Similarity vs. Possibility in measuring Fuzzy Sets Distinguishability

Corrado Mencar, Giovanna Castellano, Anna Maria Fanelli
Department of Informatics, University of Bari
Bari, Italy
e-mail: {mencar,castellano,fanelli}@di.uniba.it

Andrzej Bargiela
Department of Computing, The Nottingham Trent University
Nottingham, UK
e-mail: andre@doc.ntu.ac.uk

Abstract: *Two measures that quantify distinguishability of fuzzy sets are addressed in this paper: similarity, which exhibits sound theoretical properties but it is usually computationally intensive, and possibility, whose calculation can be very efficient but does not exhibit the same properties of similarity. It is shown that under mild conditions – usually met in interpretable fuzzy modelling – possibility can be used as a valid measure for assessing distinguishability, thus overcoming the computational inefficiencies caused by the use of similarity measures. Moreover, those procedures aimed to minimize possibility also minimize similarity and, consequently, improve distinguishability. In this sense, the use of possibility is fully justified in interpretable fuzzy modelling.*

Keywords: Distinguishability, Similarity, Possibility, Interpretable Fuzzy Modelling.

Interpretability is one of the most current issues concerning fuzzy modelling. While accuracy was the main concern of the first fuzzy model builders, in recent years interpretability has been recognized as the key feature of fuzzy models in the context of Soft Computing [1]. One of the most common interpretability constraints adopted in fuzzy modelling literature is the so-called *distinguishability* constraint. Distinguishability is a relation between fuzzy sets defined on the same Universe of Discourse. Roughly speaking, distinguishable fuzzy sets are well disjunct so they represent distinct concepts and can be assigned to semantically different linguistic labels.

Distinguishability can be formalized in different ways, the most adopted is by means of *similarity* measures. In [2] similarity measures are deeply discussed in the context of fuzzy modelling. There, similarity is interpreted as a fuzzy relation defined over fuzzy sets and corresponds to the "degree to which two fuzzy sets are equal". Similarity measures well capture all the requirements for distinguishable fuzzy sets, but their calculation is usually computationally intensive. As a consequence, most strategies for fuzzy model building that adopt similarity for interpretability are based on massive search algorithms (see, e.g. [3], [4], [5]). Alternatively, distinguishability improvement is realized in separate learning stages, often after some data driven procedure like clustering, in which similar fuzzy sets are usually merged together [2]. When distinguishability is to be considered in less time consuming learning paradigms, like neural learning, other measures are used in place of similarity, like possibility [6], [7], [8].

The possibility measure [10] has some attracting features that promote a deeper investigation in the context of distinguishability assessment. Although it is not a similarity measure, it has a clear and well-established semantics since it can be interpreted as the degree to which the flexible constraint (X is A) is satisfied. Moreover, the possibility measure between fuzzy sets can be often analytically described in terms of fuzzy sets' parameters. This makes possibility evaluation very efficient and can be effortlessly embodied in computationally inexpensive learning schemes.

The objective of the paper is to show that some significant relationships between similarity and possibility exist. Specifically, some sufficient conditions correlate possibility and similarity *in the worst case*, i.e. the lower is the possibility between two fuzzy sets, the lower is the maximum similarity that can be measured between the same fuzzy sets. Such kind of worst-case analysis is mainly due to the hypothesized conditions on the fuzzy sets, which are as general as possible so as to include a wide class of fuzzy sets. As a final result, it is shown that under some mild conditions, any transformation aimed to decrease possibility between fuzzy sets actually decreases also their similarity measure and, consequently, improves their distinguishability. In light of such theoretical results, possibility measure emerges as a good candidate for interpretability analysis and for efficient interpretable fuzzy modelling.

1 Measures to Quantify Distinguishability

In this Section similarity and possibility measures are briefly described. Any fuzzy set with a capital letter (A, B, etc.) and the corresponding membership function with μ_A, μ_B, etc. Each membership function is defined on the same Universe of Discourse U, which is assumed to be a one-dimensional closed interval $[m_U, M_U] \subset \mathbb{R}$. The set of all possible fuzzy (sub-)sets defined over U is denoted with $\mathcal{F}(U)$, while the finite family of fuzzy sets actually present in a fuzzy model is called "frame of cognition" and is denoted with $\mathbf{F}$.

