Dynamic Fuzzy Logic Leads to More Adequate "And" and "Or" Operations

Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu

1. Outline

- In the traditional (static) fuzzy logic, we select an "and"-operation (t-norm) and an "or"-operation (t-conorm).
- The result of applying these operations may differ from the expert's degrees of belief in $A \,\&\, B$ and $A \vee B$.
- Reason: the degrees $d(A \,\&\, B)$ and $d(A \vee B)$ depend:
  - not only on the expert's degrees of belief in statements A and B,
  - but also in the extent to which the statements A and B are dependent.
- We show that *dynamic* fuzzy logic enables us to automatically take this dependence into account.
- Thus, dynamic fuzzy logic leads to more adequate "and"- and "or"-operations.

2. Fuzzy Logic: Brief Reminder

- Expert rules are often formulated by using imprecise ("fuzzy") words, like "small", "medium size", or "large".
- For example, a medical recommendation depends on whether the tumor is small, medium size, or large.
- How to avoid collision with a car depends on whether the distance to the car is small, medium, or large.
- To describe such words, L. Zadeh proposed *fuzzy logic*.
- In fuzzy logic, we assign, to each value x, the degree $\mu_P(x) \in [0,1]$ to which P is satisfied; e.g.:
  - as a proportion of the experts who believe that x satisfies the given property,
  - or as a subjective probability.

3. “And” and “Or” Operations in Fuzzy Logic

- Often, an expert rule contains several conditions, e.g.:
  - If an obstacle is close *and* the car is going fast, then we need to break fast.
  - If a skin tumor is large *or* bleeding *or* has irregular shape, then we need to operate on it.
- Thus, we need to:
  - combine the degrees of confidence $a = d(A)$ and $b = d(B)$ in the corresponding component statements
  - into a single degree $d(S)$ to which the rule S is applicable.
- An algorithm $f_{\&}(a,b)$ that transforms a and b into $d(A \,\&\, B)$ is called an *“and”-operation* or a *t-norm*.
- An algorithm $f_{\vee}(a,b)$ that transforms a and b into $d(A \vee B)$ is called an *“or”-operation* or a *t-conorm*.

4. Variety of t-Norms and t-Conorms

- In fuzzy logic, there are numerous t-norms and t-conorms.
- Which one to apply depends on the relation between the statements A and B.
- This dependence can be illustrated in the probabilistic approaches, when $a = \text{Prob}(A)$.
- If A and B are independent, then the probability $f_\&(a,b)$ of $A \,\&\, B$ is equal to the product $a \cdot b = P(A) \cdot P(B)$.
- In this case, the most adequate t-norm is a product $f_\&(a,b) = a \cdot b$.
- If A and B are strongly correlated, then we should have $P(A \,\&\, B) = P(A) = P(B)$ when $A = B$.
- In this case, a t-norm $f_\&(a,b) = \min(a,b)$ is more adequate.

5. Formulation of the Problem

- The problem is that in many cases, we do not know whether A and B are correlated or not.
- In such cases, we select *some* t-norm.
- The selected t-norm may not necessarily coincide with the ideal one.
- Hence, the resulting recommendations may not be always adequate.
- The problem is with "truth-functionality":
  - the degree of confidence in $A \,\&\, B$ depends only on the degrees of confidence in A and B
  - without fully adequately taking into account the possibility of different correlations.
- This is often cited as one of the main limitations of fuzzy techniques.

6. Dynamic Fuzzy Logic

- The traditional fuzzy logic assumes that the expert's degrees of confidence do not change.
- In reality, the expert's opinions often change with time; thus:
  - to get a more adequate description of the expert opinions and rules,
  - it is necessary to take these changes into account.
- In other words,
  - to describe the expert's opinion about a statement A, instead of a single value $a \in [0,1]$,
  - we need to use a function $a(t)$ that describes how this degree changes with time t.
- Such *dynamic fuzzy logic* was proposed by Leonid Perlovsky and others.

