputs for explanation purposes. We detail the output calculation for an α-level rectangular input, which our inference algorithm will be based upon. The decomposition algorithm proposed here in 1D scales up to 2D inputs while the previous analytical expressions do not.

A. Partitioning the input space

To partition the input space, we consider supports and cores seperately. Let E_k be intervals forming a partition, obtained as a alternating sequence of cores and peripheral parts of rules conditions (see figure 8). This decomposition isolates the fuzzy set cores. The inference is straightforward from a fuzzy input lying in a core area, as, due to the strong fuzzy partition structure, only one rule is fired. In this case, the output possibility distribution is either the whole set corresponding to the fired rule conclusion for Godel or Goguen operators, or its core for Rescher-Gaines operator.

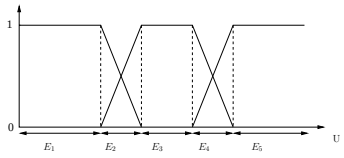


Fig. 8. Partitioning decomposition with strong fuzzy partition

B. Fuzzy input decomposition

An α-cut of A is an interval defined by:
$\forall \alpha > 0, A_\alpha = \{x \in \mathbb{R} | \mu_A(x) \geq \alpha\}$.
According to Zadeh's representation result: $A = \bigcup_{\alpha\in]0,1]} \alpha A_\alpha$. In the presence of a fuzzy input A', we first decompose A' in terms of α-cuts . Then, we decompose these cuts in terms of the above partition of the input space. In consequence, we have the identity:

$$A' = \bigcup_\alpha \left(\alpha(\bigcup_{k=1,\ldots,p} E_k \cap A_\alpha)\right)$$

where p is the number of intervals E_k.

In practice we use only a finite number of cuts with thresholds $\alpha_1 = 1 > \alpha_2 > \cdots > \alpha_n > 0$. A fuzzy set A' is then included within two inner and outer approximations (see figure 9).

$$\bigcup_{j=1,\ldots,n} \alpha_j A_{\alpha_j} \subseteq A' \subseteq \bigcup_{j=1,\ldots,n} \alpha_j A_{\alpha_{j+1}} \tag{16}$$

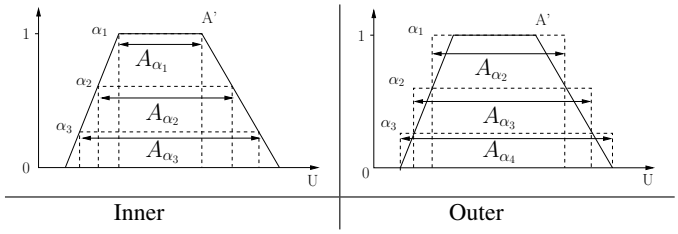


Fig. 9. α-cut decomposition

External approximations seem to be more appropriate because they include the fuzzy input. The approximated output contains the true output. It could be interesting to keep both inner and external approximations in order to reason with two approximations like for Rough Sets [31].

The double decomposition presented above will be used in the inference algorithms that follow.

C. Inference with an α-level rectangular input

Due to the partitioning of the input space, the rectangular input $(E_k \cap A_\alpha)$ overlaps on at most two fuzzy sets. If this input lies within the fuzzy set core of A_i, the result is obvious: we obtain O_i for Gödel and Goguen implications and O_i's core for Resher-Gaines implication. Figure 10 recalls inference results with a precise input and two gradual rules whose conditions form a strong partition.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

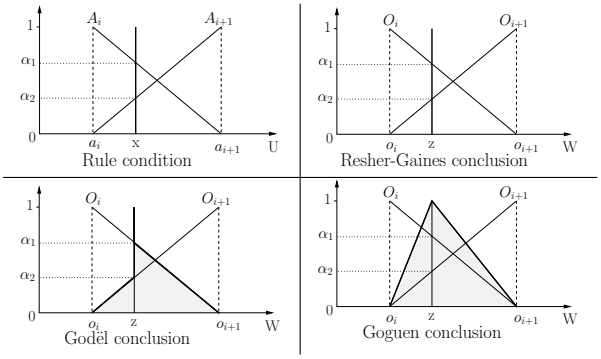


Fig. 10. Inference with two gradual implicative rules and a precise input

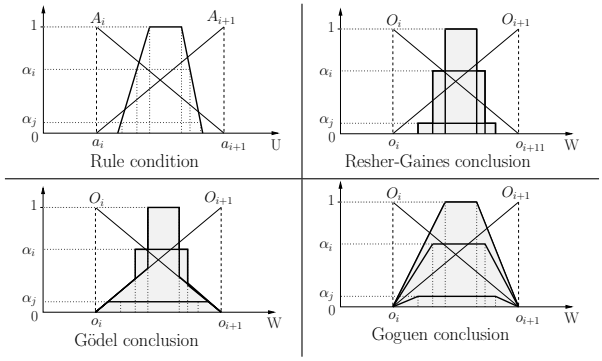


Fig. 11. Inference with two gradual implicative rules and a fuzzy input decomposed on three levels $\alpha_j < \alpha_i < 1$

Let the interval of interest $(E_k \cap A_\alpha)$ be denoted $[i_l, i_r]$. An α-level rectangular input membership function is defined by $\mu_{[i_l,i_r]}$ such that: $\mu_{[i_l,i_r]}(x) = \begin{cases} \alpha & \text{if } i_l \leq x \leq i_r \\ 0 & \text{otherwise} \end{cases}$

Since the rectangular input $[i_l, i_r]$ lies in the support of two consecutive fuzzy sets (see figure 11), the output is given by:

$$\mu_{O'}(z) = \sup_{i_l \leq x \leq i_r} \min_{i=1,\ldots,n} \Big(\alpha \top \mu_{A_i}(x) \rightarrow \mu_{O_i}(z)\Big)$$

In this specific case, it is equal to:

$$\mu_{O'}(z) = \sup_{i_l \leq x \leq i_r} \min \Big(\alpha \top \mu_{A_i}(x) \rightarrow \mu_{O_i}(z), \alpha \top \mu_{A_{i+1}}(x) \rightarrow \mu_{O_{i+1}}(z)\Big) \quad (17)$$

for some i.

Since α and $\top$ are independent of x and i, the system is equivalent to:

$$\mu_{O'}(z) = \alpha \top \sup_{i_l \leq x \leq i_r} \min \Big(\mu_{A_i}(x) \rightarrow \mu_{O_i}(z), \mu_{A_{i+1}}(x) \rightarrow \mu_{O_{i+1}}(z)\Big)$$

Next, the output behavior depends on the chosen residuated implication. We consider Resher-Gaines, Gödel and Goguen implications.