1.1 Similarity

According to [2], the similarity measure between two fuzzy sets A and B is a fuzzy relation that expresses the degree to which A and B are equal. Several similarity measures have been proposed in literature, and some of them can be found in [9]. However, in interpretability analysis, the most commonly adopted similarity measure is the following:

$$S(A, B) = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{1}$$

Distinguishability in the context of interpretable fuzzy modelling is guaranteed by imposing that similarity between any two distinct fuzzy sets must not exceed a user-given threshold σ:

$$\forall A, B \in \mathbf{F} : A \neq B \to S(A, B) \leq \sigma \tag{2}$$

The evaluation of (1) may become computationally intensive for several classes of fuzzy sets (e.g. Gaussian) because of the integration operation necessary for determining the cardinality of the involved fuzzy sets. For such reason, similarity is not used in some learning schemes, where more efficient measure are instead adopted.

1.2 Possibility

The possibility measure between two fuzzy sets A and B is defined as the degree of applicability of the fuzzy constraint $(A \,\mathrm{is}\, B)$ [10]. Possibility is evaluated according to:

$$\Pi(A, B) = \sup_{x \in U} \min\{\mu_A(x), \mu_B(x)\} \tag{3}$$

The possibility measure quantifies the extent to which A and B overlap. An overlapping threshold π means that the possibility between two any distinct fuzzy sets in the frame of cognition must not exceed π:

$$\forall A, B \in \mathbf{F} : A \neq B \to \Pi(A, B) \leq \pi \tag{4}$$

The possibility measure has two main features that are important in distinguishability analysis. First, the threshold possibility value π has a clear semantics as it can be interpreted in the context of possibility theory [10]. Conversely, the similarity threshold σ has a more arbitrary nature (it also depends on the specific definition of intersection and union operators). Second, unlike for similarity computation, numerical integration is not necessary when calculating possibility. Moreover, although the general definition (3) may require a numerical sampling of the Universe of Discourse, the possibility measure can be evaluated analytically for several classes of membership functions (e.g. triangular, Gaussian, bell-shaped, etc.,). For this second feature, possibility is envisaged to be more suitable in efficient learning schemes.

From such considerations, the possibility measure emerges as a potentially good candidate for replacing similarity measure in distinguishability assessment. To this aim, an investigation on possible relationships existing between similarity and possibility is presented in the next section.

2 Relating Similarity vs. Possibility

The main advantage deriving from using possibility in place of similarity consists in a more efficient evaluation of fuzzy set distinguishability, which can be used in on-line learning schemes. However, the adoption of possibility for quantifying distinguishability is consistent provided the existence of a monotonic relation between possibility and similarity, i.e. a relation that assures low grades of similarity for small values of possibility. We found that such relation can exist, provided that some restrictions are imposed on the involved fuzzy sets.

Generally, in interpretable fuzzy modelling, fuzzy sets are required to be interpretable, i.e. they should satisfy a number of properties. In this work, we focus on the following properties:

Convexity A fuzzy set A is convex iff the membership values of elements belonging to any interval are not lower than the membership values at the interval's extremes:

$$\forall a, b, x \in U : a \leq x \leq b \rightarrow \mu_A(x) \geq \min\{\mu_A(a), \mu_A(b)\} \tag{5}$$

Normality A fuzzy set A is normal iff there exists at least one element with full membership:

$$\exists x \in U : \mu_A(x) = 1 \tag{6}$$

Continuity A fuzzy set A is continuous if its membership function μ_A is continuous in U.

Such properties were exploited to discover possible relationships between possibility and similarity. Precisely, we derived the following theorem.

Theorem 1 *Let A and B be two fuzzy sets that are continuous, normal and convex. Let $p_A \in \arg\max \mu_A$, $p_B \in \arg\max \mu_B$ and suppose $p_A < p_B$. Let $\pi = \Pi(A, B)$ and $x_\pi \in]p_A, p_B[$ such that $\mu_A(x_\pi) = \mu_B(x_\pi) = \pi$. In addition, suppose that:*

$$\forall x \in]p_A, x_\pi[: \frac{d^2\mu_A}{dx^2}(x) \geq 0 \tag{7}$$

and:

$$\forall x \in]x_\pi, p_B[: \frac{d^2\mu_B}{dx^2}(x) \geq 0 \tag{8}$$

Then, the similarity between A and B is upper-bounded by:

$$S(A, B) \leq S_{\max} = \frac{2\pi}{r + 2\pi - r\pi} \tag{9}$$

being r the ratio between the distance $p_B - p_A$ and the length of the support of $A \cup B$[1]*:*

