An Even More Realistic (Non-Associative) Logic And Its Relation to Psychology of Human Reasoning

I. R. Goodman
Code D42215, Space & Naval Warfare Systems Center
San Diego, CA 92152, USA
goodman@spawar.navy.mil

Vladik Kreinovich, Raúl Trejo,
Jesus Martinez, and Reginaldo Gonzalez
Computer Science, Univ. of Texas at El Paso
El Paso, TX 79968, USA
vladik@cs.utep.edu

Abstract

If we know the degrees of certainty (subjective probabilities) $p(S_1)$ and $p(S_2)$ in two statements S_1 and S_2, then possible values of $p(S_1 \,\&\, S_2)$ form an interval $p = [\max(p_1 + p_2 - 1, 0), \min(p_1, p_2)]$. As a numerical estimate, it is natural to use a midpoint p of this interval; this midpoint is a mathematical expectation of $p(S_1 \,\&\, S_2)$ over a uniform (second order) distribution on all possible probability distributions.

This midpoint operation & is not associative. We show that the upper bound on the difference $a \,\&\, (b \,\&\, c) - (a \,\&\, b) \,\&\, c$ is 1/9, so if the size of the corresponding granules is $\geq 1/9$, we will not notice this associativity. This may explain the famous 7 ± 2 law, according to which we use no more than 9 granules.

1. In Expert Systems, We Need Estimates for the Degree of Certainty of $S_1 \,\&\, S_2$ and $S_1 \vee S_2$

In many areas (medicine, geophysics, military decision-making, etc.), top quality experts make good decisions, but they cannot handle all situations. It is therefore desirable to incorporate their knowledge into a decision-making computer system.

Experts describe their knowledge by statements $S_1, \ldots, S_n$ (e.g., by if-then rules). Experts are often not 100% sure about these statements S_i; this uncertainty is described by the *subjective probabilities* p_i (degrees of belief, etc.) which experts assign to their statements. The conclusion C of an expert system normally depends on several statements S_i. For example, if we can deduce C either from S_2 and S_3, or from S_4, then the validity of C is equivalent to the validity of a Boolean combination $(S_2 \,\&\, S_3) \vee S_4$. So, to estimate the reliability $p(C)$ of the conclusion, we must estimate the probability of Boolean combinations. In this paper, we consider the simplest possible Boolean combinations are $S_1 \,\&\, S_2$ and $S_1 \vee S_2$.

In general, the probability $p(S_1 \,\&\, S_2)$ of a Boolean combination can take different values depending on whether S_1 and S_2 are independent or correlated. So, to get the precise estimates of probabilities of all possible conclusions, we must know not only the probabilities $p(S_i)$ of individual statements, but also the probabilities of all possible Boolean combinations. To get all such probabilities, it is sufficient to describe 2^n probabilities of the combinations $E_1^{\varepsilon_1} \,\&\, \ldots \,\&\, E_n^{\varepsilon_n}$, where $\varepsilon_i \in \{+, -\}$, E^+ means E, and E^- means $\neg E$. The only condition on these probabilities is that their sum should add up to 1, so we need to describe $2^n - 1$ different values. A typical knowledge base may contain hundreds of statements; in this case, the value $2^n - 1$ is astronomically large. We cannot ask experts about all 2^n such combinations, so in many cases, we must estimate $p(S_1 \,\&\, S_2)$ or $p(S_1 \vee S_2)$ based only on the values $p_1 = p(S_1)$ and $p_2 = p(S_2)$.

2. Interval Estimates Are Possible, But Sometimes, Numerical Estimates Are Needed

It is known that for given $p_1 = p(S_1)$ and $p_2 = p(S_2)$:

- possible values of $p(S_1 \,\&\, S_2)$ form an interval $\mathbf{p} = [p^-, p^+]$, where $p^- = \max(p_1 + p_2 - 1, 0)$ and

$p^+ = \min(p_1, p_2)$; and

- possible values of $p(S_1 \vee S_2)$ form an interval $\mathbf{p} = [p^-, p^+]$, where $p^- = \max(p_1, p_2)$ and $p^+ = \min(p_1 + p_2, 1)$

(see, e.g., a survey [12] and references therein).