Level α has only a truncation effect on the output's height. No output element can have a higher membership than level α because the minimum is the upper bound of t-norms. According to the chosen implication, a different t-norm will be used. For Resher-Gaines and Gödel ones, the t-norm is the minimum. Then, the output is truncated at level α, but its shape is preserved. For Goguen implication, t-norm is the product. The output is also truncated at level α but the support slopes are modified (See figure 11).

Output computation for one rectangular input is straightforward depending on the chosen implication.

The approximate one-dimensional inference process is completed by performing the union of outputs inferred from each α-level rectangular input taking both decompositions into account.

D. Results of the double decomposition

The result of the inference based on a fuzzy input A' is O' of the form:

$$O' = \bigcup_{k=1,\ldots,p} \left(\bigcup_\alpha O_k'^{\overline{\alpha}}\right)$$

where $O_k'^{\overline{\alpha}} = (E_k \cap A_\alpha) oR$ is obtained in two steps. First the output possibility distribution is calculated for a level 1 rectangular input. Then the t-norm is applied to this output possibility distribution. The minimum t-norm truncates the output possibility distribution while the product t-norm also affects its slope, as illustrated in Figure 11.

E. Complexity

The inference process summary is given below. Let n be the number of α-cuts, and k the number of E_k intervals within the input partition.

- Decompose the fuzzy input by n α-cuts in order to consider it as a series of α-level rectangular inputs.
- Decompose each rectangular input according to the E_k intervals within the input partition in order to separate cores from intermediate zones.
  Then, for each α-level rectangular input, it is necessary to:
  - Infer from each bound of the α-level rectangular input.
  - Compute the convex hull of the k partial inferred sets
- Compute the union of the n convex hulls.

An analysis of the algorithm complexity follows.

- α-cut input decomposition linearly depends on n.
- Decomposition of rectangular inputs linearly depends on n and on the number of their intersections with the subsets resulting from the partition decomposition, i.e. k.
- Inference from both bounds of the rectangular input requires 2 calculations for each α-cut.
- Convex hull can be determined by considering $2k$ inferred bounds.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

- Last step is the union of n convex hulls.

As all operations linearly depend on n, this algorithm has complexity $O(n)$.

F. α-cut related approximation

The only approximation made in the one dimensional inference algorithm described above comes from the α-cut input decomposition. All the other steps include exact decompositions, they are only introduced in order to increase the algorithm efficiency. Let us give some elements to quantify the α-cut related approximation. For that purpose, we consider "identical" input and output partitions, such as the ones shown on Figure 12, with the same range $[min, max]$ and two fuzzy sets each. In that case, the analytical expression given in Equation 14 reduces to $\mu_{O_{RG}}(z) = \mu_A(z),\ \forall z \in [min, max]$. The inferred output must be identical to the fuzzy input (see Figure 12).

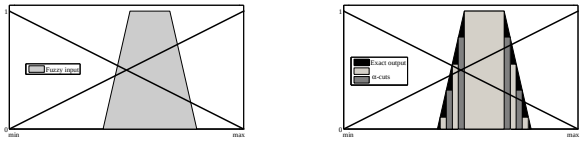


Fig. 12. α-cut related approximation

Table I gives the number of α-cuts required for reaching various accuracy levels, depending on the fuzzy input characteristics. The fuzzy input is chosen as a symmetrical trapezoidal fuzzy set.

Irrespective of the number of α-cuts, the computed output has the same core as the exact output. The accuracy level is evaluated as the ratio $\frac{computed\ output\ area}{exact\ output\ area}$. It only depends on the fuzzy input slope, which varies from 45 to 90 degrees. The results show that, whatever the slope, at most ten α-cuts are necessary for ensuring an accuracy level better than ten per cent.

	Min. Accuracy (%)				
Slope	20	15	10	5	2
45 °	6	7	10	22	81
50 °	5	6	8	15	47
60 °	3	4	5	9	22
70 °	2	3	3	5	12
80 °			2	3	6
90 °					1

TABLE I
NUMBER OF α-CUTS REQUIRED FOR A GIVEN ACCURACY

VI. 2D INFERENCE ALGORITHM

We now examine inference with fuzzy inputs in the two-dimensional case. We use the same decomposition method as in the one-dimensional case. In the sequel, we denote a rule as: $A_k \wedge B_l \rightarrow O_{k,l}$. The aim of this section is to determine the output in the presence of two fuzzy inputs. In order to reduce the complexity, a double decomposition is used again:

- α-cut decomposition: decompose each fuzzy input into a union of rectangular inputs of level α, $0 < \alpha \leq 1$. This decomposition allows to consider each fuzzy input on each dimension as a set of α-level rectangular inputs. α is identical in both dimensions.
- Partitioning decomposition: For each rectangular α-cut, a decomposition is made according to the different parts of the partition in order to handle the inference process locally. Thanks to inferential independence, the inference from the core part is obvious.

As a consequence, the inferred output is now the result of a double union:

$$O' = \bigcup_{k=1,\ldots,p} \left(\bigcup_{\alpha} O'^{\overline{\alpha}}_{k,l} \right) \tag{18}$$

$O'^{\overline{\alpha}}_{k,l}$ is the inferred output resulting from inputs $E_k \cap A_\alpha$ and $E_l \cap B_\alpha$.

The key issue to be considered is how to infer with an α-level rectangular input in each dimension. If the function to be represented by the fuzzy rule-based system is monotonic and continuous, it is sufficient to infer from each bound of the rectangular input on each dimension, in order to get the fuzzy output interval bounds. If the output is not monotonic, we need to detect the extrema of the function and deal with monotonic parts separately.

To sum up the inference process in two dimensions, it is necessary to:

- decompose the fuzzy input by α-cuts in order to consider the fuzzy input as a set of α-level rectangular inputs.
- decompose each rectangular α-cuts according to the input partition in order to separate core and overlapping zones. This allows a local inference.
- for each α-level rectangular input,
  - infer from each of the 4 vertices of the 2 α-level rectangular inputs.
  - test if there are other useful points lying inside the rectangular input, and infer from all such values.
  - the final output is the convex hull of all the outputs so inferred.
- The union of all outputs previously computed is the final result.

A. Implementation

We have three key points to study:

- *output partition*: To preserve coherence and to insure interpretability of the system, we need to choose proper output partitions.
- *Continuity*: we must insure continuity across the different areas obtained by decomposition.
- *Extremal points*: if the output is not monotonic between the two bounds of the rectangular input, we need to detect the extremal points and to consider them for the inference process.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

If X is A_1 and Y is B_1 then Z is $O_{1,1}$
If X is A_1 and Y is B_2 then Z is $O_{1,2}$
If X is A_2 and Y is B_1 then Z is $O_{2,1}$
If X is A_2 and Y is B_2 then Z is $O_{2,2}$

TABLE II
THE SET OF SIMULTANEOUSLY FIRED RULES FOR TWO INPUTS

In the sequel, we first study output partitions and the mechanism of inference for a precise input. Then, we deal with continuity and kink points.

