Probabilistic Implications

Przemysław Grzegorzewski

Systems Research Institute, Polish Academy of Sciences,
Newelska 6, 01-447 Warsaw, Poland,
and Faculty of Mathematics and Information Science,
Warsaw University of Technology,
Plac Politechniki 1, 00-661 Warsaw, Poland
pgrzeg@ibspan.waw.pl

Abstract

A new family of implication operators, called probabilistic implications, is introduced. The suggested implications are based on conditional copulas and make a bridge between probability theory and fuzzy logic. It is shown that some well-known fuzzy implications appear as a particular probabilistic implications. Therefore, it seems that this new family of implication operators might be a useful tool in approximate reasoning.

Keywords: fuzzy implication, copula, conditional probability, approximate reasoning

1. Introduction

Our knowledge of complex systems is often incomplete. It is due not only to general complexity of these systems but often is also entailed by the deficiency of our measuring instruments, including our senses. Therefore, we have to rely rather on perceptions than on precise data and on imperfect knowledge delivered by expert's statements usually formulated not in mathematical terms but in natural language. Fuzzy logic provides methodology for modelling imprecise data and make possible computation with words (see [10]).

Since a large part of expert knowledge consists of if-then rules it has been necessary to develop and formalize such if-then statements in fuzzy logic. As a result we have obtained a broad theory of fuzzy implications (see [1]).

However, one has to be aware that uncertainty appears in many systems not only because of imprecision but is rather an immanent effect of randomness. Moreover, in most cases we have to cope with imperfect knowledge which abounds with both kinds of uncertainty: imprecision and randomness. Hence our inference based on if-then rules should also comprise with these two sources of uncertainty. Therefore, another type of implication that takes into consideration both imprecision modelled by fuzzy concepts and randomness described by tools originated in probability theory would be desirable. And a construction of such concept is suggested in this very contribution.

The paper is organized as follows. In Sec. 2 we recall some information on fuzzy implications. Some remarks how to combine probability and implication are given in Sec. 3. Next in Sec. 4 we recall some information on copulas. Then, in Sec. 5 we propose the notion of probabilistic implication and explore its basic properties. Finally we show examples of probabilistic implications and indicate some open problems and directions for further research.

2. Fuzzy implications

A Boolean implication plays a central role in inference schemas, like modus ponens, modus tollens, etc. For instance, modus ponens states that given two true propositions (premises): A and conditional claim $A \to B$, the consequent B is also true. Using IF-THEN rules we may express modus ponens by the following schema:

$$\begin{array}{ll} \text{Rule:} & \text{IF } X \text{ is } A, \text{ THEN } Y \text{ is } B \\ \text{Fact:} & X \text{ is } A \\ \text{Conclusion:} & Y \text{ is } B \end{array} \tag{1}$$

In everyday life we usually employ linguistic labels expressed in natural language. Our knowledge contains uncertainties coming from various sources, like imprecision, ambiguity, vagueness, fuzziness or incomplete information, etc. This type of information forms a basis for everyday decision process. Of course, using uncertain knowledge we cannot expect certain conclusions. In particular, approximate reasoning refers to the process by which imprecise conclusions are drawn from imprecise premises. When this imprecision is modelled by fuzzy sets the term "fuzzy reasoning" is also used.

To perform approximate reasoning traditional inference schemas were appropriately generalized. In particular, the classical modeus ponens have been extended to fuzzy logic under the inference pattern called generalized modus ponens which may be expressed in the following way:

$$\begin{array}{ll} \text{Rule:} & \text{IF } X \text{ is } A, \text{ THEN } Y \text{ is } B \\ \text{Fact:} & X \text{ is } A' \\ \text{Conclusion:} & Y \text{ is } B' \end{array} \tag{2}$$

where A, A', B and B' are fuzzy sets.

To enrich capabilities of reasoning tools and approaches we may generalize not only inference schemas but also operators applied there, like the implication operator. A fuzzy implication is a generalization of the classical one and as t-norms and t-conorms belong to main operations in fuzzy logic.

In classical logic the implication can be defined in different ways depending on the mathematical framework, physical nature or philosophical background in which it appears. Anyway, despite this circumstances the so-called truth tables have to be and really are identical. It is not surprising that starting from different classical definitions and using diversity of methods we may obtain a many fuzzy implications having different properties. Anyway all of them should satisfy some basic requirements that can be perceived as a generalization of the classical truth table. These fundamental properties form the definition of fuzzy implication. Below we recall the definition given in [1] (also equivalent to those proposed in [2] or [5]).

