Probabilities with Gaps and Gluts

Abstract

Belnap-Dunn logic (BD), sometimes also known as First Degree Entailment, is a four-valued propositional logic that complements the classical truth values of *True* and *False* with two non-classical truth values *Neither* and *Both*. The latter two are to account for the possibility of the available information being incomplete or providing contradictory evidence. In this paper, we present a probabilistic extension of BD that permits agents to have probabilistic beliefs about the truth and falsity of a proposition. We provide a sound and complete axiomatization for the framework defined and also identify policies for conditionalization and aggregation. Concretely, we introduce four-valued equivalents of Bayes' and Jeffrey updating and also suggest mechanisms for aggregating information from different sources.

1 Introduction

In learning about a classical system that adheres to the laws of propositional logic, we may be faced with information that does not. Naturally, if information is scarce, our evidence may contain truth value *gaps*, neither indicating certain propositions to be true nor false. But we may also be faced with contradictory information, especially when our insights are gained by combining various bodies of evidence. This may lead to truth value *gluts*, i.e. propositions that are labelled as *both* true and false.

There has been many attempts in the literature to develop formal systems for capturing and analyzing such non-classical situations. These are generally divided into two camps. The first is motivated by adopting the philosophical position of dialetheism as defended by Priest (2006, 2007), advocating the thesis that there are true contradictions, i.e. sentences which are both true and false. Corresponding formal systems should thus allow for assigning both truth values to a sentence simultaneously. Probably the most well known example of such logical systems is the logic LP (Priest, 1979, 2002).

The second camp takes the existence of gaps and gluts as a pathological consequence of imperfect information. Crucially, one may hope than even imperfect

information would allow for at least some reliable inferences. In general, there are two ways to go here. One could either make the set of premises consistent or develop non-trivial inference rules that work on inconsistent sets of premises. Consistency of premises can be obtained by focusing on maximal consistent subsets, cf. Rescher and Manor (1970); Klein and Marra (2020), or by employing belief revision, as in AGM systems (Alchourrón et al., 1985). Mechanisms for dealing with inconsistent information, on the other hand, are developed in a variety of frameworks such as *discussive logic* (Jaskowski, 1948), *adaptive logic* (Batens, 2001), Da Costa's *logics of formal inconsistency* (1974; 1989), *relevant logic* of Anderson and Belnap (1975) and their variants.

Another well-known logical framework that falls in this last category is Belnap-Dunn logic (BD. cf. Belnap, 1977, 2019; Dunn, 1976), sometimes also going by the name of First Degree Entailment. Briefly, this system rests on two assumptions. The first is that gaps and gluts may occur even for boundedly rational agents, as information may be limited (gaps) and the question of whether a given belief set is consistent (i.e. checking for the absence of gluts) is known to be NP-hard. Building on the latter claim, BD's second assumption is that the logic of information should *not* validate the principle of explosion[1]. Just to the contrary, BD stipulates that a body of information may afford us substantial insights about some matter q, even if it contains contradictory information about some other p that is completely unrelated to q. Belnap-Dunn logic, in short, is a substructural logic, that invalidates explosion and tracks which insights can be inferred from an information base that may contain gaps and gluts.

But of course, the problem of insufficient or contradictory information does not apply to categorial *true-false* information only. Rather, probabilistic information is affected by similar arguments about gaps and gluts as those outlined above. In his 1997 paper, Jøsang puts forward a framework for three valued probabilities, incorporating uncertainty as third value that may occur naturally when evidence is ambiguous or insufficient. Notably, this framework circumvents the debated principle of insufficient reason by distinguishing situations of insufficient information from those, where equally strong evidence is available for and against some proposition. Later approaches extend this to four-valued probabilities, where the fourth value represents conflicting information, or gluts. The necessity of gluts is often argued for by considering a Bayesian agent who receives two pieces of mutually contradictory information from sources she judges highly reliable, cf. the firefighter example in Dunn and Kiefer (2019).

In short, these arguments call for a four-valued probabilistic generalization of Belnap-Dunn logic in a similar way as classical probability theory generalizes propositional logic. In a first approach to this project Michael Dunn (2010) has defined a four-valued probabilistic framework and has studied logical properties of the resulting probabilistic entailment. In a similar vein, Childers, Majer and Milne (2019) have put forward a single-valued approach to non-standard probabilities motivated by a frequentist interpretation where probability gaps and gluts may occur naturally if probabilities are derived from sampling two

---

[1]The principle of explosion states that *every* formula can be derived from a contradiction.

independent sources. They further substantiate the approach by providing a subjectivist interpretation of non-standard probabilities and the corresponding Dutch Book Argument.
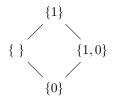
In the present paper, we offer a novel framework for non-standard probabilities that reconciles Dunn's and Childers et al.'s lines of work (Section 3). In doing so, we pursue four major goals. The first is to provide a translation mechanism between four-valued and single-valued non-standard probabilities, showing that these are different but equivalent perspectives on the same phenomenon. (Section 5). The second aim relates to an axiomatization of the system defined. While Dunn (2010) analyzes logical properties of the probabilistic inference relation ensuing from his approach, no axiomatization of the probabilistic system itself has been put forward so far. To fill this gap, we provide an axiomatization of the non-standard probabilities defined here (Section 4) and show this axiomatization sound and complete with respect to a certain class of probabilistic models (Section 6). While building on Dunn's approach, our framework slightly deviates from his in order to avoid certain conceptual problems. As we will show, both, Jøsang's three valued probabilities as well as Dunn's four-valued probability implicitly assume all events to be mutually probabilistically independent. Under this assumption, the question of conditionalization trivializes, as no proposition bears any information about any other proposition. In the present framework, we abandon this independence assumption. Consequentially, the question of conditional probabilities becomes meaningful. Defining and studying an adequate notion of conditionalization is our third goal, pursued in Section 7. The fourth goal, finally, is related to aggregation, i.e. the question of how to combine probabilistic information from various sources. Here, we will introduce various policies and study their respective properties (Section 8).

2 Logical Preliminaries

We start by giving a brief recollection of Belnap-Dunn four-valued logic before proceeding to introduce its probabilistic extensions. Belnap-Dunn four-valued logic is defined over a propositional language that is built over a set Prop of propositional variables. Formally, the logical language $\mathcal{L}_{\text{Prop}}$ is given by the Backus-Naur form:

$$\varphi :: p \,|\, \neg\varphi, \,|\, \varphi \wedge \varphi$$

Disjunction ($\vee$) is defined in the standard way. The main difference to classical propositional logic consists in the way that formulas are evaluated. In classical propositional logic, evaluations are defined as functions $v : \mathcal{L}_{\text{Prop}} \to \{0, 1\}$ that are derived from a valuation on the set of atoms Prop. For Belnap-Dunn logic there are two ways to define evaluations. One approach is to define evaluations as functions $v : \mathcal{L}_{\text{Prop}} \to \mathcal{P}(\{0, 1\})$. In other words instead of evaluating formulas on the two element lattice

3

$$\{1\}$$

$$|$$

$$\{0\}$$

they are interpreted on the four element lattice BD_4

$$\{1\}$$
$$\{\,\}\qquad\{1,0\}$$
$$\{0\}$$

Evaluating formulas in the four element lattice $\mathcal{P}(\{0,1\})$ allows for the assignment of two new truth values $\{\}$ and $\{0,1\}$. These represent so-called truth-value **gaps** and **gluts**, i.e. situations where formulas obtain neither resp. both of the classic truth values. Formally, the evaluation is defined inductively, starting from an atomic valuation $v : \text{Prop} \to \mathcal{P}(\{0,1\})$, by:

$$
\begin{array}{lll}
1 \in v(\neg\phi) & \text{iff} & 0 \in v(\phi) \\
0 \in v(\neg\phi) & \text{iff} & 1 \in v(\phi) \\
1 \in v(\phi \wedge \psi) & \text{iff} & 1 \in v(\phi) \text{ and } 1 \in v(\psi) \\
0 \in v(\phi \wedge \psi) & \text{iff} & 0 \in v(\phi) \text{ or } 0 \in v(\psi)
\end{array}
$$

An alternative approach is to use two separate classical valuations, called the positive valuation v^+ and the negative valuation v^-. Building on atomic valuations $v^+ : \text{Prop} \to \{0,1\}$ and $v^- : \text{Prop} \to \{0,1\}$, these are defined for $\phi, \psi \in \mathcal{L}_{\text{Prop}}$ as:

$$
\begin{array}{lll}
v^+(\neg\phi) = 1 & \text{iff} & v^-(\phi) = 1 \\
v^-(\neg\phi) = 1 & \text{iff} & v^+(\phi) = 1 \\
v^+(\phi \wedge \psi) = 1 & \text{iff} & v^+(\phi) = 1 \text{ and } v^+(\psi) = 1 \\
v^-(\phi \wedge \psi) = 1 & \text{iff} & v^-(\phi) = 1 \text{ or } v^-(\psi) = 1
\end{array}
$$

Both approaches yield equivalent semantics for Belnap-Dunn logic, as is easily seen. For reasons of notational convenience, we will employ the double valuation approach. Within this approach, we can define an entailment relation as $\phi \models_L \psi$ if and only if $v^+(\phi) = 1 \Rightarrow v^+(\psi) = 1$ for all double valuations (v^+, v^-). This entailment relation goes by the name of **first degree entailment**.

An important property of Belnap-Dunn logic, that we will make heavy use of later, is that it admits disjunctive (as well as conjunctive) normal forms. Just as in classical logic, a formula in disjunctive normal form is written as a disjunction of conjunctions of literals. However, unlike in classical logic, an atom might appear both positively and negatively within a conjunctive clause.

4

Theorem 1. *(Theorem 3.9 in Font (1997)) Every formula of Belnap-Dunn logic is equivalent to a formula in a conjunctive (disjunctive) normal form.*

Moreover, up to permutation of conjuncts and disjuncts, formulas in conjunctive (disjunctive) normal form may be identified with finite families of finite sets of literals (Theorem 3.15 in Přenosil (2018)).

3   Probabilistic Models

The double valuation approach's starting assumption is that positive and negative evidence are distinct. That is the absence of positive evidence for some p is not the same as negative evidence against p (or positive evidence for $\neg p$, if you will). In particular, there may be gaps, where neither evidence for p nor against p is available, and gluts, where evidence of both types is present. Within our models, we must hence treat positive and negative evidence separately. In the following, we assume Prop finite and constant. Also, we will denote the set of literals over Prop by Lit, i.e. $\mathrm{Lit} := \mathrm{Prop} \cup \{\neg p \mid p \in \mathrm{Prop}\}$.

Definition 1. A **non-standard model** is a triple $\mathcal{M} = \langle \Sigma, v^+, v^- \rangle$ where Σ is a finite or countably infinite set of states and $v^+, v^- : \Sigma \times \mathrm{Prop} \to \{0, 1\}$ are called the positive (negative) valuation function respectively. For $p \in \mathrm{Prop}$ we let $v^\pm(p) = \{s \in \Sigma \mid v^\pm(s, p) = 1\}$.

Hence, a state s of a model $\mathcal{M}$ might be assigned an inconsistent set of propositions (i.e., $s \in v^+(p) \cap v^-(p)$ for some $p \in Prop$), and may remain undecided about some propositions ($s \notin v^+(q) \cup v^-(q)$ for some $q \in Prop$).

Non-standard models provide a semantics for BD. More specifically, logical formulas of $\mathcal{L}_{\mathrm{Prop}}$ are evaluated on model-state pairs, using relations $\models^+$ and $\models^-$. From this, we then obtain the notions of a positive and negative extension.

