Least Sensitive (Most Robust) Fuzzy "Exclusive Or" Operations

Jesus E. Hernandez
Department of Electrical and Computer Engineering
University of Texas at El Paso
El Paso, TX 79968
jehernandez7@miners.utep.edu

Jaime Nava
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968
jenava@miners.utep.edu

***Abstract*—In natural language, "or" sometimes means "inclusive or" and sometimes means "exclusive or". To adequately describe commonsense and expert knowledge, it is therefore important to have not only t-conorms describing fuzzy "inclusive or" operations, but also fuzzy "exclusive or" operations $f_\oplus(a,b)$. Since the degrees of certainty are only approximately defined, it is reasonable to require that the corresponding operation be the least sensitive to small changes in the inputs. In this paper, we show that the least sensitive fuzzy "inclusive or" operation has the form $f_\oplus(a,b)=\min(\max(a,b),\max(1-a,1-b))$.**

I. Introduction

Need for fuzzy "exclusive or" operations. One of the main objectives of fuzzy logic is to formalize commonsense and expert reasoning. In commonsense and expert reasoning, people use logical connectives like "and" and "or". Depending on the context, commonsense "or" can mean both "inclusive or" – when "A or B" means that it is also possible to have both A and B, and "exclusive or" – when "A or B" means that one of the statements holds but not both.

For example, for a dollar, a vending machine can produce either a coke or a diet coke, but not both.

In mathematics and computer science, "inclusive or" is the one most frequently used as a basic operation. Because of this, fuzzy logic – an extension of usual logic to fuzzy statements characterized by "degree" of truth – is also mainly using "inclusive or" operations. However, since "exclusive or" is also used in commonsense and expert reasoning, there is a practical need for a fuzzy versions of this operation.

Comment. The "exclusive or" operation is actively used in computer design: since it corresponds to the bit-by-bit addition of binary numbers (the carry is the "and"). It is also actively used in quantum computing algorithms; see, e.g., [12].

Fuzzy versions of "exclusive or" operations are also known; see, e.g., [1]. These fuzzy versions are actively used in machine learning; see, e.g., [3], [7], [8], [14]. In particular, some of these papers (especially [8]) use a natural extension of fuzzy "exclusive or" from a binary to a k-ary operation.

A crisp "exclusive or" operation: a reminder. As usual with fuzzy operations, the fuzzy "exclusive or" operation must be an extension of the corresponding crisp operation. In the traditional 2-valued logic, with two possible truth values 0 (false) and 1 (true), the "exclusive or" operation $\oplus$ is defined as follows: $0\oplus 0=1\oplus 1=0$ and $0\oplus 1=1\oplus 0=1$. Thus, the desired fuzzy "exclusive or" operation $f_\oplus(a,b)$ must satisfy the same properties:

$$f_\oplus(0,0)=f_\oplus(1,1)=0;\quad f_\oplus(0,1)=f_\oplus(1,0)=1. \quad (1)$$

Need for the least sensitivity: reminder. Fuzzy logic operations deal with experts' degrees of certainty in their statements. These degrees are not precisely defined, the same expert can assign, say, 0.7 and 0.8 to the same degrees of belief. It is therefore reasonable to require that the result of the fuzzy operation not change much if we slightly change the inputs. A reasonable way to formalize this requirement is to require that the operation $f(a,b)$ satisfy the following property:

$$|f(a,b)-f(a',b')|\le k\cdot\max(|a-a'|,|b-b'|), \quad (2)$$

with the smallest possible value k among all operations $f(a,b)$ satisfying the given properties. Such operations are called *the least sensitive* or *the most robust*.

For t-norms and t-conorms, the least sensitivity requirement leads to reasonable operations. It is known that there is only one least sensitive t-norm ("and"-operation) $f_\&(a,b)=\min(a,b)$, and only one least sensitive t-conorm ("or"-operation) $f_\vee(a,b)=\max(a,b)$; see, e.g., [13], [9], [10].

What we do in this paper. In this paper, we describe the least sensitive fuzzy "exclusive or" operation.

