Towards a More Subtle (and Hopefully More Adequate) Fuzzy "And"-Operation: Normalization-Invariant Multi-Input Aggregation Operators

Yusuf Güven and Vladik Kreinovich

Abstract Many reasonable conditions have been formulated for a fuzzy "and"-operation: idempotency, commutativity, associativity, etc. It is known that the only "and"-operation that satisfies all these conditions is minimum, but minimum is not the most adequate description of expert's "and", and it often does not lead to the best control or the best decision. Many other more adequate "and"-operations (t-norms) have been proposed and effectively used, but they do not satisfy the natural idempotency condition. In this paper, we show that a small relaxation of the usual description of "and"-operations leads to the possibility of non-minimum idempotent operations. We also show that another natural condition – of normalization invariance – uniquely determines the resulting "and"-operation. This new "and"-operation is not only more intuitive, it leads to better application results.

1 Formulation of the Problem

Why we need fuzzy techniques and why we need a fuzzy analogue of "and": a brief reminder. To achieve good automated control and/or good automated decision making, it is desirable to incorporate as much of human knowledge as possible. The problem is that many humans express their knowledge by using imprecise ("fuzzy") words from natural language such as "small", "large", etc., words that are difficult for computers to understand. To solve this problem, Lotfi Zadeh suggested techniques – that he called *fuzzy* – that translate from words to numbers.

Yusuf Güven
Control and Automation Engineering Department
Istanbul Technical University, Istanbul, Türkiye, e-mail: guveny18@itu.edu.tr

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

At the first stage of fuzzy techniques, we ask the expert, for each natural-language word w used in the expert rules, to assign, to each possible values x of the corresponding quantity, a degree $m_w(x)$ from the interval $[0,1]$ to which this value satisfies the intended property – e.g., the degree to which x is small. The resulting function $m_w(x)$ is known as a *membership function*, or, alternatively, as a *fuzzy set*; see, e.g., [2, 4, 5, 6, 7, 8].

Often, many rules include several conditions eachs. For example, when we hire a new faculty member, we want to make sure that he/she is a good researcher *and* a good teacher *and* a good colleague, etc. Ideally, to translate these rules into numbers, we should assign, to each combination of values of the corresponding quantities, a degree to which such a combination is possible. In practice, however, this is not possible: if we have v possible values of each quantity, then for n quantities, we need v^n combinations. For realistic values v and n, e.g., for $v = 10$ and $n = 6$, we get a million combinations – and there is no way to ask million questions to the expert.

Since we cannot elicit the degree of confidence in each such combination

$$S_1 \,\&\, \ldots \,\&\, S_n,$$

we need to be able to estimate this degree based on available information – i.e., usually, based on the expert-provided degrees of confidence $c_1, \ldots, c_n$ in each of the statements $S_1, \ldots, S_n$. It is reasonable to call the algorithm that transforms the original values c_i into an estimate for the composite statement a *fuzzy "and"-operations* (or simply *"and"-operation*, for short). Let us denote this algorithm by

$$f_\&(c_1, \ldots, c_n).$$

Current approach to "and"-operations: t-norms. What properties should a reasonable "and"-operation satisfy?

From the commonsense viewpoint, $S \,\&\, S$ means the same as S – and, in general, any "and"-combination $S \,\&\, \ldots \,\&\, S$ of the same statement does not change its truth value. Thus, it is reasonable to require that for each statement S with any degree of confidence c, the resulting estimate $f(c, \ldots, c)$ of our degree of confidence in $S \,\&\, \ldots \,\&\, S$ should be the same as the original degree of confidence c in the statement S:

$$f_\&(c, \ldots, c) = c. \tag{1}$$

In mathematics, operations with this property are called *idempotent.*

Also, from the commonsense viewpoint, the "and"-combination remains the same if we swap some of the statements: $S_1 \,\&\, S_2$ means the same as $S_2 \,\&\, S_1$, and, in general, for any permutation $\pi : \{1, \ldots, n\} \mapsto \{1, \ldots, n\}$, the statement $S_1 \,\&\, \ldots \,\&\, S_n$ means the same as $S_{\pi(1)} \,\&\, \ldots \,\&\, S_{\pi(n)}$. It is therefore reasonable to require that the fuzzy "and"-operation leads to the same estimate for the two equivalent statements. For two statements, this means that

$$f_\&(c_1, c_2) = f_\&(c_2, c_1). \tag{2}$$

In the general case, we have

$$f_{\&}(c_1,\ldots,c_n) = f_{\&}(c_{\pi(1)},\ldots,c_{\pi(n)}). \tag{3}$$

This property is known as *commutativity*.

Usually, yet another property is required, a property that is not as intuitive as the first two. It is related to the fact that, e.g., the statement $S_1 \,\&\, S_2 \,\&\, S_3$ can be obtained by first forming an intermediate statement $S_1 \,\&\, S_2$ and then by combining this intermediate statement with S_3. In other words, the statement $S_1 \,\&\, S_2 \,\&\, S_3$ is equivalent to $(S_1 \,\&\, S_2) \,\&\, S_3$. In the fuzzy case, in the first approximation, it is reasonable to assume that the exact same "and"-operation is used in both cases:

- when we combine the degrees of confidence c_1 and c_2 in statements S_1 and S_2 into an estimated degree of confidence $f_{\&}(c_1,c_2)$ of the intermediate statement $S_1 \,\&\, S_2$, and
- when we combine the degrees of confidence $f_{\&}(c_1,c_2)$ in the statement $S_1 \,\&\, S_2$ with the degree of confidence c_3 of the statement S_3.

In other words, we should have

$$f_{\&}(c_1,c_2,c_3) = f_{\&}(f_{\&}(c_1,c_2),c_3). \tag{4}$$

Since similarly, we can argue that

$$f_{\&}(c_1,c_2,c_3) = f_{\&}(c_1,f_{\&}(c_2,c_3)), \tag{5}$$

we can thus conclude that

$$f_{\&}(f_{\&}(c_1,c_2),c_3) = f_{\&}(c_1,f_{\&}(c_2,c_3)). \tag{6}$$

This property is known as *associativity.*

There are also reasonable – and also commonsense-motivated – requirements of monotonicity and continuity, and a requirement that in "crisp" situations, when we are absolutely sure whether each of the statements S_i is true or false, out degree of confidence in the combination $S_1 \,\&\, \ldots \,\&\, S_n$ should be 1 or 0 depending on whether this combination is true or false.

Because of the usually implicit accepted property (4), to describe the result of a general "and"-operation, with any number of inputs, it is sufficient to describe binary operations $f_{\&}(c_1,c_2)$.

It is known that the only "and"-operation that satisfies all the properties (1)–(6) is the minimum $f_{\&}(c_1,c_2) = \min(c_1,c_2)$. For this operation, for any n, we have $f_{\&}(c_1,\ldots,c_n) = \min(c_1,\ldots,c_n)$. In many cases, this operation works well, but in some situations, its results are not intuitive.

For example, if we have an excellent researcher, an excellent teacher, etc., etc., but not very good in socializing, there is a good chance that this person will be hired. OK, how about this: a Nobel prize winner in research, a winner of numerous teaching awards, but socially awkward – I do not think any department will miss the chance to hire her. But with minimum "and"-operation, we may have 20 properties required, 19 of which are satisfied with degree 1 and one with degree 0.1 – and the resulting statement $S_1 \& \ldots \& S_n$ that this person is a good fit for the department will still have degree of confidence 0.1 – very low. In comparison, a person who is mediocre in everything – so that $c_1 = \ldots = c_n = 0.5$ – will get an overall degree of 0.5, much higher.