Definition 2. Let $\mathcal{M} = \langle \Sigma, v^+, v^- \rangle$ be a non-standard model, $s \in \Sigma$ a state and $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$ be formulas. Then
i) The **semantics** of $\mathcal{L}_{\mathrm{Prop}}$ on $(\mathcal{M}, s)$ is given by:

$$
\begin{aligned}
\mathcal{M}, s \models^+ p && \text{iff } s \in v^+(p) \\
\mathcal{M}, s \models^- p && \text{iff } s \in v^-(p) \\
\mathcal{M}, s \models^+ \varphi \wedge \psi && \text{iff } \mathcal{M}, s \models^+ \varphi \text{ and } \mathcal{M}, s \models^+ \psi \\
\mathcal{M}, s \models^- \varphi \wedge \psi && \text{iff } \mathcal{M}, s \models^- \varphi \text{ or } \mathcal{M}, s \models^- \psi \\
\mathcal{M}, s \models^+ \neg\varphi && \text{iff } \mathcal{M}, s \models^- \varphi \\
\mathcal{M}, s \models^- \neg\varphi && \text{iff } \mathcal{M}, s \models^+ \varphi
\end{aligned}
$$

ii) The **positive** and **negative extensions** of $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ are

$$
\begin{aligned}
|\varphi|_{\mathcal{M}}^+ &= \{s \in \Sigma \mid \mathcal{M}, s \models^+ \varphi\} \\
|\varphi|_{\mathcal{M}}^- &= \{s \in \Sigma \mid \mathcal{M}, s \models^- \varphi\} \quad (= \{s \in \Sigma \mid \mathcal{M}, s \models^+ \neg\varphi\})
\end{aligned}
$$

We define the entailment relation between sentences in the usual way: $\phi \models^{\pm} \psi$ if and only if for all models $\mathcal{M}$ and states σ, if $\mathcal{M}, s \models^{\pm} \phi$ then $\mathcal{M}, s \models^{\pm} \psi$. Observe the obvious connection between positive and negative extension: $|\neg\varphi|_{\mathcal{M}}^{+} = |\varphi|_{\mathcal{M}}^{-}$. Moreover, we define the **set of pure belief, pure disbelief, conflict** and **uncertainty** about φ as

$$|\varphi|_{\mathcal{M}}^{b} = |\varphi|_{\mathcal{M}}^{+} \backslash |\varphi|_{\mathcal{M}}^{-} \qquad\qquad |\varphi|_{\mathcal{M}}^{d} = |\varphi|_{\mathcal{M}}^{-} \backslash |\varphi|_{\mathcal{M}}^{+}$$
$$|\varphi|_{\mathcal{M}}^{c} = |\varphi|_{\mathcal{M}}^{+} \cap |\varphi|_{\mathcal{M}}^{-} \qquad\qquad |\varphi|_{\mathcal{M}}^{u} = \Sigma \backslash (|\varphi|_{\mathcal{M}}^{+} \cup |\varphi|_{\mathcal{M}}^{-}).$$

The terms *belief* and *disbelief*, of course, refer to the intended interpretation as doxastic state. Whenever clear by context, we omit the subscript $\mathcal{M}$.

Towards a semantics of non-standard probability theory, we expand the non-standard model defined above with a probability measure that is classic. Non-classicality of the ensuing probability assignments, then, will be derived from the underlying valuations only, i.e. from the fact that non-standard models allow for gaps and gluts of truth values.

Definition 3. A **probabilistic model** is a tuple $\mathcal{M} = \langle \Sigma, \mu, v^{+}, v^{-} \rangle$ where $\langle \Sigma, v^{+}, v^{-} \rangle$ is a non-standard model and μ is a probability measure on the full subset algebra of Σ.

Building on probabilistic models, we can derive two different probability assignments from $\mathcal{M}$, one four-valued, the other single valued. These are:

Definition 4. For a probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^{+}, v^{-} \rangle$,
i) the induced **non-standard probability function** $p_{\mu} : \mathcal{L}_{\text{Prop}} \to \mathbb{R}$ is:

$$p_{\mu}(\varphi) = \mu(|\varphi|_{\mathcal{M}}^{+})$$

ii) the induced **four-valued probability function** $\hat{p}_{\mu} : \mathcal{L}_{\text{Prop}} \to \mathbb{R}^{4}$ is

$$\hat{p}_{\mu}(\varphi) = \left(\mu(|\varphi|^{b}), \mu(|\varphi|^{d}), \mu(|\varphi|^{u}), \mu(|\varphi|^{c}) \right).^{2}$$

To end this section, we'd like to highlight a strong similarity to classic probabilistic models. Classic probability assignments can be derived from *possible worlds models* equipped with a probability function, i.e. finite classical models akin to those in Definition 3. More explicitly, for a classical model of the form $\mathcal{M} = \langle W, v, \mu \rangle$ with W a set of possible worlds, $v : W \times \text{Prop} \to \{0, 1\}$ a valuation, and $\mu : \mathcal{P}(W) \to [0; 1]$ a probability measure, the probability of some φ is given as $\mu([\varphi])$, with $[\varphi] = \{s \in \Sigma : \mathcal{M}, s \models \varphi\}$. In fact, if Prop is finite, every probability assignment to $\mathcal{L}_{\text{Prop}}$ can be obtained in this way.

Moreover, every world w of a possible worlds models W naturally corresponds to its atomic valuation, which can be represented by the subset $V \subseteq \text{Prop}$ given by $p \in V$ iff $v(w, p) = 1$ for $p \in \text{Prop}$. In the same vein, each state

---

[2]We use the following convention in naming of probability measures i) p-like names stand for syntactic measures (on languages) and μ-like names for measures on spaces. ii) Hat-superscripts are used to denote four-valued probabilities. iii) The Subscript μ may be used if a syntactic measure is derived from a space.

σ of a probabilistic model corresponds to a non-standard possible assignment $n_s \subseteq \mathcal{P}(\mathrm{Lit})$ defined by $p \in n_s$ iff $v^+(w, p) = 1$ and $\neg p \in n_s$ iff $v^-(w, p) = 1$ for $p \in At$. Hence, non standard probabilistic models are obtained from possible world models by replacing classical worlds, i.e. atomic valuations with BD-possible worlds, that is elements of $\mathcal{P}(\mathrm{Lit})$.

4   Axioms of Non-standard probability

In the following, we present a number of axioms for non-standard and four-valued probabilities. The two sets of axioms given here are easily seen to be sound w.r.t. to the semantics just presented. That they are also complete will be shown in Section 6. We can hence use these axioms for a purely syntactic definition of non-standard and four-valued probabilities.

Non-standard probabilities

We begin with axioms for single-valued non-standard probabilities, i.e. probability measures assigning each $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ a unique rational number.

Definition 5. A **non-standard probability assignment** is a function $p \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ satisfying for all $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$.

(A1)	$0 \leqslant p(\varphi) \leqslant 1$	(normalization)
(A2)	if $\varphi \models_L \psi$ then $p(\varphi) \leqslant p(\psi)$	(monotonicity)
(A3)	$p(\varphi \wedge \psi) + p(\varphi \vee \psi) = p(\varphi) + p(\psi).$	(import-export rule)

where $\models_L$ in (A2) is the entailment relation of Belnap-Dunn logic (first-degree entailment).

These axioms are strictly weaker than the classic Kolmogorov axioms (Kolmogorov, 2018). Axioms (A1)-(A3) can be derived from the Kolmogorov axioms, using that first degree entailment is a sub-relation of classical entailment. In the converse direction, however, only the non-negativity axiom ($p(\varphi) \geqslant 0$ for all φ) is derivable from (A1). Neither Kolmogorov's unit axiom ($p(\top) = 1$) nor the (σ)-additivity axioms are derivable from (A1)-(A3), as is illustrated by the fact that assigning probability .5 to *every* formula satisfies (A1)-(A3). In fact, the import-export axiom is a weak counterpart to additivity, stating that a general rule for adding probabilities that is derivable from the Kolmogorov axioms, $p(\varphi \vee \psi) = p(\varphi) + p(\psi) - p(\varphi \wedge \psi)$, continues to hold. Within the above axiomatization, the import-export axioms (A3) is the only condition regulating the relation between the probability of a formula and its negation. As a result the probabilities of φ and $\neg\varphi$ need not sum up to 1. The constraint $p(\varphi \vee \neg\varphi) + p(\varphi \wedge \neg\varphi) = p(\varphi) + p(\neg\varphi)$ allows for probabilistic gaps ($p(\varphi \vee \neg\varphi) < 1$) and gluts ($p(\varphi \wedge \neg\varphi) > 0$) to occur simultaneously. This squares with our original motivation of establishing independence between positive and negative evidence.

Four-Valued probabilities

We now turn to four-valued probability assignments. These are characterized by a total of six axioms.

Definition 6. A **four-valued probability assignment** is a function $\hat{p}$: $\mathcal{L}_{\text{Prop}} \to \mathbb{R}^4$. Writing $\hat{\varphi}$ as $(b_\varphi, d_\varphi, u_\varphi, c_\varphi)$, this function must satisfy

(D1) $0 \leqslant b_\varphi, d_\varphi, u_\varphi, c_\varphi$
(D2) $b_\varphi + d_\varphi + u_\varphi + c_\varphi = 1$
(D3) $b_{\neg\varphi} = d_\varphi,\ c_{\neg\varphi} = c_\varphi$
(D4) if $\varphi \vDash_L \psi$ then $b_\varphi + c_\varphi \leqslant b_\psi + c_\psi$
(D5) $b_{\varphi \wedge \neg\varphi} = 0,\ c_{\varphi \wedge \neg\varphi} = c_\varphi$
(D6) $b_\varphi + c_\varphi + b_\psi + c_\psi = b_{\varphi \wedge \psi} + c_{\varphi \wedge \psi} + b_{\varphi \vee \psi} + c_{\varphi \vee \psi}$

where $\vDash_L$ is first-degree entailment and $\varphi, \psi \in \mathcal{L}_{\text{Prop}}$

The four entries of $\hat{p}$ stand for pure belief (i.e. φ is true and $\neg\varphi$ is not), pure disbelief, uncertainty and conflict respectively. Let us briefly explain the axioms. The first two axioms (D1) and (D2) are classicality axioms, stating that probabilities are non-negative and that the probabilistic masses of pure belief, pure disbelief, conflict and uncertainty must add up to 1. This reflects the intuition that the four cases are mutually exclusive and jointly exhaustive, i.e. that the metatheory of gaps and gluts is classical.

Axioms (D3)-(D6) then represent structural relations between the four-valued assignments. (D3) emphasizes the strong relation between φ and $\neg\varphi$: belief in one is the same as disbelief in the other, while both share the same conflict and uncertainty. (D4) is a direct counterpart of axioms (A2) above, stating that the total belief in φ (i.e. the sum of pure belief in φ and belief in φ and $\neg\varphi$ together) must be monotonous under first degree entailment. (D5) expresses that an agent cannot have pure belief in contradictory formulas of the form $\varphi \wedge \neg\varphi$. A fortiori, the conflict about $\varphi \wedge \neg\varphi$ must be derived from (and equal to) conflict about φ alone. (D6), finally, is a counterpart to the import-export axiom (A3). Briefly, it states the total beliefs (i.e. the sum of pure belief and conflict together) of $\varphi, \psi, \varphi \vee \psi$ and $\varphi \wedge \psi$ must satisfy the import-export rule.

