Practical use of fuzzy implicative gradual rules in knowledge representation and comparison with Mamdani rules

Hazaël Jones
Cemagref, UMR ITAP
INRA, UMR ASB
Montpellier
France
jones@ensam.inra.fr

Serge Guillaume
Cemagref, UMR ITAP
361, rue JF Breton
34033 Montpellier
France
serge.guillaume@cemagref.fr

Brigitte Charnomordic
INRA, UMR ASB
2, Place Viala
34060 Montpellier
France
bch@ensam.inra.fr

Didier Dubois
IRIT-CNRS
118, Route de Narbonne
31062 Toulouse
France
didier.dubois@irit.fr

Abstract

Thanks to their ability to model natural language, fuzzy rules are very popular in expert knowledge representation. Mamdani fuzzy systems are widely used for process simulation or control. Nevertheless, fuzzy implicative rules, and especially gradual rules, provide another kind of knowledge representation, which can be very useful in approximate reasoning. In this paper, the two types of rules are compared according to their behavior in some typical situations such as rule interpolation, combination of a specific rule with a more generic one. The comparison is carried out with regard to the output possibility distribution, the crisp inferred value and the rule base consistency. Finally, we discuss the complemental aspects of these rules and we show how in certain cases gradual rules may constitute an interesting alternative to Mamdani rules.

Keywords: Fuzzy logic, Conjunctive rules, Mamdani rules, implicative rules, gradual rules, interpolation

1 Introduction

Fuzzy logic has proven to be a powerful tool to design knowledge based linguistic models, where the domain knowledge is translated into an initial structure and parameters. The model accuracy can be enhanced by using data driven search methods to tune or to automatically learn the structure and the parameters. Fuzzy controllers or fuzzy expert systems have been widely used in process control or modeling. They are most useful when modeling processes whose monitoring essentially relies on expert knowledge, due to a lack of available mathematical modeling. Fuzzy logic based decision support tools have an intrinsic explanatory power. It can be a very good reason for using them preferably to other techniques, when interpretability is at stake.

A fuzzy rule is generally written as : "If X is A then Y is B", where A and B are fuzzy sets encoding linguistic concepts. However, this common formulation may correspond to different semantics and different forms of reasoning.
Fuzzy rules can be split in two groups: conjunctive possibility rules and implicative rules [4]. The most common approach in fuzzy controllers or fuzzy expert systems, whether built from expert knowledge or learnt from data, is the Mamdani approach [5], that uses conjunctive possibility rules. Those rules can be formalized as a conjunction $\mu_A \wedge \mu_B$, and they model the occurrence of *possible* value pairs. Mamdani systems, and their counterpart zero-order Takagi-Sugeno systems, where the rule consequent is a scalar instead of a fuzzy set, are well suited to some situations where expert knowledge can be considered as an accumulation of *possible* value examples, also called *positive information*. The fuzzy inference system interpolates between these values to provide a prediction or control value.

Mamdani or Sugeno models do not apply to other cases where expert knowledge is available under a different form, corresponding to constraints or restrictions on possible values. Nevertheless, the need for this kind of knowledge representation is natural and appears in complex system modeling. For instance, in traditional food industry processes, such as cheesemaking or winemaking, many processes are composed of a sequence of several unit operations, with domain knowledge and data available for each unit.Quality prediction or defect detection applies to the final product, and decision tools must take in account all usable forms of knowledge, and be designed accordingly. For instance, traditional hard cooked type cheesemaking includes three major steps: coagulation of the milk, draining of the curd (20 hour processing) and eventually ripening of the young cheese obtained after draining (it may take several months). The intermediate product (young cheese) is characterized by some physico-chemical measurements, such as moisture content or pH.

The domain expert can formulate rules to express the relationship between these characteristics and the final quality, or the plausible appearance of a defect in the matured cheese, a few months later. These rules do not express a simple interpolation operation, and some of them must have the ability to express restrictions, in order to be instrumental when making a context driven choice between different possible situations. For instance the following rule applies: “If Moisture Content is very low then add less salt at the beginning of the ripening stage to avoid Corky Consistency”.

In this context, the second kind of fuzzy rules mentioned higher up becomes necessary. Implicative rules are modeled as $\mu_A \rightarrow \mu_B$, where $\rightarrow$ is a fuzzy implication connective. Such rules express constraints on input-output value mappings. Therefore some output values become impossible in a given context, contrary to conjunctive possibility rules, where an input-output mapping can always be derived. Implication rules represent what is called *negative information*[4].

