Should fuzzy equality and similarity satisfy transitivity? Comments on the paper by M. De Cock and E. Kerre

Frank Klawonn
Department of Electrical Engineering and Computer Science
Emden University of Applied Sciences
Constantiaplatz 4
D-26723 Emden, Germany
e-mail klawonn@et-inf.fho-emden.de

Abstract

This brief paper addresses problems that were raised in the paper [1] in connection with modelling fuzzy equality and similarity. In addition to specific comments to the cited paper, we also point out the more fundamental questions underlying this discussion.

1 Introduction

If the reader expects a definite answer to the question stated in the title, I regret that he will be disappointed by finding none in this short note. Nevertheless, the least I can do is, to elucidate important issues that are connected with the question. In the following, starting from some general comments in connection with the concept of similarity modelling, specific aspects of M. De Cock's and E. Kerre's paper are discussed and questioned.

2 Poincaré paradox, equivalence relations and transitivity

Poincaré's observation that indistinguishability might behave non-transitive in real life can be elucidated by the following example. A microscope or a telescope has a certain resolution so that we can only distinguish between two points, if their distance exceeds a certain value, say ε. Therefore, it might be the case that points A and B are indistinguishable as well as B and C, since both distances are smaller than ε. However, the distance between A and C might be greater than ε so that they can be distinguished.

The following example is more or less in the same line as the previous example, but does even better clarify what the underlying problems in handling such things are. Assume a person wants to buy a luxurious car and he sees an advertisement with a special offer for this car for $29,995. He decides to go to the car seller and buy the car. When arriving there, he discovers that the price is actually $29,996 and checking against the advertisement he realises that it was his error, he had misread the price. No one will expect the customer to reconsider his decision and not buy the car only because it was $1 more expensive than expected. So the conclusion is that it does not matter whether the price of the car is increased by $1, the customer will buy it anyway. But we must be very careful, when we iterate this argument. Even if the price was $29,997 the customer would probably not hesitate to purchase the car. But by iterating the argument that we can increase the price by $1 without influencing the customer's decision, we end up with the conclusion that we can charge any price from the customer. Speaking in terms of indistinguishability, we might say that two prices are not distinguished (w.r.t. to the decision whether to buy the car or not), if they differ by less than $1.

In both this and the previous example transitivity fails. The second example illustrates, how we can cope with the concept of magnitude – be it price of a car or any other value – in real life. We often imagine that we can consider magnitudes instead of concrete numbers in the sense that we just distinguish between magnitudes and identify each number (context dependent) with a certain magnitude. We almost pretend that the magnitudes behave like equivalence classes, but they do not correspond to a partition of the numbers. Generalising from the example of the car price, we can say that in many situations we tend to consider two numbers as indistinguishable, when their difference does not exceed a certain tolerance bound ε. Unfortunately, the relation

$$x \approx y \;\Leftrightarrow\; |x - y| \leq \varepsilon$$

is only reflexive and symmetric, but not transitive. An important aspect of equivalence relations is their one-to-one correspondence to partitions. Therefore, the relation $\approx$ does not induce a partition due to the non-transitivity.

If we want to generalise from a binary notion of indistinguishability to a fuzzy one, we should keep the duality between equivalence relations and partitions in mind. For those who are interested in a more thorough analysis of the correspondence between fuzzy equivalence relations and fuzzy partitions, the work by H. Thiele and N. Schmechel can be recommended [7, 8, 9]. Generalising the notion of a crisp partition to a fuzzy partition is not an easy task at all, meaning that the closely related concept of a fuzzy equivalence relation is not easy define as well.

3 Duality of similarity and distance

As pointed out by M. De Cock and E. Kerre ([1], section 4) and many others before, there is a close relationship between the concepts of distance and similarity (fuzzy equality or equivalence, indistinguishability). When we have to deal with numbers, distance is a canonical concept. Nevertheless, we have to be careful about a naive use of distance. First of all, we have to consider a kind of global scaling. The distance between values changes, when we measure them in milli- or kilo-units.

