“And”- and “Or”-Operations for “Double”, “Triple”, etc. Fuzzy Sets

Hung T. Nguyen, Vladik Kreinovich, and Olga Kosheleva

***Abstract*—In the traditional fuzzy logic, the expert’s degree of confidence $d(A\,\&\,B)$ in a complex statement $A\,\&\,B$ (or $A \vee B$) is uniquely determined by his/her degrees of confidence $d(A)$ and $d(B)$ in the statements A and B, as $f_\&(d(A), d(B))$ for an appropriate “and”-operation (t-norm). In practice, for the same degrees $d(A)$ and $d(B)$, we may have different degrees $d(A\,\&\,B)$ depending on the relation between A and B. The best way to take this relation into account is to explicitly elicit the corresponding degrees $d(A\,\&\,B)$ and $d(A \vee B)$, i.e., to come up with a “double” fuzzy set. If we only elicit information about pairs of statements, then we still need to estimate, e.g., the degree $d(A\,\&\,B\,\&\,C)$ based on the known values $d(A)$, $d(B)$, $d(C)$, $d(A\,\&\,B)$, $d(A\,\&\,C)$, and $d(B\,\&\,C)$. In this paper, we explain how to produce such “and”-operations for “double” fuzzy sets – and how to produce similar “or”-operations.**

I. Traditional Fuzzy Techniques: A Brief Reminder

Need for fuzzy techniques: reminder. Experts often describe their knowledge by using imprecise (“fuzzy”) words from a natural language like “small” or “fast”. One of the most widely used describe this knowledge in computer-understandable terms is to use *fuzzy techniques*, in which, for each imprecise property P and for each possible value x of the corresponding property, we store the degree $\mu_P(x)$ to which the expert believes that x satisfies the property P; see, e.g., [7], [8], [10].

Each of these values can be obtained, e.g., by asking the expert to mark his or her degree of certainty that x satisfies P by a mark on a scale from 0 to to some integer n. If the expert marks m on a scale from 0 to n, we take $\mu_P(x) = m/n$.

Another possibility is to use polling: we ask n experts and if m of them think that x satisfies the property P, we take $\mu_P(x) = m/n$. Thus obtained degree can be interpreted as a probability: namely, as a probability that a randomly selected expert thinks that x satisfies the property P.

Need for “and”- and “or”-operations. One of the main objectives of storing the expert knowledge is to enable the computer to use expert rules – rules formulated in terms of imprecise natural-language words. The conditions of such rules often include several properties: e.g., if the car in front is close *and* it is going fast, then …To figure out to what extend such rules are applicable in given situations, we need not only to describe the degree to which a given distance is close and the degree to which a given velocity is fast, we also need to find the degree to which the expert believes in the corresponding composite “and”-statement.

Ideally, we should ask the expert’s opinion about all such combinations. However, in principle, many such combinations are possible, and it is not possible to ask the expert’s opinion about all such combinations. It is therefore necessary to estimate our degree of belief in a propositional combination like $A\,\&\,B$ or $A \vee B$ in the situation when the only information that we have is the expert’s degrees of belief $d(A)$ and $d(B)$ in statements A and B. For each of these two propositional connectives $\&$ and $\vee$, we thus need to come up with an algorithm that transform the degrees $d(A)$ and $d(B)$ into a reasonable estimate for $d(A\,\&\,B)$ or $d(A \vee B)$.

Let us denote the algorithm corresponding to $\&$ by $f_\&(a, b)$, and the algorithm corresponding to $\vee$ by $f_\vee(a, b)$. Once we use these algorithms, we estimate $d(A\,\&\,B)$ as $f_\&(d(A), d(B))$ and $d(A \vee B)$ as $f_\vee(d(A), d(B))$. We want these algorithms to be reasonable. For example, since $A\,\&\,B$ is equivalent to $B\,\&\,A$, it is reasonable to require that these two formulas lead to the same estimate for $d(A\,\&\,b)$, i.e., that the equality

$$f_\&(d(A), d(B)) = f_\&(d(B), d(A))$$

be true for all possible values of $d(A)$ and $d(B)$. In mathematical terms, it is reasonable to require that the operation $f_\&(a, b)$ is *commutative.* Similarly, since $A \vee B$ also means the same as $B \vee A$, it is also reasonable to require that the operation $f_\vee(a, b)$ is commutative.