Definition 1 *A function* $I : [0,1]^2 \to [0,1]$ *is called a fuzzy implication if it satisfies the following conditions*

(I1) *if* $x_1 \leq x_2$ *then* $I(x_1, y) \geq I(x_2, y)$
(I2) *if* $y_1 \leq y_2$ *then* $I(x, y_1) \leq I(x, y_2)$
(I3) $I(0,0) = 1$
(I4) $I(1,1) = 1$
(I5) $I(1,0) = 0$.

for all $x, x_1, x_2, y, y_1, y_2 \in [0,1]$.

Note that the property (I1) means that $I(\cdot, y)$ is nonincreasing and captures the left antitonicity of the function I, i.e. a decrease in truth value of the antecedent increases its efficacy to state more about the truth value of its consequent. The axiom (I2) shows that $I(x, \cdot)$ is nondecreasing which reflects the right isotonicity of the overall truth value as a direct function of the consequent. Finally, axioms (I3) – (I5) correspond to basic properties of the classical implication. For more details on fuzzy implications we refer the reader to [1].

The family of all fuzzy implications will be denoted by $\mathcal{FI}$.

3. Implication and probability

It is worth noting that the imprecision is responsible only for aspect of uncertainty. It seems that the second main source of problems is connected with randomness. These two main grounds of uncertainty may appear together or alone but they are themselves the outputs of quite different mechanisms and hence are described and analyzed using different methodology. Thus now let us have a look on implication from the probabilistic perspective.

Roughly speaking, according to modus ponens inferential rule we may expect that provided A is "surely" true and the implication $A \to B$ is also "surely" true then B is true. However, in practice we often cannot be completely sure that given event surely occur. In many situations instead of determined premises we have only some confidence that they are true. In other words we may neither be completely sure that A nor that the the implication $A \to B$ is 100% true. We can only estimate the probability $P(A)$ that A is true and the probability $P(A \to B)$ that the implication $A \to B$ holds. Anyway, due to appropriate probabilistic inference we can transform these two probabilities into desired conclusion on the probability related to B. However, the particular inference depends on the interpretation of the probability of an implication. It seems that the most natural approach is to interpret the probability of an implication as the conditional probability $P(B|A)$, i.e.

$$P(B|A) = \frac{P(B \cap A)}{P(A)}, \tag{3}$$

where A should not be impossible, i.e. $P(A) > 0$. Hence, if we know the probability $P(A)$ of the premise A and the probability $P(B|A)$ of the implication then, due to that formula, we can find the probability $P(B \cap A)$ that both B and A are true, i.e.

$$P(B \cap A) = P(B|A) \cdot P(A).$$

Applying more sophisticated Bayesian analysis we may develop that inference. However, although conditional probability given by (3) provides a natural and convenient interpretation of the implication it cannot be treated as an implication operator. Therefore we are still looking for the adequate extension of the classical implication operator into probabilistic environment. In the next section we will show how to construct a probabilistic implication operator which is actually a fuzzy implication.

Let us also mention that the conditional probability is not the only way that combines implication with randomness. The other formalization corresponding to "material implication" interprets the probability of implication $A \to B$ as $P(B \text{ or } \neg A)$. For the comparison of these two probabilistic perspectives we refer the reader to [7]. The overview of probability of implication applications in artificial intelligence can be found e.g. in [3, 4].

4. Copulas

One of the reasons that (3) cannot serve directly for defining an implication operator is that its domain is a σ-algebra of the family of all elementary events corresponding to an experiment under study and may vary from one experiment to the other. Thus first of all, when trying to suggest a proper implication operator, we have to get rid of any particular experiment. In probability theory it is usually accessible by the use of a random variables that transform any particular space of elementary events into the real line. Since our goal is to reduce the domain of the desired operator into the unit square we will base our construction on the well-known object called a copula. Let us recall briefly its definition.