To avoid limiting ourselves to the min operation, researchers usually avoid the idempotence property and only require commutativity, associativity, monotonicity, continuity – and the desired behavior in the crisp case. The resulting binary operations are known as *t-norms.*

Natural question and what we do in this paper. A natural question is: can we preserve the intuitive property of idempotence without limiting ourselves to min? Maybe instead of avoiding idempotence, we should avoid some other, less intuitive usual assumption?

In this paper, we show that this is indeed possible. Moreover, we will use an additional commonsense property – of normalization-invariance – to describe the resulting "and"-operation, and we will show that this new operation indeed leads to better application results than the usual t-norms.

2 Our main idea

Idea. The way we described "and"-operations in the previous section, the reader have probably already guessed what we plan to do. Namely, we mentioned that in the first approximation, the same fuzzy operation is used to combine the degrees of two statements, whether they are basic statement or combinations of basic statements. All other requirements are much more intuitive, we do not want to delete them, so let us delete this less-intuitive first-approximation requirement. This way, we will get a more subtle – but hopefully, more adequate – description of "and".

To be more precise, we will allow, in general, different "and"-operations:

- an operation $f_\&^{1,1}(c_1,c_2)$ for combining the degrees of two basic statements, and
- an operation $f_\&^{2,1}(c_1,c_2)$ for combining the degrees of statements $S_1 \& S_2$ and S_3.

Together with an operation $f^{1,1,1}(c_1,c_2,c_3)$ for combining three degrees, we get the following more complex equality instead of the usual equality (4):

$$f_\&^{1,1,1}(c_1,c_2,c_3) = f_\&^{2,1}(f_\&^{1,1}(c_1,c_2),c_3). \quad (7)$$

3 Normalization: a brief reminder, and a need to fuzzy operations to be normalization-invariant

Normalization: a brief reminder. To decide which "and"-operation is most appropriate, let us recall important details of fuzzy techniques that we omitted in the above brief description of fuzzy approach.

Specifically, we mentioned that we simply solicit the values $m_w(x)$ of the membership degree from the expert. For some properties this is exactly true – namely, for the properties for which there is some quantity x that absolutely satisfies this property. For example:

- for the property "small", everyone would agree that 1 gram of bread is a small amount;
- for the property "cold", everyone would agree that -40° C is cold;
- for the property "hot", everyone would agree that 45° C is hot, etc.

However, for words describing intermediate degrees, this is not necessarily true. For example, for people living in St. Petersburg, Russia, 20° may feel most comfortable, while for people living in El Paso, Texas, 20° is somewhat chilly. So, if we combine opinions of different experts – so that 1 means everyone agrees – we will not find the value x for which $m_w(x) = 1$.

For such words, fuzzy techniques suggests *normalization*. Namely, when the largest value

$$M \stackrel{\text{def}}{=} \max_x m_w^{\text{elic}}(x)$$

of elicited membership function $m_w^{\text{elic}}(x)$ is smaller than 1, we *normalize* it, by dividing all its values by M. This way, we get a *normalized* membership function

$$m_w(x) = \frac{m_w^{\text{elic}}(x)}{M}, \tag{8}$$

for which the maximum is exactly 1.

Why fuzzy operations should be normalization-invariant. Suppose that have a selected a range – e.g., a range of temperatures – and for this range, we got the elicited membership function $m_w^{\text{elic}}(x)$. Then, we use the formula (8) to normalize it.

Suppose that now we decided to further limit the range, and the new range does not include the point at which the elicited function $m_w^{\text{elic}}(x)$ attained its maximum. Then, we will need a new normalization, by dividing by a new maximum value $\underline{M} < M$. So, for the values x from the new range, we have

$$\underline{m}_w(x) = \frac{m_w^{\text{elic}}(x)}{\underline{M}}. \tag{9}$$

One can see that the two membership functions – that described the exact same word w – differ by a multiplicative constant:

$$\underline{m}_w(x) = a \cdot m_w(x), \text{ where } a \stackrel{\text{def}}{=} \frac{M}{\underline{M}}. \quad (10)$$

Since the two functions describe the same word, it makes sense to require that the membership functions obtained by using an "and"-operation should also be equivalent in this sense – i.e., they should differ only by a multiplicative constant.

