Intuitive Idea of Implication vs. Formal Definition: How to Define the Corresponding Degree

Olga Kosheleva[1] and Vladik Kreinovich[1]
[1]University of Texas at El Paso
El Paso, Texas 79968, USA, olgak@utep.edu, vladik@utep.edu

Abstract

Formal implication does not capture the intuitive idea of "if A then B", since in formal implication, every two true statements – even completely unrelated ones – imply each other. A more adequate description of intuitive implication happens if we consider how much the use of A can shorten a derivation of B. At first glance, it may seem that the number of bits by which we shorten this derivation is a reasonable degree of implication, but we show that this number is not in good accordance with our intuition, and that a natural formalization of this intuition leads to the need to use, as the desired degree, the ratio between the shortened derivation length and the original length.

1 Formulation of the Problem

Intuitive idea of implication is different from a formal definition. A traditional formal definition of implication is that A implies B (denoted $A \to B$) if either B is true or A is false.

This formal definition is very different from the intuitive idea of what we mean when we say that A implies B: e.g., it allows formally true but intuitively meaningless implications such as "if the Moon is made out of cheese, then ghosts like to scare people." Also, in the sense of the formal definitions, all true statements imply each other, e.g., Pythagoras theorem implied the four-color theorem, which also intuitively makes no sense.

How can we formalize the intuitive idea of implication. A better formalization of the intuitive idea of implication is that if we assume A, then we can get a shorter derivation of B – e.g., a derivation that takes fewer bits when stored in the computer memory.

To describe this idea in precise terms, let us fix a formal theory T, and let us denote:

- by $L(B)$, the shortest length of a derivation of B in this theory, and
- by $L(B\,|\,A)$, the shortest length of a derivation of B from a theory $T\cup\{A\}$ (to which A is added as an additional axiom).

In these terms, we can say that A implies B if $L(B\,|\,A) < L(B)$.

Comments.

- For every true statement B and for every statement A, we can algorithmically compute the values $L(B)$ and $L(B\,\&\,A)$ – by simply trying all possible derivations of a given length. Thus, we can algorithmically check whether A intuitively implies B. Of course, this computation may take a long time – since trying all possible combinations of n steps requires time which is exponential in n and thus, not feasible for large n.
- In contrast to the formal implication, this intuitive implication has many unusual properties. For example, it is not transitive. Indeed, if true statements p and q are completely independent, then the use of p does not help us derive q: $L(q\,|\,p) = L(q)$ hence p does not imply q in this sense. On the other hand:
  - if we use p, then deriving $p\,\&\,q$ is easier then without p: we only need to derive q: $L(p\,\&\,q\,|\,p) = L(q)$ and $L(p\,\&\,q) = L(p) + L(q) > L(q)$, thus we can intuitively derive $p\,\&\,q$ from p;
  - the derivation of q from $p\,\&\,q$ is even easier: it takes just one step, so $L(q\,|\,p\,\&\,q) = 1 \ll L(q)$; thus, we can intuitively derive q from $p\,\&\,q$.

  So, in this example, p intuitively implies $p\,\&\,q$, the statement $p\,\&\,q$ intuitively implies q, but p does not intuitively imply q.

Need to consider degrees of implication. Intuitively, there is a big different between a situation in which the use of A decreases the length of the derivation of B by a small amount (e.g., by 1 step), and a situation in which the use of A drastically decreases the length of such derivation. In the second situation, A strongly implies B, while in the first situation, the degree to which A implies B is very small.

It is therefore desirable to come up with a way to measure the degree to which A implies B. This is the problem that we consider in this paper.

Comment. This conclusion is line with the general idea of Lotfi Zadeh and fuzzy logic: that everything is a matter of degree; see, e.g., [1, 2, 4, 5, 6, 7, 8].

2 How to Define the Degree of Implication: A Seemingly Reasonable Idea and Its Limitations

A seemingly reasonable idea. Intuitively, the more the use of A decreases the length of the derivation of B, the larger the degree to which A implies B. It therefore seems reasonable to define this degree as the difference $L(B)-L(B\,|\,A)$ between the corresponding lengths [3]:

- when this difference is large, A strongly implies B, while
- when this difference is small, the degree to which A implies B is very small.

Limitations of this idea. The problem with the above seemingly natural idea is that it does not fully capture the intuitive notion of implication. Indeed, let us assume that we have two independent derivations with the same degree of implication:

- with this degree, A_1 implies B_1, and
- with the same degree, A_2 implies B_2.

Then, when we simply combine the corresponding statement and the corresponding proofs, we expect to conclude that the conjunction $A_1\,\&\,A_2$ implies the conjunction $B_1\,\&\,B_2$ with the same degree of implication. What will happen if we compute this degree by using the above idea?

- Since the derivations are independent, the only way to prove $B_1\,\&\,B_2$ is to prove both B_1 and B_2, so $L(B_1\,\&\,B_2)=L(B_1)+L(B_2)$.
- Similarly, the only way to prove $B_1\,\&\,B_2$ from $A_1\,\&\,A_2$ is to prove B_1 from A_1 and B_2 from A_2, so $L(b_1\,\&\,B_2\,|\,A_1\,\&\,A_2)=L(B_1\,|\,A_1)+L(B_2\,|\,A_2)$.