We should note that the axioms presented here are weaker than those put forward in Dunn (2010). There, the probability of a conjunction $\varphi \wedge \psi$ is determined by its conjuncts through:

$$b_{\varphi \wedge \psi} = b_\varphi \cdot b_\psi \qquad\qquad d_{\varphi \wedge \psi} = d_\varphi + d_\psi - d_\varphi d_\psi + c_\varphi u_\psi + u_\varphi c_\psi$$
$$u_{\varphi \wedge \psi} = u_\varphi b_\psi + b_\varphi u_\psi + u_\varphi u_\psi \qquad c_{\varphi \wedge \psi} = b_\varphi c_\psi + c_\varphi b_\psi + c_\varphi c_\psi$$

A similar axiom for three valued probabilities (true/false/uncertain) can be found in Jøsang (1997). Notably, such definition makes conjunctions truth functional, i.e. the probability of $\varphi \wedge \psi$ is fully determined by the probabilities of φ and ψ. We take this to be too strong, especially given that no such functional dependence holds in classic probability theory. Moreover this truth functional approach implies that all propositions are mutually probabilistically independent -

8

precluding any interesting notions of conditionalization. To see this, assume that φ and ψ are classical, i.e. $\hat{p}(\varphi) = (b_\varphi, d_\varphi, 0, 0)$ and $\hat{p}(\psi) = (b_\psi, d_\psi, 0, 0)$. Then the above definition simplifies to $\hat{p}(\varphi \wedge \psi) = (b_\varphi b_\psi, d_\varphi + d_\psi - d_\varphi d_\psi, 0, 0)$. With other words, the probability (belief) in $\varphi \wedge \psi$ is the product of the probabilities of φ and ψ - which exactly is the definition of probabilistic independence.

In the following section we will show a strong correspondence between non-standard and four-valued probability assignments. Thereafter, we show axiom systems (A1)-(A3) and (D1)-(D6) to be sound and complete with respect to the class of probabilistic models defined above (Section 6). In Section 7 we then discuss approaches to conditionalization in either setting.

5 Correspondence between non-standard and four-valued probabilities

We have so far presented two different frameworks for non-standard probability, one real-valued, the other with values in $\mathbb{R}^4$. As we show now, both are different but equivalent perspectives on the same phenomenon. To this end, let $\mathbb{P}_{ns}$ and $\mathbb{P}_4$ be the set of non-standard and four-valued probability assignments respectively. That is, $\mathbb{P}_{ns}$ is the set of functions $\mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ satisfying (A1)-(A3) while $\mathbb{P}_4$ consists of all mappings $\mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ satisfying (D1)-(D6). We will show the translation map $tr^4_{ns} \colon \mathbb{P}_4 \to \mathbb{P}_{ns}$ defined by

$$tr^4_{ns}(\hat{p})(\varphi) := b_\varphi + c_\varphi \qquad \text{where } \hat{p}(\varphi) = (b_\varphi, d_\varphi, u_\varphi, c_\varphi)$$

to be a bijection. In the opposite direction, the map $tr^{ns}_4 \colon \mathbb{P}_{ns} \to \mathbb{P}_4$ is given by

$$tr^{ns}_4(p)(\varphi) := (p(\varphi) - p(\varphi \wedge \neg\varphi),\ p(\neg\varphi) - p(\varphi \wedge \neg\varphi),$$
$$1 - p(\varphi) - p(\neg\varphi) + p(\varphi \wedge \neg\varphi),\ p(\varphi \wedge \neg\varphi))$$

As expected, the maps tr^4_{ns} and tr^{ns}_4 are inverse to each other:

Theorem 2. tr^4_{ns} and tr^{ns}_4 are well-defined. Moreover $tr^{ns}_4 \circ tr^4_{ns} = id_{\mathbb{P}_4}$ and $tr^4_{ns} \circ tr^{ns}_4 = id_{\mathbb{P}_{ns}}$

Moreover, the translation maps tr^4_{ns} and tr^{ns}_4 cohere with the way we defined non-standard and four-valued assignments on a given probabilistic model.

Theorem 3. Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model and p_μ and $\hat{p}_\mu$ the induced non-standard and four-valued probability functions. Then $tr^4_{ns} \circ \hat{p}_\mu = p_\mu$ and $tr^{ns}_4 \circ p_\mu = \hat{p}_\mu$.

The remainder of this section is devoted to showing these two results.

Proof of Theorem 2. To see that tr^4_{ns} is well defined let $p' = tr^4_{ns}(\hat{p})$ for a fixed $\hat{p} \in \mathbb{P}_4$. First, note that for any $\psi \in \mathcal{L}_{\mathrm{Prop}}$ with $\hat{p}(\psi) = (b_\psi, d_\psi, u_\psi, c_\psi)$ we have $0 \leqslant b_\psi + c_\psi \leqslant 1$, showing that p' satisfies (A1). To see that p' satisfies (A2)

assume that $\varphi \models_L \psi$. By (D4), we have that $b_\varphi + c_\varphi \leqslant b_\psi + c_\psi$ and hence $p'(\varphi) \leqslant p'(\psi)$. For (A3) finally, note that by (D6) we have for any $\varphi, \psi \in \mathcal{L}_{\text{Prop}}$ that $b_\varphi + c_\varphi + b_\psi + c_\psi = b_{\varphi \wedge \psi} + c_{\varphi \wedge \psi} + b_{\varphi \vee \psi} + c_{\varphi \vee \psi}$ which immediately implies that $p'(\varphi) + p'(\psi) = p'(\varphi \wedge \psi) + p'(\varphi \vee \psi)$.

Next, we show that also tr_4^{ns} is well defined. For this fix $p \in \mathbb{P}_{ns}$. For $\psi \in \mathcal{L}_{\text{Prop}}$ denote $tr_{ns}^4(p)(\psi)$ by $(b_\psi, d_\psi, u_\psi, c_\psi)$. Using this notation, we obtain

$$\begin{aligned} b_\psi + d_\psi + u_\psi + c_\psi = \; & p(\psi) - p(\psi \wedge \neg\psi) + p(\neg\psi) - p(\psi \wedge \neg\psi) \\ & + 1 - p(\psi) - p(\neg\psi) + p(\psi \wedge \neg\psi) + p(\psi \wedge \neg\psi) \end{aligned}$$

the latter term is easily seen to equal 1, showing (D2). For (D1) note that $\psi \wedge \neg\psi \models_L \psi$ and $\psi \wedge \neg\psi \models_L \neg\psi$. By (A2), we have that $p(\psi \wedge \neg\psi) \leqslant p(\psi), p(\neg\psi)$ which, together with (A1) implies that $b_\psi, d_\psi, c_\psi \geqslant 0$. Finally, by (A1) and (A3),

$$1 - p(\psi) - p(\neg\psi) + p(\psi \wedge \neg\psi) \geqslant p(\psi \vee \neg\psi) - p(\psi) - p(\neg\psi) + p(\psi \wedge \neg\psi) = 0$$

and hence $u_\psi \geqslant 0$. The first half of (D3) follows from the fact that $b_\psi = p(\psi) - p(\psi \wedge \neg\psi) = d_{\neg\psi}$, using that $\psi =\models_L \neg\neg\psi$ and hence, by (A2), $p(\psi) = p(\neg\neg\psi)$. The second half follows from the fact that $\psi \wedge \neg\psi =\models_L \neg\psi \wedge \neg\neg\psi$ and hence, by (A2), $p(\psi \wedge \neg\psi) = p(\neg\psi \wedge \neg\neg\psi)$. Similarly, (D4) can be derived from (A2) together with the fact that $b_\psi + c_\psi = p(\psi) - p(\psi \wedge \neg\psi) + p(\psi \wedge \neg\psi) = p(\psi)$. Using the latter fact again, (D6) is an immediate consequence of (A3). For (D5), finally, note that $\psi \wedge \neg\psi =\models_L \psi \wedge \neg\psi \wedge \neg(\psi \wedge \neg\psi)$ and hence, by (A2), $p(\psi \wedge \neg\psi) = p(\psi \wedge \neg\psi \wedge \neg(\psi \wedge \neg\psi))$. This implies that $c_{\psi \wedge \neg\psi} = c_\psi$ and that $b_{\psi \wedge \neg\psi} = p(\psi \wedge \neg\psi) - p(\psi \wedge \neg\psi \wedge \neg(\psi \wedge \neg\psi)) = 0$.

Finally, we show that $tr_4^{ns} \circ tr_{ns}^4 = id_{\mathbb{P}_4}$ and $tr_{ns}^4 \circ tr_4^{ns} = id_{\mathbb{P}_{ns}}$, i.e. that tr_4^{ns} and tr_{ns}^4 are left and right inverses of each other. We begin by showing that $tr_{ns}^4(tr_4^{ns}(p)) = p$ for any $p \in \mathbb{P}_{ns}$. For $\varphi \in \mathcal{L}_{\text{Prop}}$, we have that $tr_4^{ns}(p)(\varphi)$ equals

$$\left(p(\varphi) - p(\varphi \wedge \neg\varphi), p(\neg\varphi) - p(\varphi \wedge \neg\varphi), 1 - p(\varphi) - p(\neg\varphi) + p(\varphi \wedge \neg\varphi), p(\varphi \wedge \neg\varphi)\right).$$

Hence $tr_{ns}^4(tr_4^{ns}(p))(\varphi) = p(\varphi) - p(\varphi \wedge \neg\varphi) + p(\varphi \wedge \neg\varphi) = p(\varphi)$ as desired.

For the converse direction, let $\hat{p} \in \mathbb{P}_4$. We have to show that $tr_4^{ns}(tr_{ns}^4(\hat{p})) = \hat{p}$. For this, let $\varphi \in \mathcal{L}_{\text{Prop}}$ and denote $\hat{p}(\psi)$ by $(b_\psi, d_\psi, c_\psi, u_\psi)$ for any $\psi \in \mathcal{L}_{\text{Prop}}$. By axioms (D3) and (D5) we have that $b_{\neg\varphi} = d_\varphi$, $c_{\neg\varphi} = c_\varphi$, $b_{\varphi \wedge \neg\varphi} = 0$ and $c_{\varphi \wedge \neg\varphi} = c_\varphi$. Hence, the values of $tr_{ns}^4(\hat{p})(\varphi)$, $tr_{ns}^4(\hat{p})(\neg\varphi)$ and $tr_{ns}^4(\hat{p})(\varphi \wedge \neg\varphi)$ are $b_\varphi + c_\varphi$, $d_\varphi + c_\varphi$ and c_φ respectively. We then get that

$$\begin{aligned} & tr_4^{ns}(tr_{ns}^4(\hat{p}))(\varphi) \\ = & (tr_{ns}^4(\hat{p})(\varphi) - tr_{ns}^4(\hat{p})(\varphi \wedge \neg\varphi), tr_{ns}^4(\hat{p})(\neg\varphi) - tr_{ns}^4(\hat{p})(\varphi \wedge \neg\varphi), \\ & \quad tr_{ns}^4(\hat{p})(\varphi \wedge \neg\varphi), 1 - tr_{ns}^4(\hat{p})(\varphi) - tr_{ns}^4(\hat{p})(\neg\varphi) + tr_{ns}^4(\hat{p})(\varphi \wedge \neg\varphi)) \\ = & (b_\varphi + c_\varphi - c_\varphi, d_\varphi + c_\varphi - c_\varphi, c_\varphi, 1 - (b_\varphi + c_\varphi) - (d_\varphi + c_\varphi) + c_\varphi) \\ = & (b_\varphi, d_\varphi, c_\varphi, 1 - b_\varphi - d_\varphi - c_\varphi) = (b_\varphi, d_\varphi, c_\varphi, u_\varphi) \end{aligned}$$

where the last equation employs (D2). Hence $tr_{ns}^4(tr_4^{ns}(\hat{p})) = \hat{p}$ as desired. $\quad\square$

Proof of Theorem 3. For $\varphi \in \mathcal{L}_{\text{Prop}}$ denote $\hat{p}_\mu(\varphi)$ by $(b_\varphi, d_\varphi, u_\varphi, c_\varphi)$. By Definition 3, we have

$$b_\varphi = \mu(|\varphi|_{\mathcal{M}}^b) = \mu(|\varphi|_{\mathcal{M}}^+ \backslash |\varphi|_{\mathcal{M}}^-)$$
$$c_\varphi = \mu(|\varphi|_{\mathcal{M}}^c) = \mu(|\varphi|_{\mathcal{M}}^+ \cap |\varphi|_{\mathcal{M}}^-)$$

Hence,

$$tr_{ns}^4(\hat{p}_\mu)(\varphi) = b_\varphi + c_\varphi = \mu(|\varphi|_{\mathcal{M}}^b) + \mu(|\varphi|_{\mathcal{M}}^c)$$
$$= \mu(|\varphi|_{\mathcal{M}}^+ \backslash |\varphi|_{\mathcal{M}}^-) + \mu(|\varphi|_{\mathcal{M}}^+ \cap |\varphi|_{\mathcal{M}}^-) = \mu(|\varphi|_{\mathcal{M}}^+)$$

By definition, the latter term is exactly $p_\mu(\varphi)$. Thus $tr_{ns}^4 \circ \hat{p}_\mu = p_\mu$, as desired.