In addition to the global scaling, we might want to consider a problem dependent local scaling [5]. The idea of local scaling is that the importance of distinguishing between values might depend on the magnitude of the values. If we want to purchase something in the range of $30.000, we do not really distinguish between prices that differ by $1. However, buying something for $3, $1 makes a big difference. When we consider problems with more than one variable, we need a distance function on a product space. This leads to concepts like aggregation and independence. The problems arising in this context with distance/similarity functions are not solved at all at the moment [4].

Viewing similarity or indistinguishability as a dual concept to distance, we can easily establish a similarity function E on the basis of a given pseudo-metric d, simply by using $E(x, y) = 1 - d(x, y)$. Of course, we must make sure that d never yields values greater than one. In [1] it is pointed out that *every pseudo-metric d' can be turned into a $[0, 1]$-valued pseudo-metric d by defining* $d(x, y) = \min\{1, d'(x, y)\}$. This is indeed true, although d and d' are definitely not the same pseudo-metrics. The important point is that we are interested in similarity, i.e. in small distances. And for small distances d and d' coincide.

Using the duality $E(x, y) = 1 - d(x, y)$ it is immediately clear that the conditions (PM.1) and (PM.2) in [1] enforce the reflexivity and symmetry of E. The triangular inequality (PM.3) is equivalent to the transitivity of E w.r.t. to the Łukasiewicz t-norm. Replacing the Łukasiewicz t-norm by the minimum leads to ultra-metrics. The criticism in [1] against the properties of the min-transitivity are of course justified, when we deal with numbers, since the canonical metric on real numbers is not at all an ultra-metric. Nevertheless, in other contexts min-transitivity might be useful. Consider for instance images that can be viewed with different resolutions. Let us assume that we can adjust the resolution between 0 (very rough) and 1 (high resolution). We define for two images i_1 and i_2

$$E(i_1, i_2) \;=\; \inf\{r \mid i_1 \text{ and } i_2 \text{ can be distinguished under resolution } r\}.$$

It can be easily seen that E satisfies the min-transitivity and that the corresponding metric $d(i_1, i_2) = 1 - E(i_1, i_2)$ is an ultra-metric.

The crucial question remains whether there is really a one-to-one duality between similarity and pseudo-metrics. Do we only have to find the proper

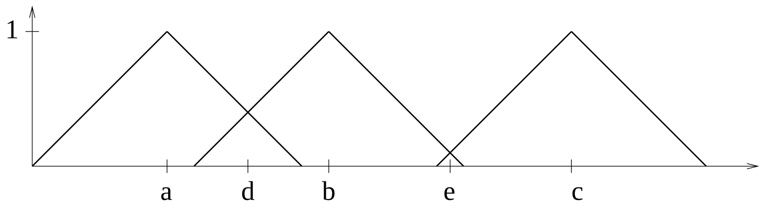

Figure 1: A fuzzy partition

(pseudo)metric in order to define a corresponding similarity relation or is there more than just that, as claimed in [1]?

One way, to have a more general concept than simply the dual of a pseudometric is provided in Definition 2 of [1]. These $[0, 1]$-valued equalities are criticised in [1], because *we feel that in the case of approximate equality it is justified to demand full reflexivity.* This argument seems to be a little bit misleading. The value $E(x, x)$ is not really interpreted as the degree of equality of x to itself, but as the degree of existence. (Therefore, this concept is also called local existence [2, 3].)