Similarly, since $A\,\&\,(B\,\&\,C)$ is equivalent to $(A\,\&\,B)\,\&\,C$, it makes sense to require that the corresponding estimates coincide, i.e., that

$$f_\&(d(A), f_\&(d(B), d(C))) = f_\&(f_\&(d(A), d(B)), d(C)).$$

In mathematical term, this means that the operation $f_\&(a, b)$ is *associative*. Similarly, it is reasonable to require that the operation $f_\vee(a, b)$ is associative. Together with additional reasonable requirements like monotonicity, continuity, etc., these properties form the definitions of *“and”-operations* (also known as *t-norms*) and *“or”-operations* (also known as *t-conorms*.

Similarly, we use a negation operation $f_\neg(a)$ to estimate the degree to which the negation $\neg A$ is true as

$$d(\neg A) \approx f_\neg(d(A)).$$

Hung T. Nguyen is with Department of Mathematical Sciences, New Mexico State University, Las Cruces, NM, 88003, USA, and with Department of Economics, Chiang Mai University, Thailand (email: hunguyen@nmsu.edu); Vladik Kreinovich and Olga Kosheleva are with the University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA (emails: olga@utep.edu, vladik@utep.edu).

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, by Grants 1 T36 GM078000-01 and 1R43TR000173-01 from the National Institutes of Health, and by a grant N62909-12-1-7039 from the Office of Naval Research.

Historical comment. Historically the first "and"- and "or"-operations – introduced in the pioneer paper [10] – are $f_\&(a,b) = \min(a,b)$, $f_\&(a,b) = a \cdot b$, $f_\vee(a,b) = \max(a,b)$, and $f_\vee(a,b) = a + b - a \cdot b$. The most widely used negation operation is $f_\neg(a) = 1 - a$.

II. Need to Go Beyond Traditional Fuzzy Techniques: Enter "Double" Fuzzy Sets

Traditional fuzzy approach: reminder. In the traditional fuzzy techniques, we base our estimate of the expert's degree of belief in a composite statement $A \,\&\, B$ only on the degrees of belief $d(A)$ and $d(B)$ in A and B, we do not take into account the relation between such statements. In reality, for the same degrees of belief in A and B, we may have different degrees of belief in $A \,\&\, B$.

First example. For example, suppose that an expert's degree of confidence in a statement A is 0.5. Then, it is reasonable to conclude that the expert's degree of confidence in the opposite statement $\neg A$ is equal to $1 - 0.5 = 0.5$.

- If we take $B = A$, then we have $d(A) = d(B) = 0.5$, and, since $A \,\&\, B$ is simply equivalent to A, we have $d(A \,\&\, B) = 0.5$.
- On the other hand, if we take $B = \neg A$, then we still have $d(A) = d(B) = 0.5$, but here, $A \,\&\, B$ is impossible, so we have $d(A \,\&\, B) = 0 \neq 0.5$.

Second example. Let us give another example, more closely related to the degrees to which different values x satisfy a given property. Namely, suppose that:

- the expert's degree of belief that a 50-years-old is old is 0.1, and
- the expert's degree of belief that a 60-years-old is old is 0.8.

What is the expert's degree of belief that 50 is old but 60 is not old? The procedure used in the traditional fuzzy logic leads to:

- $d(60 \text{ is not old}) = 1 - d(60 \text{ is old}) = 1 - 0.8 = 0.2$, and thus, to
- $d((50 \text{ is old}) \,\&\, (60 \text{ is not old})) = f_\&(0.1, 0.2)$.

Whether we use $f_\&(a,b) = \min(a,b)$ or $f_\&(a,b) = a \cdot b$, we get a positive degree – which makes no sense, since if an expert considers 50-year-olds to be old, then of course this expert should also consider 60-year-olds to be old.

Analysis of the problem. The reason for the above counterintuitive results is that the traditional fuzzy logic does not take into account the dependence between the statements.

A natural idea. A natural solution to the above problem is to explicitly elicit and store not only the expert's degree of confidence $\mu_P(x)$ that a given value x satisfies the property x but also the degree of confidence $\mu_{PP}(x, x')$ that both x and x' satisfy the property P; see, e.g., [5].

