On the Power of Fuzzy Logic

Roberto Pacheco, Alejandro Martins
Department of Production Engineering, Federal University of Santa Catarina, Florianópolis, Brazil.
Department of Industrial and Management Systems, University of South Florida, Tampa, Florida, 33620, USA.

Abraham Kandel
Department of Computer Science and Engineering, University of South Florida, Tampa, Florida, 33620 - USA.

Abstract

In this paper we address the issues brought by Elkan in his article, "The Paradoxical Success of Fuzzy Logic" [3]. Elkan's work has caused concern since purportedly reveals a Fuzzy Logic weakness regarding its theoretical foundations. A further investigation of Elkan's theorem ("theorem 1") revealed that its conclusion is not correct. After indicating the points where we disagree with Elkan, we reformulate theorem 1, calling this new version "theorem 2". Theorems 1 and 2 have the same hypotheses but different conclusions. According to theorem 2 there is a region of points that do hold the equivalence in the hypotheses of theorem 1. In other words, one does not need to change the definition of logical equivalence in theorem 1 in order to prove that Fuzzy Logic does not collapse to a two-valued logic. In a further analysis of theorem 2 we show that Elkan's work does not affect the power of Fuzzy Logic to model vagueness.

I. Introduction

Elkan's paper [3] has caused polemic among Fuzzy Logic researchers. The author criticizes Fuzzy Logic based upon two different points: first, he introduces a theorem, named "theorem 1", which states that any two logic assertions $A$ and $B$ that meet a given logical equivalence will end up in only two combinations of true values, namely the lines $t(A) = t(B)$ or $t(A) = t(\neg B)$. This means that any multivalent logic that holds the formal system described by definition 1 (see next section) collapses to only two particular relationships. Second, Elkan argues that the success of Fuzzy Logic in industrial applications (fuzzy control) is paradoxical, since it happens despite of this purported drawback.

Most answers presented by well known Fuzzy Logic researchers do not disagree with the validity of Elkan's theorem. However, the researchers argue that this theorem does not apply to Fuzzy Logic. They argue that, contrary to Fuzzy Logic, theorem 1 meets the
excluded-middle and contradiction laws (e.g., [15], [13], [2], [10]), since the definition of
logical equivalence assumed in the theorem is too restrict to be kept in Fuzzy Logic ([8],
[10], [11], [12]). They also examined the supposed reasons for Fuzzy Logic being
successful in control applications and show that Elkan's arguments are simply
unsupported. For instance, Elkan argues that the industrial applications of fuzzy control
are based on a small number of rules which causes its success rather than the fuzzy
approach itself. However, the reduced number of rules is not a coincidence but a
consequence of fuzzy predicates in the rules (e.g., [1], [5], [13], [11]).

We decided to investigate theorem 1 further. The conclusions we have drawn from this
study show that the theorem itself is not correct. There are points \( p(t(A), t(B)) \) that do hold
the equivalence relation of theorem 1 and are not on the lines \( t(A) = t(B) \) or \( t(A) = 1 - t(B) \).
We not only disagree with Elkan's opinions, but also with their mathematical basis. Even
when the theorem is reformulated, one can conclude that it is only a special case in Fuzzy
Logic. In this work we reformulate theorem 1 and analyze how Fuzzy Logic faces the new
theorem.

II. Revising Theorem 1

By revising theorem 1 we disagreed with some of the Elkan's assumptions (see Appendix
A). Here we initiate from the hypotheses of theorem 1 and conclude that its thesis is
broader than the original one proposed by Elkan.

**Definition 1**

Let \( A \) and \( B \) be arbitrary assertions. Then

\[
\begin{align*}
(a) \quad t(A \land B) &= \min \{t(A), t(B)\} \\
(b) \quad t(A \lor B) &= \max \{t(A), t(B)\} \\
(c) \quad t(\neg A) &= 1 - t(A) \\
(d) \quad t(A) &= t(B) \text{ if } A \text{ and } B \text{ are logically equivalent}^{1}
\end{align*}
\]

\(^{1}\) It is not clear if the expression (d) is a definition of *logical equivalence*, since it can be only a necessary
condition. For instance, the following definition: \( A \) and \( B \) are logically equivalent iff

\(
(i) \quad t(A) = t(B) \text{ and } (ii) \quad t(\neg A) \geq t(B).
\)

would also hold definition (d). Therefore, a better statement for definition (d) would be:

\(~(d')~ A \text{ and } B \text{ are logically equivalent iff } t(A) = t(B).~\)
**Theorem 1**

Given the formal system of definition 1, if \( \neg(A \land \neg B) \) and \( B \lor (\neg A \land \neg B) \) are logically equivalent, then for any two assertions \( A \) and \( B \), either \( t(B) = t(A) \) or \( t(B) = 1 - t(A) \).

