5-1-2022

What About Generic "And"- and "Or"-Operations: from Levels of Certainty (Philosophical-Physical-Mathematical) to a Natural Interpretation of Quantum-Like Negative Degrees of Certainty

Miroslav Svitek
Czech Technical University in Prague, svitek@fd.cvut.cz

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

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-22-53

What About Generic "And"- and "Or"-Operations: from Levels of Certainty (Philosophical-Physical-Mathematical) to a Natural Interpretation of Quantum-Like Negative Degrees of Certainty

Miroslav Svitek[1], Olga Kosheleva[2], and Vladik Kreinovich[2]
[1]Faculty of Transportation Sciences
Czech Technical University in Prague
Konviktska 20, CZ-110 00 Prague 1, Czech Republic
svitek@fd.cvut.cz
[2]University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

Abstract

Fuzzy techniques – techniques designed to convert imprecise human knowledge into precise computer-understandable terms – have many successful applications. Traditional applications of fuzzy techniques use only important general features of human reasoning and, to make an implement more efficient, ignore subtle details, details which are not important for the corresponding application. But from the more fundamental viewpoint, it is desirable to understand, in all the detail, how people actually reason. In this paper, we use general ideas of fuzzy approach to answer this question. Interestingly – and somewhat unexpectedly – the resulting analysis leads to a natural explanation of existence of several distinct levels of certainty and to a natural appearance of quantum-like negative degrees of certainty.

1 Why Study Generic "And"- and "Or"-Operations

"And"- and "or"-operations. Expert knowledge is often described by using imprecise ("fuzzy") words from natural languages such as "small'. To be able to use this knowledge in computer-based decision making, it is necessary to describe this knowledge in computer-understandable language. Computers have

been originally designed to process numbers, this is still what they do most efficiently. So, a natural idea is to describe imprecise expert knowledge by numbers.

The main idea of such description comes from fuzzy logic (see, e.g., [3, 11, 14, 15, 17, 22]), where for each imprecise property like "small" and for each possible value x of the corresponding quantity, we ask the expert to mark, on scale of 0 to 1, his/her degree of certainty that x is small. (Alternatively, we can use another scale, e.g., a scale from 0 to 10, and then divide the result by 10 to re-scale it to the interval $[0, 1]$.)

The problem is that many expert rules use logical connectives "and", "or", and "not" in their formulation. For example, a medical doctor can say that if a patient has a strong fever *and* a heavy headache then a certain medicine should be prescribed. To describe such rules in numerical form, we need to provide numerical degree of certainty not just to individual statements like "a patient has strong fever" and "a patient has a heavy headache", but also to logical combinations of these statements.

In the ideal world, we should ask the expert to estimate his/her degree of confidence in all possible logical combinations. However, the number of all such possible combinations is astronomical, in millions and billions, and there is no way that we can ask experts about all such combinations. For all other logical combinations, for which we cannot ask the expert directly, we need to be able to estimate the expert's degree of confidence in these statements based on the only information that we have, i.e., based on the expert-provided degree of confidence about combined statement.

For "and", this means that we should have an algorithm that, given the expert's degree of confidence a and b in statements A and B, will provide an estimate for the expert's degree of confidence in the "and"-combination $A \,\&\, B$ of these statements. This algorithm is called an *"and"-operation* or, for historical reasons, a *t-norm*. We will denote the result of applying this algorithm by $f_{\&}(a, b)$.

Similarly, for "or", this means that we should have an algorithm that, given the expert's degree of confidence a and b in statements A and B, will provide an estimate for the expert's degree of confidence in the "or"-combination $A \vee B$ of these statements. This algorithm is called an *"or"-operation* or, for historical reasons, a *t-conorm*. We will denote the result of this algorithm by $f_{\vee}(a, b)$.

What "and"- and "or"-operations are currently used in computer-based decision making. In the first papers that used fuzzy techniques, researchers tried to elicit the "and"- and "or'-operations that best describe actual human reasoning. However, it turned out that, taking into account that we are talking about very imprecise expert rules, adding more details to "and"- and "or"-operations did not add much to the quality of the resulting decisions and controls – while requiring a lot of time to elicit these operations and additional computation time to implement them in the actual decision-making procedures. As a result, researchers realized that from the application viewpoint, it is usually sufficient to limit ourselves to simple operations like $f_{\&}(a, b) = \min(a, b)$ or

$f_{\&}(a,b) = a \cdot b$, and leave the study of the actual "and"- and "or"-operations to psychologists and linguists.