This idea enables us to avoid the above counterintuitive conclusion. Indeed, e.g., for the property "old", once we believe that x is old, this automatically makes us believe that all larger ages correspond to "old".

For example, the degree of believe that both 50 and 60 correspond to "old" is the same as the degree of belief that 50 is old. In general, we should take $\mu_{PP}(x, x') = \mu_P(x)$ for $x < x'$.

Towards a precise description. In the traditional fuzzy approach, a property is described by a single function

$$\mu_P : X \to [0, 1].$$

In the new approach, to describe a property, we need *two* functions:

- a function $\mu_P : X \to [0, 1]$, and
- a function $\mu_{PP} : X \times X \to [0, 1]$ for which

$$\mu_{PP}(x, x') = \mu_{PP}(x', x) \text{ and } \mu_{PP}(x, x') \le \mu_P(x).$$

Since we now need two functions to describe a property, it is natural to call such pairs of functions (μ_P, μ_{PP}) *double fuzzy sets.*

From "double" to "triple" etc., fuzzy sets. In addition to asking an expert to what extent both x and x' satisfy the desired property P, we can also ask the same question about the triples (x, x', x'') etc.

III. Formulation of the Problem: We Need to Extend "And"- and "Or"-Operations to "Double", "Triple" etc. Fuzzy Sets

"And"-operations in traditional fuzzy logic: reminder. In the traditional fuzzy approach, the degree of belief that both x and x' satisfy the property P would be estimated as $f_\&(\mu_P(x), \mu_P(x'))$.

For "double" fuzzy sets, we do not need "and"-operations for pairs. In the "double" fuzzy set approach, instead of using this approximate description, we explicitly solicit, for each pair (x, x'), the expert's degree of confidence that both x and x' satisfy the property P.

For triples, we still need an appropriate "and"-operation. What if we want to estimate the expert's degree that x, x', and x'' all satisfy the property P? In the context of "double" fuzzy sets, we do not explicitly ask such questions, we only ask questions about individual elements x, x', and x'', and about pairs.

Thus, we need to estimate the desired degree $d(P(x) \,\&\, P(x') \,\&\, P(x''))$ based on the known degrees

$$\mu_P(x), \quad \mu_P(x'), \quad \mu_P(x''),$$

$$\mu_{PP}(x, x'), \quad \mu_{PP}(x, x''), \text{ and } \mu_{PP}(x', x'').$$

In other words, we still need an appropriate "and"-operation.

What we do in this paper. In this paper, we show that ideas that lead to the most popular t-norms and t-conorms can be extended to describe the desired "and"- and "or"-operations for the "double" fuzzy sets.

IV. "AND"- AND "OR"-OPERATIONS FOR TRADITIONAL FUZZY SETS: REMINDER

Degrees of confidence and subjective probabilities. Traditionally, expert's degrees of certainty are also called subjective probabilities.

In some cases, they are indeed similar to probabilities – e.g., when we determine these degrees of certainty by polling: by asking n experts whether the given statement s is true, and taking, as the degree of confidence $d(s)$, the ration m/n, where m is the number of experts who believes that s is true. In this case, the resulting degree of confidence in a statement s is equal to the probability that a randomly selected expert considers the statement s to be true.

In view of this relation between degrees of confidence and probabilities – and taking into account that probabilistic methods have been developed for many centuries now, so a lot of techniques are known – we will use probabilistic methods to derive formulas for "and"- and "or"-operations. At first glance, this may seem restrictive, but, as we will show, the most widely used "and" and "or"-operations can indeed be obtained this way.

Specifically, we formulate the following problem:

- we know the probabilities $p(s_1)$ and $p(s_2)$ of two statements s_1 and s_2;
- we want to estimate the probability $p(s_1 \,\&\, s_2)$.

Corresponding probability-related techniques: reminder. The above problem is not uniquely determined: depending on the dependence between s_1 and s_2, we may have different values of the desired probability $p(s_1 \,\&\, s_2)$. There are two main approaches to deal with this non-uniqueness:

- we can find the range of all possible values $p(s_1 \,\&\, s_2)$; and
- we can select a single "most probable" value $p(s_1 \,\&\, s_2)$.

