7-1-2013

Minimization of Average Sensitivity as a Method of Selecting Fuzzy Functions and Operations: Successes and Limitations

Riya George
University of Texas at El Paso, rmgeorge@miners.utep.edu

Suresh Subramanian
University of Texas at El Paso, ssubramanian@miners.utep.edu

Alejandro Vega
University of Texas at El Paso, avega5@miners.utep.edu

Olga Kosheleva
University of Texas at El Paso, olgak@utep.edu

Comments:
Technical Report: UTEP-CS-13-47
To appear in *Journal of Uncertain Systems*, 2014, Vol. 8.

Journal of Uncertain Systems
Vol.8, No.x, pp.xx-xx, 2014
Online at: www.jus.org.uk

Minimization of Average Sensitivity as a Method of Selecting Fuzzy Functions and Operations: Successes and Limitations

Riya George, Suresh Subramanian,
Alejandro Vega, and Olga Kosheleva
University of Texas at El Paso
500 W. University, El Paso, TX 79968, USA
rmgeorge@miners.utep.edu, ssubramanian@miners.utep.edu,
avega5@miners.utep.edu, olgak@utep.edu

Received 27 July 2012; Revised 31 July 2013

Abstract

Fuzzy logic is an extension of the standard 2-valued logic – with two possible truth values 0 ("false") and ("true") – to values (degrees of certainty) represented by arbitrary numbers from the interval $[0, 1]$. One of the main challenges in fuzzy logic is that we need to extend the usual logical operations from the set $\{0, 1\}$ to the entire interval, and there are many possible extensions. One promising technique for selecting a reasonable extension is to take into account that the fuzzy degrees of certainty are themselves only known with uncertainty; so, it makes sense to select an operation which is, on average, the least sensitive to the corresponding uncertainty. This technique has successfully worked in selecting unary and binary operations and in selecting membership functions. In this paper, we show, however, that this minimization technique does not work well for selecting ternary operations, and that in the discrete case, the results of applying this technique are somewhat counterintuitive.

Keywords: fuzzy logic, average sensitivity, "'and"-operations, "or"-operations, discrete operations, ternary operations

1 Minimization of Average Sensitivity: Description of a Method

Need for fuzzy logic. In the traditional 2-valued fuzzy logic, every statement is either true or false. The corresponding truth value is represented in the computer as, correspondingly, 1 or 0. The need to go beyond the 2-valued logic comes from the fact that experts are often uncertain ("fuzzy") about the truth values of different statements describing their knowledge: they may be confident to some extent but not fully that a given statement is true. For example, a medical expert can use rules using the word "small", but this expert is not 100% sure whether a given size tumor is small or not. To describe such imprecise ("fuzzy") statements, L. A. Zadeh proposed, in [6], to go beyond the usual "degrees of confidence" (truth values) 0 and 1 and to use numbers from the interval $[0, 1]$ to describe degrees of confidence intermediate between 0 (absolutely false) and 1 (absolutely true).

Fuzzy logic has been very successful. The resulting formalism of fuzzy logic has many successful applications; see, e.g., [2, 4].

Fuzzy logic: challenges. Once we extended the set of truth values from the 2-element set $\{0, 1\}$ to the entire interval $[0, 1]$, we need to extend operations with these truth values – e.g., logical operations such as "or" and "and" – to the entire interval. There are many different ways to extend a function to a larger set, and it is not immediately clear which of these extensions we should select.

Minimizing average sensitivity: main idea. We want to extend 2-valued operations to intermediate degrees of confidence. These intermediate degrees describe the expert's evaluation of their own uncertainty. We consider situations in which the experts are not 100% certain about their knowledge; they are similarly not 100% certain about their own degrees of certainty. The same expert can represent the same degree of certainty by slightly different values: instead of the original degree a, the same expert may produce a slightly different value $a + \Delta a$, with some unpredictable ("random") different Δa.

Since the values a and $a + \Delta a$ represent the same degree of certainty, it is reasonable to require that the results of applying the corresponding operations be the same – or at least close to each other. Since the deviations are unpredictable (random), we cannot guarantee that we always get the same results, but we should at least require that, on average, we get similar results.