7. What We Do in This Talk

- In this paper, we show that:
  - if we take this dynamics into consideration,
  - then we can get a more adequate description of "and" and "or" operations.
- Specifically, we get a description in which it is possible to distinguish between:
  - the cases when the statements are independent and
  - the cases when the statements are strongly dependent.
- This possibility will be illustrated on the example when the fuzzy degrees have a probabilistic meaning.

8. Correlation: Reminder

- In statistics:
  - the most frequent way to describe correlation between two random variables x and y is
  - to use the correlation coefficient.
- Usually:
  - the mean (expected value) of the variable x is denoted by $E[x]$, and
  - the variance $V[x]$ is defined as

$$V[x] \stackrel{\text{def}}{=} E[(x - E[x])^2] = E[x^2] - (E[x])^2.$$

- The correlation coefficient is then defined as

$$\rho = \frac{E[x \cdot y] - E[x] \cdot E[y]}{\sqrt{V[x] \cdot V[y]}}.$$

9. Relation between Correlation and the Probability $P(A \& B)$

- We consider a statement A which is true with probability a and false with the remaining probability $1-a$.
- A can be viewed as a random variable that is equal to 1 ("true") w/prob. a and to 0 ("false") w/prob. $1-a$.
- For this variable, $E[A] = 1 \cdot a + 0 \cdot (1-a) = a$ and similarly, $E[B] = b$ and $E[A \& B] = P(A \& B)$.
- Here, $A = 0$ or $A = 1$, hence $A^2 = A$, $E[A^2] = E[A]$ and thus, $V[A] = E[A^2] - (E[A])^2 = a - a^2 = a \cdot (1-a)$.
- Similarly, we can conclude that $V[B] = b \cdot (1-b)$.
- For true and false statements, "and" is simply a product, so $A \& B = A \cdot B$ and thus,

$$E[A \& B] = P(A \& B) = E[A \cdot B].$$

10. Relation between Correlation and the Probability $P(A \& B)$ (cont-d)

- In general, $\rho = \dfrac{E[A \cdot B] - E[A] \cdot E[B]}{\sqrt{V[A] \cdot V[B]}}$.
- Here, $E[A \cdot B] = P(A \& B)$, $E[A] = a$, $E[B] = b$, $V[A] = a \cdot (1 - a)$, and $V[B] = b \cdot (1 - b)$; thus:

$$\rho = \frac{P(A \& B) - a \cdot b}{\sqrt{a \cdot (1 - a) \cdot b \cdot (1 - b)}}.$$

- Thus, once we know $P(A) = a$, $P(B) = b$, and ρ, we can uniquely reconstruct $P(A \& B)$ as

$$P(A \& B) = a \cdot b + \rho \cdot \sqrt{a \cdot (1 - a) \cdot b \cdot (1 - b)}.$$

- From $P(A \& B) + P(A \vee B) = P(A) + P(B)$, we conclude that $P(A \vee B) = P(A) + P(B) - P(A \& B)$, so:

$$P(A \vee B) = a + b - a \cdot b - \rho \cdot \sqrt{a \cdot (1 - a) \cdot b \cdot (1 - b)}.$$

11. How Do We Find the Correlation Coefficient: Idea

- In the dynamic case:
  - we not only know the current expert's degrees of confidence a and b in statements A and B,
  - we also know the past degrees $a(t)$ and $b(t)$ which were, in general, different from a and b.
- When A and B are strongly correlated, it is reasonable to expect that $a(t)$ and $b(t)$ are also correlated.
- If A and B are independent, then it is reasonable to expect that $a(t)$ and $b(t)$ are also independent.
- In general:
  - to find the correlation coefficient between A and B,
  - we can use, as random variables, the values $a(t)$ and $b(t)$ corresponding to T known moments of time.