So, in principle, we can use such interval estimates and get an interval $\mathbf{p}(C)$ of possible values of $p(C)$. Sometimes, this idea leads to meaningful estimates, but often, it leads to a useless $\mathbf{p}(C) = [0, 1]$ (see, e.g., [12, 13]). In such situations, it is reasonable, instead of using the entire interval $\mathbf{p}$, to select a point within this interval as a reasonable estimate for $p(S_1 \,\&\, S_2)$ (or, correspondingly, for $p(S_1 \vee S_2)$).

3. Natural Idea: Selecting a Midpoint as the Desired Estimate

Since the only information we have, say, about the unknown probability $p(S_1 \,\&\, S_2)$ is that it belongs to the interval $[p^-, p^+]$, it is natural to select a *midpoint* of this interval as the desired estimate. In other words, if we know the probabilities p_1 and p_2 of the statements S_1 and S_2, then, as estimates for $p(S_1 \,\&\, S_2)$ and $p(S_1 \vee S_2)$, we can take the values $p_1 \,\&\, p_2$ and $p_1 \vee p_2$, where

$$p_1 \,\&\, p_2 \stackrel{\text{def}}{=} \frac{1}{2} \cdot \max(p_1 + p_2 - 1, 0) + \frac{1}{2} \cdot \min(p_1, p_2);$$

$$p_1 \vee p_2 \stackrel{\text{def}}{=} \frac{1}{2} \cdot \max(p_1, p_2) + \frac{1}{2} \cdot \min(p_1 + p_2, 1).$$

This midpoint selection is not only natural from a common sense viewpoint; it also has a deeper justification. Namely, in accordance of our above discussion, for $n = 2$ statements S_1 and S_2, to describe the probabilities of all possible Boolean combinations, we need to describe $2^2 = 4$ probabilities $x_1 = p(S_1 \,\&\, S_2)$, $x_2 = p(S_1 \,\&\, \neg S_2)$, $x_3 = p(\neg S_1 \,\&\, S_2)$, and $x_4 = p(\neg S_1 \,\&\, \neg S_2)$; these probabilities should add up to 1: $x_1 + x_2 + x_3 + x_4 = 1$. Thus, each probability distribution can be represented as a point $(x_1, \ldots, x_4)$ in a 3-D simplex

$$\mathcal{S} = \{(x_1, x_2, x_3, x_4) \,|\, x_i \geq 0 \,\&\, x_1 + \ldots + x_4 = 1\}.$$

We know the values of $p_1 = p(S_1) = x_1 + x_2$ and $p_2 = p(S_2) = x_1 + x_3$, and we are interested in the values of $p(S_1 \,\&\, S_2) = x_1$ and $p(S_1 \vee S_2) = x_1 + x_2 + x_3$. It is natural to assume that *a priori*, all probability distributions (i.e., all points in a simplex $\mathcal{S}$) are "equally possible", i.e., that there is a uniform distribution ("second-order probability") on this set of probability distributions. Then, as a natural estimate for the probability $p(S_1 \,\&\, S_2)$ of $S_1 \,\&\, S_2$, we can take the conditional mathematical expectation of this probability under the condition that the values $p(S_1) = p_1$ and $p(S_2) = p_2$:

$$E(p(S_1 \,\&\, S_2) \,|\, p(S_1) = p_1 \,\&\, p(S_2) = p_2) =$$
$$P(x_1 \,|\, x_1 + x_2 = p_1 \,\&\, x_1 + x_3 = p_2).$$

(This idea was proposed and described in [1, 4]; see also [2].)

From the geometric viewpoint, the two conditions $x_1 + x_2 = p_1$ and $x_1 + x_3 = p_2$ select a straight line segment within the simplex $\mathcal{S}$, a segment which can be parameterized by

$$x_1 \in [p^-, p^+] = [\max(p_1 + p_2 - 1, 0), \min(p_1, p_2)];$$

then, $x_2 = p_1 - x_1$, $x_3 = p_2 - x_1$, and $x_4 = 1 - (x_1 + x_2 + x_3)$. Since we start with a uniform distribution on $\mathcal{S}$, the conditional probability distribution on this segment is uniform, i.e., x_1 is uniformly distributed on the interval $[p^-, p^+]$. Thus, the conditional mathematical expectation of x_1 with respect to this distribution is equal to $(p^- + p^+)/2$, i.e., to the midpoint of this interval. Similarly, for an "or" operation, we can conclude that

$$E(p(S_1 \vee S_2) \,|\, p(S_1) = p_1 \,\&\, p(S_2) = p_2) =$$
$$\frac{1}{2} \cdot \max(p_1, p_2) + \frac{1}{2} \cdot \min(p_1 + p_2, 1).$$