The larger the average effect of the difference(s) Δa on the result of the operation, the less adequate is this operation is capturing expert reasoning. It is therefore reasonable to select an operation which is the most reasonable in this sense – i.e., for which the average sensitivity to changing the inputs degrees is the smallest possible. This is the main idea behind the method of minimizing average sensitivity proposed in [3] (see also [4]).

Minimization of average sensitivity: technical details. To formally describe the method, let us explain how we can gauge the average sensitivity of an operation $f : [0,1] \times \ldots \times [0,1] \to [0,1]$. For this operation, if we replace the original inputs $a, \ldots, b$ with modified inputs $a+\Delta a, \ldots, b+\Delta b$, then the result of the operation changes from the original value $y = f(a, \ldots, b)$ to the modified value $y_{\rm mod} = f(a+\Delta a, \ldots, b+\Delta b)$. Since the differences $\Delta a, \ldots, \Delta b$ are small, we can expand the difference $\Delta y \stackrel{\rm def}{=} y_{\rm mod} - y$ into Taylor series in $\Delta a, \ldots, \Delta b$ and keep only linear terms in this expansion. As a result, we get the expression

$$\Delta y \approx \frac{\partial f}{\partial a} \cdot \Delta a + \ldots + \frac{\partial f}{\partial b} \cdot \Delta b. \tag{1}$$

Since the differences $\Delta a, \ldots, \Delta b$ are caused by many different independent factors, it makes sense to use the Central Limit Theorem, according to which, under reasonable conditions, the distribution of the sum of many independent random variables is close to Gaussian; see, e.g., [5]. In general, a Gaussian distribution is uniquely determined by its mean E and it standard deviation σ. Since there is no reason to assume that positive deviations Δa are more frequent or more rare than negative ones, it is reasonable to assume that the mean value $E[\Delta a]$ of each deviation is 0. There is also no reason to think that different degrees have different uncertainty, so it is reasonable to assume that all the differences $\Delta a, \ldots, \Delta b$ have the same standard deviation σ. Under these assumptions, the linear combination (1) of several normally distributed random variables $\Delta a, \ldots, \Delta b$ is also normally distributed, with 0 mean and variance $V = C \cdot \sigma^2$, where

$$C \stackrel{\rm def}{=} \left(\frac{\partial f}{\partial a}\right)^2 + \ldots + \left(\frac{\partial f}{\partial b}\right)^2. \tag{2}$$

The average value of this variance can be thus described as $V_{\rm av} = C_{\rm av} \cdot \sigma^2$, where

$$C_{\rm av} \stackrel{\rm def}{=} \int_0^1 \ldots \int_0^1 \left[\left(\frac{\partial f}{\partial a}\right)^2 + \ldots + \left(\frac{\partial f}{\partial b}\right)^2\right] da \ldots db. \tag{3}$$

Thus, minimizing this measure of average sensitivity $V_{\rm av}$ is equivalent to minimizing the integral expression (3).

2 Minimization of Average Sensitivity: Successes

Selection of negation operations. In the 2-valued logic, negation $f(a)$ is described by setting $f(0) = 1$ and $f(1) = 0$. We would like to extend this operation to all possible values from the interval $[0,1]$, i.e., to consider functions $f : [0,1] \to [0,1]$ for which $f(0) = 1$ and $f(1) = 0$. Out of all such functions, we want to find the one which minimizes the expression $\int_0^1 \left(\frac{df}{da}\right)^2 da$. It turns out that this minimum is attained when $f(a) = 1 - a$ (see, e.g., [3]), which is exactly the negation operation most frequently used in fuzzy logic.

Note that we do not have to explicitly require that $f(f(x)) = x$: this property automatically follows from the requirement that the operation provide the minimum of average sensitivity.