This leads to the following definitions.

4 Definitions and the main result

Definition 1. *By an* "and"-operation*, we mean a continuous real-valued function* $f(c_1, \ldots, c_n)$ *for which the following three properties are satisfied:*

- *this function is* idempotent*, i.e.,* $f(c, \ldots, c) = c$ *for all* c*;*
- *this function is* commutative*, i.e., for every permutation* π *and for all tuples* $(c_1, \ldots, c_n)$*, we have* $f(c_1, \ldots, c_n) = f(c_{\pi(1)}, \ldots, c_{\pi(n)})$*; and*
- *this function is consistent with boolean "and", i.e.,*

$$f(0, \ldots, 0) = f(0, \ldots, 0, 1 \ldots, 1) = 0$$

and $f(1, \ldots, 1) = 1$.

Definition 2. *We say that an "and"-operation* $f(c_1, \ldots, c_n)$ *is* normalization-invariant *if for every* $a > 0$ *there exists a value* $A > 0$ *for which, for all possible tuples* $(c_1, \ldots, c_n)$*, we have*

$$f(a \cdot c_1, c_2, \ldots, c_n) = A \cdot f(c_1, c_2, \ldots, c_n). \quad (11)$$

Comment. Of course the value A may be different for different a, i.e., in effect, A is a function of a. In the following text, we will explicitly mention this by writing $A(a)$ instead of simply A.

Proposition. *The only normalization-invariant "and"-operation is*

$$f(c_1, \ldots, c_n) = \sqrt[n]{c_1 \cdot \ldots \cdot c_n}. \quad (12)$$

Proof.

1°. It is easy to show that the function (12) is a normalization-invariant "and"-operation with $A(a) = \sqrt[n]{a}$.

So, to complete the proof, it is sufficient to prove that every normalization-invariant "and"-operation has this form.

2°. Indeed, let $f(c_1,\ldots,c_n)$ be a normalization-invariant "and"-operation. By definition of an "and"-operation, this function is continuous.

From the formula (11), we conclude that

$$A(a)=\frac{f(a\cdot c_1,c_2,\ldots,c_n)}{f(c_1,c_2,\ldots,c_n)}. \tag{13}$$

For any tuple c_i, the right-hand side of the formula (13) is continuous. Thus, the left-hand side, i.e., the function $A(a)$ is continuous too.

3°. For ay a_1 and a_2, from (11), we conclude, in particular, that

$$f(a_2\cdot c_1,c_2,\ldots,c_n)=A(a_2)\cdot f(c_1,c_2,\ldots,c_n), \tag{14}$$

$$f(a_1\cdot a_2\cdot c_1,c_2,\ldots,c_n)=A(a_1)\cdot f(a_2\cdot c_1,c_2,\ldots,c_n), \tag{15}$$

$$f(a_1\cdot a_2\cdot c_1,c_2,\ldots,c_n)=A(a_1\cdot a_2)\cdot f(c_1,c_2,\ldots,c_n). \tag{16}$$

Substituting the right-hand side of the expression (14) into the formula (15) instead of $f(a_2\cdot c_1,c_2,\ldots,c_n)$, we conclude that

$$f(a_1\cdot a_2\cdot c_1,c_2,\ldots,c_n)=A(a_1)\cdot A(a_2)\cdot f(c_1,c_2\ldots,c_n). \tag{17}$$

By comparing equations (16) and (17), we conclude that

$$A(a_1\cdot a_2)\cdot f(c_1,c_2\ldots,c_n)=A(a_1)\cdot A(a_2)\cdot f(c_1,c_2\ldots,c_n). \tag{18}$$

If we divide both sides of (18) by $f(c_1,c_2\ldots,c_n)$ – which is not zero at least of values close to 1 – we conclude that

$$A(a_1\cdot a_2)=A(a_1)\cdot A(a_2) \tag{19}$$

for all possible values a_1 and a_2.