4. Problem: Midpoint Operations Are Not Associative

Any "and" operation $p_1 \,\&\, p_2$ enables us to produce an estimate for $P(S_1 \,\&\, S_2)$ provided that we know estimates p_1 for $p(S_1)$ and p_2 for $p(S_2)$. If we are interested in estimating the degree of belief in a conjunction of three statements $S_1 \,\&\, S_2 \,\&\, S_3$, then we can use the same operation twice:

- first, we apply the "and" operation to p_1 and p_2 and get an estimate $p_1 \,\&\, p_2$ for the probability of $S_1 \,\&\, S_2$;
- then, we apply the "and" operation to this estimate $p_1 \,\&\, p_2$ and p_3, and get an estimate $(p_1 \,\&\, p_2) \,\&\, p_3$ for the probability of $(S_1 \,\&\, S_2) \,\&\, S_3$.

Alternatively, we can get start by combining S_2 and S_3, and get an estimate $p_1 \,\&\, (p_2 \,\&\, p_3)$ for the same probability $p(S_1 \,\&\, S_2 \,\&\, S_3)$. Intuitively, we would expect these two estimates to coincide: $(p_1 \,\&\, p_2) \,\&\, p_3 = p_1 \,\&\, (p_2 \,\&\, p_3)$, i.e., in algebraic terms, we expect the operation $\&$ to be associative. Unfortunately, midpoint operations are *not* associative [2]: e.g., $(0.4 \,\&\, 0.6) \,\&\, 0.8 = 0.2 \,\&\, 0.8 = 0.1$, while $0.4 \,\&\, (0.6 \,\&\, 0.8) = 0.4 \,\&\, 0.5 = 0.2 \neq 0.1$.

By itself, a small non-associativity may not be so bad:

- associativity comes from the requirement that our reasoning be rational, while
- it is well known that our actual handling of uncertainty is not exactly following rationality requirements; see, e.g., [14].

So, it is desirable to find out how non-associative can these operations be.

5. How Non-Associative Are Natural (Midpoint) Operations? Main Results and Their Psychological Interpretation

We know that the midpoint operations are non-associative, i.e., that sometimes, $(a \,\&\, b) \,\&\, c \neq a \,\&\, (b \,\&\, c)$. We want to know how big can the difference $(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)$ can be.

Theorem 1. $\max\limits_{a,b,c} |(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)| = 1/9.$

Theorem 2. $\max\limits_{a,b,c} |(a \vee b) \vee c - a \vee (b \vee c)| = 1/9.$

(The proof of these two theorems use *interval computations*; for readers' convenience, the proofs are placed in the special proofs section.)

Human experts do not use all the numbers from the interval $[0, 1]$ to describe their possible degrees of belief; they use a few words like "very probable", "mildly probable", etc. Each of words is a "granule" covering the entire sub-interval of values. Since the largest possible non-associativity degree $|(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)|$ is equal to 1/9, this non-associativity is negligible if the corresponding realistic "granular" degree of belief have granules of width $\geq 1/9$. One can fit no more than 9 granules of such width in the interval $[0, 1]$. This may explain why humans are most comfortable with ≤ 9 items to choose from – the famous "7 plus minus 2" law; see, e.g., [9, 10].

This general psychological law has also been confirmed in our specific area of formalizing expert knowledge: namely, in [3], it was shown that this law explains why in intelligent control, experts normally use ≤ 9 different degrees (such as "small", "medium", etc.) to describe the value of each characteristic.

6. Interval Computations: A Tool Used in Our Proof

To prove our results, we use a technique called *interval computations* (see, e.g., [6, 7, 15]). Namely, our goal is to find the maximum of the function $|(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)|$ when $a \in [0, 1]$, $b \in [0, 1]$, and $c \in [0, 1]$. We know that the minimum of this function is 0: it is attained, e.g., if $a = b = c = 0$. Thus, what we are looking for is the *range* of the above function of three real variables.