B. Output partition coherence and interpretability

In this section, we focus on Resher-Gaines implication because its computation provides the core of outputs inferred using residuated fuzzy implications. Each input variable is associated to a strong fuzzy partition (see figure 13). The purpose of this section is to find output fuzzy sets capable of ensuring the logical coherence of the rule base system [19]. Furthermore, we need to have an interpretable output partition.

Thanks to the strong fuzzy partition a given two-dimensional precise input can trigger at most four rules, shown on table II.

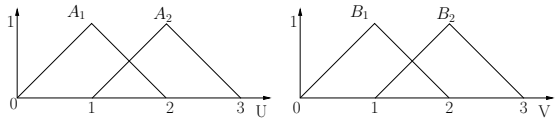


Fig. 13. Input partitions

Coherence: a rule system is coherent if for all input values, there is at most one output value totally compatible (the infered output must be normalized)

To obtain a coherent system, a necessary condition is to have $O_{1,1} \cap O_{1,2} \cap O_{2,1} \cap O_{2,2} \neq \emptyset$. Sufficient conditions are more demanding and can be found in [19]. Based on results presented in [13], we build an output coverage where $O_{1,1}$ and $O_{2,2}$ form a strong partition. In order to have an interpretable system, we choose $Support(O_{1,2}) = Support(O_{2,1}) = Support(O_{1,1}) \cap Support(O_{2,2})$. (see figure 14)

Note that this partition satisfies both system coherence and interpretability properties. According to whether the system we want to represent is symmetric or not, $O_{1,2}$ and $O_{2,1}$ may be identical or not.

C. 2D inference for a precise input

With strong input partitions, there are 3 different situations according to the location of the precise input (see figure 15).

- Case 1: both inputs lie within the fuzzy set cores of each dimension. In this situation we can directly infer the output thanks to inferential independence (see section III). Output is equal to $core(O_{k,l})$ for Resher-Gaines implication.

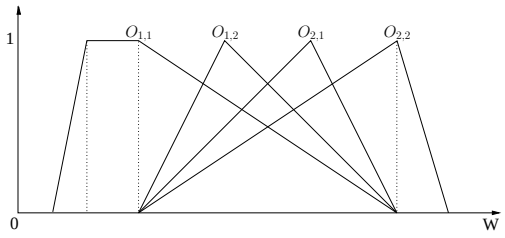


Fig. 14. Output partition for coherence and interpretability

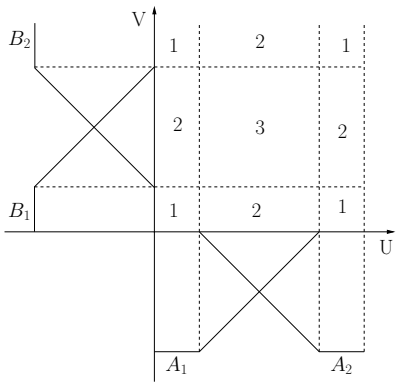


Fig. 15. Areas defined by input partitions

- Case 2: the x input lies within the fuzzy set core in a dimension and in the overlapping zone of the other dimension. For example, choose x in the core of A_1 and y between the cores of B_1 and B_2. In consequence, 2 rules are triggered: $A_1 \wedge B_1 \rightarrow O_{1,1}$ and $A_1 \wedge B_2 \rightarrow O_{1,2}$.
- Case 3: Both x and y inputs lie between the cores of adjacent fuzzy sets in U and in V (see figure 17). Four rules are triggered. This is the most complicated case.

Let us first study case 3 since case 2 is a particular case of 3. In the sequel, we denote by $[o^-(\alpha), o^+(\alpha)]$ the α-cut of the fuzzy interval O (see figure 16).

1) Case 3: Given a 2D precise input, we can compute the Resher-Gaines output [13], which is an interval defined by its lower bound z_{min} and its upper bound z_{max}. Let us denote $\alpha_i = \mu_{A_i}(x)$ and $\beta_i = \mu_{B_i}(x)$. Zones are defined on figure 17 according to the value of $m = \min(\alpha_1, \alpha_2, \beta_1, \beta_2)$, where $\alpha_1 = 1 - \alpha_2$, $\beta_1 = 1 - \beta_2$ corresponding to changes in the inference results. Table III gives the m value

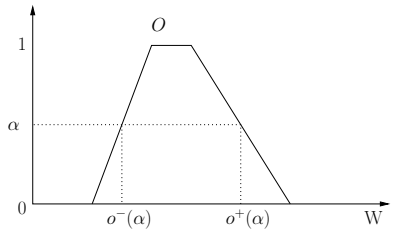


Fig. 16. Notation

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

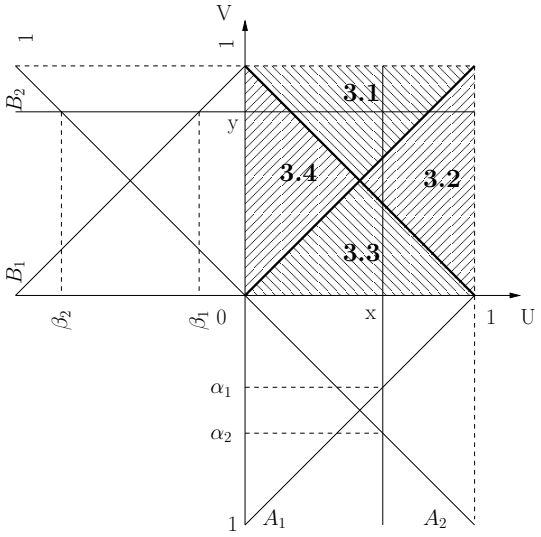


Fig. 17. Case 3: Four interesting areas

Zone	3.1	3.2	3.3	3.4
$\min(\alpha_1, \alpha_2, \beta_1, \beta_2)$	β_1	α_1	β_2	α_2

TABLE III
ZONE PROPERTIES

for each zone. Let us detail what happens for the inference in zone 3.1, which corresponds to $\beta_2 > \beta_1$ and α_1 and α_2 both greater than β_1.

In zone 3.1, the lower bound can come from four rules:

- $A_1 \wedge B_1 \rightarrow O_{1,1}$ gives us bound $o^-_{1,1}(\beta_1)$ since β_1 is less than α_1.
- $A_1 \wedge B_2 \rightarrow O_{1,2}$ gives us bound $o^-_{1,2}(\alpha_1)$ since α_1 is less than β_2.
- $A_2 \wedge B_1 \rightarrow O_{2,1}$ gives us bound $o^-_{2,1}(\beta_1)$ since β_1 is less than α_2.
- $A_2 \wedge B_2 \rightarrow O_{2,2}$ gives us bound $o^-_{2,2}(\alpha_2)$ since α_2 is less than β_2.

