Probability of Implication, Logical Version of Bayes Theorem, and Fuzzy Logic Operations

Hung T. Nguyen[1], Masao Mukaidono[2], and Vladik Kreinovich[3]
[1]Mathematics, New Mexico State University, Las Cruces, NM 88003, USA
[2]Comp. Sci., Meiji University, Kanagawa-Ken, Japan
[3]Comp. Sci., University of Texas, El Paso, TX 79968, USA
contact email vladik@cs.utep.edu

***Abstract*— Logical inference starts with concluding that if B implies A, and B is true, then A is true as well. To describe probabilistic inference rules, we must therefore define the probability of an implication "A if B". There exist two different approaches to defining this probability, and these approaches lead to different probabilistic inference rules: We may interpret the probability of an implication as the conditional probability $P(A\,|\,B)$, in which case we get Bayesian inference. We may also interpret this probability as the probability of the material implication $A \vee \neg B$, in which case we get different inference rules. In this paper, we develop a general approach to describing the probability of an implication, and we describe the corresponding general formulas, of which Bayesian and material implications are particular cases. This general approach is naturally formulated in terms of t-norms, a terms which is normally encountered in fuzzy logic.**

I. INTRODUCTION

Intuitively, when we say that an implication "A if B" ($A \leftarrow B$) is true, we mean that whenever B is true, we can therefore conclude that A is true as well. In other words, implication is what enables us to perform logical inference.

In many practical situations, we have some confidence in B, but we are not 100% confident that B is true. Similarly, we may not be 100% sure that the implication $A \leftarrow B$ is true. In such situations, we can estimate the probability $P(B)$ that B is true, and the probability $P(A \leftarrow B)$ that an implication $A \leftarrow B$ is true. How can we perform logical inference in such situations? Intuitively, we expect to be able to conclude that in this case, A should also be true with a certain probability; this probability should tend to 1 as the probabilities $P(B)$ and $P(A \leftarrow B)$ tend to 1.

How can we extend logical implication to the probabilistic case? Depending on how we interpret the probability of an implication, we get two different

There are two known answers to this question, and these answers are different because they use different formalizations of the probability of implication. The first answer from Bayesian approach, in which $P(A \leftarrow B)$ is interpreted as the conditional probability $P(A\,|\,B)$; see, e.g., [7]. The second answer comes from logical reasoning (see, e.g., [5]), where the probability $P(A \leftarrow B)$ is interpreted as the probability of the corresponding "material implication", i.e., the probability $P(A \vee \neg B)$ that either A is true or B is false.

From the purely logical viewpoint, the second answer may sound more reasonable, but there are examples where the first answer is in better accordance with common sense. Indeed, suppose that we are analyzing animals in a national park, and we are looking for a probability of the implication $A \leftarrow B$, where A is "the animal is a white", and B is "the animal is a tiger". In plain English, the probability of the statement "A if B" is naturally interpreted as a probability that a tiger is white. If out of 10,000 animals, 100 are tigers, and 10 of these tigers are white, then, in commonsense terms, the probability that a tiger is white is 10/100=0.1. This is exactly the probability provided by the Bayesian approach. However, the logical approach produces a different result: the probability of $A \vee \neg B$, i.e., the probability that an animal is either white or not a tiger is equal to 9,910/10,000=0.991 – because the statement $A \vee \neg B$ is true not only for 10 white tigers, but also for 9,900 animals which are not tigers.

This examples show that there is not a single "correct" probabilistic interpretation of an implication, but depending on the situation, different interpretations may be reasonable. It is therefore desirable to provide a comparative analysis of different interpretations.

In [3, 6], it was shown that the above two interpretations can be presented as particular cases of a more general approach, in which the difference corresponds to the difference between different t-norm-like operations (for detailed information on t-norms, see, e.g., [2, 4]).

In this paper, we describe this general approach in precise terms, and we describe all possible implication operations covered by this approach and the corresponding logical inference rules. Specifically, in Section II, we overview the main properties of Bayes formalism, in Section III, we overview how logical implication can be described in similar terms, and in Section IV, we describe the corresponding general approach to probability of implication. Several auxiliary results are presented in Section V. The main ideas of the proofs of the results from Section IV are described in Section VI.