II. Main Result

Definition 1. *A function $f:[0,1]\times[0,1]\to[0,1]$ is called a* fuzzy "exclusive or" *operation if it satisfies the following conditions: $f(0,0)=f(1,1)=0$ and $f(0,1)=f(1,0)=1$.*

Comment. We could also require other conditions, e.g., commutativity and associativity. However, our main objective is to select a single operation which is the least sensitive. The weaker the condition, the larger the class of operations that satisfy these conditions, and thus, the stronger the result that our operation is the least sensitive in this class.

Thus, to make our result as strong as possible, we selected the weakest possible condition – and thus, the largest possible class of "exclusive or" operations.

Definition 2. *Let F be a class of functions from $[0,1]\times[0,1]$ to $[0,1]$. We say that a function $f\in F$ is* the least sensitive *in the class F if for some real number k, the function f satisfies the condition*

$$|f(a,b)-f(a',b')|\le k\cdot\max(|a-a'|,|b-b'|),$$

and no other function $f\in F$ satisfies this condition.

Theorem. *In the class of all fuzzy "exclusive or" operations, the following function is the least sensitive:*

$$f_{\oplus}(a,b)=\min(\max(a,b),\max(1-a,1-b)). \quad (3)$$

Comments.

- This operation can be understood as follows. In the crisp (two-valued) logic, "exclusive or" $\oplus$ can be described in terms of the "inclusive or" operation $\vee$ as

$$a\oplus b\Leftrightarrow(a\vee b)\,\&\neg(a\,\&\,b).$$

  If we:
  - replace $\vee$ with the least sensitive "or"-operation $f_{\vee}(a,b)=\max(a,b)$,
  - replace $\&$ with the least sensitive "and"-operation $f_{\&}(a,b)=\min(a,b)$, and
  - replace $\neg$ with the least sensitive negation operation $f_{\neg}(a)=1-a$,

  then we get the expression (3) given in the Theorem.
- The above operation is associative and has a value a_0 (equal to 0.5) which satisfies the property $a\oplus a_0=a$ for all a. Thus, from the mathematical viewpoint, this operation is an example of a *nullnorm*; see, e.g., [2].

III. Proof of the Main Result

We will prove that the Theorem is true for $k=1$.

1°. First, let us prove that the operation (3) indeed satisfies the condition (2) with $k=1$. In other words, let us prove that for every $\varepsilon>0$, if $|a-a'|\le\varepsilon$ and $|b-b'|\le\varepsilon$, then $|f_{\oplus}(a,b)-f_{\oplus}(a',b')|\le\varepsilon$.

1.1°. It is known (see, e.g., [9], [10], [13]) that the functions $\min(a,b)$, $\max(a,b)$, and $1-a$ satisfy the condition (2) with $k=1$. In particular, this means that if $|a-a'|\le\varepsilon$ and $|b-b'|\le\varepsilon$, then we have

$$|\max(a,b)-\max(a',b')|\le\varepsilon \quad (4)$$

and also

$$|(1-a)-(1-a')|\le\varepsilon \text{ and } |(1-b)-(1-b')|\le\varepsilon. \quad (5)$$

1.2°. From (5), by using the property (2) for the $\max$ operation, we conclude that

$$|\max(1-a,1-b)-\max(1-a',1-b')|\le\varepsilon. \quad (6)$$

1.3°. Now, from (4) and (6), by using the property (2) for the $\min$ operation, we conclude that

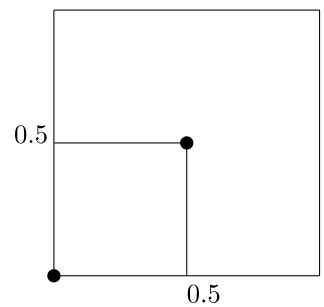
$$|\min(\max(a,b),\max(1-a,1-b))-$$
$$\min(\max(a',b'),\max(1-a',1-b'))|\le\varepsilon. \quad (7)$$

The statement is proven.