Selection of membership functions. Similar arguments can be used to select a membership function $\mu(x)$ corresponding to a certain notion (e.g., "small"), i.e., a function which describes, for each value x, to what extend we believe that this value x satisfies the given property (e.g., to what extent x is small). Here, the input x is usually estimated subjectively, so we have a similar subjective difference between two consequent estimates x and $x + \Delta x$ by the same expert. Minimizing the effect of this subjectivity leads to a similar problem of minimizing the integral $\int_0^1 \left(\frac{df}{dx}\right)^2 dx$, which results in membership functions which are linear on each of the analyzed x-segments [3]. The corresponding piece-wise linear membership functions (triangular, trapezoidal, etc.) are indeed among the most frequently used in applications of fuzzy logic.

Selection of "and" operations. Intuitively, "false and A" means false, "true and A" means simply A, and "A and B" means the same as "B and A". Thus, as "and"-operations, it is reasonable to consider functions $f : [0,1] \times [0,1] \to [0,1]$ for which, for all a, we have $f(0,a) = f(a,0) = 0$ and $f(1,a) = f(a,1) = a$. It turns out [3] that among all such operations, the minimum of average sensitivity is attained when $f(a,b) = a \cdot b$. This "algebraic product" operation is indeed among the most widely used in fuzzy logic.

Note that we do not have to explicitly require associativity or monotonicity: these properties automatically follows from the requirement that the operation provide the minimum of average sensitivity.

Selection of "or" operations. Intuitively, "false or A" means A, "true or A" means simply true, and "A or B" means the same as "B or A". Thus, as "or"-operations, it is reasonable to consider functions $f : [0,1] \times [0,1] \to [0,1]$ for which, for all a, we have $f(0,a) = f(a,0) = a$ and $f(1,a) = f(a,1) = 1$. It turns out [3] that among all such operations, the minimum of average sensitivity is attained when $f(a,b) = a+b-a\cdot b$. This "algebraic sum" operation is indeed among the most widely used in fuzzy logic.

Note that here too, we do not have to explicitly require associativity or monotonicity: these properties automatically follows from the requirement that the operation provide the minimum of average sensitivity.

Selection of "exclusive or" ("xor") operations. Intuitively, "false xxor A" means A, "true or A" means "not A", and "A xor B" means the same as "B xor A". Thus, as "xor"-operations, it is reasonable to consider functions $f : [0,1] \times [0,1] \to [0,1]$ for which, for all a, we have $f(0,a) = f(a,0) = a$ and $f(1,a) = f(a,1) = 1 - a$. It turns out [1] that among all such operations, the minimum of average sensitivity is attained when $f(a,b) = a + b - 2 \cdot a \cdot b$. This operation can be understood if we describe "xor" in terms of "not", "and", and "or" operations, as $(a \vee b) \,\&\, (\neg a \vee \neg b)$, and use $1 - a$ for negation, $a + b - a \cdot b$ for "or", and $\max(a + b - 1, 0)$ for "and".

3 Minimization of Average Sensitivity: Challenges

Selecting ternary operations: general description. So far, we have been applying the minimization of average sensitivity technique to unary and binary operations. What happens if we apply this technique to ternary operations, e.g., to an operation $f(a,b,c)$ corresponding to the triple conjunction $A \,\&\, B \,\&\, C$? Similarly to the binary case, we can set up reasonable values on the border of the corresponding unit cube $[0,1] \times [0,1] \times [0,1]$, i.e., for the triples (a,b,c) which contain 0 or 1 as one of their values. Will the above minimization technique help us to find the values $f(a,b,c)$ for the triples (a,b,c) inside the unit cube? Somewhat surprisingly, in this case, minimization does not help.