12. How Do We Find the Correlation Coefficient: Resulting Formulas

- Under this idea,

$$E[A]=\frac{1}{T}\cdot\sum_t a(t),\quad E[B]=\frac{1}{T}\cdot\sum_t b(t),$$

$$V[A]=\frac{1}{T}\cdot\sum_t a^2(t)-\left(\frac{1}{T}\cdot\sum_t a(t)\right)^2,$$

$$V[B]=\frac{1}{T}\cdot\sum_t b^2(t)-\left(\frac{1}{T}\cdot\sum_t b(t)\right)^2,$$

$$E[A\cdot B]=\frac{1}{T}\cdot\sum_t a(t)\cdot b(t),\ \text{so } \rho=\frac{E[A\cdot B]-E[A]\cdot E[B]}{\sqrt{V[A]\cdot V[B]}}.$$

- Using this value ρ, we get the desired estimates for

$$P(A\,\&\,B)=a\cdot b+\rho\cdot\sqrt{a\cdot(1-a)\cdot b\cdot(1-b)}\ \text{and}$$

$$P(A\vee B)=a+b-a\cdot b-\rho\cdot\sqrt{a\cdot(1-a)\cdot b\cdot(1-b)}.$$

13. Mathematical Comment: Ergodicity

- In producing these estimates, we implicitly assumed that:
  - averaging over time leads to the same result as
  - averaging over a sample.
- This property is called *ergodicity*.
- This property is often assumed and/or proved:
  - in statistical physics and
  - in statistical data analysis.

14. Need for Weighted Averages

- In the above formulas, we implicitly assumed that the correlation does not change in time.
- In reality, just like the expert degrees change with time, the correlation between these degrees may also change.
- It is therefore necessary to take this change into account when estimating correlation.
- One way to do that is to consider the recent values with higher weights than past values.
- In other words, we take $E[A] = \sum\limits_t w(t) \cdot a(t)$ for some weights $w(t) \geq 0$ for which $\sum\limits_t w(t) = 1$.
- A usual selection of "discount" weights is a geometric progression $w(t) = C \cdot q^t$ for some $q < 1$.
- In this case, $\sum\limits_{t=1}^{T} w_t = 1$ implies that $C = \dfrac{1-q}{1-q^{T+1}}$.

15. Weighted Averages: Resulting Formulas

- First, we compute the values

$$E[A]=\sum_t w(t)\cdot a(t),\quad E[B]=\sum_t w(t)\cdot b(t),$$

$$V[A]=\sum_t w(t)\cdot a^2(t)-\left(\sum_t w(t)\cdot a(t)\right)^2,$$

$$V[B]=\sum_t w(t)\cdot b^2(t)-\left(\sum_t w(t)\cdot b(t)\right)^2,$$

$$E[A\cdot B]=\sum_t w(t)\cdot a(t)\cdot b(t);\quad \rho=\frac{E[A\cdot B]-E[A]\cdot E[B]}{\sqrt{V[A]\cdot V[B]}}.$$

- Using this value ρ, we then compute

$$P(A\,\&\,B)=a\cdot b+\rho\cdot\sqrt{a\cdot(1-a)\cdot b\cdot(1-b)}\text{ and}$$

$$P(A\vee B)=a+b-a\cdot b-\rho\cdot\sqrt{a\cdot(1-a)\cdot b\cdot(1-b)}.$$

16. First Limitation of This Approach: Computational Complexity

- In the *static* fuzzy logic:
  - to find the degree of confidence in $A \,\&\, B$ or in $A \vee B$,
  - we simply applying a t-norm or a t-conorm to two numbers.
- In the *dynamic* case, we need to perform a large number of computations instead.
- This is unavoidable in the dynamic fuzzy logic:
  - we have *more values* for representing the expert's degree of confidence in each statement,
  - so processing these degrees takes *more computation time*.

17. Another Limitation: Non-Associativity

- Another limitation is that:
  - in contrast to the usual (*static*) fuzzy logic,
  - *dynamic* logic operations are not necessarily associative.
- In other words, the estimates for $(A \vee B) \vee C$ and for $A \vee (B \vee C)$ are, in general, different.
- We will show that this non-associativity is also a limitation
  - not of a specific *method* of extending "and"- and "or"-operations to dynamic fuzzy logic, but
  - of the very *dynamic character* of these logics.
- We will show that non-associativity occurs even if we restrict ourselves to linear operations.