Possibility theory allows to represent information by means of possibility distributions[7, 1]. A possibility distribution assigns to each value $u \in U$ a possibility degree $\pi_x(u)$ lying between 0 and 1. $\pi_x(u) = 1$ means that nothing prevents x from being equal to u. Another useful notion is the concept of guaranteed possibility $\delta_X(u)$. $\delta_X(u) = 1$ means that $x = u$ has been actually observed. δ is a measure of evidential support. Furthermore, possibility values equal to zero have a very different meaning depending on the rule type, either ignorance ($\delta_X(u) = 0$) or interdiction($\pi_x(u) = 0$). When there is both *negative* and *positive* information, both sources of information must be consistent: for instance an observed fact should not be stated as impossible. Using the information in the output possibility distributions, a defuzzified output value can be extracted, and used as the system crisp output.

Despite the need for them and though the formalism of implicative rules has been well established, they have not yet been much used. The purpose of this paper is to highlight specific features of implicative rules that make them more attractive than Mamdani rules in practical reasoning situations. First we recall in section 2 the principles of conjunctive possibility rules, then we introduce in section 3 the formalism of implicative rules, and we detail the particular case of gradual implicative rules. Section 4 compares the behavior of implicative and conjunctive possibility rules in terms of output possibility distribution, inference with crisp values, and interpolation behavior. In complex systems, all inputs are not present in all rules. They may include rules where some input variables are missing. We study the case of such a rule base including an incomplete rule. Finally we discuss some perspectives.

2 Conjunctive possibility rules

For a given variable X, a guaranteed possibility distribution associated to statement “$x \in A_i$ is possible” is such that:

$$\forall u \in U, \delta_X(u) \geq \mu_{A_i}(u)$$

Conjunctive possibility rules[2], “if X is A then Y is B”, can be understood as: “the more X is A, the more possible it is that Y lies in B”. In this approach, the operator “then” is modeled by a conjunction and the output of the rule is a guaranteed possibility distribution: $\delta_{Y|X} = \mu_A \wedge \mu_B$.

The traditional Mamdani conjunction operator is the min: $\forall (u, v) \in U \times V$,

$$\delta_{Y|X}(u,v) \geq \min(\mu_A(u), \mu_B(v))$$

The meaning of $\delta_{Y|X}(u,v)$ is: it is possible that Y is B when X is A at least to level $\min(\mu_A(u),\mu_B(v))$ (maximum of specificity).

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

Rule aggregation is disjunctive. As a rule yields a guaranteed possibility degree, when two or more rules are fired, all the corresponding degrees are guaranteed. The maximum represents a lower bound of possible values:

$$\delta_K = \max_{i=1,\dots,n} \delta^i_{Y|X} \quad (1)$$

for a knowledge base $K = \{A_i \rightarrow B_i, i = 1, ..., n\}$ of n parallel fuzzy rules[1]. $\delta_{Y|X}(u,v) = 0$ means that if $X = u$, no rule can guarantee that v is a possible value for Y. Ignorance is then represented by a null possibility distribution: $\delta_{Y|X}(u,v) = 0, \forall v$.

3 Implicative gradual rules

Implicative rules are a straightforward application of Zadeh's theories[8] of approximate reasoning. According to Zadeh, each piece of knowledge can be considered as a fuzzy restriction on a set of possible worlds. The statement " X is A_i " can be depicted as:

$$\forall u \in U, \pi_X(u) \leq \mu_{A_i}(u) \quad (2)$$

" X is A_i " now means: " X must be in A_i ", it represents a constraint, i.e., negative information. Fuzzy implicative gradual rules can extend classical logic by means of the generalized modus ponens (GMP)[6]. In classical logic, modus ponens is:

$$A \wedge (A \rightarrow B) \models B$$

where $\models$ represents the logical inference. In fuzzy logic, the generalized modus ponens gives:

$$A' \wedge (A \rightarrow B) \models B' \quad (3)$$

[1]Parallel rules have the same input space U and the same output space V.

It means that for a fact A', we are able to deduce a value B' through the implication $A \rightarrow B$. B' is the upper bound of possible values for Y. For an input A', the output value B' is given by:

$$\mu'_B(v) = \sup_{u \in U} \wedge(\mu_{A'}(u), \mu_A(u) \rightarrow \mu_B(v)) \quad (4)$$

The conjunction and implicative operators cannot be chosen independently. Equation 3 and 4 show they are interrelated. Consequently, the choice of $\wedge$ determines the implication operator.