Interval computations is a technique which allows us, given a function $y = f(x_1, \ldots, x_n)$ of several real variables and a "box" $B = \mathbf{x}_1 \times \ldots \times \mathbf{x}_n$, where $\mathbf{x}_i = [x_i^-, x_i^+]$, to compute either the range of the given function on the given box:

$$\mathbf{y} = f(\mathbf{x}_1, \ldots, \mathbf{x}_n) =$$

$$\{f(x_1, \ldots, x_n) \mid x_1 \in [x_1^-, x_1^+], \ldots, x_n \in [x_n^-, x_n^+]\},$$

or an interval $\mathbf{Y}$ which is guaranteed to contain the desired range, i.e., for which $\mathbf{y} \subseteq \mathbf{Y}$ (We cannot always compute the *exact* range because computing this exact range is intractable even for quadratic functions $f(x_1, \ldots, x_n)$: see, e.g., [8].)

This technique is based on the fact that in the computer, the computation of a function f consists of several elementary steps. For example, a compiler will translate the computation of the midpoint "and" operation

$$f(p_1, p_2) = \frac{1}{2} \cdot \max(p_1 + p_2 - 1, 0) + \frac{1}{2} \cdot \min(p_1, p_2)$$

into the following sequence of elementary steps (r_1, r_2, etc. denote the preliminary computation results):

- first, we compute $r_1 := p_1 + p_2$;
- then, we compute $r_2 := r_1 - 1$;
- compute $r_3 := \min(r_2, 0)$;
- compute $r_4 := (1/2) \cdot r_3$;
- compute $r_5 := \min(p_1, p_2)$;
- compute $r_6 := (1/2) \cdot r_5$;
- finally, compute the result as $y := r_4 + r_6$.

In this example, we have two input variables $x_1 = p_1$ and $x_2 = p_2$. In general, for each input variable x_i, we know the interval $\mathbf{x}_i = [x_1^-, x_1^+]$ of possible values. For each elementary step $h(a, b)$, if we know the intervals $\mathbf{a} = [a^-, a^+]$ and $\mathbf{b} = [b^-, b^+]$ of possible values for each of the input, then we can compute the interval $h(\mathbf{a}, \mathbf{b})$ of possible values of the results:

- $[a^-, a^+] + [b^-, b^+] = [a^- + b^-, a^+ + b^+]$;
- $[a^-, a^+] - [b^-, b^+] = [a^- - b^+, a^+ - b^-]$;
- $[a^-, a^+] \cdot [b^-, b^+] = [c^-, c^+]$, where
  $c^- = \min(a^- \cdot b^-, a^- \cdot b^+, a^+ \cdot b^-, a^+ \cdot b^+)$,
  $c^+ = \max(a^- \cdot b^-, a^- \cdot b^+, a^+ \cdot b^-, a^+ \cdot b^+)$;

- $\min([a^-, a^+], [b^-, b^+]) =$ $[\min(a^-, b^-), \min(a^+, b^+)]$;
- $\max([a^-, a^+], [b^-, b^+]) =$ $[\max(a^-, b^-), \max(a^+, b^+)]$.

These formulas are called formulas of *interval arithmetic*.

So, to find an interval that contains the desired range, we follow the original algorithm step-by-step, on each step replacing the original elementary operation with real numbers by the corresponding operation of interval arithmetic.

In particular, if we want to know the range of the values of the function $f(p_1, p_2) = p_1 \& p_2$ when $p_1 \in \mathbf{p}_1$ and $p_2 \in \mathbf{p}_2$, we do the following:

- first, we compute $\mathbf{r}_1 := \mathbf{p}_1 + \mathbf{p}_2$;
- then, we compute $\mathbf{r}_2 := \mathbf{r}_1 - [1, 1]$;
- compute $\mathbf{r}_3 := \min(\mathbf{r}_2, [0, 0])$;
- compute $\mathbf{r}_4 := [0.5, 0.5] \cdot \mathbf{r}_3$;
- compute $\mathbf{r}_5 := \min(\mathbf{p}_1, \mathbf{p}_2)$;
- compute $\mathbf{r}_6 := [0.5, 0.5] \cdot \mathbf{r}_5$;
- finally, compute the result as $\mathbf{Y} := \mathbf{r}_4 + \mathbf{r}_6$.

It is easy to prove (by induction) that at any given moment of time, the result of this procedure is guaranteed to contain the result of the interval of possible values of the corresponding quantity.