2°. Let us now assume that $f(a,b)$ is an exclusive or operation that satisfies the condition (2) with $k=1$. Let us prove that then $f(a,b)$ coincides with the function (3).

2.1°. Let us first prove that $f(0.5,0.5)=0.5$.

The proof can be illustrated by the following picture.



By the definition of the exclusive or operation, we have $f(0,0)=0$ and $f(0,1)=1$. Due to the property (2), we have

$$|f(0,0)-f(0.5,0.5)|\le\max(|0-0.5|,|0-0.5|)=0.5 \quad (8)$$

thus,

$$f(0,5,0.5)\le f(0,0)+0.5=0+0.5=0.5. \quad (9)$$

Similarly, due to the property (2), we have

$$|f(0,1)-f(0.5,0.5)|\le\max(|0-0.5|,|1-0.5|)=0.5 \quad (10)$$

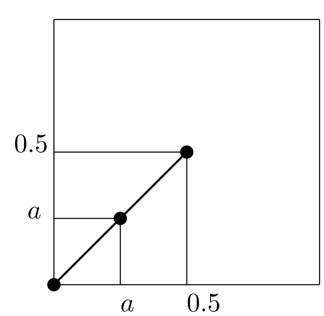
thus,

$$f(0.5,0.5)\ge f(0,1)-0.5=1-0.5=0.5. \quad (11)$$

From (9) and (11), we conclude that $f(0.5,0.5)=0.5$.

2.2°. Let us now prove that $f(a,a)=a$ for $a\le 0.5$.

This proof can be illustrated by the following picture.



Due to the property (2), we have

$$|f(0,0) - f(a,a)| \le \max(|0-a|, |0-a|) = a \tag{12}$$

thus,

$$f(a,a) \le f(0,0) + a = 0 + a = a. \tag{13}$$

Similarly, due to the property (2), we have

$$|f(0.5,0.5) - f(a,a)| \le \max(|0.5-a|, |0.5-a|) =$$

$$0.5 - a \tag{14}$$

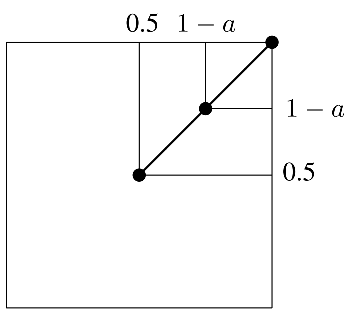
thus,

$$f(a,a) \ge f(0.5,0.5) - (0.5-a) = 0.5 - (0.5-a) = a. \tag{15}$$

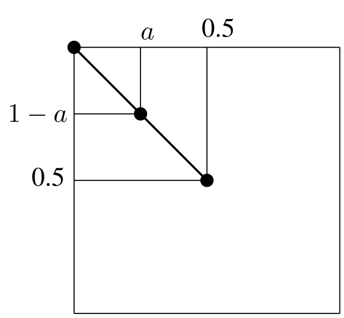
From (13) and (15), we conclude that $f(a,a) = a$.

2.3°. Similarly:

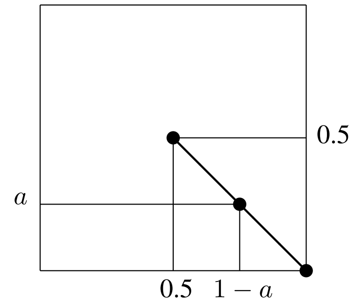
- by considering the points $(0.5, 0.5)$ and $(1,1)$, we conclude that
$$f(1-a, 1-a) = a$$
for $a \le 0.5$;



- by considering the points $(0.5, 0.5)$ and $(0,1)$, we conclude that
$$f(a, 1-a) = 1-a$$
for $a \le 0.5$;



- by considering the points $(0.5, 0.5)$ and $(1,0)$, we conclude that
$$f(1-a, a) = 1-a$$
for $a \le 0.5$.