Since rule aggregation is conjunctive, the overall lower bound is the maximum of these bounds.

$$z_{min} = \max(o^-_{1,1}(\beta_1), o^-_{1,2}(\alpha_1), o^-_{2,1}(\beta_1), o^-_{2,2}(\alpha_2))$$

$o^-_{1,1}(\beta_1)$ is always less than other bounds because its maximum is the lower bound of the core of $O_{1,1}$. Furthermore, $o^-_{2,1}(\beta_1)$ is always lower than $o^-_{2,2}(\alpha_2)$ because $\beta_1 < \alpha_2$. As a consequence, the lower bound is:

$$z_{min} = \max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$$

Similarly, we are able to compute the upper bound:

$$z_{max} = \min(o^+_{1,1}(\beta_1), o^+_{1,2}(\alpha_1), o^+_{2,1}(\beta_1), o^+_{2,2}(\alpha_2))$$

which becomes:

$$z_{max} = \min(o^+_{1,1}(\beta_1), o^+_{1,2}(\alpha_1))$$

Zone	Lower bound z_{min}	Upper Bound z_{max}
3.1	$\max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$	$\min(o^+_{1,1}(\beta_1), o^+_{1,2}(\alpha_1))$
3.2	$\max(o^-_{2,1}(\beta_1), o^-_{2,2}(\beta_2))$	$\min(o^+_{1,1}(\alpha_1), o^+_{2,1}(\beta_1))$
3.3	$\max(o^-_{2,1}(\alpha_2), o^-_{2,2}(\beta_2))$	$\min(o^+_{1,1}(\alpha_1), o^+_{2,1}(\alpha_2))$
3.4	$\max(o^-_{1,2}(\beta_2), o^-_{2,2}(\alpha_2))$	$\min(o^+_{1,1}(\beta_1), o^+_{1,2}(\beta_2))$

TABLE IV
OUTPUT INTERVALS FOR CASE 3

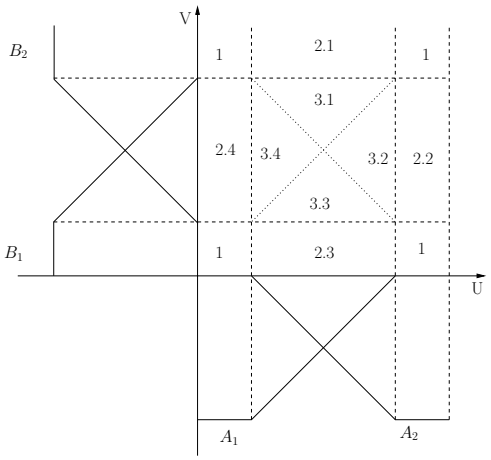


Fig. 18. Several input areas

Table IV shows results for all sub-zones of zone 3.

2) Case 2: Zone 2 can be seen as a special case of zone 3. There are four zones (2.1, 2.2, 2.3 and 2.4) adjacent to zones 3.1,3.2, 3.3 and 3.4 (see figure 18).

There are at most two rules fired in zone 2 because of the strong input partition.

For example, in case 2.1 where $\beta_1 = 0$ and $\beta_2 = 1$ only the following rules are triggered:

- $A_1 \wedge B_2 \rightarrow O_{1,2}$
- $A_2 \wedge B_2 \rightarrow O_{2,2}$

The behavior is the same as in zone 3 but less rules are triggered. z_{min} is the same as in zone 3.1 because outputs $O_{1,2}$ and $O_{2,2}$ are triggered :

$$z_{min} = \max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$$

$O_{1,1}$ is not triggered so z_{max} becomes:

$$z_{max} = o^+_{1,2}(\alpha_1)$$

Similar calculations can be made for other subzones. Outputs in zone 2 are summed up in table V.

D. The continuity of inferred outputs

In this section, we study the output continuity with respect to input variations. Figure 18 shows all possible transitions. Since zone 3 is the most general case, we first study possible transitions between its subzones. Let us examine the transition from 3.1 to 3.2. It occurs when $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$. Thus in zone 3.1, we have $\beta_1 < \beta_2$ (see table III)

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

Area	Lower bound z_{min}	Upper bound z_{max}
2.1	$\max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$	$o^+_{1,2}(\alpha_1)$
2.2	$\max(o^-_{2,1}(\beta_1), o^-_{2,2}(\beta_2))$	$o^+_{2,1}(\beta_1)$
2.3	$o^-_{2,1}(\alpha_2)$	$\min(o^+_{1,1}(\alpha_1), o^+_{2,1}(\alpha_2))$
2.4	$o^-_{1,2}(\beta_2)$	$\min(o^+_{1,1}(\beta_1), o^+_{1,2}(\beta_2))$

TABLE V
OUTPUT INTERVALS FOR CASE 2

and $\alpha_1 < \alpha_2$. The lower bound z_{min} can be computed from each of these subzones:

- $z_{min3.1} = \max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$
- $z_{min3.2} = \max(o^-_{2,1}(\beta_1), o^-_{2,2}(\beta_2))$

which gives $z_{min3.1} = o^-_{2,2}(\alpha_2)$ and $z_{min3.2} = o^-_{2,2}(\beta_2)$. Thus, we obtain $z_{min3.1} = z_{min3.2}$ because $\alpha_2 = \beta_2$.

Let us now consider the upper bound z_{max}:

- $z_{max3.1} = \min(o^+_{1,1}(\beta_1), o^+_{1,2}(\alpha_1))$
- $z_{max3.2} = \min(o^+_{1,1}(\alpha_1), o^+_{2,1}(\beta_1))$

Similarly $z_{max3.1} = o^+_{1,1}(\beta_1)$ and $z_{max3.2} = o^+_{1,1}(\alpha_1)$. Since $\alpha_1 = \beta_1$, we have $z_{max3.1} = z_{max3.2}$.

Thus, the inferred output is continuous between area 3.1 and area 3.2. In the same way, we can show that transitions from areas (3.2,3.3), (3.3,3.4) and (3.4,3.1) are continuous.

Furthermore, for the single point at the intersection of several areas, continuity is also guaranteed. Indeed, this point has levels $\alpha_1 = \alpha_2 = \beta_1 = \beta_2 = \frac{1}{2}$. The lower bound is equal to $o^-_{2,2}(\frac{1}{2})$ for all areas and the upper bound is equal to $o^+_{1,1}(\frac{1}{2})$.

This proves that the inferred output is continuous all through area 3. Since area 2 and area 1 are just particular cases of area 3, the output is also continuous in these zones.

E. Extremal points

However, we need a continuous and monotonic output to be sure that the result of the output is the convex envelope of outputs inferred from rectangular input boundaries. In the sequel, we prove that the output boundary functions defining the set-valued output are not always monotonic and we detect extremal points that need to be considered. An extremal point is typically obtained if the two local functions defining an output bound (table IV) evolve in opposite directions.