It is also easy to show that this "naive" interval computation procedure sometimes overestimates. For example, for a function $f(x_1) = x_1 \cdot (1 - x_1)$ on the interval $[0, 1]$, the computational procedure consists of the following two steps:

- $r_1 := 1 - x_1$;
- $y := x_1 \cdot r_1$,

so we get the following estimate:

- $\mathbf{r}_1 := [1, 1] - \mathbf{x}_1 = [1, 1] - [0, 1] = [1 - 1, 1 - 0] = [0, 1]$;
- $\mathbf{Y} := \mathbf{x}_1 \cdot \mathbf{r}_1 = [0, 1] \cdot [0, 1] = [0, 1]$,

while the actual range is $\mathbf{y} = [0, 0.25] \subset \mathbf{Y} = [0, 1]$.

To decrease the overestimation, we can use the following methodology of interval computations: we divide each interval $\mathbf{x}_i$ into several sub-intervals, thus dividing the original box into many sub-boxes; then, we estimate the range of the function over each of the subintervals, and then take the union of the resulting ranges as an estimate for the range over the whole original box.

If we are interested not only in the actual value of the maximum, but if we also want to know where exactly this maximum is attained, then we can use this subboxes as follows: if we have two subboxes B_1 and B_2 with range estimates $[m_1, M_1]$ and $[m_2, M_2]$, and $M_1 < m_2$, then we are guaranteed that an arbitrary value $f(x_1, \ldots, x_n)$ for $(x_1, \ldots, x_n)$ from the first subbox is smaller than every value from the second subbox. Thus, we can safely claim that the (global) maximum of the given function cannot be attained in the first subbox – hence, this first subbox can be safely removed from the list of possible location of the global maximum.

7. Proof Itself

One can easily see that the operation $\vee$ is *dual* to the operation & in the sense that $a \vee b = 1 - (1 - a) \& (1 - b)$. Because of this duality, we can easily deduce Theorem 2 from Theorem 1. Thus, it is sufficient to prove Theorem 1.

Every triple can be sorted: $a \le b \le c$. For these sorted real numbers, we want to know the relation between $t_a \stackrel{\text{def}}{=} a \& (b \& c)$, $t_b \stackrel{\text{def}}{=} b \& (a \& c)$, and $t_c \stackrel{\text{def}}{=} c \& (a \& b)$. The formulas for $a \& b$, $a \& c$, and $b \& c$ depend on the relation between $a + b$, $a + c$, $b + c$, and 1. Since $a \le b \le c$, we have $a + b \le a + c \le b + c$. Thus, there are exactly 4 possible locations of number 1 in relation to these three sums:

I. $a + b \le a + c \le b + c \le \mathbf{1}$;

II. $a + b \le a + c \le \mathbf{1} < b + c$;

III. $a + b \le \mathbf{1} < a + c \le b + c$;

IV. $\mathbf{1} < a + b \le a + c \le b + c$.

We prove by considering these cases one by one. In each case, we get expressions for $a \& b$, $a \& c$, and $b \& c$ which do not contain min and max.

We can subdivide each of these cases into subcases depending on which of the maximized and minimized terms in the expressions for $a \& (b \& c)$, $b \& (a \& c)$, and $c \& (a \& b)$ are larger.

For example, in case IV, all three sums $a + b$, $a + c$, and $b + c$ are greater than 1, so $a \& b = a + 0.5 \cdot b - 0.5$, $a \& c = a + 0.5 \cdot c - 0.5$, and $b \& c = b + 0.5 \cdot c - 0.5$.

- The value of $t_a = (b \& c) \& a$ depends on whether $(b \& c) + a \le 1$, i.e., whether $b + 0.5 \cdot c - 0.5 + a \le 1$. If we move terms which do not contain a, b, or c

to the right hand-side, and rearrange terms which do contain a, b, or c, in alphabetic order, we get an equivalent inequality $a + b + 0.5 \cdot c \le 1.5$.

- Similarly, the value of $t_b = (a \,\&\, c) \,\&\, b$ depends on whether $(a \,\&\, c) + b \le 1$, i.e., whether $a + 0.5 \cdot c - (1 - \beta) + b \le 1$, which is also equivalent to the same inequality $a + b + 0.5 \cdot c \le 1.5$.

