Probability in relevant frames

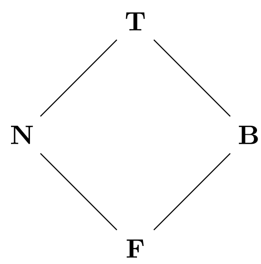
Ondrej Majer
Institute of Philosophy, Czech Academy of Sciences

1 Introduction

There were several proposals in the literature to introduce probability over Belnap-Dunn logic, which is an implication-free core of relevant logics (see [5], [4], [7]). Considerably less attention has been paid to extend probability to relevant implication in order to obtain probabilistic version of full relevant logic. The aim of this paper is to fill this gap. We will propose various possibilities of introducing probability of a relevant conditional and discuss motivations and properties of the proposed definitions with a particular attention on the problem of conditional probability vs. probability of conditional.

2 Belnap-Dunn logic

One of the main motivations for Belnap-Dunn logic (BD) [2] was to introduce a logical system that is able to deal with inconsistent / incomplete information in a non-trivial way. That requires, among other things, rejecting the classical principles Ex Falso Quodlibet and the Law of Excluded Middle. While the language $\mathscr{L}_{\mathsf{BD}}$ is built as usual from a (countable) set At of propositional variables using connectives $\{\wedge, \vee, \neg\}$, it extends the classical values **T,F** with new values **B** (Both true and false) and **N** (Neither true nor false), the designated values are **T, B** ("at least true"). They form the (de Morgan) lattice **4** also called Belnap-Dunn square (see the figure on the left). Alternatively, the truth values can be represented as subsets of $\{T, F\}$: $\mathbf{T} = \{T\}, \mathbf{F} = \{F\}, \mathbf{B} = \{T, F\}, \mathbf{N} = \emptyset$.



Semantics of BD We will use semantics for BD based on the idea of independence of positive and negative information. Formally a BD model $M = \langle S, v^+, v^- \rangle$ consists of a set of states S and two (independent) valuation relations $v^+, v^- : At \to \mathcal{P}(W)$. $v^+(p)$ denotes states containing information supporting p and analogously $v^-(p)$ denotes states rejecting p. As positive and negative valuation are independent, some states might be (with respect to some $p \in At$) *incomplete* ($s \notin v^+(p) \cup v^-(p)$) or *inconsistent* ($s \in v^+(p) \cap v^-(p)$) . The valuations extend to *positive/negative* satisfaction relation. Each satisfaction relation taken separately is in fact classical

$s \models^+ p$ iff $s \in v^+(p)$	$s \models^- p$ iff $s \in v^-(p)$
$s \models^+ \neg\varphi$ iff $s \models^- \varphi$	$s \models^- \neg\varphi$ iff $s \models^+ \varphi$
$s \models^+ \varphi \wedge \psi$ iff $s \models^+ \varphi$ and $s \models^+ \psi$	$s \models^- \varphi \wedge \psi$ iff $s \models^- \varphi$ or $s \models^- \psi$

BD-probability The probabilistic extension of BD logic, as presented in [5], applies the principle of independence to *probabilistic* positive and negative information. It works with two probability functions representing positive (negative) support of a formula respectively.

Definition 2.1. *Belnap-Dunn probability* is a function $\mathtt{p}^+ : \mathscr{L}_{\mathsf{BD}} \to [0, 1]$ which satisfies the following axioms:

(A1) normalization	$0 \leq p^+(\varphi) \leq 1$
(A2) monotonicity	if $\varphi \vdash_{BD} \psi$ then $p^+(\varphi) \leq p^+(\psi)$
(A3) inclusion/exclusion	$p^+(\varphi \vee \psi) = p^+(\varphi) + p^+(\psi) - p^+(\varphi \wedge \psi)$

The negative probability is defined by $\mathtt{p}^-(\varphi) = \mathtt{p}^+(\neg\psi)$. The main difference from the classical Kolmogorovian axioms is that the classical additivity axiom is replaced by weaker inclusion exclusion. We have $p^+(\neg\varphi) \neq 1 - p^+(\varphi)$, so the independence of positive and negative information is implemented; moreover, analogously to the propositional case, it allows for $0 < p^+(\varphi \wedge \neg\varphi)$ (positive probability of classical inconsistency), and $p^+(\varphi \vee \neg\varphi) < 1$ (classical tautologies are not certain).

