

Data Fusion with Imperfect Implication Rules

J. N. Heendeni¹, K. Premaratne¹, M. N. Murthi¹ and M. Scheutz²

¹ Elect. & Comp. Eng., Univ. of Miami, Coral Gables, FL, USA,
j.anuja@miami.edu, kamal@miami.edu, mmurthi@miami.edu

² Comp. Sci., Tufts Univ., Medford, MA, USA,
mscheutz@cs.tufts.edu

Abstract. In this paper, we develop a method to find the uncertain consequent by fusing the uncertain antecedent and the uncertain implication rule. In particular with Dempster-Shafer theoretic models utilized to capture the uncertainty intervals associated with the antecedent and the rule itself, we derive bounds on the confidence interval associated with the rule consequent. We derive inequalities for the belief and plausibility values of the consequent and with least commitment choice they become equations. We also demonstrate the consistency of our model with probability and classical logic.

Keywords: Belief, Plausibility, Fusion, Implication rule, Antecedent, Consequent, Uncertainty.

1 Introduction

Implication rules, which take the form “if A , then B ” or, as is often expressed, $A \implies B$, constitute the backbone of reasoning and inference engines. A large volume of existing work addresses the extraction of such rules from databases and their use in various application scenarios. However, most of these works assume that the evidence/information at hand is “perfect”, which, in practice, is far from the truth. Databases are rife with imperfect (e.g., ambiguous, vague, or incomplete) entries rendering the rule antecedent A , the rule \implies itself, and hence the rule consequent B to be imperfect. Even otherwise, one cannot expect to get “perfect” rules when only *finite* databases are available for rule extraction.

Probabilistic and fuzzy models are perhaps the two most commonly used approaches to capture imperfect rules [4],[10]. This current work of ours is based on *Dempster-Shafer (DS) theory* [12] which can capture a wider variety of imperfections, provide interval-based models of the underlying uncertainties, and can be considered a generalization of probability mass functions (p.m.f.s). Several previous works deal with DS theory (DST) based modeling of imperfect rules: in [6],[7], DST fusion/combination strategies are employed to get results that are most similar to ours, but general *bounds* and inequalities that we derive are absent and the approach is different; in [2], emphasis is placed on satisfying the material implications of propositional logic statements; [11] designs a complete uncertain logic framework (imperfect rules being a special case) which is compatible with classical (perfect) logic [10]. We take a different view: we do not

impose compatibility with classical logic in imperfect domains; rather, we expect compatibility only when the domain is perfect, so that our model is very general and all probability and classical logic are special cases.

We model our imperfect rules via DST Fagin-Halpern (FH) conditionals [5]. While the use of the Bayesian conditional has been criticized as a model of probabilistic imperfect rules [9],[2], we demonstrate that the DST FH conditionals can be used as an effective interval-based model of imperfect rules to fuse with imperfect antecedent. Given the uncertainty intervals associated with the rule antecedent and the rule itself, we derive explicit lower and upper bounds for the uncertainty interval of the rule consequent. Then we explicitly show its consistency with Bayesian inference and classical logic.

2 Preliminaries

Basic DST Notions. Let $\Theta = \{\theta_1, \dots, \theta_M\}$ denote the *Frame of Discernment (FoD)* which contains the discrete set of mutually exclusive and exhaustive propositions. The power set 2^Θ , i.e., the set containing all the possible subsets of Θ , is 2^Θ . For arbitrary $A \subseteq \Theta$, \bar{A} denotes those singletons that are not in A .

As usual, $m_\Theta(\cdot) : 2^\Theta \mapsto [0, 1]$ is a *basic belief assignment (BBA)* or *mass assignment* where $\sum_{A \subseteq \Theta} m_\Theta(A) = 1$ and $m_\Theta(\emptyset) = 0$. Propositions that receive non-zero mass are the *focal elements*; the set of focal elements is the *core* \mathcal{F}_Θ . The triplet $\mathcal{E} = \{\Theta, \mathcal{F}_\Theta, m_\Theta\}$ is the *body of evidence (BoE)*.

