European Centre for Soft Computing

# Why do fuzzy representations need a careful design?

Enric Trillas

Research Report nº: FSC 2008-6

# Why do fuzzy representations need a careful design?

Enric Trillas

*European Centre for Soft Computing*

This paper tries to show, only from a theoretical perspective, the importance of well designing the representation of fuzzy systems whose behaviour is known by a linguistic description of it. The process of designing the representation by means of fuzzy sets, connectives, and relations, marks an actual distinction between the fuzzy and the formal logic methodologies , two different disciplines whose agendas are not coincidental.

## 1. Introduction

In the words of John von Neumann,[1] "formal logic ... is one of the technically most refractory parts of mathematics. The reason for this is that it deals with rigid, all-on-one concepts, and has very little contact with the continuous concept of the real or of the complex number, that is, with mathematical analysis. Yet analysis is the technically most successful and best-elaborated part of mathematics." Fuzzy logic responded to the demand of von Neumann hidden in this paragraph, and thanks to the use of mathematical analysis is that in fuzzy logic there appear thresholds like, for example, that of contradiction of a fuzzy set.

Given $\mu \in [0,1]^X$ it is said that $\mu$ is self-contradictory if $\mu \leq \mu'$, $\mu(x) \leq n$, for all $x \in X$, with $n \in (0,1)$, the fixed point of the strong negation with which $\mu' = N \circ \mu$. Then, when $X$ is a closed interval of the real line and, e.g., $\mu$ is non-decreasing, the number $x_0 = \sup\{x; \mu(x) \leq \mu'(x)\}$ is the *threshold of contradiction* of $\mu$. For example, if $B = big$ in $[0,10]$ is given by $\mu_B = \frac{x}{10}$, with $N(x) = \frac{1-x}{1+x}$, it follows $n = \sqrt{2} - 1$, and $\frac{x}{10} \leq \sqrt{2} - 1$, or $x \leq 10\sqrt{2} - 10 = 4.142$. The threshold of contradiction is 4.412, that is,

- If $x \leq 4.142$, it is $\mu_B(x) \leq \mu_{not\ B}(x)$
- If $x > 4.142$, it is $\mu_{not\ B}(x) < \mu_B(x)$,

hence, only after 4.142 the degree up to which '$x$ is $B$' is greater than the degree up to which '$x$ is not $B$'. Only the points in $(4.142, 10]$ can be properly called *big*. The value 4.142, the separation point[2] of big and not big, seems to dissolve the Sorites Paradox.[3] With the different representation of $B$ given by

$$\mu_B(x) = \begin{cases} 0, & \text{if } 0 \le x \le 3 \\ x - \frac{3}{5}, & \text{if } 3 \le x \le 8 \\ 1, & \text{if } 8 \le x \le 10 \end{cases}$$

and $N = 1 - id$, it results the threshold 5.5. Notice that with $N(x) = 1 - x$, it is $n = 5$, and with $N(x) = \frac{1-x}{1+2x}$ it is $n = 0.366$. Hence *to reach an approximate enough separation point it is crucial to have a good representation not only of $\mu_B$ but also of $N$*, and this representation only can come from the context and the actual uses of $B$ and not $B$.

Formal logic only deals with logical consequences, that is, with deductive reasoning. But fuzzy logic also deals with *conjectures*, that is, with conclusions that are non-contradictory with a given body of knowledge. Let us show it with what follows. In a boolean algebra it holds the *scheme of disjunctive reasoning* $\mu+\sigma, \sigma' : \mu$, since $(\mu+\sigma)\cdot\sigma' = \mu\cdot\sigma'+\sigma\cdot\sigma' = \mu\cdot\sigma' = \mu\cdot\sigma' \le \mu$. Hence, $\mu$ *follows from the body of knowledge* $\{\mu+\sigma, \sigma'\}$, $\mu$ is a *logical consequence* of it. If $\mu, \sigma$ are in $[0,1]^X$, the inequality $(\mu+\sigma)\cdot\sigma' \le \mu$ should be translated by