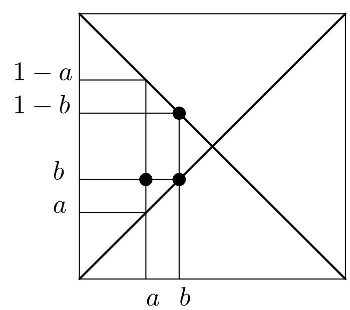


Summarizing: we have just proved that the formula (6) holds when $b = a$ and when $b = 1 - a$.

2.4°. Let us now prove that the formula (6) holds for arbitrary a and b.

In principle, we can have four cases depending on whether $b \le a$ or $b \ge a$ and on whether $b \le 1-a$ or $b \ge 1-a$. Without losing generality, let us consider the case when $b \le a$ and $b \le 1-a$; the other three cases can be proven in a similar way.

The proof for this case can be illustrated by the following picture.



For this case, we know, from Parts 2.2 and 2.3 of this proof, that $f(b,b) = b$ and $f(1-b, b) = 1-b$. Here, $b \le a \le 1-b$. Due to the property (2), we have

$$|f(a,b) - f(b,b)| \le \max(|a-b|, |b-b|) = a - b, \tag{16}$$

thus,

$$f(a,b) \le f(b,b) + (a-b) = b + (a-b) = a. \tag{17}$$

Similarly, due to the property (2), we have

$$|f(a,b) - f(1-b,b)| \le \max(|a-(1-b)|, |b-b|) =$$

$$(1-b) - a, \tag{18}$$

thus,

$$f(a,b) \ge f(1-b,b) - ((1-b)-a) =$$

$$(1-b) - ((1-b)-a) = a. \tag{19}$$

From (17) and (19), we conclude that $f(a,b) = a$.

Similarly:

- for $b \le a$ and $b \ge 1-a$, i.e., when $1-a \le b \le a$, by considering the points $(a, 1-a)$ and (a, a), we conclude that $f(a,b) = 1-a$;
- for $b \ge a$ and $b \le 1-a$, i.e., when $a \le b \le 1-a$, by considering the points (a, a) and $(a, 1-a)$, we conclude that $f(a,b) = b$;
- for $b \ge a$ and $b \ge 1-a$, i.e., when $1-b \le a \le b$, by considering the points $(1-b, b)$ and (b, b), we conclude that $f(a,b) = 1-b$.

In other words, we prove that the formula (6) holds for all a and b. The theorem is proven.

IV. Fuzzy "Exclusive Or" Operations Which Are The Least Sensitive On Average

Average sensitivity: reminder. As we have mentioned earlier, the fuzzy degrees are given with some uncertainty. In other words, different experts – and even the same expert at different times – would assign somewhat different numerical values to the same degree of certainty. In the main part of the paper, we have showed how to select fuzzy operations $c = f(a,b)$ in such a way that "in the worst case", the change in a and b would lead to the smallest possible change in the value $c = f(a,b)$.

Another reasonable possibility is to select fuzzy operations $c = f(a,b)$ in such a way that "on average", the change in a and b would lead to the smallest possible change in the value $c = f(a,b)$.

For each pair of values a and b, it is reasonable to assume that the differences Δa and Δb between the different numerical values corresponding to the same degree of certainty are independent random variables with 0 mean and small variance σ^2. Since the differences Δa and Δb are small, we can expand the difference $\Delta c = f(a+\Delta a, b+\Delta b) - f(a,b)$ in Taylor series with respect to Δa and Δb and keep only linear terms in this expansion:

$$\Delta c \approx \frac{\partial f}{\partial a} \cdot \Delta a + \frac{\partial f}{\partial b} \cdot \Delta b. \tag{20}$$

Since the variance are independent with 0 mean, the mean of Δc is also 0, and variance of Δc is equal to

$$\sigma^2(a,b) = \left(\left(\frac{\partial f}{\partial a} \right)^2 + \left(\frac{\partial f}{\partial b} \right)^2 \right) \cdot \sigma^2. \tag{21}$$

This is the variance for given a and b. To get the average variance, it is reasonable to average this value over all possible values of a and b, i.e., to consider the value

$$I \cdot \sigma^2,$$

where

$$I \stackrel{\text{def}}{=} \int_{a=0}^{a=1} \int_{b=0}^{b=1} \left(\left(\frac{\partial f}{\partial a} \right)^2 + \left(\frac{\partial f}{\partial b} \right)^2 \right) da\, db. \tag{22}$$

Thus, the average sensitivity is the smallest if, among all possible functions $f(a,b)$ satisfying the given constraints, we select a function for which the integral I takes the smallest possible value.