Given a BoE, $\mathcal{E} \equiv \{\Theta, \mathcal{F}_\Theta, m_\Theta\}$, the *belief function* $Bl_\Theta : 2^\Theta \mapsto [0, 1]$ is $Bl_\Theta(A) = \sum_{B \subseteq A} m_\Theta(B)$. The *plausibility function* $Pl_\Theta : 2^\Theta \mapsto [0, 1]$ is $Pl_\Theta(A) = 1 - Bl_\Theta(\bar{A})$. The *uncertainty interval* associated with A is $[Bl_\Theta(A), Pl_\Theta(A)]$.

Of the various notions of DST conditionals abound in the literature, the Fagin-Halpern (FH) conditional [5] possesses several attractive properties and offers a unique probabilistic interpretation and hence a natural transition to the Bayesian conditional notion [5], [3], [14].

Definition 1 (FH Conditionals). For the BoE $\mathcal{E} = \{\Theta, \mathcal{F}_\Theta, m_\Theta\}$ and $A \subseteq \Theta$ s.t. $Bl_\Theta(A) \neq 0$, the conditional belief $Bl_\Theta(B/A) : 2^\Theta \mapsto [0, 1]$ and conditional plausibility $Pl_\Theta(B/A) : 2^\Theta \mapsto [0, 1]$ of B given A are

$$Bl_\Theta(B/A) = \frac{Bl_\Theta(A \cap B)}{Bl_\Theta(A \cap B) + Pl_\Theta(A \cap \bar{B})}; \quad Pl_\Theta(B/A) = \frac{Pl_\Theta(A \cap B)}{Pl_\Theta(A \cap B) + Bl_\Theta(A \cap \bar{B})}.$$

3 Implication Rule Model Assumption

We model the implication rule by FH conditionals. Consider the implication rule $A \implies B$, where A denotes the antecedent, B denotes the consequent, and \implies denotes the rule R . Without loss of generality we assume antecedent and consequent are in same BoE. We consider the uncertainty of the rule R and it is modeled as follows;

$$Bl(R) = Bl(B/A); \quad Pl(R) = Pl(B/A) \tag{1}$$

Additionally if we have evidence on $\bar{A} \implies B$, which we model by $Bl(B/\bar{A})$ and $Pl(B/\bar{A})$, more finer results can be obtained for the consequent. Later with results, it can be seen that these conditionals are a good way to model the implication rules.

4 Fusion

The purpose of this paper is to find the uncertainty intervals of the consequent given the uncertainty intervals of antecedent and the rule. For simplicity, we will use the following notation:

$$\begin{aligned} Bl(A) = \alpha_1; \quad Pl(A) = \beta_1; \quad Bl(B) = \alpha_2; \quad Pl(B) = \beta_2; \\ Bl(B/A) = \alpha_r; \quad Pl(B/A) = \beta_r; \quad Bl(B/\bar{A}) = \bar{\alpha}_r; \quad Pl(B/\bar{A}) = \bar{\beta}_r, \end{aligned} \quad (2)$$

4.1 Results

We obtain following relations for belief and plausibility of the consequent.

If knowledge of antecedent and knowledge of both $A \implies B$ and $\bar{A} \implies B$ are available, following lower and upper bounds can be obtained for the belief and plausibility respectively.

$$\alpha_2 \geq \alpha_1 \alpha_r + (1 - \beta_1) \bar{\alpha}_r. \quad (3)$$

$$\beta_2 \leq \alpha_1 \beta_r + (1 - \beta_1) \bar{\beta}_r + (\beta_1 - \alpha_1). \quad (4)$$

We call them **General Bounds**.

Least Commitment (LC) Choice. The principle of minimum or least commitment (LC) [13], [1] dictates that we are least committed and rely on available evidence only, i.e., select the lower bound for α_2 and the upper bound for β_2 .