II. BAYESIAN APPROACH: A BRIEF REMINDER

In Bayesian approach, we interpret the probability of an implication as the conditional probability

$$P(A\,|\,B)=\frac{P(A\,\&\,B)}{P(B)}. \quad (1)$$

Due to this formula, if we know the probability $P(B)$ of B and the probability $P(A\,|\,B)$ of the implication, then we can reconstruct the probability $P(A\,\&\,B)$ that both A and B are true as follows:

$$P(A\,\&\,B)=P(A\,|\,B)\cdot P(B). \quad (2)$$

Since $A\,\&\,B$ means the same thing as $B\,\&\,A$, we therefore conclude that $P(A\,\&\,B)=P(B\,\&\,A)$, i.e., due to (1), that

$$P(A\,|\,B)\cdot P(B)=P(B\,|\,A)\cdot P(A). \quad (3)$$

This formula is the essence of the well-known Bayes theorem. In this theorem, we have a comprehensive list of n incompatible hypotheses $H_1,\ldots,H_n$, and we know the (prior) probabilities $P(H_1),\ldots,P(H_n)$ of these hypotheses. Since these hypothesis cover all all possible situations and are incompatible, we conclude that

$$P(H_1)+\ldots+P(H_n)=1. \quad (4)$$

We want to know how these prior probabilities change when we make observe some evidence E.

We assume that, for each of these hypotheses H_i, we know the conditional probability $P(E\,|\,H_i)$ that under this hypothesis, we will observe the evidence E. What we want to describe is the updated probability $P(H_i\,|\,E)$ with which the hypothesis H_i is true in the situation when the evidence E was actually observed. According to the formula (3),

$$P(H_i\,|\,E)\cdot P(E)=P(E\,|\,H_i)\cdot P(H_i), \quad (5)$$

therefore,

$$P(H_i\,|\,E)=\frac{P(E\,|\,H_i)\cdot P(H_i)}{P(E)}. \quad (6)$$

So, to determine the desired posterior probability $P(H_i\,|\,E)$, we must know $P(E\,|\,H_i)$, $P(H_i)$, and $P(E)$. We know $P(E\,|\,H_i)$ and $P(H_i)$. The only value that we do not know yet is $P(E)$, but this value is easy to determine: since the hypotheses are incompatible, and their list is comprehensive, we conclude that

$$P(E)=P(E\,\&\,H_1)+\ldots+P(E\,\&\,H_n). \quad (7)$$

Due to formula (1), we have

$$P(E\,\&\,H_i)=P(E\,|\,H_i)\cdot P(H_i),$$

hence

$$P(E)=P(E\,|\,H_1)\cdot P(H_1)+\ldots+P(E\,|\,H_n)\cdot P(H_n), \quad (8)$$

and therefore, the formula (6) take the familiar Bayes form:

$$P(H_i\,|\,E)=\frac{P(E\,|\,H_i)\cdot P(H_i)}{P(E\,|\,H_1)\cdot P(H_1)+\ldots+P(E\,|\,H_n)\cdot P(H_n)}. \quad (9)$$

III. LOGICAL APPROACH REFORMULATED IN SIMILAR TERMS

Assume now that we interpret implication as the material implication $A\vee\neg B$. In this case, the probability of the implication is interpreted as the probability $P(A\vee\neg B)$. If we know this probability and if we know $P(B)$, how can we determine $P(A\,\&\,B)$? Based on the additivity of the probability and the fact that A and $\neg A$ are incompatible, we conclude that $P(B)=P(A\,\&\,B)+P(\neg A\,\&\,B)$. Therefore,

$$P(A\,\&\,B)=P(B)-P(\neg A\,\&\,B). \quad (10)$$

The statement $\neg A\,\&\,B$ is the negation of $A\vee\neg B$, hence

$$P(\neg A\,\&\,B)=1-P(A\vee\neg B). \quad (11)$$

Substituting (11) into (10), we conclude that

$$P(A\,\&\,B)=P(B)+P(A\vee\neg B)-1. \quad (12)$$

This formula is similar to the formula (2): both formulas can be described as

$$P(A\,\&\,B)=P(A\leftarrow B)\odot P(B) \quad (13)$$

for some binary operation $a\odot b$. In the formula (2) – which corresponds to the Bayesian case – we used the function

$$a\odot b=a\cdot b. \quad (14)$$

In the formula (12) – which corresponds to the logical implication cases – we used the operation $a\odot b=a+b-1$.