Average sensitivity: known results. [11], [13]

- For negation operations, this approach select the standard function
$$f_\neg(a) = 1 - a.$$
- For "and"-operations (t-norms), this approach selects $f_\&(a,b) = a \cdot b$.
- For "or"-operations (t-conorms), this approach selects $f_\vee(a,b) = a + b - a \cdot b$.

New result: formulation. We consider *"exclusive or"* operations, i.e., functions $f(a,b)$ from $[0,1] \times [0,1]$ to $[0,1]$ for which $f(0,b) = b$, $f(a,0) = a$, $f(1,b) = 1-b$, and $f(a,1) = 1-a$.

Our main result is that among all such operations, the operation which is the least sensitive on average has the form

$$f_\oplus(a,b) = a + b - 2 \cdot a \cdot b. \tag{23}$$

Comment. This operation can be explained as follows:

- First, we represent the classical (2-valued) "exclusive or" operation $a \oplus b$ as $(a \vee b) \& (\neg a \vee \neg b)$.
- Then, to get a fuzzy analogue of this operation, we replace $p \vee q$ with $p + q - p \cdot q$, $\neg p$ with $1 - p$, and $p \,\&\, q$ with $\max(p + q - 1, 0)$.

Indeed, in this case,

$$a \vee b = a + b - a \cdot b;$$

$$\neg a \vee \neg b = (1-a) \vee (1-b) =$$

$$(1-a) + (1-b) - (1-a) \cdot (1-b) =$$

$$1 - a + 1 - b - (1 - a - b + a \cdot b) =$$

$$1 - a + 1 - b - 1 + a + b - a \cdot b = 1 - a \cdot b,$$

and thus,

$$(a \vee b) + (\neg a \vee \neg b) - 1 =$$

$$a + b - a \cdot b + 1 - a \cdot b - 1 =$$

$$a + b - 2 \cdot a \cdot b.$$

For values $a, b \in [0,1]$, we have $a^2 \le a$ and $b^2 \le b$, hence

$$(a \vee b) + (\neg a \vee \neg b) - 1 =$$

$$a + b - 2 \cdot a \cdot b \ge a^2 + b^2 - 2 \cdot a \cdot b =$$

$$(a - b)^2 \ge 0,$$

therefore, indeed

$$(a \vee b) \& (\neg a \vee \neg b) =$$

$$\max((a \vee b) + (\neg a \vee \neg b) - 1, 0) =$$

$$(a \vee b) + (\neg a \vee \neg b) - 1.$$

This replacement operation sounds arbitrary, but the resulting "exclusive or" operation is uniquely determined by the sensitivity requirement.

V. Proof of the Auxiliary Result

It is known similarly to the fact that the minimum of a function is always attained at a point where its derivative is 0, the minimum of a functional is always attained at a function where it *variational derivative* is equal to 0 (see, e.g., [5]; see also [11], [13]):

$$\frac{\delta L}{\delta f} = \frac{\partial L}{\partial f} - \sum_i \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial f_i}\right) = 0,$$

where $f_{,i} \stackrel{\rm def}{=} \dfrac{\partial f}{\partial x_i}$.