$$T(S(\mu(x), \sigma(x)), N(x)) \le \mu(x), \text{ for all } x \in X,$$

with a triplet $(T, S, N)$ verifying this inequality. The only such triplets are $(W_\varphi, W^*_\varphi, N_\varphi)$ (see[4]) and hence, in these cases $\mu$ is a logical consequence of $\{\mu+\sigma, \sigma'\}$, and in all the other cases what happens is not $(\mu+\sigma)\cdot\sigma' \le \mu$, but $(\mu+\sigma)\cdot\sigma' \not\le \mu'$, that is, $\mu$ is not contradictory with $\{\mu+\sigma, \sigma'\}$, $\mu$ is just a *conjecture* of $\{\mu+\sigma, \sigma'\}$. It only can be *induced* from this body of knowledge. For example, with $T = \min$, $S = \max$, $N = 1-\text{id}$, the inequality

$$\min(\max(\mu(x), \sigma(x)), 1 - \sigma(x)) \le 1 - \mu(x)$$

does not hold for all $\mu, \sigma$ in $[0,1]^X$, since with $\mu$ and $\sigma$ such that $\mu(x) = 1$ and $\sigma(x) = 0$, follows the absurd $\sigma(x) = 1$. Notice that, in this case, if $\mu \le \sigma$, it is $(\mu+\sigma)\cdot\sigma' \le \mu+\sigma = \mu$.

In formal logic it is almost always supposed that there is duality, that De Morgan laws do hold. But this is not the case in fuzzy logic since there are interesting cases where there is no duality. For example, the formula $(a\cdot b')' = b + a'\cdot b'$, is a law in boolean algebras though $(a\cdot b')' = a'+b$, and $b + a'\cdot b' = (b+a')\cdot(b+b') = (b+a')\cdot 1 = a'+b$. It is easy to show that this formula is neither a law in De Morgan Algebras, nor in orthomodular lattices , two structures with duality.

With fuzzy sets the presumed law es $(\mu \cdot \sigma')' = \sigma + \mu' \cdot \sigma'$, or

$$N(T(\mu(x), N(\sigma(x)))) = S(\sigma(x), T(N(\mu(x)), N(\sigma(x)))),$$

for all $x$ in $X$, and some triplet $(T, S, N)$. The only solutions of this functional equation[5] are the triplets $(Prod_\varphi, W^*_\varphi, N_\varphi)$ that do not verify $(\mu+\sigma)' = \mu' \cdot \sigma'$. Hence, $(\mu \cdot \sigma')' = \sigma + \mu' \cdot \sigma'$ is a law only in those *non-dual* fuzzy algebras.

The possibility of dissolving the Sorites Paradox, that of considering inductive reasonings, and that of not considering duality, seems to mark differences between the agendas of formal and fuzzy logic. Anyway, and because fuzzy logic applications deal many times with dynamic systems whose behaviour is known through a linguistic description of it, there are again more conspicuous differences with the methodology of formal logic. The fuzzy representation of a system is to be designed directly from its linguistic description.

## 2. On fuzzy representations

The algebra of fuzzy sets applicable to each problem is not always imported from other, let us say *generic* situations, like it is done with formal logic. It is not to be forgotten that the central subject fuzzy logic deals with is *meaning*, as conveyed by the *use* of linguistic terms, connectives, and relations appearing in the linguistic description of the system.

To pose a problem in, or with, fuzzy logic, it is basic to well understand its linguistic description and, consequently, to design its representation by means of adequate fuzzy sets, fuzzy modifiers, fuzzy connectives, and fuzzy relations. Because there are a lot of t-norms, t-conorms, strong negations, etc, the representation should be made by choosing them through a process of design according with the meaning of the involved linguistic elements. The practice of fuzzy logic does contain the art of designing fuzzy representations based on meaning.