18. Non-Associativity: Linear Restriction

- We plan to show that non-associativity occurs even if we restrict ourselves to linear operations.
- Why is such a restriction reasonable?
- One of the most frequently used probability-related fuzzy "or"-operation $f_\vee(a, b) = a + b - a \cdot b$ is:
  - approximately linear for small a and b;
  - isomorphic to $a + b$ if we appropriately re-scale the values from the interval $[0, 1]$ to $\mathbb{R}_0^+$.

19. Definitions

- For every integer t, by a *dynamical fuzzy t-value*, we mean a sequence $a = \{a_s\}_{s \le t}$, $a_s \ge 0$.
- For every t_0 and a, by a *shift* $S_{t_0}(a)$, we mean a sequence $a' = \{a'_s\}_{s \le t+t_0}$ with $a'_s = a_{s-t_0}$.
- By a *aggregation operation*, we mean an operation f that transforms t-sequences a and b into a value $c_t \ge 0$.
- An operation f is called *shift-invariant* if:
  - whenever it transforms a and b into a value c_t,
  - it transforms shifted values $S_{t_0}(a)$ and $S_{t_0}(b)$ into the same value c_{t+t_0}.
- We say that an aggregation operation f is *linear* if $c_t = Z_t + \sum\limits_{s \le t} A_{t,s} \cdot a_s + \sum\limits_{s \le t} B_{t,s} \cdot b_s$.
- By the *result* $c = f(a, b)$ of applying f to sequences a and b, we mean a sequence $c_s = f(\{a_u\}_{u \le s}, \{b_u\}_{u \le s})$.

20. Main Result about Non-Associativity

- **Proposition.** *If $c = f(a, b)$ is a shift-invariant linear commutative and associative operation, then:*
  - *the value c_t depends only on a_t and b_t and*
  - *does not depend on the values a_s and b_s for $s < t$.*
- So, any commutative linear operation that takes into account previous fuzzy estimates is *not* associative.
- Similar results are known in other application areas:
  - if we formulate natural requirements for a reasonable next step in a bargaining process,
  - then every function satisfying these requirements does not depend on the bargaining pre-history.

21. Acknowledgements

This work was supported in part:

- by the National Science Foundation grants HRD-0734825 and DUE-0926721,
- by Grant 1 T36 GM078000-01 from the National Institutes of Health, and
- by Grant 5015 “Application of fuzzy logic with operators in the knowledge based systems” from the Science and Technology Centre in Ukraine (STCU), funded by European Union.

The author is very thankful:

- to Leonid Perlovsky for his inspiring ideas and suggestions; and
- last but not the least, to the conference organizers for their invitation.

22. Proof: Meaning of Shift-Invariance

- Shift-invariance: for $a' = S_{t_0}(a)$ and $b' = S_{t_0}(b)$,

$$c_t = Z_t + \sum_{s\le t} A_{t,s}\cdot a_s + \sum_{s\le t} B_{t,s}\cdot b_s \text{ implies}$$

$$c_t = Z_{t+t_0} + \sum_{s\le t+t_0} A_{t+t_0,s}\cdot a'_s + \sum_{s\le t} B_{t+t_0,s}\cdot b'_s.$$

- Substituting $a'_s = a_{s-t_0}$ and $b'_s = b_{s-t_0}$, we get:

$$c_t = Z_{t+t_0} + \sum_{s\le t+t_0} A_{t+t_0,s}\cdot a_{s-t_0} + \sum_{s\le t} B_{t+t_0,s}\cdot b_{s-t_0}.$$

- Introducing a new variable $s' \stackrel{\text{def}}{=} s - t_0$, we get:

$$c_t = Z_{t+t_0} + \sum_{s'\le t} A_{t+t_0,s'+t_0}\cdot a_{s'} + \sum_{s\le t} B_{t+t_0,s'+t_0}\cdot b_{s'}.$$