Let us describe both approaches in detail.

Inequalities (linear programming) approach. To get a full description of the joint probability distribution on the set of two statements s_1 and s_2, we need to know the probabilities of all basic combinations $s_1 \,\&\, s_2$, $s_1 \,\&\, \neg s_2$, $\neg s_1 \,\&\, s_2$, and $\neg s_1 \,\&\, \neg s_2$. One can check that once we know the probabilities $d_1 = p(s_1)$ and $d_2 = p(s_2)$ and the probability $x \stackrel{\text{def}}{=} p(s_1 \,\&\, s_2)$, we can uniquely determine all the remaining probabilities:

$$p(s_1 \,\&\, \neg s_2) = p(s_1) - p(s_1 \,\&\, s_2) = d_1 - x,$$

$$p(\neg s_1 \,\&\, s_2) = p(s_2) - p(s_1 \,\&\, s_2) = d_2 - x, \text{ and}$$

$$p(\neg s_1 \,\&\, \neg s_2) = 1 - p(s_1) - p(s_2) + p(s_1 \,\&\, s_2) =$$

$$1 - d_1 - d_2 + x.$$

For which values x do these formulas lead to a probability distribution? In a probability distribution, all the basic probabilities are non-negative and add up to 1. It is easy to check that the values x, $d_1 - x$, $d_2 - x$, and $1 - d_1 - d_2 + x$ always add up to 1. Thus, to make sure that the value x describes a probability distribution, it is sufficient to make sure that all fours resulting values of basic probabilities are non-negative, i.e., that the following four inequalities hold:

$$x \ge 0; \quad d_1 - x \ge 0; \quad d_2 - x \ge 0; \quad 1 - d_1 - d_2 + x \ge 0.$$

In general, several possible value x satisfy these inequalities. It is reasonable to find the *range* of such values x, i.e., to find the smallest and the largest value x for which the above four expressions form a probability distribution.

From the mathematical viewpoint, we thus need to find the maximum and the minimum of x under the above four linear inequalities. The problem of optimizing a linear function under linear equalities and/or inequalities is known as *linear programming*; there exist efficient algorithms for solving such problems; see, e.g., [1], [3], [4], [9]. In view of this relation, the above approach is also known as the linear programming approach.

For the above inequalities, we can find an explicit solution if we move x to one of the sides of each inequality and all the other terms to the other side. As a result, we get the following system of four inequalities:

$$x \ge 0; \quad x \le d_1; \quad x \le d_2; \quad x \ge d_1 + d_2 - 1.$$

The inequalities $x \le d_1$ and $x \le d_2$ can be described as $x \le \min(d_1, d_2)$. Similarly, the inequalities $x \ge 0$ and $x \ge d_1 + d_2 - 1$ can be described as $x \ge \max(d_1 + d_2 - 1, 0)$. Thus, the value x determines a probability distribution if and only if

$$\max(d_1 + d_2 - 1, 0) \le x \le \min(d_1, d_2).$$

We have thus found the desired range; its lower endpoint is the value

$$\max(d_1 + d_2 - 1, 0),$$

its upper endpoint is the value $\min(d_1, d_2)$.

Both endpoints serve as possible t-norms:

- $f_\&(a, b) = \max(a + b - 1, 0)$ is the smallest possible t-norm; while
- $f_\&(a, b) = \min(a, b)$ is the largest possible t-norm; this is actually one of the most widely used t-norms.

Maximum Entropy approach. In applications of probability theory, we often encounter situations when we do not know the exact probability distribution, i.e., when several different distributions are consistent with our knowledge. Some of these distributions have smaller uncertainty, some have larger uncertainty. In this case, a reasonable idea is not to hide the possible uncertainty, i.e., to select a distribution with the largest uncertainty. There are reasonable arguments that uncertainty of a probability distribution is best described by its entropy $S = -\sum p_i \cdot \ln(p_i)$; as a result, we usually select a distribution with the largest entropy; see, e.g., [2], [6].