Definition 2 *A copula is a function* $C : [0,1]^2 \to [0,1]$ *which satisfies the following conditions:*
(a) $C(u,0) = C(0,v) = 0$ *for every* $u, v \in [0,1]$
(b) $C(u,1) = u$ *for every* $u \in [0,1]$
(c) $C(1,v) = v$ *for every* $v \in [0,1]$
(d) for every $u_1, u_2, v_1, v_2 \in [0,1]$ *such that* $u_1 \le u_2$ *and* $v_1 \le v_2$

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0. \tag{4}$$

Several interesting families of copulas, like Fŕechet's family, Farlie-Gumbel-Morgenstern's family, Marshall-Olkin's family, etc., are considered in the literature. It can be shown that every copula is bounded by the so-called Fŕechet-Hoeffding bounds, i.e. for any copula C and for all $u, v \in [0,1]$

$$W(u,v) \le C(u,v) \le M(u,v), \tag{5}$$

where

$$W(u,v) = max\{u+v-1, 0\}, \tag{6}$$
$$M(u,v) = min\{u,v\} \tag{7}$$

are also copulas. For more information on copulas we refer the reader to the famous monograph by Nelsen [6].

The importance of the copulas is clarified by the Sklar theorem [9] showing that copulas link joint distribution functions to their one-dimensional margins.

Theorem 3 *Let* X *and* Y *be random variables with joint distribution function* H *and marginal distribution functions* F *and* G*, respectively. Then there exists a copula* C *such that*

$$H(x,y) = C(F(x), G(y)) \tag{8}$$

for all $x, y \in \mathbb{R}$*. If* F *and* G *are continuous, then* C *is unique. Otherwise, the copula* C *is uniquely determined on* $Ran(F) \times Ran(G)$*. Conversely, if* C *is a copula and* F *and* G *are distribution functions, then the function* H *defined by (8) is a joint distribution function with margins* F *and* G*.*

The Sklar theorem shows that we can disregard the particular domains of the random variables under study and transform the whole problem into the unit square which is actually our goal. Actually, C captures all the information about the dependence among the components of (X,Y). Now we have to introduce a conditional structure for copulas. Namely, to describe by the copula method the conditional distribution function of Y given $X \le x$ such that $P(X \le x) > 0$ we get

$$\begin{aligned} P(Y \le y | X \le x) &= \frac{P(U \le u, V \le v)}{P(U \le u)} \\ &= \frac{H(x,y)}{F(x)} \\ &= \frac{C(F(x), G(y))}{F(x)} \\ &= \frac{C(u,v)}{u}, \end{aligned} \tag{9}$$

where $U = F(X)$ and $V = G(Y)$. An this very last concept forms the basis of our proposal for the extension of the logical implication into that probabilistic framework.

Further on let us adopt the convention that $\frac{0}{0} = 1$.

5. Probabilistic implication operator

Having in mind our discussion from the previous sections, especially eq. (9), let us consider a function $I_C : [0,1]^2 \to [0,1]$ defined by the following formula

$$I_C(u,v) = \frac{C(u,v)}{u}, \tag{10}$$

where C is any given copula. One may ask whether (10) is correctly defined for $u = 0$. However, by the Def. 2 $C(0,v) = 0$ for every $v \in [0,1]$ and hence, due to our convention that $\frac{0}{0} = 1$, we get $I_C(0,v) = 1$. So (10) is correctly defined. Anyway, to avoid any unnecessary discussions further on we will use the following definition:

Definition 4 *A function* $I_C : [0,1]^2 \to [0,1]$ *given by*

$$I_C(u,v) = \begin{cases} 1 & if\ u = 0 \\ \frac{C(u,v)}{u} & if\ u > 0, \end{cases} \tag{11}$$

where C *is a copula, is called a* ***probabilistic implication*** *(based on copula* C*).*

It is easily seen that $I_C(u,v)$ is a natural counterpart of the conditional probability (3) transformed into the general environment created by the copula theory. Let us now examine some basic properties of the function I_C to check whether it actually could serve as an implication operator.

Lemma 5 *For any copula* C *function* I_C *given by (11) satisfies the following conditions:*
(i) $I(0,0) = 1$
(ii) $I(1,1) = 1$
(iii) $I(1,0) = 0$
(iv) if $y_1 \le y_2$ *then* $I(x,y_1) \le I(x,y_2)$*.*

Proof:
Property (i) holds by definition, since according to (11) we have $I_C(0,0) = 1$.

Next two properties are just direct conclusions from 2. Actually, condition (b) in Def. 2 states that $C(u,1) = u$ for every $u \in [0,1]$. Thus $I_C(1,1) = 1$ and hence (ii) holds.

Similarly, property (iii) is fulfilled because of condition (a) in Def. 2. Since $C(u,0) = C(0,v) = 0$ for every $u, v \in [0,1]$ then $I_C(1,0) = 0$.