- Finally, the value of $t_c = (a \,\&\, b) \,\&\, c$ depends on whether $(a \,\&\, b) + c \le 1$, i.e., whether $a + 0.5 \cdot b - 0.5 + c \le 1$, which is equivalent to the inequality $a + 0.5 \cdot b + c \le 1.5$.

So, to find the expressions for t_a, t_b, and t_c, we must know where 1.5 stands in comparison with $a+b+0.5\cdot c$ and $a+0.5\cdot b+c$. Since $b \le c$, we have $0.5\cdot b \le 0.5\cdot c$, hence

$$a + b + 0.5 \cdot c = (a + b + c) - 0.5 \cdot c \le$$

$$(a + b + c) - 0.5 \cdot b = a + 0.5 \cdot b + c.$$

Due to this inequality, we have exactly three possibilities:

A. the number 1.5 can be larger than the largest of the above two expressions; in this case, both expressions are ≤ 1.5, i.e.,

$$a + b + 0.5 \cdot c \le a + 0.5 \cdot b + c \le 1.5;$$

B. the number 1.5 is in between the above two expressions; in this case,

$$a + b + 0.5 \cdot c \le 1.5 < a + 0.5 \cdot b + c;$$

C. the number 1.5 is smaller than the smallest of the above two expressions; in this case, both expressions are ≥ 1.5, i.e.,

$$1.5 < a + b + 0.5 \cdot c \le a + 0.5 \cdot b + c.$$

These subcases can be further subdivided, etc. For each of the resulting final subcases, all three combinations t_a, t_b, and t_c are described by linear expressions.

There are many such subcases, so the proof is possible but very lengthy. It turns out that interval computations can reduce this length.

Indeed, we want to find the maximum of the expression

$$|(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)|$$

when $a, b, c \in [0, 1]$. To help with the proof, we divided each interval $[0, 1]$ into 100 subintervals of length 0.01, thus generating $100^3 = 10^6$ sub-boxes. We use the above interval arithmetic (with additional operations for min and max). For each subbox, we applied the "naive" interval computations technique to get the estimate $[m_i, M_i]$ for the range of the desired function on this subbox. Then, we eliminated all subboxes for which $M_i < 1/9$.

As a result, out of the original 10^6 boxes, we have only 80 possible locations of the global maximum. For these 80 boxes, $b \in [0.54, 0.58]$, and:

- either $a \in [0.43, 0.46]$ and $c \in [0.75, 0.79]$;

- or $a \in [0.75, 0.79]$ and $c \in [0.43, 0.46]$.

When we sort a, b, and c, we get $a \in [0.43, 0.46]$, $b \in [0.54, 0.58]$, and $c \in [0.75, 0.79]$. Hence, $a + c > 1$, and we only need proofs for *half* of the cases: Cases III and IV.

Some subcases of Case IV were also eliminated. Indeed, within the above interval bounds for a, b, and c, the upper bound for $a + b + 0.5 \cdot c$ is equal to $0.46 + 0.58 + 0.5 \cdot 0.79 = 1.435 < 1.5$. Thus, to check that the value of the desired function cannot exceed 1/9, we only need to consider cases when $a+b+0.5\cdot c < 1.5$. Hence, we can dismiss Subcase C when this inequality is not satisfied, and only consider Subcases A and B in our proof.

For each final subcase, the difference

$$(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)$$

is a *linear* function, and the constraints describing this subcase are *linear* inequalities. Thus, for each subcase, we have a *linear programming* problem with rational coefficients. We can analytically solve each of these problems by computing the vertices of the corresponding polytope, and finding the vertex on which the objective function attains the largest value. As a result, we get the desired proof.

8. Auxiliary Results: Alternatives to Midpoint

Instead of selecting a midpoint, we can make a more general selection of a value in the interval $\mathbf{p}$.

By a *choice function*, we mean a function s that maps every interval $\mathbf{u} = [u^-, u^+]$ into a point $s(\mathbf{u}) \in \mathbf{u}$ so that for every c and $\lambda > 0$:

- $s([u^- + c, u^+ + c]) = s([u^-, u^+]) + c$ (*shift-invariance*);

- $s([\lambda \cdot u^-, \lambda \cdot u^+]) = \lambda \cdot s([u^-, u^+])$ (*unit-invariance*).

Proposition. [11] *Every choice function has the form* $s([u^-, u^+]) = \alpha \cdot u^- + (1-\alpha) \cdot u^+$ *for some* $\alpha \in [0,1]$.