There are several implicative rule types [2]. The main ones are: gradual rules and certainty rules. In this article, we only focus on gradual rules. Implicative gradual rules can be understood as: " The more X is A, the more Y is B ". When $\wedge = \min$, $\rightarrow$ is the Gödel implication:

$$a \rightarrow b = \begin{cases} 1 & \text{if } a \leq b \\ b & \text{else} \end{cases}$$

As we can see on figure 1, for a rule $A \rightarrow B$ and a crisp input $u_0 \in Supp(A)$, the core of B' is always larger than the core of B. Actually, this type of reasoning is driven by similarity: if the value of X is close to the core of A, then the value of Y must be close to the core of B.

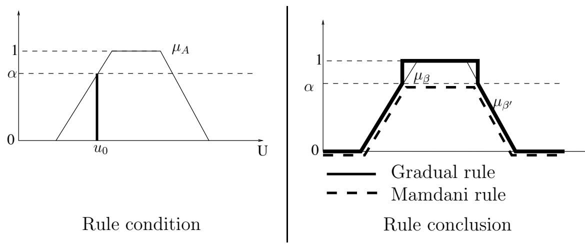

Figure 1: Inference with one rule and a crisp input

With this kind of rules, for a crisp input u_0, if $u_0 \notin Core(A)$, $Core(B')$ becomes larger than $Core(B)$. If $u_0 \notin Supp(A)$, $Core(B')$ becomes the whole output space V: it expresses ignorance. Equation 2 shows that these rules express constraints on a set of possible values. Possibility is not guaranteed: some values considered as possible by a rule can be forbidden by others. Each rule can be represented by a conditional possibility distribution $\pi^i_{Y|X} = \mu_{A_i} \rightarrow \mu_{B_i}$. The possibility distribution π^K is given by the conjunction of $\pi^i_{Y|X}$:

$$\pi^K = \min_{i=1,\dots,n} \pi^i_{Y|X}$$

For a possibility degree $\pi^K(u,v) = 0$, if $X = u$, then v is an impossible value for V. Furthermore, a possibility distribution $\pi^K(u,v)$ uniformly equal to 1 symbolises ignorance.

4 Rule base comparison

In this section, the different rules are compared according to different criteria: inference mechanism, interpolation ability and rule accumulation. Rule base behavior in the presence of incomplete rules is also examined.

4.1 Inference mechanism

The use of the sup-min composition in equation 5 can be explained by the assumption that once built, the relation obtained by the disjunction of conjunctive rules is interpreted as a whole as a model. The fuzzy system is then viewed as a constraint, using a kind of closed world assumption, although each rule was originally interpreted as a piece of imprecise data.

The output possibility distribution can be computed for each rule and then all the distributions are aggregated according to equation 1:

$$B' = A'o(\bigcup_{i=1}^{n} A_i \wedge B_i) = \bigcup_{i=1}^{n}(A'o(A_i \wedge B_i)) \quad (5)$$

because of the commutativity of o and $\bigcup$ operators. In consequence, with Mamdani rules, inference is quite easy. This method, named FITA[2], corresponds to the right part of equation 5.

For an implicative rule base, the output B' is given by:

$$B' = A'o\bigcap_{i=1}^{n} A_i \rightarrow B_i$$

When A' is a crisp input, the FITA mechanism can be used but, in the general case, as the o and $\bigcap$ operators do not commute, the previous equality no longer stands:

$$B' = A'o(\bigcap_{i=1}^{n} A_i \rightarrow B_i) \subseteq \bigcap_{i=1}^{n}(A'oA_i \rightarrow B_i)$$

[2]FITA means "First Infer Then Aggregate"

Rule aggregation has to be achieved before inference. This inference method is called FATI[3].

4.2 Interpolation between rules

The interpolation mechanism used for Mamdani rules is described in depth in [3]. Let us consider input/output partitions such as $Core(A_i) = \{a_i\}$ and $Supp(A_i) = [a_{i-1}, a_{i+1}]$.

Conjunctive possibility rules Figure 2 shows the output possibility distribution inferred by three Mamdani rules, $A_i \wedge B_i$ $(i = 1, 2, 3)$, when input u_0 moves from a_1 to a_2 (a): only truncation levels of B_1 and B_2 are affected (b). Usually, as the maximum of the distribution is an interval, a defuzzification step is 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.