So what about the actual "and"- and "or"-operations: formulation of the problem. In this paper, we go back to the original idea of eliciting the actual "and"- and "or"-operations from the experts, with the purpose of analyzing what the analysis of such actual operations can tell us about human knowledge processing.

2 Analysis of the Problem

What are actual "and"- and "or"-operations? Empirical study of the actual human reasoning consistently shows that this reasoning is very complicated and very individualized. In particular, it is known that the actual "and"- and "or"-operations are complicated, and differ from one application area to another – and even from one person to another. This was shown, e.g., when the designers of one of the world's first expert systems MYCIN (see, e.g., [9]), who spent a lot of time and effort to find the "and"- and "or"-operations most appropriate for medical reasoning tried to apply the same operations to geophysics. It was immediately discovered that geophysicists think differently, and this difference makes perfect sense:

- a medical doctor has to be very cautious, and only recommend a surgery or other major intervention when he/she is reasonably confident that this intervention will help, while
- a company prospecting for oil usually starts exploring even if there is a remaining doubt that there is oil in this area: if they wait for additional experiments that would enable them to get a perfect confidence, competitors would have gone ahead of them.

In all these cases, simple – or even more complex – "and"- and "or"-operations are only an approximation to the actual, even more complex ones. It seems like the more data we gather about the actual human decision making, the more complex the resulting model becomes. In other words, in the set of all possible "and"- and "or"-operations, the actual "and"- and "or"-operations that describe human reasoning are not some simplified elements. In terms of complexity, they are *generic* elements of this set.

So, let us study what we can say about such generic operations.

In general, what do we know about generic elements. What do we mean by saying that some element is a generic element of a set – be it set of numbers, set of functions, or set of some other objects? In a large set, we usually have a few elements with unusual properties. For example, in the set of all people, we have people who are very short, and we have people who are very tall. Both very short and very tall people are rare, so such people are clearly not generic. Generic usually means that the corresponding object only has

important properties that most elements of this set have. When we say "most", we mean that there is a natural probability measure on the corresponding set, and that a object is generic if it has the same properties as random elements in the sense of this measure.

There are several general results about random elements. In particular, there is a result by B. Tsirel'son [21] that under reasonable conditions, random elements are location on the boundary of the corresponding set – or at least close to this boundary. This general fact requires a complex proof, but it can be explained on a simple example. Suppose that you are testing a hypothesis that observations $x_1, \ldots, x_n$ are independent and normally distributed, with 0 mean and standard deviation 1. In this case, one of the most widely used tests is the t-test (see, e.g., [19]), i.e., checking whether the average $\frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2$ is smaller than a certain threshold $t_{\alpha,n} \approx 1$, where α is the corresponding level of confidence. In such a situation, the set of all observations consistent with this hypothesis has the form

$$\left\{(x_1, \ldots, x_n) : \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 \le t_{\alpha,n}\right\}.$$

However, according to the Law of Large Numbers, this mean tends to the variance – i.e., in this case, to 1. This means, informally speaking, that for large n, most observations are concentrated on the border of the above set, i.e., in the set

$$\left\{(x_1, \ldots, x_n) : \frac{1}{n} \cdot \sum_{i=1}^{n} x_i^2 \approx 1\right\},$$

for some relevant meaning of approximate equality.

Moreover, if the boundary itself has a kind of a boundary; e.g., the boundary of a square consists of four sides, and each of the sides has boundaries – vertices – then random elements are located on this "boundary of boundary", etc.

What does this general property of generic objects mean for the case of "and"-operations. In general, an "and"-operation $f_{\&}(a, b)$ is a function. There is a lot of research about random functions. This research started with Wiener measure (a mathematical description of Brownian motion), a description of functions of one variable. Still, most of the research now is about random functions of one variable – what is called random processes. There is research on random functions of two or more variables – known as random fields – but much less is done for them. Since an "and"-operation is a function of two variables, a natural idea is to look for related functions of one variable – so that more known results will be available.