To make a *description* is to describe something, that is, to present it by means of some expressions allowing to recognize, to picture, to newly build up,...., such something. A *representation* is a *new presentation* of that something by means a new form of expression, or specialized language, allowing to place in black and white the main characteristics it shows and that are relevant for some goal. That is, a representation is a model.

Consider the two statements: $p$ = It is false that John is not very tall, $q$ = It is false that John is not very short. Which one of them is more true?

Take *Heigh*(John)=$H(J) \in (0,2]$, in meters, and the membership functions $\mu_{tall} = \mu_t$, and $\mu_{short} = \mu_s$, from $(0,2]$ into $[0,1]$. As it is usual, suppose that $\mu_t$ is non-decreasing. Then, since *short* is an antonym of *tall*, it is $\mu_s(x) = \mu_t(2-x)$. As it is often usual, let us take *very*=$v$ as a non-decreasing function $[0,1] \to [0,1]$, like $v(x) = x^2$. Then,

- degree of truth($p$) = $\tau_f(\mu_{not-vt}(H(J))) = 1 - N(v(\mu_t(H(J))),$
- degree of truth($q$) = $\tau_f(\mu_{not-vs}(H(J))) = 1 - N(v(\mu_s(H(J))) = 1 - N(v(\mu_t(2 - H(J)))),$

for some strong negation $N$. To compare this two degrees, it is enough to compare $\mu_t(H(J))$ with $\mu_t(2-H(J))$. It is not restrictive to take $\mu_t(x) = 0$ for $x$ in $(0,1]$, and $\mu_t$ positive and non-decreasing for $x$ in $(1,2]$. Then

- If $0 < H(J)) \leq 1$ it is $1 \leq 2 - H(J) < 2$, and $\mu_t(H(J)) = 0 < \mu_s(H(J))$, and degree of truth $(p) <$ degree of truth $(q)$
- If $1 < H(J)) \leq 2$ it is $0 \leq 2 - H(J) < 1$, and $0 = \mu_s(H(J)) < \mu_t(H(J))$, and degree of truth $(q) <$ degree of truth $(p)$

Notice that to know which of two statements is more true it is not needed to know anything else than $\tau_f(x) = 1 - x$, $N$ is a strong negation, $v$ is non-decreasing, $\mu_t$ is non-decreasing and nul in $(0,1]$. But *for knowing the numerical degrees of p and q, we need to select N, v,* and $\mu_t$, that is, *to adequately represent* them. Different elections will give different truth values for $p$ and $q$, and not capturing correctly the meaning of *not , very*, and *tall*, can easily conduct to wrong numerical results, to solve a problem that is not that what is posed.

Let us show an example that reinforces what is just said at the end of 2.1. Consider the rule "If $x$ is small, then $y$ is big", with $X = Y = [0,10]$, $\mu_{small}(x) = \mu_S(x) = 1 - \frac{x}{10}, \mu_{big}(y) = \mu_B(y) = \frac{y}{10}, N = 1-$id, and the input $x_0 = 5$. The output is

$$\mu_{B^*}(y) = \sup_{x \in [0,10]} \min(\mu_5(x), J(1 - \frac{x}{10}, \frac{y}{10})) = J(\frac{1}{2}, \frac{y}{10}),$$

and what lacks to know $\mu_{B^*}$ is the representation $J$ of the $T$-conditional. With the following representations, or models, of $J$, it is obtained,

- Larsen's model, $J_1(1 - \frac{x}{10}, \frac{y}{10})) = (1 - \frac{x}{10})\frac{y}{10}$, $\mu_{B^*}(y) = \frac{y}{20}$
- Łukasiewicz's model, $J_2(1 - \frac{x}{10}, \frac{y}{10})) = \min(1, \frac{x+y}{10})$, $\mu_{B^*}(y) = \begin{cases} 1, & \text{if } y > 5 \\ \frac{5+y}{10}, & \text{if } y \leq 5 \end{cases}$