4°. It is known – see, e.g., [1] – that every continuous function that satisfies the condition (19) is either always equal to 0 or is equal to $A(a)=a^\alpha$ for some real value α. The function $A(a)$ cannot be identically 0, since for $a=1$, we clearly have $A(a)=1\neq 0$. So, $A(a)=a^\alpha$.

5°. From equation (11), for $a_1=1/c_1$, we can now conclude that

$$f(1,c_2,\ldots,c_n)=(1/c_1)^\alpha\cdot f(c_1,\ldots,c_n). \tag{20}$$

Multiplying both sides by c_1^α, we conclude that

$$f(c_1,c_2,\ldots,c_n)=c_1^\alpha\cdot f(1,c_2,\ldots,c_n). \tag{21}$$

Due to commutativity of the "and"-operation $f(c_1,\ldots,c_n)$, we can swap c_1 and c_2 and get the related equality:

$$f(c_1, c_2, c_3, \ldots, c_n) = c_2^{\alpha} \cdot f(c_1, 1, c_3 \ldots, c_n). \quad (22)$$

In particular, for $c_1 = 1$, this formula leads to

$$f(1, c_2, c_3, \ldots, c_n) = c_2^{\alpha} \cdot f(1, 1, \ldots, c_n). \quad (23)$$

Substituting the expression (23) into the right-hand side of the formula (21), we conclude that

$$f(c_1, c_2, c_3, \ldots, c_n) = c_1^{\alpha} \cdot c_2^{\alpha} \cdot f(1, 1, c_3 \ldots, c_n). \quad (24)$$

Similarly, by swapping c_1 with other values c_3, c_4, etc., we conclude that

$$f(c_1, c_2, c_3, \ldots, c_n) = c_1^{\alpha} \cdot c_2^{\alpha} \cdot \ldots \cdot c_n^{\alpha} \cdot f(1, 1, \ldots, 1). \quad (25)$$

By definition of "and"-operation, we have $f(1, \ldots, 1) = 1$, thus the formula (25) takes the form

$$f(c_1, c_2, c_3, \ldots, c_n) = c_1^{\alpha} \cdot c_2^{\alpha} \cdot \ldots \cdot c_n^{\alpha}. \quad (26)$$

In particular, for $c_1 = \ldots = c_n = c$, we get

$$f(c, \ldots, c) = c^{\alpha \cdot n}. \quad (27$$

On the other hand, by idempotency, we get $f(c, \ldots, c) = c = c^1$. Thus, $\alpha \cdot n = 1$ and $\alpha = 1/n$.

The proposition is proven.

5 The new operation leads to better application results

Ok, so maybe the new "and"-operation is more intuitive, but it is better in applications?

Analysis described in [3] on three datasets –Parkinson diseases, while wine, and AIDS – shows that in all three cases, the use of the new operation leads to more accurate predictions

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), HRD-1834620 and HRD-2034030 (CAHSI Includes), EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES), and by the AT&T Fellowship in Information Technology.

It was also supported by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

References

1. J. Aczél and J. Dhombres, *Functional Equations in Several Variables*, Cambridge University Press, 2008.
2. R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.
3. Y. Güven, A. Köklü, and T. Kumbasar, "Zadeh's type-2 fuzzy logic systems: precision and high-quality prediction intervals", *Proceedings of the 2024 IEEE World Congress on Computational Intelligence WCCI*, Yokohama, Japan, June 30–July 5, 2024, to appear.
4. G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
5. J. M. Mendel, *Explainable Uncertain Rule-Based Fuzzy Systems*, Springer, Cham, Switzerland, 2024.
6. H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.
7. V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.
8. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.