The idea behind local existence in the context of fuzzy systems is the following. Consider the 'fuzzy partition' (without discussing whether this can really be interpreted as a partition) in Fig. 1. It does indeed make sense to interpret the first fuzzy set as approximately a or as the fuzzy set of values that are similar to a [6]. Correspondingly, the other two fuzzy sets represent the fuzzy sets of values that are similar to b and c, respectively. When using these fuzzy sets in rules like 'if x is approximately a (b, c, resp.), then ...', one of the rules applies fully, when x is either equal to a, b or c. However, if x is equal to d or – even worse – to e, none of the rules is applicable to a sufficiently high degree. As mentioned before, we can easily define a fuzzy M-equality relation in the sense of Definition 1 in [1] so that the fuzzy sets can be interpreted as fuzzy sets of values that are similar to a, b and c, respectively. However, if we allow for the more general Definition 2 in [1], we would again have $E(a, a) = E(b, b) = E(c, c) = 1$, but $E(d, d) = 0.4$ and $E(e, e) = 0.2$. In this sense, the degree of local existence $E(x, x)$ indicates, how good x is actually covered by the given fuzzy partition. In this way, we have some additional information about how good our conclusions can be expected for the value x, i.e. a low degree of local existence of x warns us to be careful about the information we have on x. Although obviously quite appealing, this concept is seldom used in fuzzy systems.

4 Are resemblance relations what we need?

There is a link between indistinguishability/fuzzy equality and distance/pseudometrics. It might be argued that the approach requiring $E(x, y) = 1 - d(x, y)$ for an indistinguishability relation and a pseudo-metric might be too restrictive. So it might be worth to take a closer look at the concept of a resemblance relation as specified in Definition 6 in [1].

First of all, it should be remarked that as a direct consequence of Proposition 5 in [1] property (R3) already implies the symmetry (R2).

Another question is, what is the intuitive idea behind the mapping g? I would find it more intuitive, if the mapping g would be completed dropped and a pseudo-metric $d^{(X)}$ on the considered universe X is required. This is all what is needed in the definition and g is only used to compute the induced pseudo-metric $d^{(X)}$ by $d^{(X)}(x,y) = d(g(x), g(y))$. Therefore, in the following, in order to refrain from unnecessary complexity, I will always assume that g is the identity mapping and no longer refer to g.

Another interesting question is: Can we find for any reflexive and symmetric fuzzy relation E a pseudo-metric d on X s.t. E becomes a d-resemblance relation. The answer is yes.

Proposition 1 *Let E be a reflexive and symmetric fuzzy relation on X. Then with*

$$d(x,y) \;=\; \begin{cases} 2 - E(x,y) & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases} \tag{1}$$

E is a d-resemblance relation.

(Note that, if it is required that d is bounded by one, we can use $\hat{d}(x,y) = 0.5 \cdot d(x,y)$.)

Proof. First we prove that d is a pseudo-metric. We have $d(x,x) = 0$ by definition and from the symmetry of E we obtain $d(x,y) = d(y,x)$. d satisfies the triangular inequality, since

$$\begin{aligned} d(x,y) + d(y,z) &= \left\{ \begin{array}{ll} 0 & \text{if } x = y = z \\ 4 - E(x,y) - E(y,z) & \text{if } x \neq y \text{ and } y \neq z \\ 2 - \max\{E(x,y), E(y,z)\} & \text{otherwise} \end{array} \right\} \\ &\geq \left\{ \begin{array}{ll} 0 & \text{if } x = y = z \\ 2 & \text{if } x \neq y \text{ and } y \neq z \\ 2 - E(x,z) & \text{otherwise} \end{array} \right\} \\ &\geq \left\{ \begin{array}{ll} 0 & \text{if } x = z \\ 2 - E(x,z) & \text{if } x \neq z \end{array} \right\} \\ &= d(x,z). \end{aligned}$$

Finally, we have to show that $d(x,y) \leq d(z,u)$ implies $E(x,y) \geq E(z,u)$.

Case 1: $x = y$. Then we have $E(x,y) = 1 \geq E(z,u)$.

Case 2: $x \neq y$, $u = z$. This case leads to the contradiction $2 - E(x,y) = d(x,y) \leq d(z,u) = 0$.