3 Relevant logics

There are several systems called relevant logic (see, e.g. [8], [9] for an overview). Most of them build on the conjunction/disjunction fragment of BD logic and differ with respect to defining negation and implication. From the point of view of motivation we will have in mind the most well known system R, but from the technical point of view we avoid choosing a particular relevant logic and build on a very weak framework of the full commutative Lambek calculus (see [10]).

We present our framework semantically; we will in fact extend the frame semantics given in the previous section. All relevant logics interpret the implication using a ternary relation R on a set of states S in the following way:

$$\phi \to \psi \text{ is supported in a state } x \in S \text{ iff } y \models \phi \text{ implies } z \models \psi \text{ for all } y, z \text{ s. t. } Rxyz.$$

Formally a (positive) relevant frame contains two more elements – a set $L \subseteq S$ responsible for validity and an information order $\leq$ on S satisfying the following conditions:

Definition 3.1. *Positive relevant frame* is a quadruple $\mathcal{F} = \langle S, L, \leq, R \rangle$, where S, L are a nonempty sets, R is a ternary relation on S and
i) $\leq$ is a partial order on S
ii) $Rxyz$ and $x' \leq x, y' \leq y, z' \geq z$ implies $Rx'y'z'$,
iii) item L is a non-empty, upwards closed subset of $(S, \leq)$, satisfying
$y \leq z$ iff there is an x in L such that $Rxyz$.
iv) if $Rxyz$ then $Ryxz$,

Negation There are essentially two ways of treating negation in relevant logics. The first of them is based on the notion of compatibility of states; two states are compatible if they not contradict each other with respect to the support of some formula. Compatibility is formally represented as a binary relation on states. If we have sCs', then s supports truth of $\neg\phi$ iff s' doesn't support ϕ.

Definition 3.2. A *compatibility frame* is a tuple $\langle S, L, \leq, C, R \rangle$ such that $\langle S, L, \leq, R \rangle$ is a positive relevant frame C is a binary relation on S such that i) xCy implies yCx ii) xCy and $x' \leq x, y' \leq y$ then $x'Cy'$. A compatibility model is a compatibility frame plus a (single) valution v.

Another option is to extend BD models with two ternary relations R_1, R_2: the first one is responsible for the truth of a conditional, the second one for its falsity:

$$\neg(\phi \to \psi) \text{ is supported in a state } x \in S \text{ iff } y \models \phi \text{ and } z \not\models \psi \text{ for some } y, z \text{ s. t. } Sxyz.$$

Definition 3.3. *A Routley frame* is a tuple $\langle S, L, \leq, R_1, R_2 \rangle$ such that $\langle S, L, \leq, R_1 \rangle$ is a positive relevant frame and R_2 is a ternary relation on S satisfying $Rxyz$ and $x \leq x', y \leq y', z' \leq z$ implies $Rx'y'z'$. A Routley model is a Routley frame equipped with a positive and a negative valuation.

4 Probability and conditionals

The problem of defining the probability of a conditional in classical logic has been extensively discussed in the literature. One of the central points of these discussions was the validity of the Adams thesis [1] claiming, that probability of a conditional should be identified with the corresponding conditional probability. The infamous triviality result by David Lewis [6] showed that the applicability of the Adams thesis is limited to a certain fragment of the classical logic. We will discuss this question in the context of relevant frames.

Probability in relevant frames We assume that each state s in a relevant frame is equipped with a "local" probability function $\mathtt{p}_s$ on implication-free formulas that satisfies the axioms (A1) to (A3) of BD-probability, we call such a frame probabilistic. Depending on a particular way of treating negation there might be some more conditions probability should satisfy. We can interpret $\mathtt{p}_s(\phi)$ either as a relative frequency of ϕ observed in the state s or as a subjective probability of an observer who's information state is s. In the rest of this section we discuss various ways to extend the BD-probability to conditionals. One of the main difficulties here is to connect the probability of implication, which is an intensional connective to the local probabilities dealing with extensive connectives (conjunctions and disjunctions).

One straightforward way of defining the probability of a relevant conditional $\varphi \to \psi$ is to read the truth conditions of implication probabilistically: consider all couples $\langle y, z \rangle$ connected to a state x via the relevance relation R such that $y \models \varphi$ and check the proportion of those which support truth of $\varphi \to \psi$:

$$\mathtt{p}_x(\phi \to \psi) = \frac{|\{\langle y, z\rangle|\ Rxyz, y \models \phi, z \models \psi|}{|\{\langle y, z\rangle|\ Rxyz, y \models \phi\}|}$$

Intuitively, this definition tells us what is the probability of reaching a ψ-state in the third R coordinate for a randomly chosen φ-state in the second coordinate. The disadvantage of this solution is that it works only for finite frames and for formulas, where the implication is the main connective. Moreover, it ignores the information given by the local probability.