Moreover, the latter formula implies that $tr_4^{ns} \circ tr_{ns}^4 \circ \hat{p}_\mu = tr_4^{ns} \circ p_\mu$. By Theorem 2, we have have $tr_4^{ns} \circ tr_{ns}^4 = id_{\mathbb{P}_4}$. Hence, the last equation reduces to $\hat{p}_\mu = tr_4^{ns} \circ p_\mu$, proving the second part of the theorem. $\square$

6 A Completeness Result.

Having shown that non-standard and four-valued probability assignments are equivalent, as witnessed by the bijection $tr_{ns}^4 \colon \mathbb{P}_4 \to \mathbb{P}_{ns}$, we now turn our attention to the class of probability functions that are induced by probabilistic models. As it turns out, these are fully characterized by our axioms (A1)-(A3). More specifically, we will show that axioms (A1)-(A3) are a sound and complete characterization of the induced non-standard probability functions of probabilistic models. Of course, by Theorems 2 and 3, this implies that also (D1)-(D6) are a sound and complete characterization of the induced four-valued probability functions of probabilistic models. In fact, the soundness part is easy to check:

Lemma 1. *Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model and p_μ the induced non-standard probability function. Then p_μ satisfies (A1)-(A3).*

Towards completeness, we will show a stronger result. Recall that *completeness* expresses that every $p \in \mathbb{P}_{ns}$ is the induced non-standard probability function of *some* probabilistic model $\mathcal{M}$. This $\mathcal{M}$ may, however, not be unique as p may be not expressive enough to completely determine all properties of $\mathcal{M}$. As we will show $\mathcal{M}$, is almost unique. More specifically, we determine a class $\mathbb{M}^{can}$ of canonical models such that every $p \in \mathbb{P}_{ns}$ is the induced non-standard probability function of *exactly one* $\mathcal{M} \in \mathbb{M}^{can}$.

Definition 7. *i)* We call a probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ **canonical** iff $\Sigma = \mathcal{P}(\text{Lit})$ and v^+, v^- satisfy

$$v^+(p) = \{\sigma \in \mathcal{P}(\text{Lit}) \mid p \in \sigma\} \qquad v^-(p) = \{\sigma \in \mathcal{P}(\text{Lit}) \mid \neg p \in \sigma\}$$

ii) $\mathbb{M}^{\textbf{can}}$ is the set of canonical probabilistic models.

Remark: The set $\mathbb{M}^{can}$ is representative of the set of all models in the following sense: For any probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$, there is a unique canonical model $\mathcal{M}_c = \langle \mathcal{P}(\mathrm{Lit}), \mu_c, v_c^+, v_c^- \rangle$ and a unique function $f \colon \mathcal{M} \to \mathcal{M}_c$ such that $x \in v^\pm(p) \Leftrightarrow f(x) \in v_c^\pm(p)$ and $\mu_c(\sigma_c) = \mu(f^{-1}(\sigma_c))$ for all $\sigma_c \in \mathcal{P}(\mathrm{Lit})$. In particular, $p_\mu(\varphi) = p_{\mu_c}(\varphi)$ for all $\varphi \in \mathcal{L}_{\mathrm{Prop}}$. The main theorem of this section is:

Theorem 4. *For any $p \in \mathbb{P}_{ns}$ there is a unique canonical model $\mathcal{M}_p = \langle \mathcal{P}(\mathrm{Lit}), \mu, v^+, v^- \rangle$ with induced non-standard probability function p_μ such that $p = p_\mu$.*

Corollary 1. *Axioms (A1)-(A3) are sound and complete with respect to the class of induced non-standard probability functions of probabilistic models.*

By Theorems 2 and 3, the previous result readily translates to the level of four-valued probability functions.

Theorem 5. *For any $\hat{p} \in \mathbb{P}_4$ there is a unique canonical model $\mathcal{M}_{\hat{p}} = \langle \mathcal{P}(\mathrm{Lit}), \mu, v^+, v^- \rangle$ with induced four-valued probability function $\hat{p}_\mu$ such that $\hat{p} = \hat{p}_\mu$.*

Corollary 2. *Axioms (D1)-(D6) are sound and complete with respect to the class of induced four-valued probability functions of probabilistic models.*

Proof of Theorem 4. Fix $p \in \mathbb{P}_{ns}$. Let $\Sigma = \mathcal{P}(\mathrm{Lit})$ and let $v^\pm \colon \mathrm{Prop} \to \mathcal{P}(\Sigma)$ be defined as $v^+(q) = \{\sigma \in \Sigma \mid q \in \sigma\}$ and $v^-(q) = \{\sigma \in \Sigma \mid \neg q \in \sigma\}$ respectively. We will construct a classic probability function $\mu \colon \mathcal{P}(\Sigma) \to [0;1]$ such that the canonical model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ satisfies $p_\mu = p$. It suffices to construct the underlying probability mass function $W \colon \Sigma \to [0;1]$, i.e. the function satisfying $W(x) = \mu(\{x\})$ for $x \in \Sigma$. We will do so by induction on $|x|$ for $x \in \Sigma = \mathcal{P}(\mathrm{Lit})$. The construction proceeds in three steps. As an induction base, we set $\mu(x_{max})$ with x_{max} the unique element in Σ with $|x_{max}| = |\mathrm{Lit}|$. In the induction step, we define $\mu(x)$ for all x with $|x| = k \geqslant 1$, assuming that $\mu(y)$ has already been defined for all y with $|y| > k$. In the last step, finally, we define $\mu(\varnothing)$, where $\varnothing$ is the unique element of Σ of cardinality 0.

We will need to ensure that that $\mu([\varphi]) = p(\varphi)$ for all $\varphi \in \mathcal{L}_{\mathrm{Prop}}$, where $[\varphi]$ denotes the truth set of φ in the non-standard model $\langle \Sigma, v^+, v^- \rangle$, i.e. $[\varphi] = \{x \subseteq \mathrm{Lit} \mid \bigwedge_{q \in x} q \models_L \varphi\}$. Note that by the normal form theorem (Theorem 1) and axiom (A2), it suffices to show this property for all $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ that are in disjunctive normal form. Moreover note that for any $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$ in disjunctive normal form, we have that $\mu([\varphi \vee \psi]) = \mu([\varphi]) + \mu([\psi]) - \mu([\varphi \wedge \psi])$, as witnessed by

$$\mu([\varphi \vee \psi]) = \sum_{x \models \varphi \vee \psi} \mu(x) = \sum_{x \models \varphi} \mu(x) + \sum_{x \models \psi} \mu(x) - \sum_{x \models \varphi \wedge \psi} \mu(x)$$
$$= \mu([\varphi]) + \mu([\psi]) - \mu([\varphi \wedge \psi]).$$

By (A3), hence, knowing that $\mu([*]) = p(*)$ for $* \in \{\varphi, \psi, \varphi \wedge \psi\}$ guarantees that $\mu([\varphi \vee \psi]) = p(\varphi \vee \psi)$. It thus suffices to show that $\mu([\varphi]) = p(\varphi)$ whenever φ

12

is a conjunction of literals, i.e. of the form $\bigwedge_{q \in x} q$ with $x \subseteq \mathrm{Lit}$. We will show this property to hold alongside our inductive construction.

For the first step, let x_{max} be $\bigwedge_{q \in \mathrm{Lit}} q$, the unique element in $\mathcal{P}(\Sigma)$ of maximal cardinality. Note that $\{x_{max}\}$ is the truth set of the formula $\bigwedge_{q \in \mathrm{Lit}} q$. We thus set $W(x_{max}) := p(\bigwedge_{q \in \mathrm{Lit}} q)$. By axiom (A1) we have that $0 \leqslant W(x_{max}) \leqslant 1$.

For the inductive step let $k \geqslant 1$ and assume that $W(y)$ has already been defined for all y with $|y| > k$. We simultaneously define $W(x)$ for all $x \subseteq \mathrm{Lit}$ with $|x| = k$. Let such x be given. Note that the truth set of $\bigwedge_{q \in x} q$ is $\{y \subseteq \mathrm{Lit} \mid x \subseteq y\}$. By induction assumption $W(y)$ is already defined for all $\{y \subseteq \mathrm{Lit} \mid x \subset y\}$. We can hence define

$$W(x) := p(\bigwedge_{q \in x} q) - \sum_{\{y \subseteq \mathrm{Lit} \mid x \subset y\}} W(y). \tag{1}$$

We have that $W(x) \leqslant p(\bigwedge_{q \in x} q)$ and thus $W(x) \leqslant 1$. On the other hand, note that $\{y \subseteq \mathrm{Lit} \mid x \subset y\}$ is the truth set of $\bigvee_{y \supset x} \bigwedge_{q \in y} q$. Hence, by induction assumption,

$$\sum_{\{y \subseteq \mathrm{Lit} \mid x \subset y\}} W(y) = p(\bigvee_{y \supset x} \bigwedge_{q \in y} q). \tag{2}$$

Moreover, note that $\bigvee_{y \supset x} \bigwedge_{q \in y} q \models \bigwedge_{q \in x} q$ and hence, by (A2), $p(\bigvee_{y \supset x} \bigwedge_{q \in y} q) \leqslant p(\bigwedge_{q \in x} q)$. Combining this inequality with (1) and (2) yields $W(x) \geqslant 0$.

For the last step, finally, assume that $W(x)$ is already defined for all $x \neq \varnothing$. We then set $W(\varnothing) = 1 - \sum_{x \neq \varnothing} W(x)$. It follows immediately that $\sum_{x \in \Sigma} W(x) = 1$. Moreover, by our induction, $W(x) \geqslant 0$ for all $x \neq \varnothing$, hence $W(\varnothing) \leqslant 1$. On the other hand, note that $\{x \subseteq \mathrm{Lit} \mid x \neq \varnothing\}$ is the truth set of $\bigvee_{q \in \mathrm{Lit}} q$. By induction assumption, $\sum_{x \neq \varnothing} W(x) = p(\bigvee_{q \in \mathrm{Lit}} q)$. By axiom (A1), hence, $W(\varnothing) \geqslant 0$.