Since the meaning of the operation $\odot$ is to transform probabilities into a new probability, and probabilities only take values from the interval $[0,1]$, it is reasonable to require that the operation $a\odot b$ always takes the values from the interval $[0,1]$. The operation $a\odot b=a+b-1$ does not always satisfy this requirement, because when a and b are both, say, less than 0.5, we have $a+b-1<0$. This does not affect our application because we always have $P(B)+P(A\vee\neg B)\geq 1$. However, to make the operation $a\odot b$ everywhere defined as a function from probabilities to probabilities, it is reasonable to set its value to 0 when $a+b-1<0$, i.e., to consider a new operation

$$a\odot b=\max(0,a+b-1). \quad (15)$$

Both operations (14) and (15) are examples of *t-norms*, operations describing "and" in fuzzy logic. Informally, the appearance of a t-norm makes sense because $A \,\&\, B$ is true if B is true *and* the implication $A \to B$ is true, so it is reasonable to conclude that our degree of belief $P(A \,\&\, B)$ in $A \,\&\, B$ is equal to the result of an "and"-operation (t-norm) $a \odot b$ applied to the degrees of belief $P(B)$ that B is true and the degree of belief $P(A \leftarrow B)$ that the implication is true. This justification is informal. In the following text, we will make a more formal justification.

Meanwhile, since the formula (12) is similar to the formula (2), we will use this analogy to deduce the logical inference analogue of the Bayes formula. Since $A \,\&\, B$ means the same thing as $B \,\&\, A$, we therefore conclude that $P(A \,\&\, B) = P(B \,\&\, A)$, i.e., due to (13), that

$$P(A \leftarrow B) \odot P(B) = P(B \leftarrow A) \odot P(A). \tag{16}$$

In particular, for the exhaustive list of n incompatible hypotheses $H_1, \ldots, H_n$, and for an evidence E, we conclude that

$$P(H_i \leftarrow E) \odot P(E) = P(E \leftarrow H_i) \odot P(H_i). \tag{17}$$

Therefore,

$$P(H_i \leftarrow E) = (P(E \leftarrow H_i) \odot P(H_i)) \oslash P(E), \tag{18}$$

where $a \oslash b$ is the inverse operation to $\odot$, i.e., an operation for which $(a \oslash b) \odot b = a$.

It is worth mentioning that for a general t-norm $\odot$, the corresponding inverse operation $\oslash$ is usually called a fuzzy implication [2, 4].

For multiplication (14), the inverse operation is division $a \oslash b = a/b$ (as used in the formula (6)). To make sure that the values of this operation stays within the interval $[0,1]$, we should replace it with

$$a \oslash b = \min(a/b, 1).$$

For our particular operation (15), the inverse operation is

$$a \oslash b = \min(1 + a - b, 1);$$

here, similarly to the case of division, we added $\min(1, \ldots)$ to make sure that the value of this operation stays within the interval $[0,1]$.

Due to formula (7) and the fact that

$$P(E \,\&\, H_i) = P(E \leftarrow H_i) \odot P(H_i),$$

we get an expression for $P(E)$:

$$P(E) = P(E \leftarrow H_1) \odot P(H_1) + \ldots + P(E \leftarrow H_n) \odot P(H_n).$$

So, we conclude that

$$P(H_i \leftarrow E) = (P(E \leftarrow H_i) \odot P(H_i)) \oslash \tag{19}$$

$$(P(E \leftarrow H_1) \odot P(H_1) + \ldots + P(E \leftarrow H_n) \odot P(H_n)).$$

This is a direct logical analogue of the Bayes theorem.

IV. GENERAL APPROACH TO DESCRIBING PROBABILITY OF AN IMPLICATION AND ITS RELATION WITH FUZZY LOGIC

A. General Definition

Let us describe a general definition of the probability $P(A \leftarrow B)$. This probability should only depend on the events A and B. Thus, our first requirement is that once we know the probabilities of all possible Boolean combinations A and B, we should be able to determine the desired probability $P(A \leftarrow B)$.