Now let us prove (iv) that $I_C(u,\cdot)$ is nondecreasing. But it is obvious, because any copula is a nondecreasing function for each argument and hence if $v_1 \leq v_2$ then we get $I_C(u,v_1) = \frac{C(u,v_1)}{u} \leq \frac{C(u,v_2)}{u} = I_C(u,v_2)$ for any $u > 0$. If $u = 0$ then $I_C(u,v_1) = 1 = I_C(u,v_2)$, which completes the proof. ■

Our Lemma 5 states that our function I_C fulfills requirements (I2)-(I5) specified in Def. 1 of a fuzzy implication. Thus an immediate question is whether $I_C(\cdot,v)$ is nonincreasing for any copula C, i.e. if $u_1 \leq u_2$ then $I_C(u_1,v) \geq I_C(u_2,v)$. If $u_1 = u_2 = 0$ it is obvious by (11). Unfortunately, if $0 < u_1 < u_2$ then for some copulas I_C may not satisfy (I1).

Example 1
Let us consider the lower Fŕechet-Hoeffding bound (6). Suppose $u_1 = 0.2 < u_2 = 0.3$ and $v = 0.9$. Then $I_W(u_1,v) = \frac{1}{2} < I_W(a_2,v) = \frac{2}{3}$ which shows that I_W is not nonincreasing. Thus $I_W \notin \mathcal{FI}$. ■

Therefore, a probabilistic implication is not - in general - a fuzzy implication. But operators having properties like I_C are often called *engineering implications* (see, e.g. [8]).

We can formulate the following theorem.

Theorem 6 *If for every $u_1, u_2, v \in [0,1]$ such that $u_1 \leq u_2$ a copula C satisfies*

$$C(u_1,v)u_2 \geq C(u_2,v)u_1 \tag{12}$$

then the probabilistic implication I_C given by (11) is a fuzzy implication.

Proof:
By Lemma 5 function I_C fulfills requirements (I2)-(I5) specified in Def. 1 for any copula C. Condition (12) assures that $I_C(\cdot,v)$ is nonincreasing so (I1) is also satisfied and hence I_C is a fuzzy implication, i.e. $I_C \in \mathcal{FI}$. ■

Further on probabilistic implications satisfying condition (12) will be called *probabilistic fuzzy implications*.

Taking into account the last theorem and the whole discussion given in the previous section we may conclude that probabilistic fuzzy implications make a bridge between probability and the theory of fuzzy implication operators. Here are some examples of probabilistic fuzzy implications. It is worth noting that choosing some specific copulas we may obtain well known fuzzy implications.

Example 2
As a first example let us consider the product copula $\Pi(u,v) = uv$, which characterizes independent random variables when the distribution functions are continuous. Therefore, the probabilistic implication based on the product copula is

$$I_\Pi(u,v) = \begin{cases} 1 & \text{if } u = 0 \\ v & \text{if } u > 0. \end{cases} \tag{13}$$

One can easily check that (13) satisfies (12) so the probabilistic implication based on the product copula is a fuzzy implication, i.e. $I_\Pi \in \mathcal{FI}$.

Let us also consider the Rescher implication I_{RS} (see [1]) defined by

$$I_{RS}(u,v) = \begin{cases} 1 & \text{if } u \leq v \\ 0 & \text{if } u > v. \end{cases} \tag{14}$$

As it is known (see [1]), if $I \in \mathcal{FI}$ and $J \in \mathcal{FI}$ then $I \vee J \in \mathcal{FI}$. Hence, in particular, $I_\Pi \vee I_{RS} \in \mathcal{FI}$, where

$$I_\Pi \vee I_{RS}(u,v) = \begin{cases} 1 & \text{if } u \leq v \\ v & \text{if } u > v. \end{cases} \tag{15}$$

Hence one may notice that $I_\Pi \vee I_{RS}(u,v) = I_{GD}(u,v)$, i.e. the operator (15) obtained as a supremum of the probabilistic implication based on the product copula I_Π and the Rescher implication I_{RS} is equivalent to the famous Gödel implication I_{GD} (see [1]). ■

Example 3
The Gödel implication is not the only well known implication that might be obtain from the probabilistic implication. Let us now consider another specific copula, i.e. the upper Fŕechet-Hoeffding bound M given by (7). The probabilistic implication based on $M(u,v)$ is

$$\begin{aligned} I_M(u,v) &= \frac{\min\{u,v\}}{u} \\ &= \begin{cases} 1 & \text{if } u \leq v \\ \frac{v}{u} & \text{if } u > v \end{cases} \end{aligned} \tag{16}$$

which is nothing else than the Goguen implication, i.e. $I_M(u,v) = I_{GG}(u,v)$. ■