Along the lines of the proof, we have ensured that $\mu([\varphi]) = p(\varphi)$ for all φ of the form $\bigwedge_{q \in x} q$ for some $x \subseteq \mathrm{Lit}$. By the above remark, this ensures that $\mu([\varphi]) = p(\varphi)$ for all φ, i.e. that $p_\mu = p$. $\qquad\square$

To end the static parts of this paper, we provide a graphical overview over the relationships identified so far. By Theorems 2 to 5, the diagram in Figure 1 commutes. Moreover, each pair of opposite arrows in the upper half of the diagram, i.e. the pairs (tr_{ns}^4, tr_4^{ns}), $(p \to \mathcal{M}_p, \mu \to p_\mu)$ and $(\hat{p} \to \mathcal{M}_{\hat{p}}, \mu \to \hat{p}_\mu)$ are left- and right inverses to each other.

7 Conditioning

In a classic setting, Bayesian conditioning on a formula φ describes a situation, where φ is learned to be true with probability 1 – and hence $\neg\varphi$ true with probability 0. A generalization of this rule is Jeffrey conditioning, where an
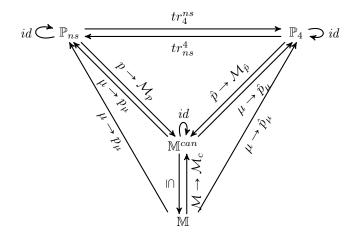
Figure 1: The relationships identified so far. By Theorems 2 to 5, this diagram commutes.
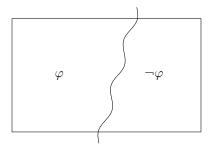

Figure 2: Classic Conditioning

agent may learn the probability of φ to be any value in $q \in [0; 1]$, rather than only the extremal value of 1 (or 0, when $\neg\varphi$ is learned) permitted in Bayes' conditioning.

Either method is best illustrated semantically. Within a classical setting, any formula φ defines a binary partition $\{[\varphi], [\neg\varphi]\}$ on the state space, cf. Figure 2. Jeffrey conditioning is then executed by linearly expanding or contracting the original measure μ on $[\varphi]$ and $[\neg\varphi]$ to some new $\overline{\mu}$ in such a way that $\overline{\mu}([\varphi]) = q$ and $\overline{\mu}([\neg\varphi]) = 1 - q$. We hence get for any $\varphi \in \mathcal{L}_{\text{Prop}}$ that

$$\overline{\mu}([\psi]) = \mu([\psi \wedge \varphi])\frac{q}{\mu([\varphi])} + \mu([\psi \wedge \neg\varphi])\frac{1-q}{\mu([\neg\varphi])} \tag{3}$$

which, in the case of Bayesian condition (i.e. $q = 1$) reduces to the well-known formula $\overline{\mu}([\psi]) = \frac{\mu([\psi \wedge \varphi])}{\mu([\varphi])}$.

Conditionalization in our extended setting follows a similar idea. However,

note that both Bayes' and Jeffrey conditioning implicitly rest on the facts that $p(\varphi) + p(\neg\varphi) = 1$ and that $p(\varphi \wedge \neg\varphi) = 0$, i.e. that there are no gaps and gluts. As this fact no longer holds, conditioning will behave differently in a non-standard setting. In fact, we will show that non-standard probabilities allow for two different notions of Jeffrey updating, one where a new value for the probability of φ, i.e. $p(\varphi)$ is learned, the other where a new value of the four-valued vector $\hat{p}(\varphi)$ is acquired. The former version of Jeffrey updating is best described on the level of non-standard probability assignments, the latter on the level of four-valued assignments. Yet, using the maps tr_{ns}^4 and tr_4^{ns}, both versions of updating can naturally be applied to either non-standard or four-valued probability assignments.

Just as in the standard case, non-normal Bayes conditioning can be defined as extremal case of Jeffrey updates. In fact, non-normal Bayes conditioning has been studied independently, for instance in Mares (1997). The current framework generalizes the latter's approach by also incorporating Jeffrey updating and by identifying a number of different Bayes like updates, containing the one put forward by Mares.

7.1 Updating on non-standard information

In our first notion of updating, the agent's update proscribes her to set the probability of φ to some $q \in [0;1]$. Notably, within a non-standard setting, this does not carry any information about the value of $\neg\varphi$ - the agent may or may not leave $p(\neg\varphi)$ unchanged in her update. In line with classic Jeffrey updating, **non-standard Jeffrey updating** is best illustrated semantically. For any set $\psi \in \mathcal{L}_{\mathrm{Prop}}$, we can dissect the state space of a probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ in two sets – the truth set $[\varphi]$ of φ and it's complement $\Sigma \backslash [\varphi]$. Unlike in the classic case, however, $\Sigma \backslash [\varphi]$ is not the truth set of $[\neg\varphi]$, nor of any other $\psi \in \mathcal{L}_{\mathrm{Prop}}$. Yet, we can define Jeffrey updating as in the classic case.

Definition 8. Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model. Let $q \in [0;1]$ and $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ such that $\mu([\varphi]) \in (0;1)$. Then the **semantic non-standard Jeffrey update** for updating the probability of φ to be q on $\mathcal{M}$ is the probabilistic model $\mathcal{M}^{\varphi,q} = \langle \Sigma, \mu^{\varphi,q}, v^+, v^- \rangle$ determined by:

$$\mu^{\varphi,q}(\{x\}) = \begin{cases} \mu(\{x\}) \cdot \frac{q}{\mu([\varphi])} & \text{iff } x \in [\varphi] \\ \mu(\{x\}) \cdot \frac{1-q}{1-\mu([\varphi])} & \text{else.} \end{cases}$$

Fact 1. *Non-standard Jeffrey updating is successful, i.e. for any probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$, any $q \in [0;1]$ and $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ such that $\mu([\varphi]) \in (0;1)$ the non-standard Jeffrey update on $\mathcal{M}$ updating the probability of φ to q satisfies $\mu^{\varphi,q}(x)([\varphi]) = q$.*

Despite the fact that the set $\Sigma \backslash [\varphi]$ is not definable, we can give a syntactic characterization of non-standard Jeffrey-updating. The following is a non-standard equivalent to classic Jeffrey's updating, cf. Formula (3).

Lemma 2. *Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model. Let $q \in [0;1]$ and $\varphi \in \mathcal{L}_{\text{Prop}}$ such that $\mu([\varphi]) \in (0;1)$. Then for any $\psi \in \mathcal{L}_{\text{Prop}}$, the non-standard Jeffrey update $\mathcal{M}^{\varphi,q} = \langle \Sigma, \mu^{\varphi,q}, v^+, v^- \rangle$ of $\mathcal{M}$ satisfies:*

$$\mu^{\varphi,q}([\psi]) = \mu([\psi \wedge \varphi]) \cdot \frac{q}{\mu([\varphi])} + (\mu(\psi) - \mu(\psi \wedge \varphi))\frac{1-q}{1-\mu([\varphi])}$$

Notably, after translating the previous fact into its induced non-standard probability assignments p_μ and $p_{\mu^{\varphi,q}}$, we obtain a fully syntactic characterization of non-standard Jeffrey updating.

Definition 9. Let $p \colon \mathcal{L}_{\text{Prop}} \to \mathbb{R}$ be a non-standard probability assignment, let $q \in [0;1]$ and $\varphi \in \mathcal{L}_{\text{Prop}}$ with $p(\varphi) \in (0;1)$. Then the **syntactic non-standard Jeffrey update** setting the probability of φ to q is the probability function $p^{\varphi,q} \colon \mathcal{L}_{\text{Prop}} \to \mathbb{R}$ defined by

$$p^{\varphi,q}(\psi) = p(\psi \wedge \varphi) \cdot \frac{q}{p(\varphi)} + (p(\psi) - p(\psi \wedge \varphi))\frac{1-q}{1-p(\varphi)}$$

By construction, semantic and syntactic non-standard Jeffrey updating coincide in the following sense.

Fact 2. *Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model, let $q \in [0;1]$ and $\varphi \in \mathcal{L}_{\text{Prop}}$ with $p(\varphi) \in (0;1)$. Then $p_{\mu^{\varphi,q}} = p_\mu^{\varphi,q}$.*

We will hence omit the labels and only speak of non-standard Jeffrey updating. We end this section with three facts about non-standard Jeffrey updating.

Fact 3. *Assume that the non-standard probability function $p \colon \mathcal{L}_{\text{Prop}} \to \mathbb{R}$ is classic, i.e. satisfies the Kolmogorov axioms. Moreover, let $\varphi \in \mathcal{L}_{\text{Prop}}$ with $p(\varphi) \in (0;1)$ and $q \in [0;1]$. Then the non-standard and the classic Jeffrey update for setting the probability of φ to q coincide, i.e. for all $\psi \in \mathcal{L}_{\text{Prop}}$*

$$p^{\varphi,q}(\psi) = p(\psi \wedge \varphi)\frac{q}{p(\varphi)} + p(\psi \wedge \neg\varphi)\frac{q}{p(\neg\varphi)}.$$

From this, it follows directly that

Fact 4. *Non-standard Jeffrey updating is not commutative. That is, there is a non-standard probability function $p \colon \mathcal{L}_{\text{Prop}} \to \mathbb{R}$ and $\varphi, \psi \in \mathcal{L}_{\text{Prop}}$ and $q, r \in [0;1]$ with $p(\varphi), p(\psi), p^{\varphi,q}(\psi), p^{\psi,r}(\varphi) \in (0;1)$ such that $(p^{\varphi,q})^{\psi,r} \neq (p^{\psi,r})^{\varphi,q}$.*

Non-standard Bayesian updating

Just as in the classic case, we will define non-standard Bayesian updating as special case of non-standard Jeffrey updating where the probability of φ is set to 1. In this case, the formula of Definition 9 simplifies to the same formula as in the classical case. Note that this is also the first of two approaches to Bayes updating proposed by Mares (1997). The second proposal by Mares, in contrast is not related to any version of Bayes updating presented here, as it strives to actively minimize conflict.

Definition 10. Let $p\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ be a non-standard probability function and let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ with $p(\varphi) > 0$. Then the **(positive) non-standard Bayesian update** on φ is the function $p^{\varphi,pos}$:

$$p^{\varphi,pos}(\psi) = \frac{p(\psi \wedge \varphi)}{p(\varphi)} \qquad \text{for } \varphi \in \mathcal{L}_{\mathrm{Prop}}.$$

Unlike in the classical setting, however, non-standard Bayesian updating does not cover all extremal cases. Setting the probability of φ to 0 is *not* the same as setting the probability of $\neg\varphi$ to 1, hence this case needs to be treated separately.

Definition 11. Let $p\colon \mathcal{L}_{\mathrm{Prop}} \to [0;1]$ be a non-standard probability function and let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ with $p(\varphi) < 1$. Then the **negative non-standard Bayesian update** on φ is the function $p^{\varphi,neg}$:

$$p^{\varphi,neg}(\psi) = \frac{p(\psi) - p(\psi \wedge \varphi)}{1 - p(\varphi)} \qquad \text{for } \varphi \in \mathcal{L}_{\mathrm{Prop}}.$$

As their classic counterpart, positive and negative non-standard Bayesian conditioning are order independent:

Lemma 3. *Let* $p\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ *and let* $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$ *with* $p(\varphi), p(\psi), p^{\varphi}(\psi), p^{\psi}(\varphi) \in (0;1)$. *Then* $(p^{\varphi,*})^{\psi,\times} = (p^{\psi^{\times}})^{\varphi,*}$ *for* $*, \times \in \{pos, neg\}$.