Applying this *variational equation* to the functional $I = \int L\,da\,db$, with $L = \left(\dfrac{\partial f}{\partial a}\right)^2 + \left(\dfrac{\partial f}{\partial b}\right)^2$, we conclude that

$$-\frac{\partial}{\partial a}\left(2\cdot\frac{\partial f}{\partial a}\right) - \frac{\partial}{\partial b}\left(2\cdot\frac{\partial f}{\partial b}\right) = 0,$$

i.e., we arrive at the equation

$$\nabla^2 f = 0, \tag{24}$$

where $\nabla \stackrel{\rm def}{=} \left(\dfrac{\partial f}{\partial a}, \dfrac{\partial f}{\partial b}\right)$ and

$$\nabla^2 f = \frac{\partial^2 f}{\partial a^2} + \frac{\partial^2 f}{\partial b^2}.$$

The equation (24) is known as the *Laplace equation*, and it is known (see, e.g., [4]) that a solution to this equation is uniquely determined by the boundary conditions – i.e., in our case, by the values on all four parts of the boundary of the square $[0,1]\times[0,1]$: lines segments $a = 0$, $a = 1$, $b = 0$, and $b = 1$. One can easily show that the above function $f(a,b) = a + b - 2\cdot a\cdot b$ satisfies the Laplace equation – since both its second partial derivatives are simply 0s. It is also easy to check that for all four sides, this function coincides with our initial conditions:

- when $a = 0$, we get $f(a,b) = 0 + b - 2\cdot 0\cdot b = b$;
- when $a = 1$, we get $f(a,b) = 1 + b - 2\cdot 1\cdot b = 1 - b$;
- when $b = 0$, we get $f(a,b) = a + 0 - 2\cdot a\cdot 0 = a$;
- when $b = 1$, we get $f(a,b) = a + 1 - 2\cdot 1\cdot a = 1 - a$.

Thus, due to the above property of the Laplace equation, the function $f(a,b) = a + b - 2\cdot a\cdot b$ is the only solution to this equation with the given initial condition – therefore, it coincides with the desired *the least sensitive on average* "exclusive or" operation (which satisfies the same Laplace equation with the same boundary conditions.

The theorem is proven.

Acknowledgments.

The authors would like to thank Drs. Patricia Nava and Vladik Kreinovich for their help.

References

[1] B. C. Bedregal, R. H. S. Reiser, and G. P. Dimuro, "Xor-implications and e-implications: classes of fuzzy implications based on fuzzy xor", *Electronic Notes in Theoretical Computer Science (ENTCS)*, 2009, Vol. 247, pp. 5–18.

[2] G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, Springer-Verlag, Berlin-Heidelberg-New York, 2007.

[3] L. Y. Cai and H. K. Kwan, "Fuzzy Classifications Using Fuzzy Inference Networks", *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, 1998, Vol. 28, No. 3, pp. 334–347.

[4] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, Providence, Rhode Island, 1998.

[5] I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Dover Publ., New York (2000)

[6] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Upper Saddle River, New Jersey: Prentice Hall, 1995.

[7] H. Le Capitaine and C. Frélicot, "A new fuzzy 3-rules pattern classifier with reject options based on aggregation of membership degrees", *Proceedings of the 12th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'08*, Malaga, Spain, 2008.

[8] L. Mascarilla, M. Berthier, and C. Frélicot, "A k-order fuzzy OR operator for pattern classification with k-order ambiguity rejection", *Fuzzy Sets and Systems*, 2008, Vol. 159, No. 15, pp. 2011–2029.

[9] H. T. Nguyen and V. Kreinovich, "Towards theoretical foundations of soft computing applications", *International Journal on Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS)*, 1995, Vol. 3, No. 3, pp. 341-373.

[10] H. T. Nguyen, V. Kreinovich, and D. Tolbert, "On robustness of fuzzy logics". *Proceedings of the 1993 IEEE International Conference on Fuzzy Systems FUZZ-IEEE'93*, San Francisco, California, March 1993, Vol. 1, pp. 543–547.

[11] H. T. Nguyen, V. Kreinovich, and D. Tolbert, "A measure of average sensitivity for fuzzy logics", *International Journal on Uncertainty, Fuzziness, and Knowledge-Based Systems*, 1994, Vol. 2, No. 4, pp. 361–375.

[12] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[13] H. T. Nguyen and E. A. Walker, *First Course on Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

[14] W. Pedrycz and G. Succi, "fXOR fuzzy logic networks", *Soft Computing*, 2002, Vol. 7, pp. 115–120.