In the above case, we have four probabilities

$$x, \quad d_1 - x, \quad d_2 - x, \text{ and } 1 - d_1 - d_2 + x,$$

so the entropy takes the form

$$S = -x \cdot \ln(x) - (d_1 - x) \cdot \ln(d_1 - x) - (d_2 - x) \cdot \ln(d_2 - x) -$$

$$(1-d_1-d_2+x)\cdot\ln(1-d_1-d_2+x).$$

To find the value x for which entropy is the largest, we differentiate this expression relative to x and equate the derivative to 0. As a result, we get

$$-\ln(x)+\ln(d_1-x)+\ln(d_2-x)-\ln(1-d_1-d_2+x)=0.$$

Moving all negative terms to the right-hand side, we get

$$\ln(d_1-x)+\ln(d_2-x)=\ln(x)+\ln(1-d_1-d_2+x).$$

Raising e to the power of both sides, and taking into account that $e^{a+b}=e^a\cdot e^b$ and that $e^{\ln(z)}=z$, we conclude that

$$(d_1-x)\cdot(d_2-x)=x\cdot(1-d_1-d_2+x).$$

Opening parentheses, we get

$$d_1\cdot d_2-x\cdot(d_1+d_2)+x^2=x-x\cdot(d_1+d_2)+x^2.$$

Canceling similar terms in both sides, we get $x=d_1\cdot d_2$.

The corresponding "and"-operation is indeed one of the most widely used in fuzzy logic.

How to derive the corresponding "or"-operations: idea. The corresponding "or"-operations can be derived from the "and"-operations if we take into account that $s_1\vee s_2$ is equivalent to $\neg((\neg s_1)\,\&\,(\neg s_2))$.

For probabilities, $p(\neg s)=1-p(s)$. Thus, $p(\neg s_1)=1-p(s_1)$ and $p(\neg s_2)=1-p(s_2)$. So, once we have selected the "and"-operation $f_\&(a,b)$, we can determine the probability $p((\neg s_1)\,\&\,(\neg s_2))$ as

$$p((\neg s_1)\,\&\,(\neg s_2))=f_\&(p(\neg s_1),p(\neg s_2))=$$

$$f_\&(1-p(s_1),1-p(s_2)).$$

Hence, the desired probability $p(s_1\vee s_2)$ can be estimated as

$$p(s_1\vee s_2)=p(\neg((\neg s_1)\,\&\,(\neg s_2)))=$$

$$1-p((\neg s_1)\,\&\,(\neg s_2)))=1-f_\&(1-p(s_1),1-p(s_2)).$$

In other words, once we have defined an "and"-operation $f_\&(a,b)$, we can determine the corresponding "or"-operation as

$$f_\vee(a,b)=1-f_\&(1-a,1-b).$$

Let us show what we get when we apply this idea to the above "and"-operations.

"Or"-operations: inequalities (linear programming) approach.

- For $f_\&(a,b)=\max(a+b-1,0)$, we get
$$f_\vee(a,b)=\min(a+b,1).$$
- For $f_\&(a,b)=\min(a,b)$, we get
$$f_\vee(a,b)=\max(a,b);$$
this is actually one of the most widely used "or"-operations (t-conorms).

"Or"-operation: Maximum Entropy approach. For $f_\&(a,b)=a\cdot b$, we get

$$f_\vee(a,b)=a+b-a\cdot b.$$

This "or"-operation is indeed one of the most widely used in fuzzy logic.

V. "And"- and "Or"-Operations for "Double" Fuzzy Sets

Analysis of the problem. Let us apply the above approaches to estimate

$$x=d(s_1\,\&\,s_2\,\&\,s_3)$$

for double fuzzy sets.

To fully describe the probability distribution for the case of three statements, we need to find the probabilities of all eight possible basic combinations:

$$p(s_1\,\&\,s_2\,\&\,s_3),\quad p(s_1\,\&\,s_2\,\&\,\neg s_3),$$

$$p(s_1\,\&\,\neg s_2\,\&\,s_3),\quad p(s_1\,\&\,\neg s_2\,\&\,\neg s_3),$$

$$p(\neg s_1\,\&\,s_2\,\&\,s_3),\quad p(\neg s_1\,\&\,s_2\,\&\,\neg s_3),$$

$$p(\neg s_1\,\&\,\neg s_2\,\&\,s_3),\ \text{and}\ p(\neg s_1\,\&\,\neg s_2\,\&\,\neg s_3).$$