For example, in figure 19, an extremal point appears at the lower bound in zone 3.1. In this area, the lower bound is equal to $\max(o^-_{1,2}(\alpha_1), o^-_{2,2}(\alpha_2))$, where $o^-_{1,2}(\alpha_1)$ increases and $o^-_{2,2}(\alpha_2)$ decreases. Thus, there is an extremal point when $o^-_{1,2}(\alpha_1) = o^-_{2,2}(\alpha_2)$. As we know fuzzy sets $O_{1,2}$ and $O_{2,2}$, we can easily find the α_1 level that corresponds to this extremal point.

For each zone, an extremal point can appear on only one bound as we can see on table VI. When necessary, we split the non monotonic output in order to restrict ourself to monotonic outputs.

The complexity analysis can be done in a similar way to the one dimensional case. All steps described for one dimension still hold for each input. One additional step is needed: extremal point detection. This operation requires two tests per α-cut. Thus the two dimensional algorithm has complexity $O(n)$.

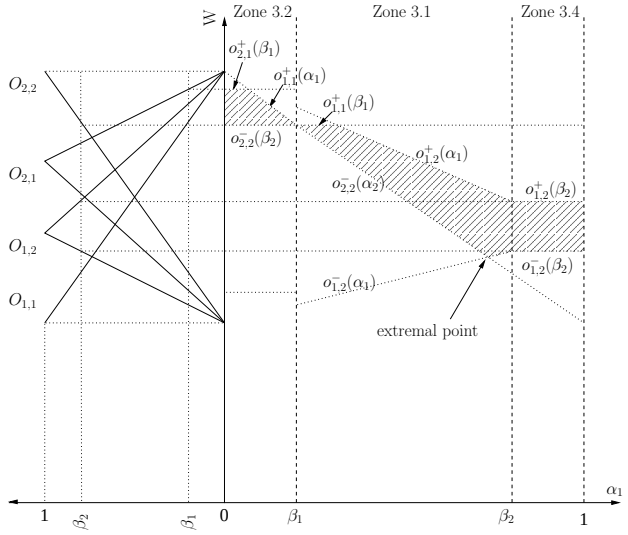


Fig. 19. Output evolution according to α_1 level

Zone	lower bound	upper bound
3.1	$o^-_{1,2}(\alpha_1) = o^-_{2,2}(\alpha_2)$	No
3.2	$o^-_{2,1}(\beta_1) = o^-_{2,2}(\beta_2)$	No
3.3	No	$o^+_{1,1}(\alpha_1) = o^+_{2,1}(\alpha_2)$
3.4	No	$o^+_{1,1}(\beta_1) = o^+_{1,2}(\beta_2)$

TABLE VI
CONDITIONS FOR EXTREMAL POINTS ACCORDING ZONE

F. Comparison with a naive sampling procedure

To demonstrate the efficiency of the proposed algorithm, we now give some results comparing it with inference from a naive sampling of the support. Input partitions and fuzzy inputs are shown on Figure 20. The chosen fuzzy inputs are symmetric triangles having a reasonable width with respect to the partition fuzzy sets. Figure 20 also displays the output partition and the inference results, for 1000 naive samples (*reference output*), 6 α-cuts and 6 naive samples. The rules are the ones given in Table II, with $O_{1,2} = O_{2,1}$.

Table VII summarizes the comparison between our algorithm, based on α-cut decomposition, and a naive sampling strategy. For each row, the number given in the first column is either the number of α-cuts or the sample size. For the α-cut based algorithm, the first α-cut is of level 1, and the following ones are regularly spaced in the unit interval. For

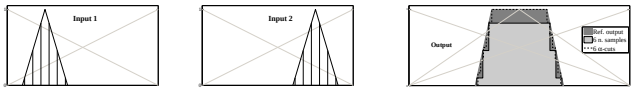


Fig. 20. Comparison of naive sampling and α-cut sampling

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

n	max (naive s.)	%area (naive s.)	%area (α-cut)
2	0.1	11.8	88.5
3	1	81.1	92.3
4	0.7	66.3	94.2
5	1	90.4	95.4
6	0.82	79.2	96.2
7	1	93.8	96.7
8	1	85.2	97.2
10	0.9	88.3	97.7
15	1	97.2	98.5
20	0.953	94.2	98.9
50	0.982	97.6	99.5
60	0.985	99.6	99.8

TABLE VII
COMPARISON OF NAIVE SAMPLING AND α-CUT SAMPLING

the naive sampling algorithm, samples are regularly spaced in the 0.1 level α-cut, and combinations of all samples are considered. For each row of the table, the second column, labeled max gives the maximum possibility degree of the output distribution for naive sampling. Obviously, this degree is not given for the α-cut based algorithm, as it is always equal to 1. The last two columns show the $\frac{inferred\ area}{reference\ area}$ ratio, the reference area being computed by taking 1000 alpha-cuts. Let us first point out that the complexity is not the same for the two algorithms. n α-cuts result in $2n+2$ strict inferences, while n naive samples require n^2 strict inferences. An examination of this table then shows that, for naive sampling, the maximum possibility degree does not have a monotonic behaviour when n increases, causing a non monotonic behaviour of the inferred output possibility distribution area. Furthermore, this phenomenon is amplified by the random handling of extremal points with the naive sampling. We also note that an accuracy of 95% is obtained with 5 α-cuts, i.e. 12 strict inference operations, while the same accuracy requires more than 20 naive samples, i.e. 400 operations. To conclude this discussion, we can say that the α-cut decomposition based algorithm provides an "intelligent" sampling by the means of α-cuts.

VII. nD INFERENCE

The extension to fuzzy implicative rule inference with high dimensional precise inputs is straightforward. In that case, the fuzzy rule base has a FITA (*First Infer then Aggregate*) behavior, and the inference result is given by Equation 9.

When dealing with fuzzy inputs, extending the approach is a complex task. Finding extremal points will be cumbersome in a high dimensional input space, so we suggest a working alternative.

1) We propose to use the same double decomposition method as in the two-dimensional case. In the sequel, we denote a rule as: $A_k \wedge B_l \wedge C_m \ldots \wedge Z_z \rightarrow O_{k,l,\ldots,z}$.
   - α-cut decomposition: decompose each fuzzy input into a union of rectangular inputs of level α, $0 < \alpha \leq 1$. This decomposition allows to consider each fuzzy input on each dimension as a set of α-level rectangular inputs. α is identical in all dimensions.
   - Partitioning decomposition: For each rectangular α-cut, a decomposition is made according to the different parts of the partition in order to handle the inference process locally.