A natural way to form a function of one variables is to consider the case when $a = b$, in this case we have a function $f(a) \stackrel{\text{def}}{=} f_{\&}(a, a)$. In general, the natural limitations on possible "and"-operations $f_{\&}(a, b)$ are:

- that this operation coincides with the usual "and"-operation when we know for sure whether each of the statements is true (1) or false (0), i.e.,

$f_\&(0,0) = f_\&(0,1) = f_\&(1,0) = 0$ and $f_\&(1,1) = 1$;

- that a person's degree of confidence $f_\&(a,b)$ in $A \,\&\, B$ is the same as this person's degree of confidence $f_\&(b,a)$ in $B \,\&\, A$: $f_\&(a,b) = f_\&(b,a)$;
- that a person's degree of confidence $f_\&(a,b)$ in $A \,\&\, B$ cannot exceed his/her degree of confidence in A or in B: $f_\&(a,b) \le a$ and $f_\&(a,b) \le b$; and
- that if the person becomes more confident in A and/or in B, then his/her degree of confidence in $A \,\&\, B$ will either increase or remain the same, but it cannot decrease: if $a \le a'$ and $b \le b'$, then $f_\&(a,b) \le f_\&(a',b')$.

In relation to our function $F(a) = f_\&(a,a)$, this means that $F(0) = 0$, $F(1) = 1$, $F(a) \le a$, and that if $a < a'$, then $F(a) \le F(a')$.

In general, for a set determined by inequalities, its boundary is where some of these inequalities become equalities. For example, for an interval

$$[0,1] = \{x : 0 \le x \,\&\, x \le 1\}$$

its boundary consists of the points in which one of these two inequalities $0 \le x$ or $x \le 1$ becomes equalities, i.e., when either $x = 0$ or $x = 1$.

Similarly, for the square

$$\{(x,y) : 0 \le x \,\&\, x \le 1 \,\&\, 0 \le y \,\&\, y \le 1\},$$

the boundary consists of four sides: the side where $x = 0$, the side where $x = 1$, the side where $y = 0$, and the side where $y = 1$.

In our case, this means that on the boundary, some of the inequalities become equalities, i.e., that we must have $F(a) = a$ for some $a \in (0,1)$ and/or that we must have $F(a) = F(a')$ for some $a < a'$.

From this viewpoint, "boundary of boundary" corresponds to a situation when many inequalities become equalities, i.e., in particular, that we have $F(a) = a$ for several values $a \in (0,1)$.

Which "and"-operations have these properties? To answer this question, let us recall that there is a known general description of all possible "and"-operations; see, e.g., [3, 11, 14, 15, 17]. This description starts with describing so-called *Archimedean* "and"-operations, i.e., operations for which either $f_\&(a,b) = f^{-1}(f(a) \cdot f(b))$ for some strictly monotonic function $f(a)$ or $f_\&(a,b) = f^{-1}(\max(f(a) + f(b) - 1, 0))$. For each of these operations, for all $a \in (0,1)$, we have $f_\&(a,a) < a$.

To describe a general "and"-operation, on the interval $[0,1]$, we select several intervals $I_i = [\underline{a}_i, \overline{a}_i]$ whose interiors do not intersect. On each of these intervals, the "and"-operation is isomorphic to an Archimedean one, and in all other cases – i.e., when a and b do not belong to the same interval I_i – we have $f_\&(a,b) = \min(a,b)$. In this arrangement, the equality $F(a) = a$ is satisfied if and only if the point a is one of the following two types:

- it is either an endpoint $\underline{a}_i$ or $\overline{a}_i$ of one of the selected intervals,
- or this point a does not belong to any of the selected intervals I_i at all.

So, we can conclude that in the actual generic "and"-operation, there are most probably several such intervals I_i and maybe an area that does not belong to any of these intervals.

This may sound as a somewhat unexpected conclusion, since in the usual pragmatic approach to fuzzy logic, such weird general "and"-operations are viewed as practically useless purely mathematical constructions – to the extent that many textbooks on fuzzy techniques do not even mention such operations or mention them in passing. What we showed is that these seemingly exotic over-complicated operations are the ones that describe the actual human reasoning!

OK, interesting mathematics, but so what? At first glance, our excitement sounds unjustified: OK, we have some interesting mathematical construction, but the goal of fuzzy techniques is to describe human reasoning. What can this mathematical construction tell us about human reasoning? Let us try to answer this question.