- Reichenbach's model $J_3(1-\frac{x}{10}, \frac{y}{10})) = \frac{x}{10}(1-\frac{y}{10})) + \frac{y}{10}$, $\mu_{B^*}(y) = \frac{1}{2} + \frac{y}{20}$
- Kleene-Dienes's model, $J_4(1-\frac{x}{10}, \frac{y}{10})) = \frac{1}{10}\max(x, y)$, $\mu_{B^*}(y) = \frac{1}{20}\max(1, \frac{y}{5})$

that are very different curves. Hence, the correct representation of the rule's meaning (the way it is currently used) is very important. The semantics of the problem, that is, how all the linguistic terms and the rule in itself are used, is crucial.

Notice that the system's behaviour is linguistically described by the rule, and to represent it is necessary to know what is meant by it. In the cases of Łukasiewicz, Reichenbach, and Kleene-Dienes, the meaning is '*Not* $x$ is *small*' *or* '$y$ is *big*', but in the Larsen's case is '$x$ is *small*' *and* '$y$ is *big*', that correspond to the two different protoforms $\mu \to \sigma = \mu' + \sigma$, and $\mu \to \sigma = \mu \cdot \sigma$. With the Early-Zadeh model, $J_5(a, b) = \max(1-a, \min(a, b))$, it results $\mu_{B^*}(y) = J_5(\frac{1}{2}, \frac{y}{10}) = \max(\frac{1}{2}, \min(\frac{1}{2}, \frac{y}{10})) = \frac{1}{2}\max(1, \frac{y}{5})$, again a different curve and that corresponds to a protoform of the type $\mu \to \sigma = \mu' + \mu \cdot \sigma$.

Of the more than forty functions $J$ that are used in fuzzy logic to represent the rules,[6] only those that are R-implications and S-implications are *implication functions*, that is, verify a set of properties directly imported from the boolean material conditional. Since, in general, $J$ is to make forwards inference, $J$ should verify the inequality of *Modus Ponens* $T(a, J(a, b)) \leq b$, for all $a, b$ in $[0, 1]$ and some continuous t-norm $T$.

In a boolean algebra, $a \cdot z \leq b$ is equivalent to $z \leq a' + b$. Hence, $\sup\{z; a \cdot z \leq b\} = a' + b$. In the fuzzy case, with a continuous t-norm $T$, it is $\sup\{z \in [0, 1]; T(a, z) \leq b\} = J_T(a, b)$, that is an R-implication, and a $T$-conditional since it verifies $T(a, J_T(a, b)) = \min(a, b) \leq b$. For any other $T$-conditional $J$, from $T(a, J(a, b)) \leq b$, for all $a, b$ in $[0, 1]$, it follows $J(a, b) \leq J_T(a, b)$, that is, $J_T$ *is the biggest* $T$*-conditional.* Nevertheless, since $J_T$ does not correspond with any protoform with the connectives $\cdot, +, '$, like it happens in boolean algebras, the election of a $J_T$ to represent a rule is a problem of a different nature because the translation of the rule's meaning by $J_T$ does not appear directly. Even more, because of $J \leq J_T$, it follows $T(c, J(a, b)) \leq T(c, J_T(a, b))$, and the output with $J$ is always smaller than that with $J_T$. Hence, there is always the risk of obtaining a too big output with $J_T$. For example, in the case of 2.2, with $T = \min$, the output is

$$J_{\min}(\frac{1}{2}, \frac{y}{10}) = \begin{cases} 1, & \text{if } 5 \leq y \\ \frac{y}{10}, & \text{if } 5 > y, \end{cases}$$

clearly bigger than the output $\frac{y}{20}$ obtained with the Larsen's model (also a min-conditional), and with a discontinuity in $x = 5$. By taking $T$ =prod (Larsen's is also a prod-conditional) the output is

$$J_{\text{prod}}(\frac{1}{2}, \frac{y}{10}) = \begin{cases} 1, \text{ if } 5 \leq y \\ \frac{y}{5}, \text{ if } 5 > y \end{cases}$$

also clearly bigger than $\frac{y}{20}$, but without any discontinuity. The case with R-implications $J_T$ is peculiar since is a representation that does not come from a protoform [a] directly obtained from what the rule means.