7.2 Updating on four-valued information

Within non-standard probability, knowing the probability of φ does not provide any information about the probability of $\neg\varphi$. Hence, in learning about φ, two cases are to be distinguished. In the first case, the agent only receives information about φ, without learning anything about $\neg\varphi$ or $\varphi \wedge \neg\varphi$. In the second case, the agent learns the full probabilistic information about φ, that is, the probabilities of φ and $\neg\varphi$, but also the size of the corresponding gap and glut. As discussed above, this information can be encoded in a vector $(b, d, u, c) \in \mathbb{R}^4$ specifying the new pure belief (i.e. belief without conflict), pure disbelief (belief in $\neg\varphi$ without conflict), uncertainty and conflict about φ.

Again, the notion of four-valued Jeffrey updating is best illustrated semantically. As shown in Figure 3, for any $\varphi \in \mathcal{L}_{\mathrm{Prop}}$, the sets of pure belief, pure disbelief, uncertainty and conflict about φ jointly form a partition $([\varphi]\backslash[\varphi \wedge \neg\varphi], ([\neg\varphi]\backslash[\varphi \wedge \neg\varphi], \Sigma\backslash[\varphi \vee \neg\varphi], [\varphi \wedge \neg\varphi])$ of a probabilistic model $\mathcal{M}$. Hence, a similar idea as in classic Jeffrey updating can be applied, linearly expanding or shrinking the measure on each of these four cells to their appropriate size. Notably, linear expansion (to a larger size) is only well defined if the cell to be expanded has a strictly positive measure. We capture this with the notion of admissibility of a vector (b, d, u, c):

Definition 12. Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and denote $\hat{p}_\mu(\varphi)$ by $(b_\varphi, d_\varphi, u_\varphi, c_\varphi)$. We call a vector $(b, d, u, c) \in [0;1]^4$ with $b + d + u + c = 1$ **admissible for** φ if it satisfies that $b = 0$ if $b_\varphi = 0$, $d = 0$ if $d_\varphi = 0$, $u = 0$ if $u_\varphi = 0$ and $c = 0$ if $c_\varphi = 0$.
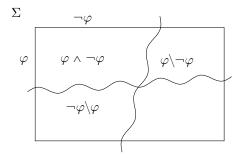
Figure 3: Four-valued conditioning

Definition 13. Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and let $(b, d, u, c) \in [0; 1]^4$ admissible for φ. Then **four-valued Jeffrey updating** on φ to (b, d, u, c) is the model $\mathcal{M}^{\varphi,(b,d,u,c)} = \langle \Sigma, \mu^{\varphi,(b,d,u,c)}, v^+, v^- \rangle$ with:

$$\mu^{\varphi,q}(x) = \begin{cases} \mu(x) \cdot \frac{b}{\mu([\varphi] - \mu[\varphi \wedge \neg\varphi])} & \text{iff } x \in [\varphi] \backslash [\varphi \wedge \neg\varphi] \\ \mu(x) \cdot \frac{d}{\mu([\neg\varphi]) - \mu([\varphi \wedge \neg\varphi])} & \text{iff } x \in [\neg\varphi] \backslash [\varphi \wedge \neg\varphi] \\ \mu(x) \cdot \frac{c}{\mu([\varphi \wedge \neg\varphi])} & \text{iff } x \in [\varphi \wedge \neg\varphi] \\ \mu(x) \cdot \frac{u}{1 - \mu([\varphi \vee \neg\varphi])} & \text{else} \end{cases}$$

Fact 5. *Four-valued Jeffrey updating is successful, i.e. for any probabilistic model $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$, any $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and any $(b, d, u, c) \in [0; 1]^4$ that is admissible for φ, the non-standard Jeffrey update on $\mathcal{M}$ setting the probability of φ to (b, d, u, c) satisfies $\hat{p}_{\mu^{\varphi,(b,d,u,c)}}(\varphi) = (b, d, u, c)$.*

Just as in the case of non-standard Jeffrey conditioning, we obtain a purely syntactic characterization of four-valued Jeffrey updating. Unfortunately, the drop in elegance with respect to standard Jeffrey updating is significant.

Lemma 4. *Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and let $(b, d, u, c) \in [0; 1]^4$ be admissible for φ. Then non-standard Jeffrey update on*

$\mathcal{M}$ *setting the probability of* φ *to* (b, d, u, c) *satisfies for any* $\psi \in \mathcal{L}_{\mathrm{Prop}}$ *that*

$$b'_\psi = \frac{b}{b_\varphi}(b_{\varphi,\psi}) + \frac{d}{d_\varphi}(b_{\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$
$$d'_\psi = \frac{b}{b_\varphi}(b_{\varphi,\overline{\psi}}) + \frac{d}{d_\varphi}(b_{\overline{\varphi},\overline{\psi}})$$
$$+ \frac{u}{u_\varphi}(d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$
$$u'_\psi = \frac{b}{b_\varphi}(b_\varphi - b_{\varphi,\psi} - b_{\varphi,\overline{\psi}} - c_{\varphi,\psi} + c_{\varphi,\overline{\varphi},\psi}) + \frac{d}{d_\varphi}(d_\varphi - b_{\overline{\varphi},\psi} - b_{\overline{\varphi},\overline{\psi}} - c_{\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(1 - d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_\varphi - c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$
$$c'_\psi = \frac{b}{b_\varphi}(c_{\varphi,\psi} - c_{\varphi,\overline{\varphi},\psi}) + \frac{d}{d_\varphi}(c_{\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(c_\psi - c_{\varphi,\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$

where $(b_\psi, d_\psi, u_\psi, c_\psi)$ *and* $(b'_\psi, d'_\psi, u'_\psi, c'_\psi)$ *denote the four-valued probability vector of* ψ *before and after the update. In the above equations,* $\overline{\psi}$ *is shorthand for* $\neg\psi$*, while* φ, ψ *stands for* $\varphi \wedge \psi$*. For ease of notation, this formula uses the convention that* $\frac{0}{0} = 0$*.*

Proof. Consider the propositions φ and ψ as well as the labeling of areas in the top row of Figure 4. By definition of updating, the mass of areas 1-4 need to be multiplied by $\frac{b}{b_\varphi}$, that of areas 5-8 by $\frac{d}{d_\varphi}$, the weight of areas 9-12 by $\frac{u}{u_\varphi}$ and that of areas 13-16 by $\frac{b}{b_\varphi}$. Moreover, the agent's pure belief in ψ, i.e. $\overline{b}_\psi$ is the joint mass of areas 1, 5, 9 and 13, her disbelief in ψ the joint mass of areas 2, 6, 10 and 14, her uncertainty is the joint weight of areas 3, 7, 11 and 15 and her conflict set the sum of areas 4, 8, 12 and 16.

To check correctness of the above equations, it then suffices to verify that the formulas pick out the respective fields, i.e. that $b_{\varphi,\psi}$ is the size of field 1, $b_{\overline{\varphi},\psi}$ is the size of field 5, $d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}$ is the size of field 9 and so on. That this is the case follows from the pictures in Figure 4, showing the belief and disbelief sets for certain composites of φ and ψ. $\qquad\square$

Again, the latter set of equations can be read purely syntactically. Thus, we get a syntactic counterpart to semantic four-valued Jeffrey updates.

Definition 14. Let $\hat{p}: \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ be a four-valued probability function and let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$. Moreover, let $(b, d, u, c) \in [0; 1]^4$ be admissible for φ. Then **(syntactic) four-valued Jeffrey updating** with the vector (b, d, u, c) yields a four-
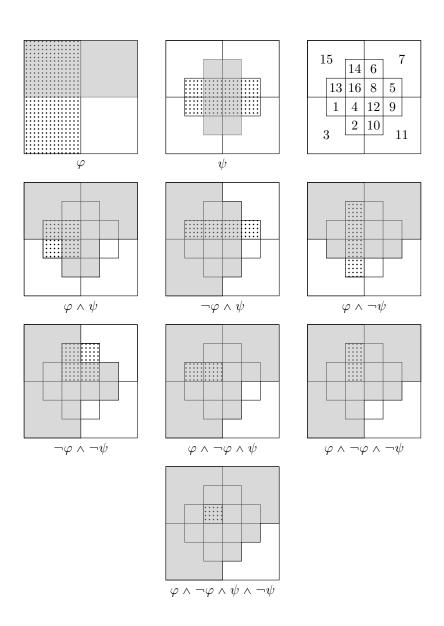
Figure 4: Belief and disbelief sets of φ (top left), ψ (top center) and various combinations thereof. Belief sets are dotted, disbelief set shaded. The diagrams fall in 16 sections that are labelled as shown on the top right.

valued probability function $\hat{p}^{\varphi,(b,d,u,c)}$ defined by $\hat{p}^{\varphi,(b,d,u,c)}(\psi) = (b'_\psi, d'_\psi, u'_\psi, c'_\psi)$ with:

$$b'_\psi = \frac{b}{b_\varphi}(b_{\varphi,\psi}) + \frac{d}{d_\varphi}(b_{\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$

$$d'_\psi = \frac{b}{b_\varphi}(b_{\varphi,\overline{\psi}}) + \frac{d}{d_\varphi}(b_{\overline{\varphi},\overline{\psi}})$$
$$+ \frac{u}{u_\varphi}(d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$

$$u'_\psi = \frac{b}{b_\varphi}(b_\varphi - b_{\varphi,\psi} - b_{\varphi,\overline{\psi}} - c_{\varphi,\psi} + c_{\varphi,\overline{\varphi},\psi}) + \frac{d}{d_\varphi}(d_\varphi - b_{\overline{\varphi},\psi} - b_{\overline{\varphi},\overline{\psi}} - c_{\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(1 - d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_\varphi - c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}})$$

$$c'_\psi = \frac{b}{b_\varphi}(c_{\varphi,\psi} - c_{\varphi,\overline{\varphi},\psi}) + \frac{d}{d_\varphi}(c_{\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi})$$
$$+ \frac{u}{u_\varphi}(c_\psi - c_{\varphi,\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}) + \frac{c}{c_\varphi}(c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}).$$

By construction, semantic and syntactic non-standard Jeffrey updating coincide in the following sense.

Fact 6. *Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a probabilistic model, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and let $(b,d,u,c) \in [0;1]^4$ be admissible for φ. Then $\hat{p}_{\mu^{\varphi,(b,d,u,c)}} = \hat{p}_\mu^{\varphi,(b,d,u,c)}$.*

We will hence omit the distinction between semantic and syntactic and only speak of four-valued Jeffrey updating. We end this section with three facts about this updating.