Alternatively, we might read the ternary relation in relevant frames dynamically as a process of composing pieces of information from R-related states and assume that this relation is non-deterministic, i.e. $R : S^3 \to [0, 1]$ and $\sum R(x, y, z) = 1$ for every $x \in S$. The definition of probability would be given in a way similar to the standard transition systems.

$$\mathtt{p}_x(\varphi \to \psi) = \sum_{y,z \in S} Rxyz \cdot \mathtt{p}_y(\varphi)\mathtt{p}_z(\psi)$$

This corresponds to the chance that my R relation takes me to a couple of states such that I observe φ in the first state and ψ in the second. It is similar to the previous definition, but all steps are random here.

Previous proposals were attempting to represent probability of a conditional. Now we turn to the problem of conditional probability. Having BD probabilities defined at every state we can always define it locally. But conditional probability is a powerful tool of learning new information and we now consider this dynamic reading. Imagine I want to predict from the point of view of a state x probability of ψ in the state z connected with x via the R relation on the condition that I learn probability $\mathtt{p}_y(\varphi)$ at the second state y. I assume the proportion between φ- and ψ-states remains constant during my gathering information hence I can use is my local conditional probability $\mathtt{p}_x(\psi|\varphi) = \frac{\mathtt{p}_x(\psi \wedge \varphi)}{\mathtt{p}_x(\varphi)}$. The only additional input I need is the probability $\mathtt{p}_y(\varphi)$ of φ observed at y. If I define conditional probability as the value of my prediction I obtain:

$$\mathtt{p}_x(\psi|\varphi) = \frac{\mathtt{p}_x(\psi \wedge \varphi)}{\mathtt{p}_x(\varphi)} \cdot \mathtt{p}_y(\varphi)$$

Obviously, we need to prove that conditional probability defined in this way is indeed a probability function. This requires imposing more constraints on the local probabilities regulating their interaction with the basic structures of relevant frames. Also our prediction given by dynamic conditional probability matches the local probability in the targeted state only if the local probabilities are in an appropriate relation with R.

We defined dynamic conditional probability only for triples $\langle x, y, z\rangle$ related by R. In order to define dynamic conditional probability for a particular state x we need to take into account all such values for different y, z. We do it in a way standard in many-valued modal logics, which gives us a lower bound of the predicted value:

$$\mathtt{p}_z(\psi|\varphi) = inf\{\mathtt{p}_{xyz}(\psi|\varphi)|y, z \text{ such that } Rxyz\}$$

.

We proposed several ways of introducing probability to the relevant logic with a special attention to relevant implication. We defined probability of a relevant conditional as well as dynamic conditional probability based on relevant frames. Our proposals still have some drawbacks, namely the applicability to a limited class of formulas, but this is also a drawback of some classical solutions. In the talk we discuss in a more depth the properties and mutual relations of the definitions we introduced. We will also briefly comment on their relation to the framework of relevant epistemic logic presented in [3].

References

[1] Ernest Adams. The logic of conditionals. *Inquiry*, 8(1-4):166–197, 1965.

[2] N.D. Belnap. How a computer should think. In H. Omori and H. Wansing, editors, *New Essays on Belnap-Dunn Logic*, volume 418 of *Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science)*, pages 35–53. Springer, Cham, 2019.

[3] Marta Bílková, Ondrej Majer, Michal Pelis, and Greg Restall. Relevant agents. In *Advances in Modal Logic 8*, pages 22–38, 2010.

[4] J. Michael Dunn. Contradictory information: Too much of a good thing. *Journal of Philosophical Logic*, 39:425–452, 2010.

[5] D. Klein, O. Majer, and S. Rafiee Rad. Probabilities with gaps and gluts. *Journal of Philosophical Logic*, 50(5):1107–1141, October 2021.

[6] David Lewis. Probabilities of conditionals and conditional probabilities. In *IFS: Conditionals, Belief, Decision, Chance and Time*, pages 129–147. Springer, 1976.

[7] E.D. Mares. Paraconsistent probability theory and paraconsistent bayesianism. *Logique et analyse*, pages 375–384, 1997.

[8] Edwin D Mares. *Relevant logic: A philosophical interpretation*. Cambridge University Press, 2004.

[9] Stephen Read. *Relevant logic*. Oxford: Blackwell, 1988.

[10] Greg Restall. *An introduction to substructural logics*. Routledge, 2002.