The combination $p = \alpha \cdot p^- + (1-\alpha) \cdot p^+$ (first proposed by Hurwicz [5]) has been successfully used in areas ranging from submarine detection to petroleum engineering [11]; in [16], this approach is applied to second-order probabilities.)

With this approach, we get the following formulas which generalize the above definitions:

$$p_1 \,\&\, p_2 \stackrel{\text{def}}{=} \alpha \cdot \max(p_1 + p_2 - 1, 0) + (1-\alpha) \cdot \min(p_1, p_2);$$

$$p_1 \vee p_2 \stackrel{\text{def}}{=} \alpha \cdot \max(p_1, p_2) + (1-\alpha) \cdot \min(p_1 + p_2, 1).$$

Theorem 3.

$$\max_{a,b,c} |(a \,\&\, b) \,\&\, c - a \,\&\, (b \,\&\, c)| = \frac{\alpha \cdot (1-\alpha)}{2 + \alpha \cdot (1-\alpha)}.$$

$$\max_{a,b,c} |(a \vee b) \vee c - a \vee (b \vee c)| = \frac{\alpha \cdot (1-\alpha)}{2 + \alpha \cdot (1-\alpha)}.$$

Comment. This non-associativity degree is the smallest ($= 0$) when $\alpha = 0$ or $\alpha = 1$, and the largest ($= 1/9$) for midpoint operations ($\alpha = 0.5$).

In our proof, it was useful to first show that the new operations have *some* properties of associativity: namely, it turns out that for every α, both operations are *semi-associative* in the sense that $a \le b \le c$ implies that $a * (b * c) \ge b * (a * c) \ge c * (a * b)$.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, by Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grant number F49620-00-1-0365, and by Grant No. W-00016 from the U.S.-Czech Science and Technology Joint Fund.

The authors are very thankful to H.T. Nguyen and T. Whalen for helpful discussions.

References

[1] D. Bamber, "Entailment with near surety of scaled assertions of high conditional probability", *Journal of Philosophical Logic*, 2000 (to appear).

[2] A.D.C. Bennett, J.B. Paris, and A. Vencovská, "A new criterion for comparing fuzzy logics for uncertain reasoning", *Journal of Logic, Language, and Information*, 2000, Vol. 9, pp. 31–63.

[3] L. Godo, R. Lopez de Mantaras, C. Sierra, and A. Verdaguer, "MILORD: The Architecture and management of Linguistically expressed Uncertainty", *International Journal of Intelligent Systems*, 1989, Vol. 4, pp. 471–501.

[4] I.R. Goodman and H.T. Nguyen, "Probability updating using second-order probabilities and conditional event algebra", *Information Sciences*, 2000.

[5] L. Hurwicz, *A criterion for decision-making under uncertainty*, Technical Report 355, Cowles Commission, 1952.

[6] R.B. Kearfott, *Rigorous global search: continuous problems*, Kluwer, Dordrecht, 1996.

[7] R.B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

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

[9] G.A. Miller, "The magical number seven plus or minus two: some limits on our capacity for processing information", *Psychological Review*, 1956, Vol. 63, pp. 81–97.

[10] P.M. Milner, *Physiological psychology*, Holt, NY, 1970.

[11] H.T. Nguyen and V. Kreinovich, "Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications", In [7], pp. 245–290.

[12] H.T. Nguyen, V. Kreinovich, and Q. Zuo, "Interval-valued degrees of belief: applications of interval computations to expert systems and intelligent control", *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems (IJUFKS)*, 1997, Vol. 5, No. 3, pp. 317–358.

[13] N.J. Nilsson, "Probabilistic logic", *Artificial Intelligence*, 1996, Vol. 28, pp. 71–87.

[14] P. Suppes, D.M. Krantz, R.D. Luce, and A. Tversky, *Foundations of measurement*, Vol. I-III, Academic Press, San Diego, CA, 1989.

[15] Website on interval computations: http://www.cs.utep.edu/interval-comp.

[16] T. Whalen, "Interval probabilities induced by decision problems", In: R.R. Yager, J. Kacprzyk, and M. Pedrizzi (Eds.), *Advances in the Dempster-Shafer Theory of Evidence*, Wiley, N.Y., 1994, pp. 353–374.