The corresponding LC choice;

$$\alpha_2 = \alpha_1 \alpha_r + (1 - \beta_1) \bar{\alpha}_r. \quad (5)$$

$$\beta_2 = \alpha_1 \beta_r + (1 - \beta_1) \bar{\beta}_r + (\beta_1 - \alpha_1). \quad (6)$$

If knowledge of antecedent and knowledge of only $A \implies B$ are available the relations become;

$$\alpha_2 \geq \alpha_1 \alpha_r. \quad (7)$$

$$\beta_2 \leq 1 - \alpha_1 (1 - \beta_r). \quad (8)$$

We call them **Relaxed Bounds**. Since we are using only the implication $A \implies B$, these are the bounds for the **imperfect implication**. With the LC choice these bounds become following equations.

$$\alpha_2 = \alpha_1 \alpha_r. \quad (9)$$

$$\beta_2 = 1 - \alpha_1(1 - \beta_r). \quad (10)$$

Uncertainty. The upper bound for the uncertainty of the General Bounds;

$$\beta_2 - \alpha_2 \leq (\beta_1 - \alpha_1) + \alpha_1(\beta_r - \alpha_r) + (1 - \beta_1)(\bar{\beta}_r - \bar{\alpha}_r). \quad (11)$$

The above term for the upper bound of the uncertainty has an interesting intuitive interpretation: the uncertainty interval of the consequent is bounded above by the uncertainty of the antecedent plus the uncertainties of the rules $A \implies B$ and $\bar{A} \implies B$ weighted by their corresponding belief terms; $Bl(A) = \alpha_1$ and $Bl(\bar{A}) = 1 - \beta_1$. And it can be easily shown that this term, the upper bound for the uncertainty, is always less than or equal to 1, which is correct and intuitive. The lower bound for the uncertainty is 0 since by definition $\beta_2 \geq \alpha_2$. (Note that we define α_2 and β_2 as belief functions and then find relations for them). And also it is clear that the relations which were obtained for β_2 are always greater than or equal to the relations which were obtained for α_2 .

The inequalities can be written in one line;

$0 \leq \alpha_1\alpha_r \leq \alpha_1\alpha_r + (1 - \beta_1)\bar{\alpha}_r \leq \alpha_2 \leq \beta_2 \leq \alpha_1\beta_r + (1 - \beta_1)\bar{\beta}_r + (\beta_1 - \alpha_1) \leq 1 - \alpha_1(1 - \beta_r) \leq 1$ and it is apparent that when more knowledge (The term knowledge refers to the knowledge of belief and plausibility values.) is available the bounded uncertainty interval of the consequent gets narrower. When there is no knowledge of implication rules, the uncertainty interval of the consequent is $[0, 1]$, when knowledge of antecedent and $A \implies B$ is available the interval becomes $[\alpha_1\alpha_r, 1 - \alpha_1(1 - \beta_r)]$, when antecedent and both $A \implies B$, $\bar{A} \implies B$ are available the interval becomes $[\alpha_1\alpha_r + (1 - \beta_1)\bar{\alpha}_r, \alpha_1\beta_r + (1 - \beta_1)\bar{\beta}_r + (\beta_1 - \alpha_1)]$ and the knowledge of consequent is available it is $[\alpha_2, \beta_2]$.

4.2 Proofs

Without loss of generality we assumed that A and B are in same BoE. Therefore;

$$Bl(B) = Bl(B \cap A) + Bl(B \cap \bar{A}) + \sum_{\substack{\emptyset \neq P \subseteq (B \cap A) \\ \emptyset \neq Q \subseteq (B \cap \bar{A})}} m(P \cup Q). \quad (12)$$

$$Bl(B) \geq Bl(B \cap A) + Bl(B \cap \bar{A}), \quad (13)$$

We know that [8] $Bl(A) \leq Bl(B \cap A) + Pl(\bar{B} \cap A)$. This, together with the FH conditionals (where $Bl(A) \neq 0$) can then be used to write

$$Bl(A) Bl(B/A) \leq Bl(B \cap A). \quad (14)$$

This inequality holds true for $Bl(A) = 0$ as well. Substitute \bar{A} for A in (14):

$$Bl(\bar{A}) Bl(B/\bar{A}) \leq Bl(B \cap \bar{A}). \quad (15)$$

Use (13), (14) and (15), and use the fact that $Bl(\bar{A}) = 1 - Pl(A)$ with notation (2) to get (3);

$$\alpha_2 \geq \alpha_1\alpha_r + (1 - \beta_1)\bar{\alpha}_r.$$

Substitute \bar{B} for B in (13),(14), and (15) to get,

$$Bl(\bar{B}) \geq Bl(A) Bl(\bar{B}/A) + Bl(\bar{A}) Bl(\bar{B}/\bar{A}). \quad (16)$$