It is well known that in order to determine the probabilities of all possible Boolean combinations of A and B, it is sufficient to know the probabilities $P(A \,\&\, B)$, $P(A \,\&\, \neg B)$, $P(\neg A \,\&\, B)$, and $P(\neg A \,\&\, \neg B)$ of four *atomic* statements $A \,\&\, B$, $A \,\&\, \neg B$, $\neg A \,\&\, B$, and $\neg A \,\&\, \neg B$. For simplicity, in the following text, we will denote the corresponding probabilities by P_{11}, P_{10}, P_{01}, and P_{00}. The corresponding four atomic statements make up a comprehensive list of incompatible events, so their sum should be equal to 1:

$$P(A \,\&\, B) + P(A \,\&\, \neg B) + P(\neg A \,\&\, B) + P(\neg A \,\&\, \neg B) = 1.$$

Thus, we can define a general implication operation as a function of these four probabilities:

Definition 1. *By a probability distribution P, we mean a quadruple of non-negative values P_{11}, P_{10}, P_{01}, and P_{00} for which $P_{11} + P_{10} + P_{01} + P_{00} = 1$. The set of all probability distributions is denoted by $\mathcal{P}$.*

Definition 2. *By an probabilistic logical operation, we mean a function $F : \mathcal{P} \to [0,1]$. For every two events A and B, the result of applying the probabilistic logical operation F is then defined as*

$$F(P) \stackrel{\text{def}}{=}$$

$$F(P(A \,\&\, B), P(A \,\&\, \neg B), P(\neg A \,\&\, B), P(\neg A \,\&\, \neg B)).$$

For the Bayesian definition (1), we have $P(A \,\&\, B) = P_{11}$, and $P(B) = P(A \,\&\, B) + P(\neg A \,\&\, B) = P_{11} + P_{01}$, hence

$$F(P_{11}, P_{10}, P_{01}, P_{11}) = \frac{P_{11}}{P_{11} + P_{01}}. \tag{20}$$

For the logical definition $P(A \leftarrow B) \stackrel{\text{def}}{=} P(A \vee \neg B)$, we have (due to (11)) $P(A \vee \neg B) = 1 - P(\neg A \,\&\, B) = 1 - P_{01}$, hence

$$F(P_{11}, P_{10}, P_{01}, P_{11}) = 1 - P_{01}. \tag{21}$$

B. What Does It Mean to be an Implication?

Definition 1 is a general definition of a probabilistic logical operation, it does not distinguish between implication, conjunction, disjunction, etc. What makes an operation an implication operation? One thing that is true for implication $A \leftarrow B$ and not for other operations is that the implication depends only on what happens when B is true and should not be affected by what happens when B is false. In other words, if for two distributions, we have the same values of $P(A\,\&\,B)$ and $P(\neg A\,\&\,B)$, then for these two distributions, we should get exactly the same value of $P(A \leftarrow B)$.

Another condition describing implication is that if when B always implies A, i.e., when $\neg A\,\&\,B$ is impossible (i.e., when $P(\neg A\,\&\,B) = P_{01} = 0$), then $A \leftarrow B$ must be true with probability 1.

Let us describe these conditions formally:

Definition 3. *We say that two probability distributions P and P' are equivalent when B is true if $P_{11} = P'_{11}$ and $P_{01} = P'_{01}$.*

Definition 4. *We say that a probabilistic logical operation F is an implication operation if the following two conditions hold:*

- *$F(P) = F(P')$ for all pairs P and P' that are equivalent when B is true;*
- *if $P_{01} = 0$, then $F(P) = 1$.*

Both operations (20) and (21) are implication operations in this sense. In general, the following simple propositions provides a complete descriptions of such implication operations:

Proposition 1. *A probabilistic logical operation F is an implication operation if and only if F depends on only two variables P_{11} and P_{01}, i.e., if $F(P) = f(P_{11}, P_{01})$ for some function f of two variables for which $f(P_{11}, 0) = 1$ for all values P_{11}.*

Since

$$P_{01} = P(\neg A\,\&\,B) = P(B) - P(A\,\&\,B) = P_{*1} - P_{11},$$

where we denoted $P_{*1} \stackrel{\text{def}}{=} P(B)$, we can reformulated Proposition 1 as follows:

Proposition 1′. *A probabilistic logical operation F is an implication operation if and only if F depends on only two variables P_{11} and P_{*1}, i.e., if $F(P) = g(P_{11}, P_{*1})$ for some function g of two variables for which $g(P_{11}, P_{11}) = 1$ for all values P_{11}.*

Thus, to describe all possible implication operations, we must describe the corresponding functions of two variables.