**Reviewing the Proof**

In theorem 1, there are two assertions assumed as logically equivalent. According to definition 1, if two assertions are logically equivalent they have the same true value. Therefore:

\[
t(\neg(A \land \neg B)) = t(B \lor (\neg A \land \neg B))
\]

Considering definition 1, we have:

\[
1 - t(t(A) \land t(-,B)) = \max[t(B), t(\neg A \land \neg B)]
\]

\[
1 - \min[t(A), t(\neg B)] = \max[t(B), \min(t(\neg A), t(\neg B))]
\]

\[
1 - \min[t(A), 1 - t(B)] = \max\{t(B), \min[1 - t(A), 1 - t(B)]\} = \max\{t(B), \min[1 - t(A), 1 - t(B)]\}
\]

(1)

the equivalence (1) is true depending upon the relation between the assertions \( A \), \( B \) and their complements. Given a point \( p(t(A), t(B)) \), the right and left hands of expression (1) are different iff:

\[
t(B) < 1 - t(A) \quad \text{and} \quad 1 - t(B) < 1 - t(A)
\]

which led us to

\[
t(A) < t(B) < 1 - t(A) \quad \text{(region R1)}
\]

The inequalities above define the region of points \( p(t(A), t(B)) \) that do not hold expression (1). This region is shown in Figure 1a.

The hypothesis about the assertions \( A \) and \( B \) in theorem 1 is clear: "... any two assertions \( A \) and \( B \)". That means no presumption for them. Region R1 was defined considering points \( p(t(A), t(B)) \), that is \( t(A) \) and \( t(B) \) are in a specific order. However, given any two true values (e.g., 0.6 and 0.4) we can not assume a fixed order to verify theorem 1 (i.e. 0.6 can be either \( t(A) \) or \( t(B) \)). In other words, the points that hold expression (1) must either \( p(t(A), t(B)) \) or \( p(t(B), t(A)) \). By switching \( t(A) \) by \( t(B) \) (and vice-versa) in (1), one derives the following expression:

\[
\max[1 - t(B), t(A)] = \max\{t(A), \min[1 - t(B), 1 - t(A)]\}
\]

(2)
Now, the left and right hands are different iff:

\[ t(B) < t(A) < 1 - t(B) \]  \hspace{1cm} \text{(region R2)}

[FIGURE 1 ABOUT HERE]

Figure 1b shows region R2. The union between regions R1 and R2 defines all points that do not hold expression (1) of theorem 1. Thus, the complement of this region defines all points that do hold expression (1). It is possible to show that the complement of the set \((R1 \cup R2)\) is given by:

\[ \{t(A) \geq 1 - t(B)\} \cup \{t(A) = t(B)\} \]  \hspace{1cm} \text{(region R3)}

[FIGURE 2 ABOUT HERE]

Figure 2 shows region R3. The “arrow” R3 is a consequence of the fact the set \((R1 \cup R2)^c\) includes the line \(t(A) = t(B)\). Contrary to Elkan’s proof, there are points \((t(A), t(B))\) that do hold the equivalence of theorem 1 without meeting the expressions \(t(A) = t(B)\) or \(t(A) = 1 - t(B)\). Table 1 has examples of points that verify (1) and (2) (points in region R3 and bolded in Table 1) and points that hold only one of these equivalencies (points in region R1 or R2).

[TABLE 1 ABOUT HERE]

The main conclusion is that the thesis of theorem 1 is incomplete. In the following, we propose a restatement of the theorem.

**Theorem 2**

Given the formal system of definition 1, if \(- (A \land -B)\) and \(B \lor (\neg A \land \neg B)\) are logically equivalent, then for any two assertions \(A\) and \(B\), either \(t(B) = t(A)\) or \(t(B) \geq 1 - t(A)\).