3 First Consequence: Levels of Certainty Naturally Appear

Usual continuous approach to certainty vs. the existence of levels of certainty: a seeming contradiction. In the traditional fuzzy logic, certainty is described by a number from the interval $[0, 1]$. Alternative, we can use a probabilistic approach to describing uncertainty, in which case certainty of a statement can be described by the probability that this statement is true – and this probability is also a number from the same interval $[0, 1]$. In both approaches, we have a continuous scale and, thus, a *continuous* transition from less certain statements to more certain one. In general, if we use, e.g., the usual product "and"-operation $f_{\&}(a, b) = a \cdot b$, then, even if we start with several highly confidence statements with $a = b = \ldots = c = 0.99$, then, by combining these statements, we can get statement $A\ B\,\&\ \ldots\ \&\,C$ whose degree of certainty is as close to 0 as we want.

In reality, however, practical uncertainties fall into several clearly distinct levels. On the top, there is mathematical level, where we only consider absolutely correct, rigorously proved statements. Next is the level of reasoning in physics, where we can often make conclusions without rigorously proving existence, continuity, or differentiability – as long as the resulting predictions make physical sense. This is not just because of lack of mathematical ingenuity and the resulting inability to prove the corresponding existence – the situation is much more complex: e.g., in general quantum field theory, there is still no consistent precise mathematical description (not to mention quantum gravity and other similar phenomena); see, e.g., [10, 20]. There are some mathematical

formulas, but it is known that in some cases, they lead to physically useless infinite values for appropriate quantities.

Then there is a level of other natural sciences like geo- or biosciences, where there is even fewer rigor – and thus, fewer certainty, etc. Once we assign a numerical value describing our degree of certainty in each statement, each level can be described by the range $[\underline{d}, \overline{d}]$ of acceptable degrees of uncertainty. Everything mathematician do should have to satisfy some minimal certainty requirements, and the same is true for physicists – mathematicians may make fun of the physicists' non-perfectly-strict conclusions, but there is a certain high level of rigor that a physical paper must satisfy.

We just mentioned the main levels of certainty, but within each level, there are sub-levels. For example, in mathematics, there is a level of constructive mathematics (seem, e.g., [1, 2, 4, 5, 7, 8, 12, 13, 18]) when the statement $\exists x\, P(x)$ that there exists an object x satisfying a property $P(x)$ is considered proven only if we have an algorithm that effectively constructs such an object. Another sublevel is when, in principle, we allow non-algorithmic constructions but do not allow Axiom of Choice, etc.

The interesting thing about each level (and sub-level) of certainty is that, in contrast to the above example of the product "and"-operation on the interval $[0, 1]$, each level is self-sufficient: mathematicians only use mathematical level of certainty, physicists only use physical level of certainty, etc. In other words, any "and"-combination of statements from the same level belongs to this same level. If we have statement A, B, ..., C from some level, for which the degrees of certainty are at least as high as the minimum $\underline{d}$ acceptable at this level, then our degree of certainty in the combines statement $A \,\&\, B \,\&\, \ldots \,\&\, C$ cannot get lower than this level $\underline{d} > 0$.

So, at first glance, there seems to be contradiction between the usual fuzzy description of uncertainty and the observed existence of clearly distinct levels of certainty.

But is there a contradiction? Let us show that the seeming contradiction comes not from the essence of fuzzy logic, but from the fact that we are using a simplified example of an "and"-operation – and that for the actual, more complex "and"-operation, not only there is no contradiction, but the existence of layers naturally follows.

Indeed, according to the general description of "and"-operations, for any two values $v < v'$ for which $F(v) = f_{\&}(v, v) = v$ and $F(v') = f_{\&}(v', v') = v'$ – and, as we mentioned, such levels exist for the actual "and"-operation – if we have any two statements A and B with degrees of certainty a and b from the interval $[v, v']$, then the degree of certainty $f_{\&}(a, b)$ in the statement $A \,\&\, B$ also belongs to the same interval. So, for the actual "and"-operation, the values v for which $F(v) = v$ naturally divide the whole range $[0, 1]$ of possible values of degree of certainty into sub-ranges $[\underline{d}, \overline{d}]$ which are self-sufficient – in the sense that:

- if we have statement A, B, ..., C from this sub-range,
- then our degree of certainty in the combines statement $A \,\&\, B \,\&\, \ldots \,\&\, C$

also belongs to this sub-range.