Fact 7. *Assume that the four-valued probability function $\hat{p} \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ is classic, i.e. $\hat{p}(\psi) \in \mathbb{R}^2 \times \{0\}^2$. Moreover, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and $(b,d,0,0) \in [0;1]^4$ be admissible for φ, i.e. $b = 0$ if $p(\varphi) = 0$ and $d = 0$ if $p(\neg\varphi) = 0$. Then the non-standard and the classic Jeffrey update setting the probability of φ to q coincide, i.e. for all $\psi \in \mathcal{L}_{\mathrm{Prop}}$*

$$\hat{p}^{\varphi,(b,d,0,0)}(\psi) = \left(b_{\psi \wedge \varphi}\frac{q}{b_\varphi} + b_{\psi \wedge \neg\varphi}\frac{q}{b_{\neg\varphi}},\ 1 - b_{\psi \wedge \varphi}\frac{q}{b_\varphi} - b_{\psi \wedge \neg\varphi}\frac{q}{b_{\neg\varphi}},\ 0,\ 0 \right)$$

From this, it follows directly that

Fact 8. *Non-standard Jeffrey updating is not commutative. That is, there is a four-valued probability function $\hat{p} \colon \mathcal{L}_{\mathrm{Prop}} \to [0;1]^4$, some $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$ and $(b,d,u,c), (b',d',u',c') \in [0;1]$ such that (b,d,u,c) is admissible for φ in $\hat{p}$ and in $\hat{p}^{(b',d',u',c')}$, while (b',d',u',c') is admissible for ψ in both $\hat{p}$ and $\hat{p}^{(b,d,u,c)}$, such that*

$$(p^{\varphi,(b,d,u,c)})^{\psi,(b',d',u',c')} \neq (p^{\psi,(b',d',u',c')})^{\varphi,(b,d,u,c)}$$

Four-valued Bayesian updating

Just as in the classical case, we can define four-valued Bayesian updating as a special instance of Jeffrey updating where the information acquired is extremal. Here, we focus on three cases. In the first, the agent learns the vector $(1,0,0,0)$, i.e. she acquires full pure belief in φ. In the second and third case, the agent learns the vectors $(0,0,1,0)$ or $(0,0,0,1)$ respectively, acquiring full belief in uncertainty or conflict about φ. The remaining case, learning $(0,1,0,0)$, follows from these, as it corresponds to updating on information $(1,0,0,0)$ about $\neg\varphi$. In either of our three cases, the above definition of four-valued Jeffrey updating simplifies to:

Definition 15. *i*) Let $\hat{p}\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ be a four-valued probability function such that $b_\varphi > 0$, where $\hat{p}(\varphi) = (b_\varphi, d_\varphi, u_\varphi, c_\varphi)$. Then **positive four-valued Bayesian updating** on φ yields the function $\hat{p}^{\varphi,+}$ defined by $\hat{p}^{\varphi,+} = (b'_\psi, d'_\psi, u'_\psi, c'_\psi)$ with

$$b'_\psi = \frac{b_{\varphi,\psi}}{b_\varphi}$$

$$d'_\psi = \frac{b_{\varphi,\overline{\psi}}}{b_\varphi}$$

$$u'_\psi = \frac{b_\varphi - b_{\varphi,\psi} - b_{\varphi,\overline{\psi}} - c_{\varphi\wedge\psi} + c_{\varphi,\overline{\varphi},\psi}}{b_\varphi}$$

$$c'_\psi = \frac{c_{\varphi,\psi} - c_{\varphi,\overline{\varphi},\psi}}{b_\varphi}$$

ii) Let $\hat{p}\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ be a four-valued probability function such that $u_\varphi > 0$, where $\hat{p}(\varphi) = (b_\varphi, d_\varphi, u_\varphi, c_\varphi)$. Then **uncertainty Bayesian updating** about φ is defined as: $\hat{p}^{\varphi,u}(\psi) = (b'_\psi, d'_\psi, u'_\psi, c'_\psi)$ with

$$b'_\psi = \frac{d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{u_\varphi}$$

$$d'_\psi = \frac{d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - d_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{u_\varphi}$$

$$u'_\psi = \frac{1 - d_{\varphi,\overline{\varphi},\psi,\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{u_\varphi}$$

$$c'_\psi = \frac{c_\psi - c_{\varphi,\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\overline{\varphi},\psi} + c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{u_\varphi}$$

iii) Let $\hat{p}\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ be a four-valued probability function such that $c_\varphi > 0$, where $\hat{p}(\varphi) = (b_\varphi, d_\varphi, u_\varphi, c_\varphi)$. Then **conflict Bayesian updating** about φ is

defined as: $\hat{p}^{\varphi,c}(\psi) = (b'_\psi, d'_\psi, u'_\psi, c'_\psi)$ with

$$b'_\psi = \frac{c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{c_\varphi}$$

$$d'_\psi = \frac{c_{\varphi,\overline{\varphi},\overline{\psi}} - c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{c_\varphi}$$

$$u'_\psi = \frac{c_\varphi - c_{\varphi,\overline{\varphi},\psi} - c_{\varphi,\overline{\varphi},\overline{\psi}} + c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{c_\varphi}$$

$$c'_\psi = \frac{c_{\varphi,\overline{\varphi},\psi,\overline{\psi}}}{c_\varphi}$$

Just as its classic counterpart, four-valued Bayesian conditioning in all three flavors is order independent:

Lemma 5. *Let* $\hat{p}\colon \mathcal{L}_{\mathrm{Prop}} \to [0;1]$ *and let* $\varphi, \psi \in \mathcal{L}_{\mathrm{Prop}}$ *such that* $\hat{p}^{\varphi,a}$, $\hat{p}^{\psi,b}$, $(\hat{p}^{\varphi,a})^{\psi,b}$ *and* $(\hat{p}_b^{\psi,b})^{\varphi,a}$ *are all defined. Then* $(\hat{p}^{\varphi,a})^{\psi,b} = (\hat{p}_b^{\psi,b})^{\varphi,a}$. *for* $a,b \in \{+,u,c\}$

7.3 Interaction Principles

Using the translation functions tr_{ns}^4 and tr_4^{ns}, both notions of Jeffrey conditioning, non-standard and four-valued, work on both types of probability functions defined, non-standard and four-valued. However, the notions of updating do not correspond to each other. While non-standard Jeffrey conditioning applies to situations where only the probability of φ is set, without any mention of the probabilities of $\neg\varphi$ or $\varphi \wedge \neg\varphi$, four-valued Jeffrey conditioning covers cases where new probabilities of $\varphi, \neg\varphi$ and the corresponding gap and glut are all proscribed simultaneously. Hence, even after appropriate transformations of their domains with tr_{ns}^4 and tr_4^{ns}, the two types of Jeffrey updates are *not* inter-definable. This, however, changes if we move to non-standard and four-valued Bayesian updating. Each of the three types of four-valued Bayesian updating is equivalent to a composition of two steps of non-standard Bayesian updating. Moreover, the order of these two steps does not matter.

Lemma 6. *Let* $\hat{p}\colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$ *be a four-valued probability assignment and let* $\varphi \in \mathcal{L}_{\mathrm{Prop}}$.

    i) if $b_\varphi > 0$, *then* $tr_{ns}^4(\hat{p}^{\varphi,+}) = (tr_{ns}^4(\hat{p})^{\varphi,pos})^{\neg\varphi,neg} = (tr_{ns}^4(\hat{p})^{\neg\varphi,neg})^{\varphi,pos}$

    ii) if $u_\varphi > 0$, *then* $tr_{ns}^4(\hat{p}^{\varphi,u}) = (tr_{ns}^4(\hat{p})^{\varphi,neg})^{\neg\varphi,neg} = (tr_{ns}^4(\hat{p})^{\neg\varphi,neg})^{\varphi,neg}$

    iii) if $c_\varphi > 0$, *then* $tr_{ns}^4(\hat{p}^{\varphi,u}) = (tr_{ns}^4(\hat{p})^{\varphi,pos})^{\neg\varphi,pos} = (tr_{ns}^4(\hat{p})^{\neg\varphi,pos})^{\varphi,pos}$

Proof. i) By Theorem 5, there is a unique canonical model $\mathcal{M} = \langle \mathcal{P}(\mathrm{Lit}), \mu, v^+, v^- \rangle$ such that $\hat{p}_\mu = \hat{p}$. By Facts 2 and 6, it hence suffices to show the claim for semantic four-valued Jeffrey updating on $\mathcal{M}$. Note that the result of positive Bayesian updating, i.e. the updated four-valued probability function $\mu^{\varphi,(1,0,0,0)}$ of $\mathcal{M}^{\varphi,(1,0,0,0)} = \langle \mathcal{P}(\mathrm{Lit}), \mu^{\varphi,(1,0,0,0)}, v^+, v^- \rangle$ is uniquely determined by the conditions

(1) $\mu^{\varphi,(1,0,0,0)}(x) = 0$ whenever $x \notin [\varphi]\backslash[\varphi \wedge \neg\varphi]$
(2) $\mu^{\varphi,(1,0,0,0)}(x)/\mu^{\varphi,(1,0,0,0)}(y) = \mu(x)/\mu(y)$ whenever $x, y \in [\varphi]\backslash[\varphi \wedge \neg\varphi]$
with $\mu(y) > 0$.

By construction, the two-valued Bayesian updates, i.e. Jeffrey updates $\mu^{\varphi,1}$ and $\mu^{\neg\varphi,0}$ both satisfy (2). Moreover, $\mu^{\varphi,1}(x) = 0$ whenever $x \notin [\varphi]$ and $\mu^{\neg\varphi,0}(x) = 0$ whenever $x \in [\neg\varphi]$. Thus both $(\mu^{\varphi,1})^{\neg\varphi,0}$ and $(\mu^{\neg\varphi,0})^{\varphi,1}$ also satisfy (1). Hence, both $(\mu^{\varphi,1})^{\neg\varphi,0}$ and $(\mu^{\neg\varphi,0})^{\varphi,1}$ satisfy conditions (1) and (2) and, hence, are identical to $\mu^{\varphi,(1,0,0,0)}$. This implies that $tr_{ns}^4(\hat{p}^{\varphi,+}) = (tr_{ns}^4(\hat{p}^{\varphi,pos})^{\neg\varphi,neg}) = (tr_{ns}^4(\hat{p}^{\neg\varphi,neg})^{\varphi,pos})$.
The proofs of $ii)$ and $iii)$ follow similarly. $\square$

7.4 Conditioning on Partial Information

In the previous sections we investigated updating a probability function with a generalized Jeffery rule by learning either only a new value for the belief in φ (Section 7.1) or the entire four-valued probability vector assigned to φ (Section 7.2). However, there may be other contexts where the agent acquires partial information about the (four-valued) probability of φ, e.g. only a new value for pure belief or pure disbelief in φ.

The idea for conditioning on partial information proceeds along the same lines as for complete information, i.e. by a modified version of Jeffery conditioning. The only difference is that the partiality of information, say about φ does not permit to work with the full partition induced by φ on a model $\mathcal{M}$, i.e. the partition into $\{[\varphi]\backslash[\neg\varphi], [\neg\varphi]\backslash[\varphi], \Sigma - [\varphi \vee \neg\varphi], [\varphi \wedge \neg\varphi]\}$, cf Figure 3, but with a coarsening thereof.

By obtaining partial information we mean that the agent learns the values of a *partial* assignment $a : \{b, d, u, c\} \rightharpoonup [0;1]$, i.e. an assignment proscribing new values for *some* of the agent's pure belief, pure disbelief, uncertainty and conflict, but not necessarily for all. Let us denote the domain of a, i.e. those $x \in \{b, d, u, c\}$ for which $a(x)$ is defined, by $dom(a)$. For simplicity, we assume that $\varnothing \subset dom(a) \subset \{b, d, u, c\}$ with both inclusions strict. Following the same intuitions as in the four-valued case, we can define conditioning on the partial information a by setting the new pure belief, disbelief, uncertainty and conflict in φ to be $a(b), a(d), a(u)$ and $a(c)$ respectively whenever this is defined and afterwards rescaling the probabilistic mass on the remaining area appropriately.

Formally, to ensure that the corresponding operation is well-defined, we need to assume that $\sum_{y \in dom(a)} y \leq 1$. Denoting the prior four-valued probability vector of φ with $(b_\varphi, d_\varphi, u_\varphi, c_\varphi)$, the Jeffery updating sketched above will lead to the posterior four-valued probability vector $(\bar{b}_\varphi, \bar{d}_\varphi, \bar{u}_\varphi, \bar{c}_\varphi)$ with:

$$\bar{x}_\varphi = \begin{cases} a(x) & \text{iff } x \in dom(a) \\ x \cdot \dfrac{1 - \sum_{y \in dom(a)} y}{1 - \sum_{y \in dom(a)} y_\varphi} & else \end{cases}$$

for $x \in \{b, d, u, c\}$. With this, we can formally define partial Jeffery updating.