   Consequently, the inferred output is again the result of a double union (similar to equation 18), where $O'^{\overline{\alpha}}_{k,l,\ldots,z}$ is the inferred output resulting from inputs $E_k \cap A_\alpha$, $E_l \cap B_\alpha \ldots E_z \cap Z_\alpha$.
2) Inferring from inputs within cores
   Whatever the space dimension, when the input data are located within a given core in all input dimensions, the result is the corresponding rule conclusion or its core for Rescher-Gaines implication. This is a consequence of inferential independence.
3) Inferring from inputs outside cores
   We propose an "intelligent sampling" for these subareas. All α-cut parts located in subareas outside cores will be approximated by means of a set of sample points, with corresponding membership grades to be used as weights. Inference will be done for all combinations of points, using inference for precise inputs as explained at the beginning of the section. The resulting weighted intervals will be merged using fuzzy union. In order not to miss extremal points, fine-grained sampling will be performed. Compared to "naive sampling", α-cut sampling will not miss the fuzzy input core part and the approximation will be better.
4) Coherence and interpretability
   Special care should be given to the output partition design. As 2^n rules are likely to be simultaneously fired, the partition may count 2^n overlapping fuzzy sets with a non empty intersection. This may harm the system interpretability, even with small values of n. Fortunately, as previously mentioned in the comment on figure 14, some of them may be identical.
   When considering higher dimensional rules, one must not forget the nature of implicative rules, i.e. that they represent constraints. Therefore, it is not always easy for an expert to express constraints simultaneously relating many variables.

VIII. ILLUSTRATION: DIAGNOSING A CHEESE-MAKING PROCESS

To show the interest of our method, we will consider a problem of predictive diagnosis for a hard-cooked type cheese-making process. Two parameters are important to determine cheese firmness: MC (Moisture Content), the cheese moisture content percentage at the end of the making process and DEE (Dry Extract Evolution), the loss of water during the first 15 days of the maturation process. The goal is to predict the cheese firmness at the end of maturation (4 to 10 months or longer) according to these two parameters. The two measurements (MC and DEE) come from sensors tainted with significant imprecision. So, we need to use fuzzy inputs in our system in order to correctly represent these

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

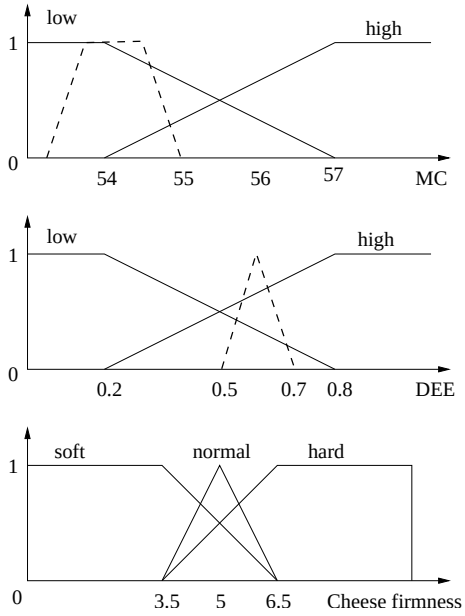


Fig. 21. Fuzzy sets for prediction of firmness - A fuzzy input is plot in dash lines

measurements. The firmness is a crisp real value ranging between 0 and 10, supplied by an expert sensory panel, and cannot be measured by a mechanical device. Input and output expert partitions are shown on figure 21. Let us refer to the typical output partition shown on Figure 14, we note that $O_{1,2}$ and $O_{2,1}$ are identical and represented by the fuzzy set *Normal*. Experts know some relations between MC, DEE and cheese firmness. This rule system is a simplified system that does not take into account the whole complexity of the process:

- If MC is high and DEE is low then the cheese will be soft
- If MC is high and DEE is high then the cheese will be normal
- If MC is low and DEE is low then the cheese will be normal
- If MC is low and DEE is high then the cheese will be hard

Some explanations follow. When the cheese is very wet, if it does not lose enough water, the cheese will be soft, but if it loses a lot of water, the cheese firmness will be normal. Similarly, if moisture content is low and if a lot of water is lost, the cheese will be hard.

A. Inference from a fuzzy input

1) Implicative rules: Fuzzy inputs are shown in dashed lines on figure 21. MC is modelled by a trapezoidal fuzzy set, due to two kinds of imprecision (sensor error plus calculation error) to take into account. DEE only suffers from sensor error. We apply our algorithm as follows:

- Alpha-cut decomposition: for this example, we decide to choose 3 α-cuts for the decomposition, as shown on figure 22.

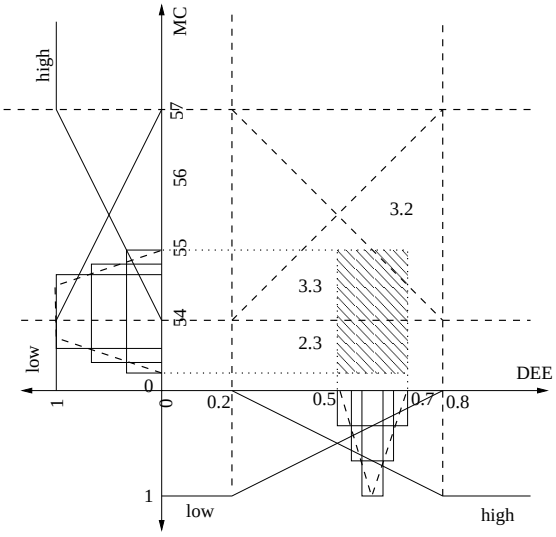


Fig. 22. Partitioning decomposition

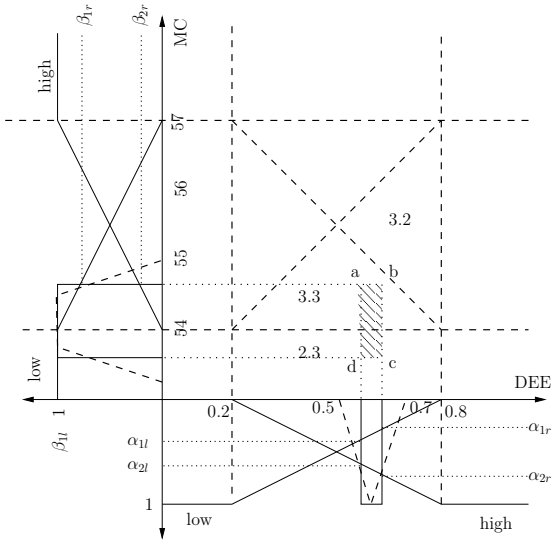


Fig. 23. Partitioning decomposition for a level 1 α-cut

- Partitioning decomposition: We decompose MC and DEE inputs according to partitions as shown on figure 22. We see the 3 zones activated by the corresponding values of $MC = 54.12 \pm 0.75$ and $DEE = 0.6 \pm 0.1$.
- Inference: for a two-dimensional α-cut rectangular input, we need to infer the four vertices a, b, c and d. We denote right and left rectangular input α levels by α_r, β_r and α_l, β_l. Figure 23 shows level 1 rectangular inputs on each dimension. Points a and b are in zone 3.3 and points c and d are in zone 2.3. The intervals inferred from each point are :
  - Point a: $[hard^-(\alpha_{2l}), soft^+(\beta_{2r})] = [5.3, 5.8]$. The interval is deduced from table IV. For example, the lower bound is $\max(o^-_{2,1}(\alpha_2), o^-_{2,2}(\beta_2)) = \max(hard^-(\alpha_{2l}), normal^-(\beta_{2r}) = hard^-(\alpha_{2l})$ in this case. The upper bound and the bounds of other intervals are similarly computed.