Selecting ternary "and" operations. Let us show that among all ternary "and" operations $f(a,b,c)$, i.e., among all the functions with $f(a,b,0) = f(a,0,b) = f(0,a,b) = 0$ and $f(1,1,1) = 1$, the smallest possible value $(= 0)$ of the expression (3) is attained for the "crisp" function which is equal to 0 for all $(a,b,c) \neq (1,1,1)$ and to 1 for $(a,b,c) = (1,1,1)$. This can be explained if we consider functions which are equal to 0 except for a small cube $[1-\varepsilon,1] \times [1-\varepsilon,1] \times [1-\varepsilon,1]$ and which, on this cube, rapidly increase to 0. On the corresponding interval $[1-\varepsilon,1]$ of width ε, the function f increases from 0 to 1, thus, its slope on this integral is equal to $\frac{1}{\varepsilon}$: $\frac{\partial f}{\partial a} \approx \frac{1}{\varepsilon}$. Hence, the integrated expression in the formula (3) is proportional to $\frac{1}{\varepsilon^2}$, and its integral over the small cube of size $\varepsilon \times \varepsilon \times \varepsilon$ and of volume ε^3 is proportional to $\frac{1}{\varepsilon^2} \cdot \varepsilon^3 = \varepsilon$. Thus, when ε tends to 0, the value of the expression (3) also tends to 0.

Selecting ternary "or" operations. Let us show that among all ternary "or" operations $f(a,b,c)$, i.e., among all the functions with $f(a,b,1) = f(a,1,b) = f(1,a,b) = 1$ and $f(0,0,0) = 0$, the smallest possible value ($= 0$) of the expression (3) is attained for the "crisp" function which is equal to 1 for all $(a,b,c) \neq (0,0,0)$ and to 0 for $(a,b,c) = (0,0,0)$. This can be explained if we consider functions which are equal to 1 except for a small cube $[0,\varepsilon] \times [0,\varepsilon] \times [0,\varepsilon]$ and which, on this cube, rapidly decrease to 0. On the corresponding interval $[0,\varepsilon]$ of width ε, the function f increases from 0 to 1, thus, its slope on this integral is equal to $\frac{1}{\varepsilon}$: $\frac{\partial f}{\partial a} \approx \frac{1}{\varepsilon}$. Hence, the integrated expression in the formula (3) is proportional to $\frac{1}{\varepsilon^2}$, and its integral over the small cube of size $\varepsilon \times \varepsilon \times \varepsilon$ and of volume ε^3 is proportional to $\frac{1}{\varepsilon^2} \cdot \varepsilon^3 = \varepsilon$. Thus, when ε tends to 0, the value of the expression (3) also tends to 0.

Discrete case: general description. Instead of considering all possible values from the interval $[0,1]$, we can consider only finitely many values. For example, we can pick some value n, and consider only values $a_i = \frac{i}{n}$ with $i = 0,1,\ldots,n$, i.e., only values $0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1$. In this case, we need to find the values $f_{i,j} \stackrel{\text{def}}{=} f(a_i,a_j)$. In the discrete case, instead of partial derivatives, we have differences, and instead of the integral, we get the sum. Thus, a natural analogue of the formula (3) is to minimize the expression

$$\sum_{i=0}^{n}\sum_{j=0}^{n-1}(f_{i,j+1} - f_{i,j})^2 + \sum_{i=0}^{n-1}\sum_{j=0}^{n}(f_{i+1,j} - f_{i,j})^2. \quad (4)$$

Selecting discrete "and" operations: success. Everything works OK if we require that $f_{0,i} = f_{0,i} = 0$ and $f_{n,i} = f_{i,n} = \frac{i}{n}$ for all i. In this case, we get the same formula $f_{i,j} = \frac{i}{n} \cdot \frac{j}{n}$ as in the continuous case.

Selecting discrete "and" operations: challenge. In the case of negation, we simply fixed the values for the 2-valued logic, and used the minimization of average sensitivity to find all the other values. Why not try the same approach here? Let us fix only the values $f(0,0) = f(0,1) = f(1,0) = 0$ and $f(1,1) = 1$, i.e., only the values $f_{0,0} = f_{0,n} = f_{n,0} = 0$ and $f_{n,n} = 1$, and among all such matrices $f_{ij,j}$, let us select the one which minimizes the average sensitivity (4).