If we know the values

$$d_1=p(s_1),\quad d_2=p(s_2),\quad d_3=p(s_3),$$

$$d_{12}=p(s_1\,\&\,s_2),\quad d_{13}=p(s_1\,\&\,s_3),$$

$$d_{23}=p(s_2\,\&\,s_3),\ \text{and}$$

$$x=p(s_1\,\&\,s_2\,\&\,s_3),$$

then we can uniquely reconstruct all remaining seven probabilities:

$$p(s_1\,\&\,s_2\,\&\,\neg s_3)=p(s_1\,\&\,s_2)-p(s_1\,\&\,s_2\,\&\,s_3)=d_{12}-x;$$

$$p(s_1\,\&\,\neg s_2\,\&\,s_3)=p(s_1\,\&\,s_3)-p(s_1\,\&\,s_2\,\&\,s_3)=d_{13}-x;$$

$$p(\neg s_1\,\&\,s_2\,\&\,s_3)=p(s_2\,\&\,s_3)-p(s_1\,\&\,s_2\,\&\,s_3)=d_{23}-x;$$

$$p(s_1\,\&\,\neg s_2\,\&\,\neg s_3)=p(s_1)-p(s_1\,\&\,s_2)-p(s_1\,\&\,s_3)+$$

$$p(s_1\,\&\,s_2\,\&\,s_3)=$$

$$d_1-d_{12}-d_{13}+x;$$

$$p(\neg s_1\,\&\,s_2\,\&\,\neg s_3)=p(s_2)-p(s_1\,\&\,s_2)-p(s_2\,\&\,s_3)+$$

$$p(s_1\,\&\,s_2\,\&\,s_3)=$$

$$d_2-d_{12}-d_{23}+x;$$

$$p(\neg s_1\,\&\,\neg s_2\,\&\,s_3)=p(s_3)-p(s_1\,\&\,s_3)-p(s_2\,\&\,s_3)+$$

$$p(s_1\,\&\,s_2\,\&\,s_3)=$$

$$d_3-d_{13}-d_{23}+x;$$

$$p(\neg s_1\,\&\,\neg s_2\,\&\,\neg s_3)=$$

$$1-p(s_1)-p(s_2)-p(s_3)+$$

$$p(s_1\,\&\,s_2)+p(s_1\,\&\,s_3)+p(s_2\,\&\,s_3)-$$

$$p(s_1 \,\&\, s_2 \,\&\, s_3) =$$

$$1 - d_1 - d_2 - d_3 + d_{12} + d_{13} + d_{23} - x.$$

Inequalities approach. Let us start with the inequalities approach.

Similar to the case of two statements, these eight probabilities add up to one, so the only requirement is that all these eight expressions are non-negative:

$$x \ge 0; d_{12} - x \ge 0; \quad d_{13} - x \ge 0; \quad d_{23} - x \ge 0;$$

$$d_1 - d_{12} - d_{13} + x \ge 0; \quad d_2 - d_{12} - d_{23} + x \ge 0;$$

$$d_3 - d_{12} - d_{23} + x \ge 0;$$

$$1 - d_1 - d_2 - d_3 + d_{12} + d_{23} + d_{13} - x \ge 0.$$

By moving x to one side and all other terms to another side, we get an equivalent set of inequalities:

$$x \ge 0; \quad x \le d_{12}; \quad x \le d_{13}; \quad x \le d_{23};$$

$$x \ge d_{12} + d_{13} - d_1; \quad x \ge d_{12} + d_{23} - d_2; \quad x \ge d_{13} + d_{23} - d_3;$$

$$x \le 1 - d_1 - d_2 - d_3 + + d_{12} + d_{13} + d_{23}.$$

These inequalities provide several lower and upper bounds for x. The value x is larger than or equal to several lower bounds if and only it is larger than or equal than the largest of these lower bounds. Similarly, the value x is smaller than or equal to several upper bounds if and only it is smaller than or equal than the smallest of these upper bounds. Thus, the above eight inequalities are equivalent to the following inequality:

$$\max(d_{12} + d_{13} - d_1, d_{12} + d_{23} - d_2, d_{13} + d_{23} - d_3, 0) \le x \le$$

$$\min(d_{12}, d_{13}, d_{23}, 1 - d_1 - d_2 - d_3 + d_{12} + d_{13} + d_{23}).$$