$$r = \frac{p_B - p_A}{|\mathrm{supp}\, A \cup B|} \tag{10}$$

Proof. *The maximally similar fuzzy sets that are normal, convex and with possibility π must be built so that the cardinality of the intersection is the highest possible, while the cardinality of the union is the smallest possible. The following two fuzzy sets $A_{\max}$ and $B_{\max}$ satisfy such requirements (see fig. 1 for an example):*

$$\mu_{A_{\max}}(x) = \begin{cases} \pi & \textit{if} \quad x \in [\min \mathrm{supp}\, A \cup B, p_A[\\ \frac{x(\pi-1)+x_\pi - p_A\pi}{x_\pi - p_A} & \textit{if} \quad x \in [p_A, x_\pi] \\ \pi & \textit{if} \quad x \in]x_\pi, \max \mathrm{supp}\, A \cup B] \\ 0 & \qquad \textit{elsewhere} \end{cases} \tag{11}$$

$$\mu_{B_{\max}}(x) = \begin{cases} \pi & \textit{if} \quad x \in [\min \mathrm{supp}\, A \cup B, x_\pi[\\ \frac{x(\pi-1)+x_\pi - p_B\pi}{x_\pi - p_B} & \textit{if} \quad x \in [x_\pi, p_B] \\ \pi & \textit{if} \quad x \in]p_B, \max \mathrm{supp}\, A \cup B] \\ 0 & \qquad \textit{elsewhere} \end{cases} \tag{12}$$

The intersection and the union of such membership functions coincide in all points of the support except the interval $[p_A, p_B]$, where the membership functions have null second derivative. As a consequence, any fuzzy sets satisfying the hypothesis will have:

$$\forall x \in [p_A, x_\pi] : \mu_A(x) \geq \mu_{A_{\max}}(x) \tag{13}$$

and

$$\forall x \in [x_\pi, p_B] : \mu_B(x) \geq \mu_{B_{\max}}(x) \tag{14}$$

In this way, the part of fuzzy sets that are involved in the union but not in the intersection is minimized. More specifically, the intersection of the two fuzzy sets has the following membership function:

$$\mu_{A_{\max} \cap B_{\max}}(x) = \begin{cases} \pi & \textit{if} \quad x \in \mathrm{supp}\, A \cup B \\ 0 & \qquad \textit{elsewhere} \end{cases} \tag{15}$$

The union of the two fuzzy sets has the following membership function:

$$\mu_{A_{\max} \cup B_{\max}}(x) = \begin{cases} \pi & \textit{if} \quad x \in [\min \mathrm{supp}\, A \cup B, p_A[\\ \frac{x(\pi-1)+x_\pi - p_A\pi}{x_\pi - p_A} & \textit{if} \quad x \in [p_A, x_\pi] \\ \frac{x(\pi-1)+x_\pi - p_B\pi}{x_\pi - p_B} & \textit{if} \quad x \in [x_\pi, p_B] \\ \pi & \textit{if} \quad x \in]p_B, \max \mathrm{supp}\, A \cup B] \\ 0 & \qquad \textit{elsewhere} \end{cases} \tag{16}$$

The similarity of the two fuzzy sets is:

$$S(A_{\max}, B_{\max}) = \frac{|A_{\max} \cap B_{\max}|}{|A_{\max} \cup B_{\max}|} = \frac{\pi\, |\mathrm{supp}\, A \cup B|}{\pi\, |\mathrm{supp}\, A \cup B| + \frac{1}{2}(1-\pi)(p_B - p_A)} \tag{17}$$

By defining r as in (10), the similarity is shown to be equal to (9). Note that $A_{\max}$ and $B_{\max}$ are not continuous. However, continuous fuzzy sets may be defined so as to be arbitrary similar to $A_{\max}$ and $B_{\max}$. Hence, by defining $S_{\max} = S(A_{\max}, B_{\max})$, the maximal similarity measure is the upper-bound of the actual similarity between the original fuzzy sets A and B. ■