Thus, for the difference, we have

$$L(B_1\,\&\,B_2)-L(B_1\,\&\,B_2\,|\,A_1\,\&\,A_2)=$$

$$L(B_1)+L(B_2)-L(B_1\,|\,A_1)-L(B_2\,|\,A_2)=$$

$$(L(B_1)-L(B_1\,|\,A_1))+(L(B_2)-L(B_2\,|\,A_2)).$$

Thus, if both derivations had the same degree of implication d, i.e., if we had

$$L(B_1)-L(B_1\,|\,A_1)=L(B_2)-L(B_2\,|\,A_2)=d,$$

then for the degree $L(B_1\,\&\,B_2)-L(B_1\,\&\,B_2\,|\,A_1\,\&\,A_2)$, we get not the same degree d, but the much larger value $2d$.

So, if we combine the two independent derivations with low degrees of implication, then magically the degree of implication increases – this does not make sense.

Challenge and what we do in this paper. Since the seemingly natural definition contradicts our intuition, it is desirable to come up with another definition, a definition that will be consistent with our intuition.

This is what we will do in this paper.

3 Towards a New Definition of Degree of Implication

Towards a formal definition. Let us fix a certain degree of implication. This degree occurs for some examples, when the length of the original derivation $L(B)$ is b and the length of the conditional derivation $L(B\,|\,A)$ is a. The smaller a, the larger the degree of implication. Thus, not for every length b, we can have the given degree of implication: e.g.,

- for $b = 2$, we only have two options $a = 0$ and $a = 1$, i.e., only two possible degrees of implication; while
- for $b = 10$, we have ten possible values $a = 0, \ldots, 9$ corresponding to ten possible degrees of implication.

For each possible degree d, let $f_d(b)$ describe the length of the conditional derivation corresponding to the length b of the original derivation and this particular degree. As we have mentioned, this function is not necessarily defined for all b.

Let us assume that the function $f_d(b)$ is defined for some value b. This means that there exists an example with this particular degree-of-implication when the original derivation has length b and the conditional derivation has length $a = f_d(b)$. Then, as we have mentioned in the previous section, if we have k different independent pairs of statements (A_i, B_i) of this type, for which $L(B_1) = \ldots = L(B_k) = b$ and $L(B_1\,|\,A_1) = \ldots = L(B_k\,|\,A_k) = a$, then, for derivation of $B_1 \,\&\, \ldots \,\&\, B_k$ from $A_1 \,\&\, \ldots \,\&\, A_k$ of the same degree of implication, we have

$$L(B_1 \,\&\, \ldots \,\&\, B_k) = L(B_1) + \ldots + L(B_k) = k \cdot b$$

and

$$L(B_1 \,\&\, \ldots \,\&\, B_k\,|\,A_1 \,\&\, \ldots \,\&\, A_k) = L(B_1\,|\,A_1) + \ldots + L(B_k\,|\,A_k) = k \cdot a.$$

Thus, if $f_d(b) = a$, then for each integer $k > 1$, we have $f_d(k\cdot b) = k\cdot a = k\cdot f_d(b)$. So, we arrive at the following definition.

Definition 1. *By a* function corresponding to a given degree of implication*, we mean a partial function $f_d(b)$ from natural numbers to natural numbers that satisfies the following property:*

- *if this function is defined for some d,*
- *then for each natural number $k > 2$, it is also defined for $k \cdot d$, and*

$$f_d(k \cdot b) = k \cdot f_d(b).$$

Proposition 1. *Let f_d be a function corresponding to a given degree of implication. Then, for every two values b and b' for which f_d is defined, we have*

$$\frac{f_d(b)}{b} = \frac{f_d(b')}{b'}.$$

Proof. Since the function f_d is defined for b, then, by taking $k = b'$, we conclude that this function is also defined for $b \cdot b'$, and $f_d(b \cdot b') = b' \cdot f_d(b)$. Similarly, from the fact that this function is defined for b', by taking $k = b$, we get $f_d(b \cdot b') = b \cdot f_d(b')$.

Thus, we conclude that $b' \cdot f_d(b) = b \cdot f_d(b')$. Dividing both sides by $b \cdot b'$, we get the desired equality.

Conclusion. Different degrees of implication correspond to different values of the ratio $\dfrac{f_d(b)}{b}$. Thus, we can use the corresponding ratio $\dfrac{L(B\,|\,A)}{L(B)}$ as the desired degree to what A implies B.

Comment. It may be better to use a slightly different measure

$$1 - \frac{L(B\,|\,A)}{L(B)} = \frac{L(B\,|\,A) - L(B)}{L(B)},$$

since this measure is 0 if the use of A does not help at all and 1 if it decreases the length of the proof maximally.

Acknowledgments

This work was supported in part by the National Science Foundation via grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science) and HRD-1242122 (Cyber-ShARE Center of Excellence).

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

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

[3] V. Kreinovich, L. Longpré, and H. T. Nguyen, "Towards formalization of feasibility, randomness, and commonsense implication: Kolmogorov complexity, and the necessity of considering (fuzzy) degrees", In: Nguyen Hoang Phuong and Ario Ohsato (eds.), *Proceedings of the Vietnam-Japan Bilateral Symposium on Fuzzy Systems and Applications VJFUZZY'98*, HaLong Bay, Vietnam, 30th September-2nd October, 1998, pp. 294–302.

[4] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[5] H. T. Nguyen and V. Kreinovich, "Nested intervals and sets: concepts, relations to fuzzy sets, and applications", In: R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[6] H. T. Nguyen, C. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[7] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

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