Thus, we get the formulas for the lower and upper estimations for $p(s_1 \,\&\, s_2 \,\&\, s_3)$:

- as the lower estimate, we can take
  $$\max(d_{12} + d_{13} - d_1, d_{12} + d_{23} - d_2, d_{13} + d_{23} - d_3, 0);$$
- as the upper estimate, we can take
  $$\min(d_{12}, d_{13}, d_{23}, 1 - d_1 - d_2 - d_3 + d_{12} + d_{13} + d_{23}).$$

Maximum Entropy approach. For each value x form the corresponding range, we get a probability distribution with probabilities x, $d_{12} - x$, $d_{13} - x$, $d_{23} - x$, $d_1 - d_{12} - d_{13} + x$, $d_2 - d_{12} - d_{23} + x$, $d_3 - d_{12} - d_{23} + x$, and

$$1 - d_1 - d_2 - d_3 + d_{12} + d_{23} + d_{13} - x.$$

The entropy of this distribution is equal to

$$S = -x \cdot \ln(x) - (d_{12} - x) \cdot \ln(d_{12} - x) - (d_{13} - x) \cdot \ln(d_{13} - x) -$$

$$(d_{23} - x) \cdot \ln(d_{23} - x) -$$

$$(d_1 - d_{12} - d_{13} + x) \cdot \ln(d_1 - d_{12} - d_{13} + x) -$$

$$(d_2 - d_{12} - d_{23} + x) \cdot \ln(d_2 - d_{12} - d_{23} + x) -$$

$$(d_3 - d_{13} - d_{23} + x) \cdot \ln(d_3 - d_{13} - d_{23} + x) -$$

$$(1 - d_1 - d_2 - d_3 + d_{12} + d_{23} + d_{13} - x) \cdot$$

$$\ln(1 - d_1 - d_2 - d_3 + d_{12} + d_{23} + d_{13} - x).$$

Differentiating this expression with respect to x and equating the derivative to 0, we conclude that

$$-\ln(x) + \ln(d_{12} - x) + \ln(d_{13} - x) + \ln(d_{23} - x) -$$

$$\ln(d_1 - d_{12} - d_{13} + x) - \ln(d_2 - d_{12} - d_{23} + x) -$$

$$\ln(d_3 - d_{13} - d_{23} + x) +$$

$$\ln(1 - d_1 - d_2 - d_3 + d_{12} + d_{23} + d_{13} - x) = 0.$$

If we raise e to the power of both side, we get a 4-th order equation (actually 3rd order since terms x^4 cancel out). In this case, however, we do not have a closed form solution, we have to use numerical methods to solve this equation.

Comment. Similar ideas can be used to describe "or"-operations and to describe "and"- and "or"-operations for "triple" etc. fuzzy sets.

References

[1] M. Ceberio, V. Kreinovich, S. Chopra, L. Longpré, H. T. Nguyen, B. Ludäscher, and C. Baral, "Interval-type and affine arithmetic-type techniques for handling uncertainty in expert systems", *Journal of Computational and Applied Mathematics*, 2007, Vol. 199, No. 2, pp. 403–410.

[2] B. Chokr and V. Kreinovich, "How far are we from complete knowledge: complexity of knowledge acquisition in Dempster-Shafer approach", In: R. R. Yager, J. Kacprzyk, and M. Pedrizzi (Eds.), *Advances in the Dempster-Shafer Theory of Evidence*, Wiley, New York, 1994, pp. 555–576.

[3] S. Chopra, "Affine arithmetic-type techniques for handling uncertainty in expert systems", *International Journal of Intelligent Technologies and Applied Statistics*, 2008, Vol. 1, No. 1, pp. 59–110.

[4] Th. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, MA, 2009.

[5] I. R. Goodman, "Fuzzy sets as equivalence classes of random sets", In: R. Yager et al. (eds.), *Fuzzy Sets and Possibility Theory*, Pergamon Press, Oxford, UK, 1982, pp. 327–432.

[6] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, Massachusesst, 2003.

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

[8] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.

[9] N. J. Nilsson, "Probabilistic logic", *Artifical Intelligence*, 1986, Vol. 28, No. 1, pp. 71–87.

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