C. Natural Implication Operations

Since we are considering the probabilistic uncertainty, it is reasonable to consider not only individual events A, A', etc., but also "composites" (probabilistic combinations) of such events. The general idea behind such combinations is that we take a lottery with a certain probability p and then pick A if the lottery succeeds and A' otherwise. According to the probability theory, the probability of the resulting event $\widetilde{A}$ is equal to

$$P(\widetilde{A}) = p \cdot P(A) + (1-p) \cdot P(A'). \tag{22}$$

It is also true that

$$P(\widetilde{A}\,\&\,B) = p \cdot P(A\,\&\,B) + (1-p) \cdot P(A'\,\&\,B). \tag{23}$$

It is natural to require that in this case, if we keep the same condition B, then the probability of an implication with the conclusion $\widetilde{A}$ should also be equal to the similar probabilistic combination:

$$P(\widetilde{A} \leftarrow B) = p \cdot P(A \leftarrow B) + (1-p) \cdot P(A' \leftarrow B). \tag{24}$$

This requirements can be formulated as follows:

Definition 5. *An implication operation $g(P_{11}, P_{*1})$ is called natural if for every four real numbers P_{11}, P'_{11}, P_{*1}, and p, we have*

$$g(p \cdot P_{11} + (1-p) \cdot P'_{11}, P_{*1}) =$$
$$p \cdot g(P_{11}, P_{*1}) + (1-p) \cdot g(P'_{11}, P_{*1}). \tag{25}$$

Proposition 2. *An implication operation is natural if and only if it has the form*

$$g(P_{11}, P_{*1}) = 1 - \frac{P_{*1} - P_{11}}{h(P_{*1})} \tag{26}$$

for some function $h(P_{*1})$ of one variable.

Both formulas (20) and (21) can be thus represented: the Bayes case corresponds to $h(z) = z$, and the logical case corresponds to $h(z) = 1$.

D. Final Result: Natural Implication Operations Corresponding to Commutative Aggregation Rule

Due to Proposition 2, for each natural implication g, if we know $a \stackrel{\text{def}}{=} P(A \leftarrow B) = g(P_{11}, P_{*1})$ and $b \stackrel{\text{def}}{=} P(B) = P_{*1}$, then we can reconstruct the probability $t \stackrel{\text{def}}{=} P(A\,\&\,B) = P_{11}$. Indeed, in terms of a, b, and t, the formula (26) has the form

$$a = 1 - \frac{b-t}{h(b)};$$

hence

$$1 - a = \frac{b-t}{h(b)};$$

so $(1-a) \cdot h(b) = b - t$ and therefore,

$$P(A\,\&\,B) = t(P(A \leftarrow B), P(B)), \tag{27}$$

where we denoted

$$t(a,b) = b - (1-a) \cdot h(b). \tag{28}$$

The function $t(a,b)$ describe an aggregation operation whose intuitive meaning (as we mentioned earlier) is "and". Since "A and B" means the same as "B and A", it is reasonable to require that this aggregation operation be commutative:

Definition 6. *We say that a natural implication operation* (26) *corresponds to a commutative aggregation rule if the corresponding aggregation operation* (28) *if commutative.*

Proposition 3. *A natural implication corresponds to a commutative aggregation rule if and only if has the following form:*

$$P(A \leftarrow B) = 1 - \frac{P(\neg A \,\&\, B)}{\alpha + (1-\alpha)\cdot P(B)}. \quad (29)$$

E. Conclusions and Discussions

Both the Bayes formula and the logical implication are covered by the formula (29):

- the Bayes formula corresponds to $\alpha = 0$, and
- the logical implication formula corresponds to $\alpha = 1$.

Thus, our conclusion is that natural requirements determine a 1-parametric family of formulas for the probability of implication, formulas which are intermediate between the two extreme cases: Bayesian and logical.

In general, for $h(b) = \alpha + (1-\alpha)\cdot b$, the corresponding aggregation operation (28) has the form

$$t(a,b) = a \odot b = \alpha \cdot (a+b-1) + (1-\alpha)\cdot a \cdot b. \quad (30)$$

One can easily check that this operation is always associative.

For this operation $\odot$, the inverse operation $c = a \oslash b$ can be determined from the equation $a = b \odot c$, i.e., $a = \alpha\cdot(b+c-1) + (1-\alpha)\cdot b\cdot c$. This is a linear equation in terms of c, from which we conclude that

$$a \oslash b = \frac{a + \alpha\cdot(1-b)}{\alpha + (1-\alpha)\cdot b}. \quad (31)$$

One can easily check that for $\alpha = 0$, we get the Bayes' inverse a/b, and for $\alpha = 1$, we get the inverse operation $1 + a - b$ corresponding to the logical implication.