In other words, levels of certainty indeed naturally appear.

4 Below 0 and Above 1?

What about the future? In the previous section, we talked about the current situation. A natural question is: how will it evolve? To answer this question, let us recall what we usually interpret as levels 0 and 1, and what happened with the corresponding levels of certainty in the past.

What do we choose as levels 0 and 1? Naturally, we identify 1 with the level which, at present, corresponds to the largest degree of certainty, and 0 with the level that corresponds to the lowest possible degree of certainty.

Resulting problem. This sounds reasonable but the problem is that, historically, new levels appear all the time – and will probably appear again. For example, the current mathematical level of rigor appeared, it is final form, only in the 19 century, with the "revolution of rigor" that lead to formal definitions of continuity, limits, etc.; see, e.g., [6]. Constructive mathematics – one of the highest level of rigor – appeared only in the 20th century.

Negative and larger-than-1 degrees naturally appear. It is natural to expect that new even higher levels of certainty will appear in the future. How will we be able to describe them – taking into account that the degree 1 is already taken by some level which is lower than this forthcoming one? Naturally, we will end up with degree of certainty which are higher than 1 – and, similarly, degree of certainty which are lower than 0. In the probabilistic approach, this would mean having subjective probabilities smaller than 0 or larger than 1.

Can this be related to quantum computing? Such negative and larger-than-1 subjective probabilities may provide a new interpretation for negative value in quantum physics [10, 20] – and especially in quantum computing [16], where, in contrast to the general quantum physics that uses general complex numbers $a+b\cdot \mathrm{i}$, only real numbers are used. In this case, real numbers from the interval $[0,1]$ can be, in principle, interpreted in the usual probabilistic sense, so all that remains is to interpret numbers which are smaller than 0 and numbers which are larger than 1 – and this is exactly what actual fuzzy operations provide.

On the qualitative level, this interpretation makes sense. Indeed, as we have mentioned, both negative values and values which are larger than 1 correspond to future knowledge. Naturally, if we use some elements of future knowledge in our computations, this can make computations faster – this may explain why many quantum computing algorithms are indeed faster and more efficient than the corresponding non-quantum ones: e.g., the well-known Grover's quatum algorithm can find an element with the desired property in a n-element list in $\sqrt{n}$ computations steps, which a similar non-quantum algorithm requires at

least n steps which is much longer: e.g., for n =1,000,000, it is 1,000 times longer.

Instead of a conclusion: reminder. And all this – explaining levels of certainty, relating to quantum algorithms – naturally follows from the analysis of the actual "and"-operations!

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), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The authors are thankful to the editors of the special issue for their encouragement.

References

[1] O. Aberth, *Precise Numerical Analysis Using C++*, Academic Press, New York, 1998.

[2] M. J. Beeson, *Foundations of constructive mathematics*, Springer-Verlag, N.Y., 1985.

[3] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[4] E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, 1967.

[5] E. Bishop and D. S. Bridges, *Constructive Analysis*, Springer, N.Y., 1985.

[6] C. B. Boyer and U. C. Merzbach, *A History of Mathematics*, Wiley, New York, 1991.

[7] D.S. Bridges, *Constructive Functional Analysis*, Pitman, London, 1979.

[8] D. S. Bridges and S. L. Vita, *Techniques of Constructive Analysis*, Springer-Verlag, New York, 2006.

[9] B. G. Buchanan and E. H. Shortliffe, *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, Massachusetts, 1984.

[10] R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

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

[12] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1998.

[13] B. A. Kushner, *Lectures on Constructive Mathematical Analysis*, Amer. Math. Soc., Providence, Rhode Island, 1984.

[14] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[15] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press: Cambridge, U.K., 2000

[17] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[18] M. Pour-El and J. Richards, *Computability in Analysis and Physics*, Springer-Verlag, New York, 1989.

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

[20] K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2017.

[21] B. S. Tsirel'son, "A geometrical approach to maximum likelihood estimation for infinitedimensional Gaussian location. I", *Theory of Probability and its Applications*, 1982, Vol. 27, pp. 411–418.

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