Example 4
The product copula discussed in Example 2 appears as a special case in many families of copulas. For instance, Farlie-Gumbel-Morgenstern's family given by

$$C_\theta(u,v) = uv + \theta uv(1-u)(1-v) \tag{17}$$

is a one-parameter family of copulas, where $\theta \in [-1,1]$ is responsible for the dependence structure. In particular, for $\theta = 0$ we obtain the product copula. Indeed, $C_\theta(u,v)|_{\theta=0} = uv = \Pi(u,v)$.

A probabilistic implication based on the Farlie-Gumbel-Morgenstern copula is then

$$I_{FGM(\theta)}(u,v) = \begin{cases} 1 & \text{if } u = 0 \\ v + \theta v(1-u)(1-v) & \text{if } u > 0. \end{cases} \tag{18}$$

Since by (13) and (17) $I_{FGM(0)} = I_{\Pi}$, hence probabilistic implication (18) for $\theta = 0$ is a fuzzy implication. Now let us check whether we can obtain a probabilistic implication based on the Farlie-Gumbel-Morgenstern copula with parameter $\theta \neq 0$.

Thus, by Th. 6 we have to examine whether condition (12) holds for any $\theta \in [0, 1]$. Taking any $u_1, u_2, v \in [0, 1]$ such that $u_1 \leq u_2$ we can see that (12) is satisfied when

$$[u_1 v + \theta u_1 v(1 - u_1)(1 - v)]u_2 \geq$$
$$\geq [u_2 v + \theta u_2 v(1 - u_2)(1 - v)]u_1$$

which is equivalent to

$$\theta(1 - u_1) \geq \theta(1 - u_2)$$

and holds if and only if $\theta \geq 0$. Therefore, we may conclude that

$$I_{FGM(\theta)} \in \mathcal{FI} \Leftrightarrow \theta \geq 0,$$

i.e., the probabilistic implication (18) based on the Farlie-Gumbel-Morgenstern copula is a fuzzy implication not for all possible values of parameter θ but only for $\theta \geq 0$. ■

6. Conclusions

A new family of implication operators, called probabilistic implications were introduced in the paper. It was discussed when probabilistic implications are fuzzy implications. Some examples illustrating these new tools were also given. As they are called, probabilistic implications give a promising link from probability to theory of fuzzy implications that might be useful in approximate reasoning. However, many problems and questions are still open. One of the most important is to characterize a family of the copulas leading to probabilistic fuzzy implications. It would be desirable to find mathematical formulae which are not only efficient but also providing a clear and natural interpretation. Next problem is to examine properties of the probabilistic implications based on different families of copulas.

References

[1] M. Baczynski and Y. Jayaram, *Fuzzy Implications*, Springer, 2008.

[2] J.C. Fodor, M. Roubens, *Fuzzy Preference Modelling and Multicriteria Decision Support*, Kluwer, Dordrecht, 1994.

[3] N. Greene, An overview of conditionals and biconditionals in probability, In *proceedings of the American Conference on Applied Mathematics* (MATH '08), Harvard, Massachusetts, USA, March 24-26, 2008.

[4] N. Greene, Methods of assessing and ranking probable sources of error, In *proceedings of the Applied Computing Conference* (ACC '08), Istanbul, Turkey, May 27-30, 2008.

[5] L. Kitainik, *Fuzzy Decision Procedures with Binary Relations*, Kluwer, Dordrecht, 1933.

[6] R. B. Nelsen, *An Introduction to Copulas*, Springer, New York, 1999.

[7] H. T. Nguyen, M. Mukaidono and V. Kreinovich, Probability of Implication, Logical Version of Bayes' Theorem, and Fuzzy Logic Operations, In *proceedings of the FUZZ-IEEE'2002*, Honolulu, Hawaii, May 12-17, 2002, pp. 530-535.

[8] L. Rutkowski, *Artificial Intelligence - Methods and Techniques*, PWN, Warsaw, 2006 (in Polish).

[9] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, Elsevier, New York, 1983.

[10] L. Zadeh, From computing with numbers to computing with words - From manipulation of measurements to manipulation of perceptions, *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, 45, 105-119, 1999.