- Two linear functions coincide if and only if all their coefficients coincide, so:

$$Z_t = Z_{t+t_0},\quad A_{t,s} = A_{t+t_0,s+t_0},\ \text{and } B_{t,s} = B_{t+t_0,s+t_0}.$$

23. Meaning of Shift-Invariance (cont-d)

- *Reminder:*

$$Z_t = Z_{t+t_0}, \quad A_{t,s} = A_{t+t_0,s+t_0}, \text{ and } B_{t,s} = B_{t+t_0,s+t_0}.$$

- For every two values t and t', we can take $t_0 = t' - t$, then $t + t_0 = t'$ hence $Z_t = Z_{t'}$.
- Thus, Z_t does not depend on t: $Z_t = Z$.
- From $A_{t,s} = A_{t+t_0,s+t_0}$, by taking $t_0 = -s$, we conclude that $A_{t,s} = A_{t-s,0}$.
- Thus, $A_{t,s} = A_{t-s}$, where $A_t \stackrel{\text{def}}{=} A_{t,0}$.
- Similarly, we conclude that $B_{t,s} = B_{t-s}$, for $B_t \stackrel{\text{def}}{=} B_{t,0}$.
- Thus, a shift-invariant linear operation has the form

$$c_t = Z + \sum_{s \le t} A_{t-s} \cdot a_s + \sum_{s \le t} B_{t-s} \cdot b_s.$$

24. Meaning of Commutativity

- Reminder: $c_t = Z + \sum\limits_{s \le t} A_{t-s} \cdot a_s + \sum\limits_{s \le t} B_{t-s} \cdot b_s$.
- Commutativity means that

$$Z + \sum_{s \le t} A_{t-s} \cdot a_s + \sum_{s \le t} B_{t-s} \cdot b_s =$$

$$Z + \sum_{s \le t} A_{t-s} \cdot b_s + \sum_{s \le t} B_{t-s} \cdot a_s.$$

- Here again, the fact that the two linear functions coincide means that all their coefficients must coincide.
- So, we conclude that $A_t = B_t$ for all t.
- Thus, the above formula for c_t takes the form

$$c_t = Z + \sum_{s \le t} A_{t-s} \cdot (a_s + b_s).$$

25. Meaning of Associativity

- *Reminder:* $c_t = \sum\limits_{s \le t} A_{t-s} \cdot (a_s + b_s)$.
- Associativity means that $f(f(a,b),c) = f(a, f(b,c))$.
- In $f(f(a,b),c)$, we first combine a and b into $d = f(a,b)$, and then combine d and c into $e = f(d,c)$.
- If we keep track only of the dependence on a_t, b_t, and c_t, we get $d_t = A_0 \cdot (a_t + b_t) + \ldots$ and thus:

$$e_t = A_0 \cdot (d_t + c_t) + \ldots = A_0^2 \cdot (a_t + b_t) + A_0 \cdot c_t + \ldots$$

- Associativity implies that

$$A_0^2 \cdot (a_t + b_t) + A_0 \cdot c_t = A_0^2 \cdot (b_t + c_t) + A_0 \cdot a_t.$$

- Since the two linear functions coincide, their coefficients must coincide, i.e., we must have $A_0 = A_0^2$.
- Thus, we have $A_0 = 0$ or $A_0 = 1$.

26. How We Will Prove Non-Associativity

- *Reminder:* $c_t = \sum\limits_{s \le t} A_{t-s} \cdot (a_s + b_s)$.
- We have proven that in the associativity case, $A_0 = 0$ or $A_0 = 1$.
- We will show that in both cases $A_0 = 0$ and $A_0 = 1$, we have $A_1 = A_2 = \ldots = 0$.
- This will prove that c_t depends only on a_t and b_t and does not depend on the previous values a_s and b_s.
- In both cases, we will prove it by contradiction.
- We will assume that $A_j \ne 0$ for some $j \ge 1$.
- In this proof, k will denote the smallest index $k \ge 0$ for which $A_k \ne 0$.