Use the facts; $Bl(\bar{B}) = 1 - Pl(B)$, $Bl(\bar{B}/A) = 1 - Pl(B/A)$ and $Bl(\bar{B}/\bar{A}) = 1 - Pl(B/\bar{A})$ with notation (2) to get (4);

$$\beta_2 \leq \alpha_1 \beta_r + (1 - \beta_1) \bar{\beta}_r + (\beta_1 - \alpha_1).$$

If the knowledge of $\bar{A} \implies B$ is unavailable, we relax the rule by assuming total uncertainty which is $[\bar{\alpha}_r, \bar{\beta}_r] = [0, 1]$ in (3) and (4) to get (7) and (8);

$$\alpha_2 \geq \alpha_1 \alpha_r.$$

$$\beta_2 \leq 1 - \alpha_1(1 - \beta_r).$$

5 Consistency with Probability and Classical Logic

5.1 Consistency with Probability

For p.m.f.s, we have (a) $Pr(B) = Pr(B \cap A) + Pr(B \cap \bar{A})$, which corresponds to (12), except that the additional summation term vanishes and the inequality (13) reduces to an equality; and (b) $Pr(B \cap A) = Pr(A) Pr(B|A)$, which corresponds to (14), except that the inequality reduces to an equality. Therefore, instead of the bounds for α_2 and β_2 , we get equalities identical to the LC choice.

$$\alpha_2 = \alpha_1 \alpha_r + (1 - \beta_1) \bar{\alpha}_r; \quad \beta_2 = \alpha_1 \beta_r + (1 - \beta_1) \bar{\beta}_r + (\beta_1 - \alpha_1). \quad (17)$$

When the belief and plausibility are equal for each proposition, DST models reduce to p.m.f.s. The belief and plausibility (which are now identical) of each proposition then yield the probability of that same proposition. Suppose the antecedent and the rules $A \implies B$ and $\bar{A} \implies B$ are probabilistic, i.e.,

$$\alpha_1 = \beta_1 = Pr(A); \quad \alpha_r = \beta_r = Pr(B/A); \quad \bar{\alpha}_r = \bar{\beta}_r = Pr(B/\bar{A}). \quad (18)$$

Substitute in (17) to get $\alpha_2 = \beta_2 = \alpha_1 \alpha_r + (1 - \alpha_1) \bar{\alpha}_r$. This corresponds to $Pr(B) = Pr(A) Pr(B|A) + Pr(\bar{A}) Pr(B|\bar{A})$.

5.2 Consistency with Classical Logic

Note that $\alpha_1 = \beta_1 = 1$ and $\alpha_1 = \beta_1 = 0$ imply the occurrence or non-occurrence of proposition A with 100% confidence. Though FH conditionals are not defined when $\alpha_1 = \beta_1 = 0$ [5] we have shown that our relations are valid even $\alpha_1 = \beta_1 = 0$, further if we take a limiting argument and let $\alpha_1 = \beta_1$ tend to 0 in the limit and the result will be same as if we substitute 0 for both α_1 and β_1 in the equations. We associate the two cases $\alpha_1 = \beta_1 = 1$ and $\alpha_1 = \beta_1 = 0$ with the logical ‘Truth’ and logical ‘False’ in classical logic. For example, with $\alpha_1 = \beta_1 = \{0, 1\}$ and $\alpha_2 = \beta_2 = \{0, 1\}$, Table 1 shows the truth table of $A \implies B$. To see the