**III. What is Implicit in the Hypotheses of Theorem 2?**

Once theorem 1 has been reformulated, there is an important question to be answered: what is the implication of the logical equivalence in the theorem? We can address this
issue by considering the following properties of the operators ∪ (union),
∩ (intersection) and ◇ (complement) of multivalent logic [7]:

a) Commutativity: \[ A ∪ B = B ∪ A, A ∩ B = B ∩ A. \]
b) Distributivity: \[ A ∪ (B ∩ C) = (A ∪ B) ∩ (A ∪ C), A ∩ (B ∪ C) = (A ∩ B) ∪ (A ∩ C). \]
c) De Morgan's Laws: \[ (A ∩ B)^\complement = A^\complement ∪ B^\complement, (A ∪ B)^\complement = A^\complement ∩ B^\complement. \]
d) Involution: \[ (A^\complement)^\complement = A. \]
e) Identity: \[ A ∪ ∅ = A, A ∩ X = A, \] where \( X \) is the universe of domain
f) Equivalence: \[ (A^\complement ∪ B) ∩ (A ∪ B^\complement) = (A^\complement ∩ B^\complement) ∪ (A ∩ B). \]

Now, let us rewrite the logical equivalence in theorems 1 and 2 using the notation above:

\[ \neg(A ∧ \neg B) \iff B ∨ (\neg A ∧ B) \]

\[ (A ∩ B^\complement)^\complement = B ∪ (A^\complement ∩ B^\complement) \quad (i) \]

according to properties (c) and (d), the left hand of this expression can be rewritten by:

\[ (A ∩ B^\complement)^\complement = A^\complement ∪ (B^\complement)^\complement \]

\[ (A ∩ B^\complement)^\complement = A^\complement ∩ B \quad (ii) \]

now, considering properties (a) and (b), the right hand of the equivalence (i) can also be rewritten by:

\[ B ∪ (A^\complement ∩ B^\complement) = (A^\complement ∪ B) ∩ (B ∪ B^\complement) \quad (iii) \]

Therefore, the logical equivalence (i) becomes the relation between the right hands of (ii) and (iii):

\[ A^\complement ∪ B = (A^\complement ∪ B) ∩ (B ∪ B^\complement) \quad (iv) \]

this expression is the same to:

\[ C = C ∩ D, \quad \text{where } C = A^\complement ∪ B \text{ and } D = B ∪ B^\complement. \]

Hence, expression (iv) is true (and consequently (i) is true) iff:

\[ C \subseteq D \]

which is the same to:

\[ A^\complement ∪ B \subseteq B ∪ B^\complement \quad (v) \]

since \( B \subseteq B ∪ B^\complement \), (v) is true iff:

\[ A^\complement \subseteq B ∪ B^\complement \quad (vi) \]

equation (vi) holds depending on the following relationships between \( A^\complement \), \( B \), and \( B^\complement \):

• \( A^\complement \subseteq B \)
in this case, the left hand of the expression (iv) becomes: \( A^c \cup B = B \), and the right hand of (iv) becomes: \((A^c \cup B) \cap (B \cup B^c) = B \) (TRUE)

- \( A^c \subseteq B^c \)
  
  now, the left hand of (iv) is unknown: \( A^c \cup B = ? \), but, since the right hand of (iv) comes from the right hand of (i):
  
  \[ B \cup (A^c \cap B^c) = B \cup A^c = A^c \cup B \] (TRUE according to propriety (a))

- \( B \cup B^c = X \), where \( X \) is the universe of domain. (which is the excluded-middle law)

  in this case, according to the expression (vi): \( A^c \subseteq X \) (TRUE according to (e))

Now, in order to assume that \( A \) and \( B \) are any assertions, as stated in theorem 2, it is necessary to obtain the three new expressions derived when the analysis starts from the equivalence: \((B \cap A^c)^c = A \cup (B^c \cap A^c) \) (i'). The three equivalent expressions are:

\[ B^c \subseteq A; B^c \subseteq A^c \] and \( A \cup A^c = X \).