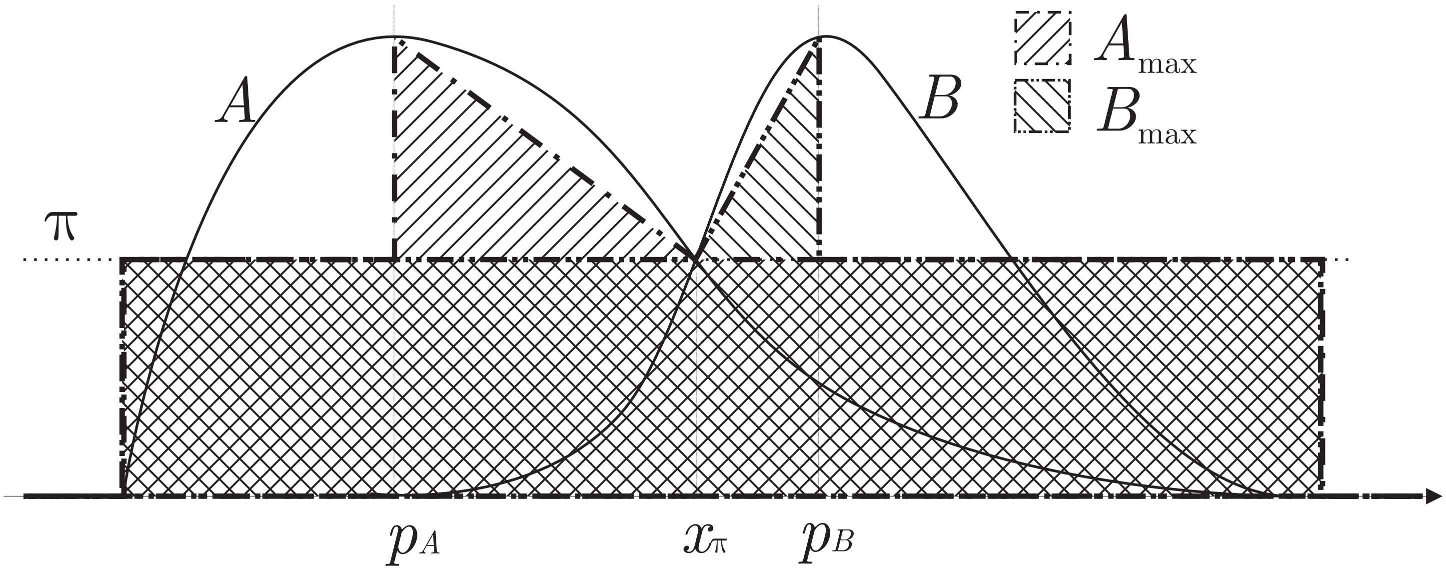

Figure 1: Example of fuzzy sets with maximal similarity for a given possibility measure

Note that the additional requirement for the second derivatives is not a significant limitation, since commonly used fuzzy set shapes (triangular, trapezoidal, Gaussian, bell-shaped, etc.) satisfy such requirement. Care must be paid for Gaussian fuzzy sets, where the point x_π should lay before the inflexion point, (i.e. should be smaller than center+width for fuzzy set A and higher than center-width for fuzzy set B). Nevertheless, the theorem establishes only sufficient conditions, hence partial violation of the requirements do not invalidate the results.

The relationship between possibility and similarity established by the theorem holds only for the upper-bound of the similarity measure, while the actual value is strictly related to the shape of the membership function. Paradoxically, the actual similarity measure between two low-possibility fuzzy sets may be higher than two high-possibility fuzzy sets. However, *relation (9) assures that the similarity measure does not exceed a defined threshold that is monotonically related to the possibility measure.* As a consequence, any modelling technique that assures small values of possibility between fuzzy sets, indirectly provides small values of similarity and, hence, good distinguishability between fuzzy sets. Thus, relation (9) justifies the use of possibility measure in interpretable fuzzy modelling.

3 Possibility and similarity minimization

Minimizing similarity between fuzzy sets is a classical approach to improve their distinguishability. However, the definition of similarity calls for computationally intensive methods or separate stages, hence it is not convenient to use similarity for on-line learning schemes. When efficient learning schemes are necessary, other measures are adopted in place of similarity, hence an interesting issue concerns how reducing non-similarity measures effectively reduces similarity. Here we focus on possibility measure as an alternative measure to quantify distinguishability. The following lemma characterizes a wide class of procedures for possibility minimization.