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

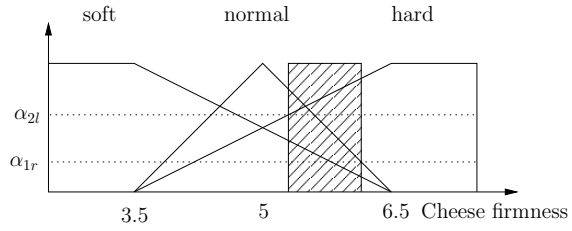


Fig. 24. Output for a level 1 α-cut

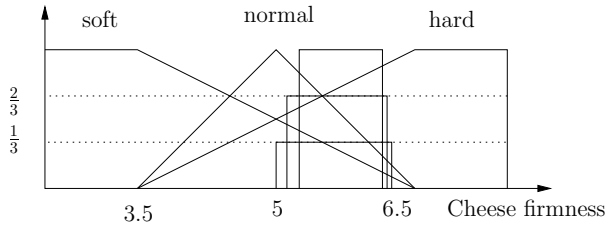


Fig. 25. Output union of the different α-cuts

- Point b: $[hard^-(\alpha_{2r}), soft^+(\beta_{2r})] = [5.8, 5.8]$
- Point c: $[hard^-(\alpha_{2r}), normal^+(\alpha_{1r})] = [5.8, 6.1]$
- Point d: $[hard^-(\alpha_{2l}), normal^+(\alpha_{1l})] = [5.3, 5.9]$

There are no extremal points within that zone. Indeed the point where $hard^-(\alpha_2)$ is equal to $normal^+(\alpha_1)$ is not in the range of variation of α_1 and α_2. Consequently, the level 1 output is the interval: $[hard^-(\alpha_{2l}), normal^+(\alpha_{1r})] = [5.3, 6.1]$ as we can see on figure 24.

In the same way, it is possible to compute inferred intervals for the other two α-level rectangular inputs.

- Final result: The final output result is the union of all α-level inferred outputs (see figure 25).

This example shows how the imprecision is propagated while being maintained within reasonable bounds through the inference process. The double decomposition gives a discrete approximation of the real output. The higher the number of α-cuts, the better the approximation. Let us point out that inferences for all α-cuts are exact. The approximation only concerns the input decomposition into α-cuts.

The inferred output interval may intersect several output fuzzy sets. If it belongs to a single fuzzy set, the inferred output is considered as precise. If it belongs to two fuzzy sets (*soft* and *normal* for example), it is considered as imprecise.

2) Conjunctive rules: The output obtained from Mamdani inference [32] using the same data is shown on figure 26. Note that the output partition is a strong partition. The inferred output overlaps the three output fuzzy sets. Consequently, it is difficult to interprete this result without defuzzification. Centroid defuzzification gives us a firmness equal to 6.0. Note that, as we saw on section II-C, defuzzification is influenced by the fuzzy sets shape. The imprecision of the fuzzy input is not respected in the defuzzified inference result.

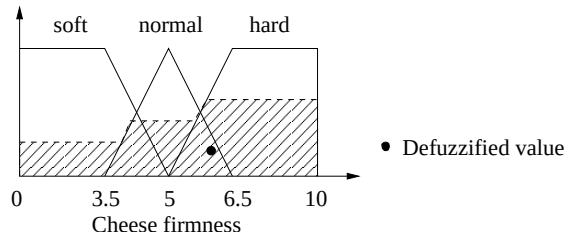


Fig. 26. Inference result with a conjunctive rule system

B. Numerical results

We tested the rule system given above with precise inputs for the two different kinds of rules. Since this example is a simplified rule system, the quality of the prediction is not very good. However, it is sufficient to demonstrate the difference between conjunctive rules and implicative rules. A representative sample of 103 cheeses was studied.

Inference results are analyzed at a symbolic level: The inferred output is considered as *good* if it mainly belongs to the same output fuzzy set than the crisp reference output,*wrong* otherwise.

1) Implicative rules:

- 33 wrong predictions
- 49 good but imprecise predictions, meaning that the inferred output contains the observed value but overlaps two output fuzzy sets.
- 21 good and precise predictions.

These results show a lot of imprecise predictions. This behavior was expected since the rule system is a simplified one. However, only 33 wrong predictions are made by this system. As we saw in section II-C, by adding more rules (and more input variables), the implicative rule system could be more precise and the output quality improved for 49 imprecise prediction.

2) Conjunctive rules:

- 56 wrong predictions
- 47 good and precise predictions

With conjunctive rules, there are many wrong predictions because of the defuzzification process.

Each inferred output is then an artificially precise value. With conjunctive rules it is impossible to refine the inference result because adding more rules will only increase the output imprecision because of the disjunctive agregation.

This example shows us the negative side effects of defuzzification. It also points out the ability of implicative rules to respect the input imprecision and thus to obtain a better prediction quality.

IX. Conclusion

This paper lays the foundation for a practical inference method with a system of implicative fuzzy rules and fuzzy inputs. For a fuzzy input, we can get an exact discretization of the result using α-cuts and a partitioning decomposition of inputs. Inferring with this kind of fuzzy system is especially

Author-produced version of the article published in IEEE Transactions on Fuzzy Systems, 2009, N°17(1), p.61-78.
The original publication is available at http://ieeexplore.ieee.org
Doi: 10.1109/TFUZZ.2008.2007851

appropriate when modeling expert knowledge expressing constraints (as opposed to Mamdani rules). The interest of the method has been shown on a simplified predictive diagnosis case-study of cheese production process, for which expert rules with two dimensional input conditions are available. In the future, more rules will be introduced to improve the results. Variables will also be added to refine the fuzzy rule systems according to needs. Nevertheless constraint management in highly dimensional spaces may be problematic, and be the true limit to the use of implicative rule systems in real world modelling. It might be better to consider an alternate way of dealing with larger systems while keeping in mind their interpretability: combination of various systems of lower dimension. Unlike conjunctive rule bases, implicative ones may be combined in either a parallel or a sequential way. In the former case, both rule bases use the same output universe and the result is their intersection: this is in full agreement with implicative rule agregation. In the latter case, the output is used to feed the next system. As the algorithm is able to manage fuzzy inputs, no defuzzification step is needed. Data imprecision is properly taken into account at all steps. More generally a perspective to this work is to relate higher dimensional fuzzy rule-based reasoning to possibilistic networks [33] where the idea of decomposition of a large fuzzy relation in lower dimension entities is at work.