By applying definition 1, one can show that the first expressions \((A^c \subseteq B)\) and its equivalent \((B^c \subseteq A)\) hold for any sets \(A\) and \(B\). The same expressions can be derived directly from one of the conclusions of theorem 2 (i.e., \( t(B) \geq 1 - t(A) \)). The expressions \((A^c \subseteq B^c)\) and \((B^c \subseteq A^c)\) hold only if \( t(B) = t(A) \), which is the other conclusion of theorem 2.

Finally, when the expressions \((B \cup B^c = X)\) and \((A \cup A^c = X)\) hold (i.e., excluded-middle law holds), the true values of \( A \) and \( B \) must be either 1 or 0.

As Elkan argued, the excluded-middle law is not required to find points that hold the theorem, since its conclusions \( t(B) \geq 1 - t(A) \) or \( t(B) = t(A) \) do not necessarily presuppose this law.

And how does Fuzzy Logic face theorem 2? As we will see in the next section all cases that hold the logical equivalence in theorem 2 are just instances in Fuzzy Logic. By accepting only assertions that hold the theorem, one would unnecessarily constrain its power of representation.

**IV. Theorem 2 and Fuzzy Logic**

Theorem 2 states that either \( t(B) = t(A) \) or \( t(B) \geq 1 - t(A) \). Here we show that these expressions are only special cases of Fuzzy Logic. In order to discuss this point, we present an example of a group of fuzzy sets which describe linguistic variables that qualify temperatures. Figure 3 presents the fuzzy sets COLD, WARM and HOT.
As an example, we can ask someone to classify the temperature of a glass of “hot” water according to the categories “cold,” “warm,” and “hot”. The person should hold the glass for a minute and classify the temperature. If the glass is “very hot”, the answer will certainly be “hot” ($\mu_{\text{HOT}}(t) = 1$). If we decrease the temperature, eventually the categorization will pass to “warm” ($\mu_{\text{WARM}}(t) > 0$). Now, let us think about these two classifications. We can keep inquiring the person about the temperature with questions like “is it ‘hot’ or ‘warm’? How much?”. Even grading the truth of the answers, the person is forced to think in “hot” and “warm” as two mutually exclusive concepts. Thus, if one says “the glass is 0.6 ‘hot’” ($\mu_{\text{HOT}}(f) = 0.6$), implicitly one also believes with some grade of truth in the sentence “the glass is ‘not warm’” ($\mu_{\text{NOT WARM}}(f) \leq 0.6$).

Figure 3 shows the temperatures that somehow meet the assumed dichotomy between HOT and WARM. Curiously, the interval of temperatures represents one of the two sets of points $p(\mu_{\text{HOT}}(f), \mu_{\text{NOT WARM}}(f))$ that hold theorem 2 (i.e., $\mu_{\text{HOT}}(f) \geq \mu_{\text{NOT WARM}}(f)$). In Fuzzy Logic, these temperatures are not the only ones that can be classified with some grade as “hot” or “warm” (implicitly “not warm”). In fact, these temperatures represent points that one can distinguish with some grade between the semantic concepts “hot” and “warm”. But a continuous variation in the temperature would eventually reach points when our perception is not so clear. Humans do not have a black-white definition regarding concepts as “hot”, “warm”, “good”, “fair”, and so on. There is no reason for forcing such categorization. By representing linguistic terms taking into account only theorem 2, one is implicitly compelling the model to this all-nothing world. Formally, Fuzzy Logic allows us to escape from this constraint by ignoring the excluded-middle and contradiction laws. As Figure 3 shows, points that hold theorem 2 are only instances in Fuzzy Logic. There is no plausible reason for being confined to those values. In fact, the intent to avoid this limitation leaving it as a particular case, was the main inspiration for the creation of Fuzzy Logic [14].

The Notion of Logical Equivalence

Another crucial issue when we are checking how Fuzzy Logic faces theorem 2 is the notion of logical equivalence stated in definition 1. Again, this definition is only a particular

---

2 The other set of points (where $\mu_{\text{HOT}}(f) = \mu_{\text{WARM}}(f)$) occurs whenever the person agrees with the assertion “temperature $t$ is both ‘hot’ and ‘warm’ with the same grade”. In Figure 3, there are only two such temperatures $t$. Therefore, even being a particular case in Fuzzy Logic these points in the proposition of theorem 2 are not so relevant.
References


APPENDIX A: Disagreements with Elkan's Proof

Here we analyze the reasons that led Elkan to his conclusions.

a) What does it mean “any assertions A and B”?