27. Case $A_0 = 0$

- *Reminder:* $c_t = \sum\limits_{s \le t} A_{t-s} \cdot (a_s + b_s)$, with $A_0 = A_1 = \ldots = A_{k-1} = 0, \quad A_k \ne 0.$
- For $d = f(a, b)$ and $e = f(d, c)$, we get

$$d_t = Z + A_k \cdot (a_{t-k} + b_{t-k}) + \ldots; \;\; e_t = Z + A_k \cdot (d_{t-k} + c_{t-k}) + \ldots$$

- Here, $d_{t-k} = Z + A_k \cdot (a_{t-2k} + b_{t-2k}) + \ldots$; thus, we have
$$e_t = Z + A_k \cdot Z + A_k^2 \cdot (a_{t-2k} + b_{t-2k}) + A_k \cdot c_{t-k} + \ldots$$
- Similarly, $f(a, f(b, c))$ leads to
$$e_t = Z + A_k \cdot Z + A_k^2 \cdot (b_{t-2k} + c_{t-2k}) + A_k \cdot a_{t-k} + \ldots$$
- So, $Z + A_k \cdot Z + A_k^2 \cdot (a_{t-2k} + b_{t-2k}) + A_k \cdot c_{t-k} + \ldots = Z + A_k \cdot Z + A_k^2 \cdot (b_{t-2k} + c_{t-2k}) + A_k \cdot a_{t-k} + \ldots$
- The left-hand side of this equality does not depend on a_{t-k}, while the right-hand side does ($A_k \ne 0$).
- Thus, the equality is indeed impossible.

28. Case $A_0 = 1$

- We find $d = f(a, b)$ and $e = f(d, c)$, w/ $A_0 = 1$, $A_1 = \ldots = A_{k-1} = 0$, $A_k \neq 0$, $c_t = \sum\limits_{s \le t} A_{t-s} \cdot (a_s + b_s)$.
- Here, $d_t = Z + a_t + b_t + A_k \cdot (a_{t-k} + b_{t-k}) + \ldots$ so $e_t = Z + d_t + c_k + A_k \cdot (d_{t-k} + c_{t-k}) + \ldots$
- Here, $d_{t-k} = Z + a_{t-k} + b_{t-k} + \ldots$; thus, we have

$$e_t = Z + (Z + a_t + b_t + A_k \cdot (a_{t-k} + b_{t-k}) + \ldots) + c_t +$$
$$A_k \cdot ((Z + a_{t-k} + b_{t-k} + \ldots) + c_{t-k}) + \ldots =$$
$$2Z + a_t + b_t + c_t + A_k \cdot (2a_{t-k} + 2b_{t-k} + c_{t-k}) + \ldots$$

- Similarly, the expression $f(a, f(b, c))$ leads to

$$e_t = 2Z + a_t + b_t + c_t + A_k \cdot (2b_{t-k} + 2c_{t-k} + a_{t-k}) + \ldots$$

- Thus $2Z + a_t + b_t + c_t + A_k \cdot (2a_{t-k} + 2b_{t-k} + c_{t-k}) + \ldots = 2Z + a_t + b_t + c_t + A_k \cdot (2b_{t-k} + 2c_{t-k} + a_{t-k}) + \ldots$

29. Case $A_0 = 1$ (cont-d)

- We have concluded that:

$$2Z + a_t + b_t + c_t + A_k \cdot (2a_{t-k} + 2b_{t-k} + c_{t-k}) + \ldots =$$

$$2Z + a_t + b_t + c_t + A_k \cdot (2b_{t-k} + 2c_{t-k} + a_{t-k}) + \ldots$$

- *Reminder:* $A_k \neq 0$.
- Here:
  - The left-hand contains a_{t-k} with a coefficient $2A_k$, while
  - the right-hand side has this variable with a different coefficient $A_k \neq 2A_k$.
- Thus, the equality is impossible in this case as well.
- The proposition is proven.