REFERENCES

[1] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International journal on man-machine studies*, vol. 7, pp. 1–13, 1975.

[2] D. Dubois and H. Prade, "What are fuzzy rules and how to use them," *Fuzzy Sets and Systems*, vol. 84(2), pp. 169–185, 1996.

[3] J. F. Baldwin and N. C. F. Guild, "Modelling controllers using fuzzy relations," *Kybernetes*, vol. 9, pp. 223–229, 1980.

[4] A. D. Nola, W. Pedrycz, and S. Sessa, "An aspect of discrepancy in the implementation of modus ponens in the presence of fuzzy quantities," *Int. J. Approx. Reasoning*, vol. 3, no. 3, pp. 259–265, 1989.

[5] J. Mendel, "Fuzzy logic systems for engineering: a tutorial," in *Proc. IEEE*, vol. 83, 1995, pp. 345–377.

[6] H. Jones, S. Guillaume, B. Charnomordic, and D. Dubois, "Practical use of fuzzy implicative gradual rules in knowledge representation and comparison with mamdani rules," in *EUSFLAT, Barcelone*. EUSFLAT, 2005.

[7] D. Dubois and H. Prade, "Gradual inference rules in approximate reasoning," *Information Sciences*, vol. 61, pp. 103–122, 1992.

[8] S. Galichet, D. Dubois, and H. Prade, "Imprecise specification of ill-known functions using gradual rules." *Int. J. Approx. Reasoning*, vol. 35, no. 3, pp. 205–222, 2004.

[9] L. Ughetto, D. Dubois, and H. Prade, "Implicative and conjunctive fuzzy rules - A tool for reasoning from knowledge and examples," in *AAAI-99, Orlando, Floride (USA)*. California: AAAI Press/The MIT Press, 1999, pp. 214–219.

[10] N. Karnik, J. Mendel, and Q. Liang, "Type-2 fuzzy logic systems," *IEEE Transactions on Fuzzy Systems*, vol. 7, no. 6, pp. 643 – 658, 1999.

[11] D. Dubois, H. Prade, and L. Ughetto, "A new perspective on reasoning with fuzzy rules," *International Journal of Intelligent Systems*, vol. 18, no. 5, pp. 541–567, 2003.

[12] L. A. Zadeh, "A theory of approximate reasoning," *Machine Intelligence*, vol. 9, pp. 149–194, 1979.

[13] D. Dubois, H. Prade, and M. Grabisch, "Gradual rules and the approximation of control laws," *Theoretical aspects of fuzzy control*, pp. 147–181, 1995.

[14] R. Martin-Clouaire, "Semantics and computation of the generalized modus ponens: The long paper," *International journal of Approximate Reasoning*, vol. 3, pp. 195–217, 1987.

[15] E. Trillas and L. Valverde, "On some functionally expressible implications for fuzzy set theory," in *Proc. Third Internat. Seminar on Fuzzy Set Theory*, E. Klement, Ed., Linz, Austria, 1981, pp. 173–190.

[16] D. Dubois, H. Prade, and L. Ughetto, "Fuzzy logic, control engineering and artificial intelligence," in *Fuzzy Algorithms for control*, H. Verbruggen, H.-J. Zimmermann, and R. Babuska, Eds. Dordrecht, Pays-Bas: Kluwer, 1999, pp. 17–57.

[17] B. Schott and T. Whalen, "Nonmonotonicity and discretization error in fuzzy rule-based control using coa and mom defuzzification," in *Fifth IEEE International Conference on Fuzzy Systems*, 1996, pp. 450–456.

[18] E. Van Broekhoven, "Monotonicity aspects of linguistic fuzzy models," Ph.D. dissertation, University of Gent, Belgium, March 30,2007.

[19] D. Dubois, H. Prade, and L. Ughetto, "Checking the coherence and redundancy of fuzzy knowledge bases," *IEEE Transactions on Fuzzy Systems*, vol. 5, pp. 398–417, 1997.

[20] D. Dubois, R. Martin-Clouaire, and H. Prade, *Practical computing in fuzzy logic*. Fuzzy Computing, 1988, pp. 11–34.

[21] L. Ughetto, D. Dubois, and H. Prade, "Efficient inference procedures with fuzzy inputs," in *FUZZY-IEEE'97*, Barcelona, 1997, pp. 567–572.

[22] M. Sugeno and T. Takagi, "Multi-dimensional fuzzy reasoning," *Fuzzy Sets and Systems*, vol. 9, pp. 313–325, 1983.

[23] D. Dubois and H. Prade, "Basic issues on fuzzy rules and their application to fuzzy control," in *Fuzzy Logic and Fuzzy Control-Lecture Notes in Artificial Intelligence*, R. A. Driankov D., Eklund P.W., Ed. Springer-Verlag, 1994, vol. 833, pp. 3–14.

[24] L. Koczy and K. Hirota, "Interpolative reasoning with insufficient evidence in sparse fuzzy rule bases," *Information Sciences*, vol. 71, pp. 169–201, 1993.

[25] S. Jenei, E.-P. Klement, and R. Konzel, "Interpolation and extrapolation of fuzzy quantities revisited - an axiomatic approach," *Soft Computing*, vol. 5(3), pp. 179–193, 2001.

[26] ——, "Interpolation and extrapolation of fuzzy quantities - the multiple-dimensional case," *Soft Comput.*, vol. 6(3-4), pp. 258–270, 2002.

[27] L. Ughetto, "Inferential independence of fuzzy rules." in *ECAI*, 1998, pp. 605–609.

[28] N. N. Morsi and A. A. Fahmy, "On generalized modus ponens with multiple rules and a residuated implication," *Fuzzy Sets Syst.*, vol. 129, no. 2, pp. 267–274, 2002.

[29] N. N. Morsi, "A small set of axioms for residuated logic," *Inf. Sci.*, vol. 175, no. 1-2, pp. 85–96, 2005.

[30] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, 1988.

[31] ——, "Rough fuzzy sets and fuzzy rough sets," *International Journal of General Systems*, vol. 17, pp. 191–209, 1990.

[32] S. Guillaume, B. Charnomordic, and J.-L. Lablée, *FisPro: An open source portable software for fuzzy inference systems*, 2002.

[33] J. Gebhardt and R. Kruse, "Background to and perspectives on possibilistic graphical models," in *Applications of Uncertainty Formalisms*, ser. Lecture Notes in Computer Science, A. Hunter and S. Parsons, Eds., vol. 1455. Springer, 1998, pp. 397–414.