The function $J$, the $T$-conditional representing the linguistic rule, not always can be induced from the way the rule is used (its meaning, *à la* Wittgenstein[7]), but from some contextual information. For example, if it is known that '$J(a, b) = 1$ is equivalent with $a \leq b$', it should be $J = J_T$ for some $T$ to be determined. If '$J(a, b) = 1$ is equivalent to $a = b = 1$', it is $J(a, b) = T(a, b)$ for some $T$ to be determined. If $J(0, b) \neq 1$ it can be chosen $J(a, b) = T(a, b)$, with $T = W_\varphi$ if $J(0, b) > 0$.

Sometimes, what can be supposed is a law coming from the boolean case for $\rightarrow$. For example,[8] with the law $\mu \rightarrow (\sigma \rightarrow \mu) = \mu_1$, only R-implications and some S-implications can be used. With the law $(\mu \rightarrow \mu') \rightarrow \mu = \mu$, only S-implication $J(a, b) = \max(N(a), b)$ can be used. With the law $(\mu \rightarrow \mu') \rightarrow \mu' = \mu_1$, no operator of the basic known models can be used.

What is clear is that there is not a universal class of connectives. There are the particular features of the given problem, as well as those shown by the involved knowledge, than can eventually lead to choose the fuzzy connectives. The linguistic expressions describe a concrete physical system and mean something that should be correctly represented. The process of designing the fuzzy representation is essential for not solving a problem different from that currently considered.

## 3. A case-example of representation.

In the novel "Desde la ciudad nerviosa" (From the nervous city) of 2004, by the Spanish writer Enrique Vila-Matas ([9]), it appears the paragraph

> *"I had always told myself that if life has no sense neither has reading, but suddenly it seemed to me that the process of reading to*

[a] In algebraic logic, a conditional is called *material* when it is expressed by a formula with connectives. In this sense, in boolean algebras, also $a \rightarrow b = a \cdot b$ is material. We preferred to call them *expressible by a protoform.*

> *search for artists of the not, did have a lot of sense. Unexpectedly, I felt that the search for bartlebys gave sense to my life"*,

here translated from Spanish into English. Hidden in this paragraph is the reasoning,

> If "*If life has no sense neither has reading*" then "*If reading has sense either has life*", (∗∗), that indeed involves three conditionals.

To know if it is contextually true, how to model it? Let us start from the following working hypotheses,

1. "Sense" is gradable. If not, why "*have a lot of sense*"?
2. The author seems to be tolerant with truth. If not, why (1)?
3. The author distinguishes between contrasymmetric conditionals, since he simultaneously considers "*If life has no sense, neither has reading*", and "*If reading has sense, either has life*".
4. The author does not consider the conditionals being expressed by connectives. He never writes, for example, "*Life has no sense, or reading has no sense*" instead of "*If live has no sense, neither has reading*".
5. The author is a passionate reader and his literary style is complex, as it is shown in all his books.

Then([10]),

a. Let us take the statements' degree in $[0, 1]$, with *not* represented by $N = 1 - id$.
b. Because of 5, $t_0$ =Degree (reading has sense) will be taken as a parameter that verifies $t_0 > 1 - t_0 \Leftrightarrow t_0 > 0.5$, and we can suppose that $t_0$ is big: $t_0 > 0.8$.
c. The value r=Degree (Life has sense) is taken as a variable.

Hence, for some $T$ conditional $J$, is $\alpha$=Degree (If life has no sense, neither has reading) $= J(1 - r, 1 - t_0)$ (see ([11])), and by (2), (3) and (4), it only will be considered the case of $J$ being an R- implication $J_T$ with $T \neq W_\varphi$.