Case 3: $x \neq y$, $u \neq z$. This implies $2 - E(x,y) = d(x,y) \leq d(z,u) = 2 - E(z,u)$ and therefore $E(x,y) \geq E(z,u)$. □

It is not my task to provide an intuitive interpretation on which the idea of a resemblance relation is based. The above proposition shows at least that any reflexive and symmetric fuzzy relation can be viewed as a resemblance relation by choosing the pseudo-metric constructed in the proof. If use (1) in order to understand how E is constructed from d (if d would be given), then we simply have $E(x,y) = \min\{2 - d(x,y), 1\}$, i.e. E does not distinguish between elements whose distance is less than 1 and otherwise it treats elements just dual as d. Of course, the constructed d in Proposition 1 is not unique at all. Nevertheless, since E must satisfy the – in some sense – order preserving property (R3), E might in addition only 'scale' the values of any appropriate d a little bit. Is this the intended interpretation of a resemblance relation? The examples provided in [1] are all more or less in this line.

5 Conclusions

When we try to define a useful concept for a notion like similarity, indistinguishability, fuzzy equality or fuzzy equivalence, it is important to first clarify the semantic background and then introduce the definitions. In this sense, I consider Example 4 in [1] as slightly misleading. First of all, I am a little bit confused by the quite complex representation: The relation/similarity/equality of v_i and v_j is assumed to be represented by a whole fuzzy relation, not only by an equality degree, but by a complete fuzzy relation. If we assume that the degree of equality between 20 and 23 as well as between 23 and 26 is 0.9, then we end up with an equality degree of 0.8 between 20 and 26, when we apply the Łukasiewicz t-norm. This does not sound contradictory. But we must be aware, when we apply this type of compositional rule of inference that we are computing lower bounds. This means that we only know that the degree of equality between 20 and 26 is *at least* 0.8. If we for instance apply the compositional rule if inference to the pairs (23,26) and (26,23), we conclude that the degree of equality between 23 and 23 is at least 0.8, which is indeed true. But we know, of course, that the correct degree is one.

Unless we do not specify, in which way we want to interpret degrees of equality, membership or whatsoever and how we want to handle them (estimating lower/upper bounds ...), it is nothing more then number crunching, when we propose new ways to operate with fuzzy concepts.

References

[1] M. De Cock, E. Kerre, On (Un)Suitable Relations to Model Approximate Equality. Fuzzy Sets and Systems (this issue).

[2] U. Höhle, M-valued Sets and Sheaves over Integral Commutative CL-Monoids. In: S.E. Rodabaugh, E.P. Klement, U. Höhle (eds.), Applications of Category Theory to Fuzzy Subsets. Kluwer Academic Publishers, Dordrecht (1992), 33-72.

[3] U. Höhle, L.N. Stout, Foundations of Fuzzy Sets. Fuzzy Sets and Systems 40 (1991), 257-296.

[4] F. Höppner, F. Klawonn, P. Eklund, Learning indistinguishability from data. Soft Computing (to appear)

[5] F. Klawonn, Fuzzy Sets and Vague Environments. Fuzzy Sets and Systems 66 (1994), 207-221.

[6] R. Kruse, J. Gebhardt, F. Klawonn: Foundations of Fuzzy Systems. Wiley, Chichester (1994).

[7] H. Thiele, On the Mutual Defineability of Fuzzy Tolerance Relations and Fuzzy Tolerance Coverings. Proc. 25th International Symposium on Multiple-Valued Logic. IEEE Computer Society, Los Alamitos, CA (1995), 140-145.

[8] H. Thiele, On Similarity Based Fuzzy Clusterings. Proc. 17th Linz Seminar on Fuzzy Set Theory, Linz (1996).

[9] H. Thiele, N. Schmechel, The Mutual Defineability of Fuzzy Equivalence Relations and Fuzzy Partitions. Proc. Intern. Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and the Second International Fuzzy Engineering Symposium, Yokohama (1995), 1383-1390.