Selecting discrete "and" operation: resulting equations. Differentiating the expression (4) wit respect to $f_{i,j}$ and equating the derivatives to 0, we get the following equations:

- When $0 < i < n$ and $0 < j < n$, i.e., for the points (a_i,a_j) inside the square for which there are four immediate neighbors, we conclude that the value $f_{i,j}$ is equal to the arithmetic average of the values of these three neighbors:
$$f_{i.j} = \frac{1}{4} \cdot (f_{i,j-1} + f_{i,j+1} + f_{i-1,j} + f_{i+1,j}). \quad (5)$$
- For pairs (i,j) on the border, there are only three neighbors, and the corresponding equation states that the value $f_{i,j}$ is equal to the arithmetic average of the three neighboring values.

The simplest case of $n = 2$: computations. 2-valued logic corresponds to $n = 1$, the simplest non-trivial case is the case $n = 2$. In this case, we fix the values $f_{0,0} = f_{0,2} = f_{2,0} = 0$ and $f_{2,2} = 1$, and we need to determine values $f_{0,1}$, $f_{1,0}$, $f_{1,1}$, $f_{1,2}$, and $f_{2,1}$. Since the above system of linear equations does not change if we swap i and j, the solution also should not change under such a swap. So, we must have $f_{0,1} = f_{1,0}$ and $f_{1,2} = f_{2,1}$, and, in effect, we only have three unknowns: $f_{0,1}$, $f_{1,1}$, and $f_{1,2}$. After we take into account this symmetry, for these three unknowns, the above equal-to-the-average equations take the following form:

$$f_{1,1} = \frac{1}{4} \cdot (2f_{0,1} + 2f_{1,2}); \quad (6)$$

$$f_{0,1} = \frac{1}{3} \cdot (f_{0,0} + f_{2,0} + f_{1,1}) = \frac{f_{1,1}}{3}; \tag{7}$$

$$f_{1,2} = \frac{1}{3} \cdot (f_{1,1} + f_{0,2} + f_{2,2}) = \frac{1 + f_{1,1}}{3}. \tag{8}$$

Substituting $f_{0,1} = \frac{f_{1,1}}{3}$ and $f_{0,1} = \frac{1 + f_{1,1}}{3}$ into the formula (6), we conclude that

$$f_{1,1} = \frac{1}{2} \cdot (f_{0,1} + f_{1,2}) = \frac{1}{2} \cdot \left(\frac{f_{1,1}}{3} + \frac{1 + f_{1,1}}{3}\right) = \frac{f_{1,1}}{3} + \frac{1}{6}. \tag{9}$$

Thus, $\frac{2}{3} \cdot f_{1,1} = \frac{1}{6}$, and hence, $f_{1,1} = \frac{1}{4}$. Substituting this value $f_{1,1}$ into the formulas (6) and (7), we get $f_{0,1} = \frac{1}{12}$ and $f_{1,2} = \frac{5}{12}$. So here, the matrix $f_{i,j}$ has the form

$$f_{i,j} = \begin{pmatrix} f_{0,2} = 0 & f_{1,2} = \frac{5}{12} & f_{2,2} = 1 \\ f_{0,1} = \frac{1}{12} & f_{1,1} = \frac{1}{4} & f_{2,1} = \frac{5}{12} \\ f_{0,0} = 0 & f_{1,0} = \frac{1}{12} & f_{2,0} = 0 \end{pmatrix}, \tag{10}$$

i.e.,

&	0	0.5	1
1	0	0.4167	1
0.5	0.0833	0.2500	0.4167
0	0	0.0833	0

The simplest case of $n = 2$: discussion. Somewhat surprisingly, we do not get the expected values $f(0, a) = 0$ and $f(1, a) = a$. Moreover, we lose monotonicity, since here:

$$f(0, 0) = 0 < f(0, 0.5) = \frac{1}{12} > f(0, 1) = 0.$$

For "or" operations, the results are similar. For "or" operation $f_{i,j}$, if we require that $f_{0,i} = f_{i,0} = \frac{i}{n}$ and $f_{i,n} = f_{n,i} = 1$ for all i, then we get a good result $f_{i,j} = \frac{i}{n} + \frac{j}{n} - \frac{i}{n} \cdot \frac{j}{n}$. However, if we only require that $f_{0,0} = 0$ and that $f_{0,n} = f_{n,0} = f_{n,n} = 1$, then, for $n = 2$, a similar minimization leads to