Table 1. Truth Table for $A \implies B$ in Classical Logic

$\alpha_1 = \beta_1$	$\alpha_2 = \beta_2$	$A \implies B$
0	0	1
0	1	1
1	0	0
1	1	1

consistency with classical logic, we now use $\alpha_1 = \beta_1 = \{0, 1\}$, $\alpha_r = \beta_r = \{0, 1\}$, and $\bar{\alpha}_r = \bar{\beta}_r = \{0, 1\}$ with (5) and (6) (Note that (3) and (4) become (5) and (6) in classical logic case as in probabilistic case). See Table 2. The results for $\alpha_2 = \beta_2$ in Table 2 can be expressed as $(A \wedge (A \implies B)) \vee (\neg A \wedge (\neg A \implies B)) = (A \wedge (\neg A \vee B)) \vee (\neg A \wedge (A \vee B)) = B$.

Table 2. When we have both $A \implies B$ and $\bar{A} \implies B$ for the Classical Logic Case obtained from (5) and (6)

$\alpha_1 = \beta_1$	$\alpha_r = \beta_r$	$\bar{\alpha}_r = \bar{\beta}_r$	α_2	β_2	$\alpha_2 = \beta_2$
0	0	0	0	0	\implies 0
0	0	1	1	1	\implies 1
0	1	0	0	0	\implies 0
0	1	1	1	1	\implies 1
1	0	0	0	0	\implies 0
1	0	1	0	0	\implies 0
1	1	0	1	1	\implies 1
1	1	1	1	1	\implies 1

Now consider the implication case; (9) and (10). (Note that (7) and (8) become (9) and (10) in classical logic case).

Table 3. When we have only $A \implies B$ for the Classical Logic Case obtained from (9) and (10)

$\alpha_1 = \beta_1$	$\alpha_r = \beta_r$	α_2	β_2	$\alpha_2 \neq \beta_2$ in general	
				α_2	β_2
0	0	0	1	\implies 0	1
0	1	0	1	\implies 0	1
1	0	0	0	\implies 0	0
1	1	1	1	\implies 1	1

Let us compare the entries of Table 3 (obtained from (9) and (10)) and Table 1 (truth table for $A \implies B$ in classical logic). (a) *Antecedent is true*: See lines 3-4 where both tables show identical behavior. (b) *Antecedent is false*: See lines 1-2 where the tables behave differently. (b.1) *When rule is true*: the consequent can take 0 or 1 in both tables. Note that the information in lines 1-2 of Table 1 are captured in line 2 of Table 3 which explains when antecedent is false though implication rule is true the consequent can be true or false. (b.2) *When rule is false*: this is not in Table 1 whereas line 1 of Table 3 not only allows this, but it also allows the consequent to take either 0 or 1 value which explains when antecedent is false and the implication rule is false the consequent can be true or false.

From these tables it is clear that the relations we developed are consistent with classical logic as well as they better explain the classical logic behaviour of implication rules than conventional implication truth table.

6 Illustrative Example

As an illustrative simple example, consider red and black balls, and 3 urns A, B, and C: urn A has 3 red, 5 black, plus 2 additional balls; urn B has 5 red, 2 black, plus 3 additional balls; and urn C has 3 red, 5 black, plus 2 additional balls. The additional balls could be any combination of red and black balls.

First we select urn A and randomly take out a ball (first trial). If we get a red ball (RB), we select urn B; otherwise, if we get a black ball (BB), we select urn C. Then we take out a ball from the selected urn (second trial). What are the belief and plausibility values of getting a RB in the second trial?

In the DST framework, let $[\alpha_1, \beta_1]$, $[\alpha_r, \beta_r]$, and $[\alpha_2, \beta_2]$ denote the belief and plausibility values corresponding to getting a RB in the first trial, getting a RB in the second trial given that the first trial yields a RB, and getting a RB in the second trial, respectively. Therefore, $[\alpha_1, \beta_1] = [0.3, 0.5]$, $[\bar{\alpha}_1, \bar{\beta}_1] = [0.5, 0.7]$, $[\alpha_r, \beta_r] = [0.5, 0.8]$, $[\bar{\alpha}_r, \bar{\beta}_r] = [0.3, 0.5]$, and $[\alpha_2, \beta_2] = [0.36, 0.65]$ (computed by accounting for all the possibilities).