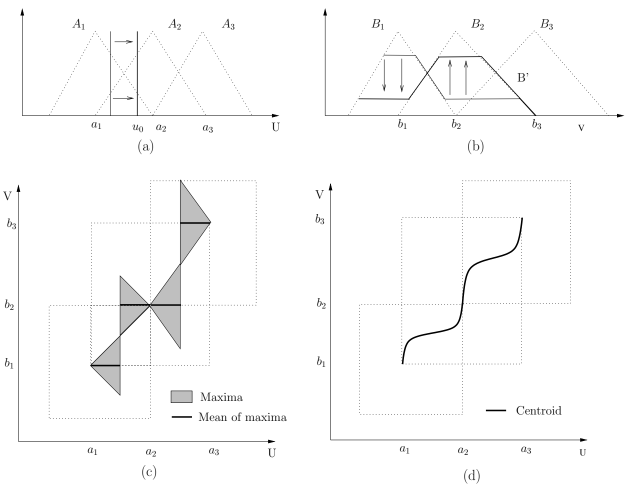

Figure 2: Interpolation with Mamdani rules

Gradual implicative rules Figure 3 illustrates the case of three gradual rules $A_i \rightarrow B_i (i = 1, 2, 3)$. 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.

[3]FATI means "First Aggregate Then Infer"

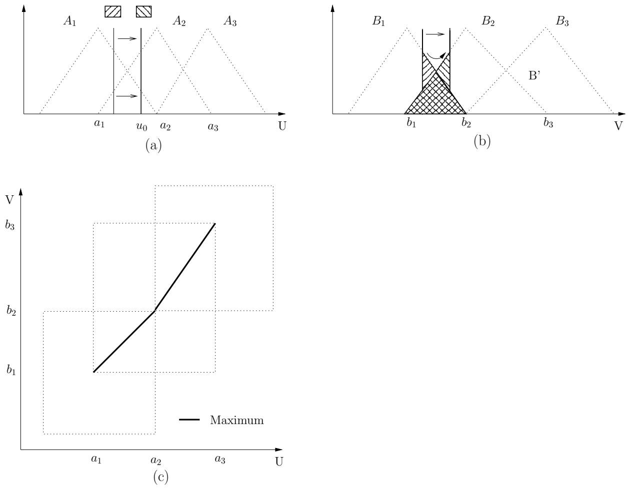

Figure 3: Interpolation with gradual implicative rules

4.3 Influence of the fuzzy set parameters

Let us consider two rules triggered at the same level.

Conjunctive possibility rules When two trapezoidal output fuzzy sets have equal widths, the inferred value (mean of maxima or centroid) is equal to y such as $\mu_{B_1}(y) = \mu_{B_2}(y)$. 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 4.

Gradual implicative rules This behavior is impossible with gradual implicative rules because rules are aggregated in a conjunctive way.

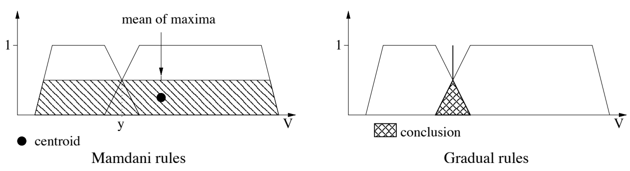

Figure 4: Parameter influence

4.4 Incomplete rules

This section examines the interesting case of a rule base which includes incomplete rules. Let us consider a system of two rules, with a specific rule: "if X is A_1 and if Y is B, then Z is C_1"and a more general one: "if X is A_2, then Z is C_2". In the following examples (figures 5 and 6), an input value b is set for Y such as $\mu_B(b) = \beta$ and the X input value varies from a_1 to a_2. Let a' the U value such as $\mu_{A_1}(a') = \beta$ and c' the Z value such as $\mu_{C_1}(c') = \beta$.

Conjunctive possibility rules Figure 5 shows the output possibility distribution with a centroid defuzzification. Between a_1 and a', the C_1' output level is truncated to level β, while the C_2' level increases. Then, when the input moves from a' to a_2, B is immaterial for the inference process. Subfigure (c) shows the centroid defuzzification result. The dashed line recalls the inference result after defuzzification with two rules $A_1 \wedge C_1$ and $A_2 \wedge C_2$. The specific rule inhibits the output variation: in the illustrated case, it is impossible to reach the C_1 value.