Probably the most polemic point in Elkan's proof is his understanding of the sentence “any assertion A and B”. Elkan correctly states that every possible values of \( t(A) \) and \( t(B) \) must be checked in order to prove theorem 1. Then, he assumes that every point \( p(t(A), t(B)) \) must be in some of the regions delimited by the following equivalencies:

\[
\begin{align*}
(a) & \quad A \rightarrow B \\
(b) & \quad \neg A \rightarrow \neg B \\
(c) & \quad A \rightarrow B \\
(d) & \quad A \rightarrow \neg B \\
(e) & \quad B \rightarrow A \\
(f) & \quad \neg B \rightarrow \neg A \\
(g) & \quad \neg B \rightarrow A \\
(h) & \quad B \rightarrow \neg A
\end{align*}
\]

Given definition 1 (where \( t(-A) = 1 - t(A) \)), checking all implications above implies in considering the points in square \([1,0] \times [1,0]\) twice. Given the hypothesis “A and B are any assertions”, the only assumption regarding \( t(A) \) and \( t(B) \) is that both points \( p_1(t(A), t(B)) \) and \( p_2(t(B), t(A)) \) must hold the theorem (i.e., the order is irrelevant). There is no need to check their complements, since there is nothing assumed for these sets.

b) There Is a Region Neglected in Elkan's Proof

During his proof, Elkan neglected a region of points when he defined the domain of \( p(t(A), t(B)) \) that do not verify the logical equivalence in theorem 1. Let us revise this step:

By analyzing expression (1), Elkan concluded that the numerical expressions are different if:

\[ t(B) < 1 - t(B) < 1 - t(A) \]

which leads to

\[ t(A) < t(B) < 0.5 \]  \quad (region R4)

This partially true because there are other points that do not hold expression (1). As we showed before, all points in region R1 (Fig. 1a) do not hold (1) and Elkan's region is only half of R1. The neglected region is given by:

\[ 0.5 < t(B) < 1 - t(A) \]  \quad (region R5)

Regions R4 and R5 are showed in Figure A1 and their union in Figure 1a. The negligence of region R5 did not affect Elkan's proof due to his interpretation of the sentence “any assertions A and B” (see item (a) in this appendix).
region R1: \( t(A) < t(B) < 1 - t(A) \)

region R2: \( t(B) < t(A) < 1 - t(B) \)
Figure 2: Points that violate Elkan’s theorem 1.
Table 1: Points in R1, R2, and R3 and the respective values of expressions (1) and (2).

<table>
<thead>
<tr>
<th>R_i</th>
<th>t(A)</th>
<th>t(B)</th>
<th>\neg(A \land \neg B)</th>
<th>\Leftrightarrow</th>
<th>B \lor (\neg A \land \neg B)</th>
<th>\neg(B \land \neg A)</th>
<th>\Leftrightarrow</th>
<th>A \lor (\neg B \land \neg A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>0.3</td>
<td>0.6</td>
<td>0.7</td>
<td>n</td>
<td>0.6</td>
<td>0.4</td>
<td>y</td>
<td>0.4</td>
</tr>
<tr>
<td>R2</td>
<td>0.7</td>
<td>0.1</td>
<td>0.3</td>
<td>y</td>
<td>0.3</td>
<td>0.9</td>
<td>n</td>
<td>0.7</td>
</tr>
<tr>
<td>R3</td>
<td>0.8</td>
<td>0.6</td>
<td>0.6</td>
<td>y</td>
<td>0.6</td>
<td>0.8</td>
<td>y</td>
<td>0.8</td>
</tr>
<tr>
<td>R3</td>
<td>0.3</td>
<td>0.9</td>
<td>0.9</td>
<td>y</td>
<td>0.9</td>
<td>0.3</td>
<td>y</td>
<td>0.3</td>
</tr>
</tbody>
</table>
Figure 3: Fuzzy Numbers that describe linguistic values of temperatures.
Figure A1: Points that do not verify the logical equivalence (1).