How can we interpret these new aggregation operations? One can show that if we "renormalize" the probabilities by using a transformation

$$P \to s(P) = (1-\alpha)\cdot P + \alpha, \quad (31)$$

then in this new scale, the aggregation operation (30) becomes a simple product: $s(a\odot b) = s(a)\cdot s(b)$.

The rescaling (31) makes perfect sense (see, e.g., [1]). Indeed, one of the natural methods to ascribe the subjective probability $P(A)$ to a statement A is to take several (N) experts, and ask each of them whether he or she believes that A is true. If $N(A)$ of them answer "yes", we take $d(A) = N(A)/N$ as the desired certainty value. If all the experts believe in A, then this value is 1 (=100%), if half of them believe in A, then $t(A) = 0.5$ (50%), etc.

Knowledge engineers want the system to include the knowledge of the entire scientific community, so they ask as many experts as possible. But asking too many experts leads to the following negative phenomenon: when the opinion of the most respected professors, Nobel-prize winners, etc., is known, some less self-confident experts will not be brave enough to express their own opinions, so they will rather follow the majority. How does their presence influence the resulting subjective probability?

Let N denote the initial number of experts, $N(A)$ the number of those of them who believe in A, and M the number of shy experts added. Initially, $d(A) = N(A)/N$. After we add M conformist experts, the number of experts who believe in A becomes $N(A)+M$ out of the total of $M+N$. So the new value of the subjective probability is

$$P'(A) = \frac{N(A)+M}{N+M} = (1-\alpha)\cdot P(A) + \alpha,$$

where we denoted $\alpha \stackrel{\text{def}}{=} M/N$.

Thus, each new operation can be simply interpreted as the Bayesian operation but in a different probability scale natural for expert systems.

V. AUXILIARY RESULT

In the previous text, we described how, knowing $P(B)$ and $P(A \leftarrow B)$, we can reconstruct the probability $P(A \,\&\, B)$ as $P(a \leftarrow B) \odot P(B)$. A natural question is: what if we want to reconstruct the probability $P(A)$ instead?

Since we know $A(B)$ and $P(A \,\&\, B)$, this means that we know, out of all cases in which B is true, in what portion of them A is also true. To get the probability of A, we must also know when A is true for cases in which B is false, i.e., we must also know the probability $P(A \,\&\, \neg B)$. Then, we will be able to reconstruct the total probability $P(A)$ of A as

$$P(A) = P(A \,\&\, B) + P(A \,\&\, \neg B). \quad (32)$$

By definition of an implication operation, we do not have any information on whether A is true or not in cases when B is false.

- It may be that A is always true when B is false. In this case, $P(A \,\&\neg B) = P(\neg B) = 1 - P(B)$.
- It may also happen that A is never true when B is false. In this case, $P(A \,\&\, \neg B) = 0$.
- It is also possible for $P(A \,\&\, \neg B)$ to take any value from the interval $[0, 1-P(B)]$.

Thus, due to formula (B1), the only information that we have about the probability $P(A)$ is that this probability belongs to the interval

$$P(A) \in \mathbf{P}(A) \stackrel{\text{def}}{=} [P^-(A), P^+(A)],$$

where:

$$P^-(A) = P(A \leftarrow B) \odot P(B);$$
$$P^+(A) = P(A \leftarrow B) \odot P(B) + 1 - P(B). \quad (33)$$

In particular, for the Bayes implication, we have

$$P^-(A) = P(A \leftarrow B) \cdot P(B);$$

$$P^+(A) = P(A \leftarrow B) \cdot P(B) + 1 - P(B). \quad (34)$$

For the logical implication, we have

$$P^-(A) = P(A \leftarrow B) + P(B) - 1;$$

$$P^+(A) = P(A \leftarrow B). \quad (35)$$

For the general implication operation, we have

$$P^-(A) = \alpha \cdot (P(A \leftarrow B) + P(B) - 1) +$$

$$(1 - \alpha) \cdot P(A \leftarrow B) \cdot P(B);$$

$$P^+(A) = \alpha \cdot P(A \leftarrow B) + (1 - \alpha) \cdot (1 - P(B)) +$$

$$(1 - \alpha) \cdot P(A \leftarrow B) \cdot P(B). \quad (36)$$

VI. PROOFS: MAIN IDEAS

A. Proof of Proposition 2

In the formula (25), let us take $P_{11} = 1$ and $P'_{11} = 0$. Then, this formula turns into the following one:

$$g(p, P_{*1}) = p \cdot g(1, P_{*1}) + (1 - p) \cdot g(0, P_{*1}). \quad (37)$$

Thus, for $p = P_{11}$, we have

$$g(P_{11}, P_{*1}) = P_{11} \cdot g(1, P_{*1}) + (1 - P_{11}) \cdot g(0, P_{*1}). \quad (38)$$