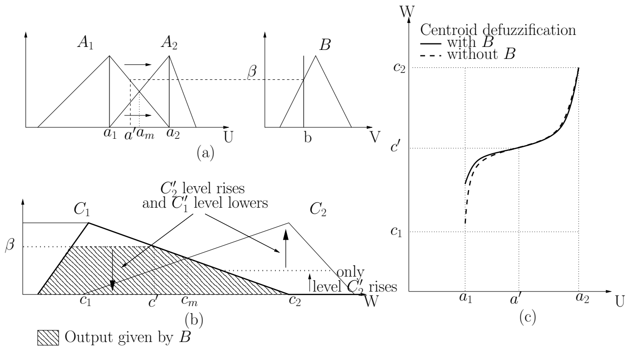

Figure 5: Conjunctive possibility rules: inference with an incomplete rule

Gradual implicative rules When a varies from a_1 to a', the influence of the specific rule is restricted by β. This leads to more imprecision: the maximum value of the possibility distribution is not unique anymore. It corresponds to the interval: $[c_{inf}, c']$, c_{inf} such as $\mu_{C_1}(c_{inf}) = \mu_{A_1}(a)$. The more the input value is close to a', the narrower the interval. Applying a mean of maxima defuzzification operator gives an output similar to the one given by Mamdani rules, but here piecewise linear. Between a' and a_2, B is immaterial for the inference process, the maximum is unique.

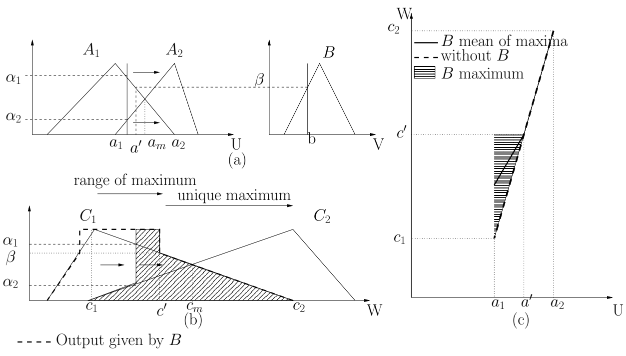

Figure 6: Implicative gradual rules: inference with an incomplete rule

4.5 Rule accumulation

Adding a conjunctive rule enlarges the output possibility distribution. Then a rule system will always have a solution even if the rule base conflicts with knowledge representation. If an infinity of rules is 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 could think that adding new rules to the knowledge base would lead to a more accurate system.

Implicative rules formulate constraints on possible input/output mappings. The more rules in a rule base, the more precise the output fuzzy set at the risk of reaching inconsistency. Inconsistency arises when for a given input $\pi_{Y/X}(u, v) = 0, \forall v$. This feature is interesting because it allows to check logical consistency of the rule base.

5 Conclusion

Mamdani rules are widely used in knowledge representation even if in some typical situations the system may exhibit an unexpected behavior. Gradual rules, which are implicative rules, may constitute an interesting alternative to Mamdani rules as they also provide an interpolation between rule conclusions.

Moreover, as the resulting possibility distribution is the upper bound of the authorized values according to the knowledge base, a null degree means the corresponding value is impossible. Thus, this kind of rules can be used to model range restrictions.

A related topic is to find the proper partition of the input and output spaces, which ensures a good behavior of the rule base. For instance one may require that if the input is equal to the condition of one rule, the output should be precisely the conclusion of this rule, despite the presence of other rules. This behavior is generally impossible to observe using Mamdani systems.

References

[1] Didier Dubois and Henri Prade. Fuzzy sets in approximate reasoning. *Fuzzy Sets and Systems*, 40(1):143–202, 1991.

[2] Didier Dubois and Henri Prade. What are fuzzy rules and how to use them. *Fuzzy Sets and Systems*, 84(2):169–185, 1996.

[3] Didier Dubois, Henri Prade, and Laurent Ughetto. Fuzzy logic, control engineering and artificial intelligence. In H.B. Verbruggen, H.-J. Zimmermann, and R. Babuska, editors, *Fuzzy Algorithms for control*, pages 17–57. Kluwer, Dordrecht, Pays-Bas, 1999.

[4] Didier Dubois, Henri Prade, and Laurent Ughetto. A new perspective on reasoning with fuzzy rules. *International Journal of Intelligent Systems*, 18(5):541–567, 2003.

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

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

[7] L A Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1(1):3–28, 1978.

[8] L A Zadeh. A theory of approximate reasoning. *Machine Intelligence*, 9:149–194, 1979.