Let us take $T = \min$ and $T = prod$. With $T = \min$, and $J_{\min}$ it results

$$\alpha = \begin{cases} 1, & \text{if } t_0 \leq r \\ 1 - t_0, & \text{if } t_0 > r, \end{cases}$$

which is non acceptable since it is not continuous, and only takes two values. With $T = prod$, and $J_{prod}$, it results

$$\alpha = \begin{cases} 1, & \text{if } t_0 \le r \\ \frac{1-t_0}{1-r}, & \text{if } t_0 > r, \end{cases}$$

a function that is never 0, distinguishes countersymmetry, and is not linear. Let us take this model.

Call, $\gamma$=Degree (Reasoning (∗∗)), with respective degrees $\alpha$ and $\beta$ for its components. With the same model than that accepted for $\alpha$, it is

$$\beta = \begin{cases} 1, & \text{if } t_0 \le r \\ \frac{r}{t_0}, & \text{if } t_0 > r. \end{cases}$$

Hence,

$$\gamma = \begin{cases} 1, & \text{if } 1 - t_0 \le r \\ \frac{r(1-r)}{t_0(1-t_0)}, & \text{if } 1 - t_0 > r, \end{cases}$$

a function showing that, except when $t_0$ is not close to 1, the reasoning (∗∗) is contextually true.

It should be pointed out how an adequate context-modeling, or design, from what is known of both the current text and the author's work, allows to design a reasonably good representation of the given linguistic description.

## 4. Conclusion

The flexible subjects fuzzy logic deals with (that are in contraposition to the typically rigid or formal logic) force a different methodology than that of formal sciences to approach the problem. This is like the case of Physics, whose methodology in not that strictly formal of Mathematics, even though mathematical models play an important role in Physics. But these models are to be experimentally tested against the word. Like it happens with the mathematical models in fuzzy logic, that are important in the measure that they allow to well representing the linguistic description of system an/or processes. In this sense, to say that the agenda of fuzzy logic is similar to that of formal logic is like taking similar those of Physics and Mathematics.

Despite its name fuzzy logic is, in the first place, for representing some reasonings involved with imprecision and uncertainty. Fuzzy logic is closer to an experimental science than to a formal one. In part, because of the use of real numbers and continuity properties as well as the apparition of thresholds. Because fuzzy logic mainly deals with gradable properties and there are a lot of different t-norms, t-conorms, strong negations, and T-conditionals, a careful design is necessary at each case.

## Acknowledgements

## References

1. J. von Neumann, *The general and logical theory of automata*, in *Cerebral Mechanisms in Behavior. The Hixon Symposium*, ed. L. A. Jeffress (John Wiley & Sons, New York, 1951), New York, pp. 1–41.
2. M. Black, *Dialogue* **2**, 1 (1963).
3. E. Trillas and L. A. Urtubey, Forthcoming in *Applied Soft Computing.*
4. E. Trillas and C. Alsina, *On disjunctive reasoning with fuzzy sets*, tech. rep. (2006), European Centre for Soft Computing.
5. E. Trillas and C. Alsina, *International Journal of Approximate Reasoning* **37**, 87 (2004).
6. O. Cordón, F. Herrera and A. Peregrín, *Fuzzy Sets and Systems* **86**, 15 (1997).
7. L. Wittgenstein, *Philosophical Investigations*, third edn. (Basil Blackwell, Oxford, 1981).
8. E. Trillas, M. Mas, M. Monserrat and J. Torrens, Forthcoming in *Int. Jour. of Approx. Reason.*
9. E. Vila-Matas, *Desde la ciudad nerviosa* (Alfaguara, Barcelona, 2004).
10. E. Trillas, On some features of fuzzy logic, in *Proceedings Fuzzy Information Processing Society, NAFIPS'08*, 2008. (Forthcoming).
11. E. Trillas, C. Alsina and A. Pradera, *Rev. R. Acad. Cien. Serie A. Mat.* **98**, 259 (2004).