Substituting $P_{11} = P_{*1} - P_{01}$ into the formula (38), we conclude that

$$g(P_{11}, P_{*1}) =$$

$$(P_{*1} - P_{01}) \cdot g(1, P_{*1}) + (1 - P_{*1} + P_{01}) \cdot g(0, P_{*1}). \quad (39)$$

Combining together terms proportional to P_{01}, we conclude that

$$g(P_{11}, P_{*1}) = g_1(P_{*1}) - P_{01} \cdot g_2(P_{*1}), \quad (40)$$

where we denoted

$$g_1(P_{*1}) \stackrel{\text{def}}{=} P_{*1} \cdot g(1, P_{*1}) + (1 - P_{*1}) \cdot g(0, P_{*1});$$

$$g_2(P_{*1}) \stackrel{\text{def}}{=} g(1, P_{*1}) - g(0, P_{*1}).$$

By definition of an implication operation, its result is 1 if $P_{01} = 0$. Thus, in the formula (40), we have $g_1(P_{*1}) = 1$, so

$$g(P_{11}, P_{*1}) = 1 - P_{01} \cdot g_2(P_{*1}). \quad (41)$$

If we denote $1/g_2(z)$ by $h(z)$, then, since $P_{01} = P_{*1} - P_{11}$, we get the desired formula (26). The proposition is proven.

B. Proof of Proposition 3

Commutativity $t(a, b) = t(b, a)$ of the operation (27) means that for every two real numbers a and b, we have

$$b - (1 - a) \cdot h(b) = a - (1 - b) \cdot h(a). \quad (42)$$

In particular, for $a = 0$, this equality leads to

$$b - h(b) = -(1 - b) \cdot h(0),$$

hence

$$h(b) = b + \alpha \cdot (1 - b),$$

where we denoted $\alpha \stackrel{\text{def}}{=} h(0)$. Combining terms proportional to b, we conclude that

$$h(b) = \alpha + (1 - \alpha) \cdot b.$$

Substituting this expression into (26), we get the desired formula. The proposition is proven.

ACKNOWLEDGMENTS

This work was supported in part by NASA under cooperative agreement NCC5-209 and grant NCC 2-1232, by the Air Force Office of Scientific Research grants F49620-95-1-0518 and F49620-00-1-0365, by grant No. W-00016 from the U.S.-Czech Science and Technology Joint Fund, and by NSF grants CDA-9522207, ERA-0112968, and 9710940 Mexico/Conacyt.

REFERENCES

[1] B. Bouchon-Meunier, V. Kreinovich, A. Lokshin, and H. T. Nguyen, "On the formulation of optimization under elastic constraints (with control in mind)", *Fuzzy Sets and Systems*, 1996, Vol. 81, No. 1, pp. 5–29.

[2] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ, 1995.

[3] M. Mukaidono and Y. Yamauchi, "Logical version of Bayes' theorem", *Proceedings of the Second Vietnam-Japan Bilateral Symposium on Fuzzy Systems and Applications VJFUZZY'2001*, Hanoi, Vietnam, December 7–8, 2001, pp. 17–21.

[4] H. T. Nguyen and E. A. Walker, *First Course in Fuzzy Logic*, CRC Press, Boca Raton, FL, 1999.

[5] N. J. Nilsson, "Probabilistic logic", *Artificial Intelligence*, 1986, Vol. 28, No. 1, pp. 71–78.

[6] Y. Yamauchi and M. Mukaidono, "probabilistic inference and Bayesian theorem based on logical implication", In: N. Zhong, A. Skowron, and S. Ohsuga (eds.), *New Directions in Rough Sets, Data Mining, and Granular-Soft Computing*, Springer-Verlag Lecture Notes in Artificial Intelligence, Vol. 1711, 1999, pp. 334–342.

[7] H. M. Wadsworth, Jr. (eds.), *Handbook of statistical methods for engineers and scientists*, McGraw-Hill Publishing Co., New York, 1990.