Let us now see what our results yield: the general bounds yield $0.30 \leq \alpha_2 \leq \beta_2 \leq 0.69$; the relaxed bounds yield $0.15 \leq \alpha_2 \leq \beta_2 \leq 0.94$. Both these contain $[\alpha_2, \beta_2] = [0.36, 0.65]$. Also note that the general bounds are much tighter than the relaxed bounds (which ignore the information in $[\bar{\alpha}_r, \bar{\beta}_r]$).

7 Conclusion

We have derived mathematical relations to belief and plausibility values of a consequent when the belief and plausibility values of corresponding antecedent and implication rule are given, by modelling the implication rule with DST FH conditionals. The results and their consistency with probability and classical logic demonstrate the reasonability of modelling the uncertain implication rules by DST FH conditionals.

Further the results are more general and flexible than the previous works since the derivations are not imposed by any probabilistic or classical logic relations which are special cases of this model.

Acknowledgment. This work is based on research supported by the US Office of Naval Research (ONR) via grants #N00014-10-1-0140. and #N00014-11-1-0493, and the US National Science Foundation (NSF) via grant #1038257.

References

1. Aregui A, Denoeux T (2008) Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities. *Int. J. of Approx. Reasoning* 49(3):574–594
2. Benavoli A, et al (2008) Modelling uncertain implication rules in evidence theory. In: *Proc. FUSION*, Cologne, Germany, pp 1–7
3. Denneberg D (1994) Conditioning (updating) non-additive measures. *Ann. of Oper. Res.* 52(1):21–42
4. Dubois D, Hullermeier E, Prade H (2001) Toward the representation of implication-based fuzzy rules in terms of crisp rules. In: *Proc. Joint IFSA World Congress and NAFIPS Int. Conf.*, vol 3, pp 1592–1597
5. Fagin R, Halpern JY (1991) A new approach to updating beliefs. In: Bonissone PP, Henrion M, Kanal LN, Lemmer JF (eds) *Proc. UAI*, Elsevier Science, New York, NY, pp 347–374
6. M.L. Ginsberg (1984) Nonmonotonic reasoning using Dempsters rule. *Proc. National Conf. Artificial Intell.* 1984, pp. 126-129.
7. Hau HY, Kashyap RL (1990) Belief combination and propagation in a lattice-structured inference network. *IEEE Trans. on Syst., Man and Cyber.*, 20(1):45–57
8. Kulasekera EC, et al (2004) Conditioning and updating evidence. *Int. J. of Approx. Reasoning* 36(1):75–108
9. Lewis D (1976) Probabilities of conditionals and conditional probabilities. *The Phil. Rev.* LXXXV(3):297–315
10. Nguyen H, Mukaidono M, Kreinovich V (2002) Probability of implication, logical version of Bayes theorem, and fuzzy logic operations. In: *Proc. IEEE Int. Conf. on Fuzzy Syst. (FUZZ-IEEE)*, Honolulu, HI, vol 1, pp 530–535
11. Nunez RC, et al (2013) Modeling uncertainty in first-order logic: a Dempster-Shafer theoretic approach. In: *Proc. Int. Symp. on Imprecise Prob.: Theories and Appl. (ISIPTA)*, Compiègne, France
12. Shafer G (1976) *A Mathematical Theory of Evidence*. Princeton Univ. Press, Princeton, NJ
13. Smets P (1994) What is Dempster-Shafer’s model? In: Yager RR, Fedrizzi M, Kacprzyk J (eds) *Advances in the Dempster-Shafer Theory of Evidence*, John Wiley and Sons, New York, NY, pp 5–34
14. Wickramaratne TL, Premaratne K, Murthi MN (2013) Toward efficient computation of the Dempster-Shafer belief theoretic conditionals. *IEEE Trans. on Cyber.* 43(2):712–724