$$f_{i,j} = \begin{pmatrix} f_{0,2} = 1 & f_{1,2} = \frac{11}{12} & f_{2,2} = 1 \\ f_{0,1} = \frac{7}{12} & f_{1,1} = \frac{3}{4} & f_{2,1} = \frac{11}{12} \\ f_{0,0} = 0 & f_{1,0} = \frac{7}{12} & f_{2,0} = 1 \end{pmatrix}, \tag{11}$$

i.e.,

$\vee$	0	0.5	1
1	1	0.9167	1
0.5	0.5833	0.7500	0.9167
0	0	0.5833	1

Here also, we do not get the expected values $f(0,a)=a$ and $f(1,a)=1$, and we also lose monotonicity:

$$f(1,0)=1>f(1,0.5)=\frac{11}{12}<f(1,1)=1.$$

Similar results for $n=3$, $n=4$, and $n=5$. For larger values of n, we can use software to solve the corresponding systems of linear equations. The resulting solutions (presented in the Appendix) is also non-monotonic.

References

[1] J. E. Hernandez and J. Nava, "Least sensitive (most robust) fuzzy 'exclusive or' operations", *Proceedings of the 30th Annual Conference of the North American Fuzzy Information Processing Society NAFIPS'2011*, El Paso, Texas, March 18–20, 2011.

[2] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[3] 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.

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

[5] D. J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman and Hall/CRC Press, Boca Raton, Florida, 2011.

[6] L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.

A Discrete "And" and "Or" Operations with the Smallest Average Sensitivity: Cases of $n=3$, $n=4$, and $n=5$

Case of $n=3$.

&	0	1/3	2/3	1
1	0	0.2588	0.5490	1
2/3	0.1049	0.2273	0.3882	0.5490
1/3	0.0874	0.1574	0.2273	0.2588
0	0	0.0874	0.1049	0

∨	0	1/3	2/3	1
1	1	0.8951	0.9126	1
2/3	0.7413	0.7727	0.8427	0.9126
1/3	0.4510	0.6119	0.7727	0.8951
0	0	0.4510	0.7413	1

Case of $n=4$.

&	0	0.25	0.50	0.75	1
1	0	0.1926	0.3777	0.6160	1
0.75	0.1074	0.2000	0.3245	0.4702	0.6160
0.50	0.1223	0.1755	0.2500	0.3245	0.3777
0.25	0.0840	0.1298	0.1755	0.2000	0.1926
0	0	0.0840	0.1223	0.1074	0

$\vee$	0	0.25	0.50	0.75	1
1	1	0.8926	0.8777	0.9160	1
0.75	0.8074	0.8000	0.8245	0.8702	0.9160
0.50	0.6223	0.6755	0.7500	0.8245	0.8777
0.25	0.3840	0.5298	0.6755	0.8000	0.8926
0	0	0.3840	0.6223	0.8074	1

Case of $n = 5$.

&	0	0.2	0.4	0.6	0.8	1
1	0	0.1578	0.2941	0.4476	0.6573	1
0.8	0.1048	0.1793	0.2768	0.3941	0.5244	0.6573
0.6	0.1352	0.1777	0.2424	0.3169	0.3914	0.4476
0.4	0.1231	0.1540	0.1982	0.2424	0.2768	0.2941
0.2	0.0800	0.1170	0.1540	0.1777	0.1793	0.1578
0	0	0.0800	0.1231	0.1352	0.1048	0

$\vee$	0	0.2	0.4	0.6	0.8	1
1	1	0.8952	0.8648	0.8769	0.9200	1
0.8	0.8422	0.8207	0.8223	0.8460	0.8830	0.9200
0.6	0.7059	0.7232	0.7576	0.8018	0.8460	0.8769
0.4	0.5524	0.6086	0.6831	0.7576	0.8223	0.8648
0.2	0.3427	0.4756	0.6086	0.7232	0.8207	0.8952
0	0	0.3427	0.5524	0.7059	0.8422	1