Lemma 2 *Let A and B two fuzzy sets defined on the Universe of Discourse U, which are continuous, normal and convex. Let $p_A \in \arg\max \mu_A$, $p_B \in \arg\max \mu_B$ and suppose $p_A < p_B$. Let $\Phi : \mathcal{F}(U) \to \mathcal{F}(U)$ be a transformation such that $B' = \Phi(B)$ is a continuous, normal and*

[1]The support of a fuzzy set is the (crisp) set of all elements with non-zero membership, i.e. $\operatorname{supp} X = \{x \in U : \mu_X(x) > 0\}$. For convex fuzzy sets, the support is an interval.

convex fuzzy set, $\forall x \in U : x \leq p_B \rightarrow \mu_{B'}(x) \leq \mu_B(x)$*. Then,*

$$\Pi\left(A, B'\right) \leq \Pi(A, B) \tag{18}$$

Conversely, if B' *is such that* $\forall x \in U : x \leq p_B \rightarrow \mu_{B'}(x) \geq \mu_B(x)$*, then*

$$\Pi\left(A, B'\right) \geq \Pi(A, B) \tag{19}$$

Two very common examples of transformations satisfying the lemma's hypothesis are the translation and the contraction of the membership function. Such transformations can be effectively used to reduce possibility between two fuzzy sets, but do such transformations effectively reduce similarity? The following corollary gives the sufficient conditions that guarantee such reduction.

Corollary 3 *Any transformation* $\Phi : \mathcal{F}(U) \rightarrow \mathcal{F}(U)$ *such that* $B' = \Phi(B)$ *preserves lemma's hypothesis with* $\pi' = \Pi(A, B') < \pi$ *and* $r' = \frac{p_{B'} - p_A}{|\operatorname{supp} A \cup B'|} \geq r$*, produces a decrease of the maximal similarity* $S_{\max}$*.*

As a consequence of the corollary, every method aimed to minimize possibility (to a user-given threshold) actually reduces (maximal) similarity, thus improving distinguishability. In this sense, the adoption of possibility as a measure of distinguishability is fully justified. The additional constraint required in the corollary (the ratio r must not decrease) is always fulfilled by any translation that lengthens the distance between the modal points. Attention must be paid for contraction, as the support may be reduced. However, if multiplicative contraction is adopted (i.e. a contraction such that $\mu_{B'}(x) > 0$ if $\mu_B(x) > 0$), then the corollary can be still applied.

4 Conclusions

Similarity can be considered as the most representative measure for distinguishability of fuzzy sets. Besides similarity, possibility can be considered as an effective alternative to quantify distinguishability. The key features of possibility measures are a sound semantical meaning of its values and the computationally efficiency of the calculation procedure. Also, a possibility measure can be expressed in most cases analytically in terms of fuzzy sets parameters, so it can be used in many learning schemes without resorting computationally intensive algorithms. In this paper we proved that under mild conditions (always satisfied by interpretable fuzzy sets) possibility and similarity are related monotonically, so that procedures aimed to minimize possibility also minimize similarity and, consequently, improve distinguishability. This suggests that possibility can be successfully adopted in interpretable fuzzy modelling.

References

[1] P. Bonissone, Y.-T. Chen, K. Goebel and P.S. Khedkar, "Hybrid Soft Computing Systems: Industrial and Commercial Applications," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1641-1667, 1999

[2] M. Setnes, R. Babuska, U. Kaymak, and H. R. Van Nauta Lemke, "Similarity measures in fuzzy rule base simplification," *IEEE Trans. on Sys., Man and Cybernetics, part B*, 28(3):376-386, 1998.

[3] H. Roubos and M. Setnes, "Compact and transparent fuzzy models and classifiers through iterative complexity reduction," *IEEE Trans. on Fuzzy Systems*, 9(4):516-524, 2001.

[4] P. Meesad and G. G. Yen, "Quantitative measures of the accuracy, comprehensibility, and completeness of a fuzzy expert system," in *Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ'02)*, Honolulu, Hawaii, 2002, pp. 284-289.

[5] Y. Jin, "Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement," *IEEE Trans. on Fuzzy Systems*, 8(2):212-221, 2000.

[6] Y. Jin and B. Sendhoff, "Extracting interpretable fuzzy rules from RBF networks," *Neural Processing Letters*, 17:149-164, 2003.

[7] J. Espinosa and J. Vandewalle, "Constructing fuzzy models with linguistic integrity from numerical data - AFRELI algorithm," *IEEE Trans. on Fuzzy Systems*, 8(5):591-600, 2000.

[8] G. Castellano, A.M. Fanelli and C. Mencar, "A Neuro-Fuzzy Network to Generate Human Understandable Knowledge from Data," *Cognitive Systems Research Journal, Special Issue on Computational Cognitive Modeling*, 3(2):125-144, 2002.

[9] M. Setnes, "Fuzzy rule-base simplification using similarity measures," Master's thesis, Dept. of Electronic Engineering, University of Delft, 1995.

[10] D. Dubois and H. Prade, "Fuzzy Sets and Systems: Theory and Applications," Academic Press, New York, 1980.