24

Definition 16. Let $a : \{b, d, u, c\} \rightharpoonup [0; 1]$ be a partial assignment such that $\sum_{y \in dom(a)} y \leqslant 1$. Let $\mathcal{M} = \langle \Sigma, \mu, v^+, v^- \rangle$ be a model, let $\varphi \in \mathcal{L}_{\mathrm{Prop}}$ and let the vector $(\bar{b}_\varphi, \bar{d}_\varphi, \bar{u}_\varphi, \bar{c}_\varphi)$ defined above be admissible for φ. Then the **four-valued Jeffrey update of φ on the partial information** a is defined as the four-valued Jeffrey update on φ to $(\bar{b}_\varphi, \bar{d}_\varphi, \bar{u}_\varphi, \bar{c}_\varphi)$.

8 Aggregation

Assume two agents informed you about their credences in φ. You take both agents as similarly competent and equally informed. Yet, they equip you with different assessments of φ. How, then, should you combine these judgments towards forming your own belief about φ? Within standard probability theory, your options are fairly limited. You may, for instance, decide to follow one of the agents, or build a weighted average between the two. A broad number of approaches in the literature on peer disagreement, for instance, promotes to split the difference equally see for instance Elga (2007); Christensen (2007) on conciliationism, but also Kelly (2010) for an opposing opinion.

8.1 Aggregating non-standard probabilities

Within the non-standard probabilities studied here, further options open up. First, note that within classic probability theory, learning about the agents credence in ψ also informs us about her degree of belief in $\neg\psi$. This does not hold true within the current non-standard setting. Hence, let us assume for the current analysis that agents inform us about both their positive and negative attitude towards φ, that is about $p(\varphi)$ and $p(\neg\varphi)$, or even about their four-valued vector $\hat{p}(\varphi)$. Of course, we may follow the previous strategies and form weighted averages between the agents' assessments of φ. If needed, this policy could be specified to also taking a weighted average on the agents conflict and uncertainty and, more general, their remaining belief set.

Definition 17. Let $k \in [0, 1]$.
i) Assume agents A and E provide their non-standard assessments of φ, i.e. $p_A(\varphi)$, $p_E(\varphi), p_A(\neg\varphi)$ and $p_E(\neg\varphi)$. Then their k-**weighted non-standard aggregate belief** $p^k_{\{A,E\}}$ is defined by

$$p^k_{\{A,E\}}(\varphi) = kp_A(\varphi) + (1-k)p_E(\varphi) \text{ and } p^k_{\{A,E\}}(\neg\varphi) = kp_A(\neg\varphi) + (1-k)p_E(\neg\varphi).$$

ii) For agent A and E's four-valued probabilitiy assessments $(b, d, u, c)_A$ and $(b, d, u, c)_E$ for φ, i.e. $\hat{p}_A(\varphi)$ and $\hat{p}_E(\varphi)$ their k-**weighted four-valued aggregate belief** $\hat{p}^k_{\{A,E\}}$ is:

$$\hat{p}^k_{\{A,E\}}(\varphi) = k\hat{p}_A(\varphi) + (1-k)\hat{p}_E(\varphi).$$

Lemma 7. *Weighted averaging can be applied to an entire belief base simultaneously. That is, when agents A and E both provide their full subjective non-standard probability functions $p_A, p_E \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ (resp. $\hat{p}_A, \hat{p}_E \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}^4$),*

Updating rule	$p_{\{A,E\}}(\varphi)$	$p_{\{A,E\}}(\neg\varphi)$
k-weighted	$kp_A(\varphi) + (1-k)p_E(\varphi)$	$kp_A(\neg\varphi) + (1-k)p_E(\neg\varphi)$
credulous	$max(p_A(\varphi), p_E(\varphi))$	$max(p_A(\neg\varphi), p_E(\neg\varphi))$
cautious	$min(p_A(\varphi), p_E(\varphi))$	$min(p_A(\neg\varphi), p_E(\neg\varphi))$
optimist	$max(p_A(\varphi), p_E(\varphi))$	$min(p_A(\neg\varphi), p_E(\neg\varphi))$
pessimist	$min(p_A(\varphi), p_E(\varphi))$	$max(p_A(\neg\varphi), p_E(\neg\varphi))$

Table 1: Different rules for aggregating agent A and E's non-standard beliefs in φ and $\neg\varphi$, i.e. $p_A(\varphi)$, $p_E(\varphi), p_A(\neg\varphi)$ and $p_E(\neg\varphi)$.

a weighted average belief $p^k_{\{A,E\}} \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ can be defined by $kp_A + (1-k)p_E$. Likewise, $\hat{p}^k_{\{A,E\}} \colon \mathcal{L}_{\mathrm{Prop}} \to \mathbb{R}$ can be defined by $k\hat{p}_A + (1-k)\hat{p}_E$. Moreover, these policies commute with tr^4_{ns}, that is

$$tr^4_{ns}(\hat{p}^k_{\{A,E\}}) = p^k_{\{A,E\}} \ and \ tr^{ns}_4(p^k_{\{A,E\}}) = \hat{p}^k_{\{A,E\}}.$$

Non-Standard beliefs, however, allow for further aggregation policies that do not have classic counterparts. **Credulous** agents, for instance, could opt for the maximal values of their input in terms of belief and disbelief simultaneously. That is, they could set their updated belief and disbelief in φ to be $max(p_A(\varphi), p_E(\varphi))$ and $max(p_A(\neg\varphi), p_E(\neg\varphi))$ respectively. Likewise, **cautious** agents may rather chose to belief and disbelief φ only to an amount supported by all input information. Such agents would set their belief and disebelief in φ to $min(p_A(\varphi), p_E(\varphi))$ and $min(p_A(\neg\varphi), p_E(\neg\varphi))$ respectively.

In special situations, further policies are conceivable. When testing the safety of a new drug, for example, agents may be extremely vary of false positives while being much less concerned with false negatives. Such an agent might decide to set her new belief in φ to $min(p_A(\varphi), p_E(\varphi))$ while adopting $max(p_A(\neg\varphi), p_E(\neg\varphi))$ as new disbelief in φ. Likewise, also the combination of $max(p_A(\varphi), p_E(\varphi))$ with $min(p_E(\neg\varphi), p_E(\neg\varphi))$ are conceivable. In some sense, the latter two policies are aggregation functions that minimize type I and type II errors. For a lack of a better name we call these **pessimist** and **optimist** updating rules respectively. See Table 1 for an overview.
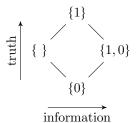
Unlike weighted average, none of these four policies can be applied to an entire belief set simultaneously.

Fact 9. *Let p_A and p_E be such that $p_A(\varphi) = 1$ and $p_A(\neg\varphi) = p_A(\varphi \wedge \neg\varphi) = 0$, while $p_E(\neg\varphi) = 1$ and $p_E(\varphi) = p_E(\varphi \wedge \neg\varphi) = 0$. Then p_A and p_E are consistent, but the function $p_{\{A,E\}}$ defined by $p_{\{A,E\}}(*) = max(*)$ for $* \in \{\varphi, \neg\varphi, \varphi \wedge \neg\varphi\}$ is not.*

Proof. To see that p_A and p_E are consistent consider a nonstandard model with three worlds, x, y, z and $v^+(p) = \{x, y\}, v^-(p) = \{x, z\}$. The measure μ_A putting all weight on y is such that $p_{\mu_A}(*) = p_A(*)$ for $* \in \{\varphi, \neg\varphi, \varphi \wedge \neg\varphi\}$, showing p_A consistent by Lemma 1. Likewise, the measure μ_E putting all weight

on z shows p_E consistent. For the inconsistency of $p_{\{A,E\}}$, finally, note that $p_{\{A,E\}}(\varphi \wedge \neg\varphi) = 0$ and $p_{\{A,E\}}(\varphi) = p_{\{A,E\}}(\neg\varphi) = 1$. Plugging these three values into (A3) yields $0 + p_{\{A,E\}}(\varphi \vee \neg\varphi) = 2$, contradicting (A1). □

Likewise, the missing conditions for cautious updates cannot be retrieved by extending the policy of taking minima to the agents' assessments of $\varphi \vee \neg\varphi$, as can be seen from the previous Fact. In particular, there is no counterpart to Lemma 7 for credulous or cautious update. Neither can be performed for all $\varphi \in \mathcal{L}_{\text{Prop}}$ simultaneously.

Before proceeding to four-valued updating, we compare the above policies to operations in non-probabilistic Belnap-Dunn logic. For this, recall the classic Belnap-Dunn bi-lattice of truth values BD_4.

This bi-lattice can be interpreted in two directions relating to *truth values* and the available *information*. We denote meet and join of the truth lattice operations by $\wedge$ and $\vee$ while meet and join for the information lattice operations are $\sqcap$ and $\sqcup$. Note that we can identify an assignment of BD_4-values to some formula φ with a non-standard probability assignment of $p(\varphi)$ and $p(\neg\varphi)$ into $\{0,1\}$. More specifically, assigning $\{1,0\}$ to some φ corresponds to $p(\varphi) = p(\neg\varphi) = 1$, while assigning $\{1\}$, resp $\{0\}$ to φ corresponds to $p(\varphi) = 1$, $p(\neg\varphi) = 0$ and $p(\varphi) = 0$, $p(\neg\varphi) = 1$ respectively. Value $\{\}$, finally, corresponds to $p(\varphi) = p(\neg\varphi) = 0$. For a probability assignment $p(\varphi), p(\neg\varphi) \in \{0,1\}$, we denote the corresponding BD_4 value by $t_p(\varphi)$. Applying this correspondence, we obtain the following characterization of the four updating policies introduced above:

Lemma 8. *Assume when asked about their credences in φ, agents A and E provide extremal assignments, i.e. $p_A(\varphi), p_A(\neg\varphi), p_E(\varphi), p_E(\neg\varphi) \in \{0,1\}$. Then*

Credulous update	*yields beliefs in φ and $\neg\varphi$*	$t_{p_A}(\varphi) \sqcup t_{p_E}(\varphi)$
Cautious update	*that are equal to*	$t_{p_A}(\varphi) \sqcap t_{p_E}(\varphi)$
Opimistic update		$t_{p_A}(\varphi) \vee t_{p_E}(\varphi)$
Pessimistic update		$t_{p_A}(\varphi) \wedge t_{p_E}(\varphi).$

Finally, we consider the special case where both agents input *classic* probability values, i.e. values such that $p(\varphi) + p(\neg\varphi) = 1$.

Fact 10. *When p_A and p_E are classic, i.e. $p_A(\varphi) + p_A(\neg\varphi) = p_E(\varphi) + p_E(\neg\varphi) = 1$, then the same holds for the aggregated belief when aggregation follows weighted averaging, optimistic or pessimistic updates. That is, these three rules preserve classicality. This does not hold for credulous and cautious updating. The latter two rules turn classic inputs beliefs for agent A and E into non-classic aggregate values as soon as A and E disagree about $p(\varphi)$.*

8.2 Aggregating four-valued probabilities.

So far, we have assumed aggregation to operate on non-standard probability assignments. Within the above framework, agents provide their subjective non-standard beliefs in both φ and $\neg\varphi$, which the various aggregative mechanisms described above then merge into aggregate belief values for φ and $\neg\varphi$. But of course, our agents might also provide their subjective four-valued probabilities $\hat{p}_A(\varphi) = (b_\varphi^A, d_\varphi^A, u_\varphi^A, c_\varphi^A)$ and $\hat{p}_E(\varphi) = (b_\varphi^E, d_\varphi^E, u_\varphi^E, c^A E_\varphi)$ instead. Naturally, we could then hope to obtain an aggregate four-valued probability

$$\hat{p}_{\{A,E\}}(\varphi) = (b_\varphi^{\{A,E\}}, d_\varphi^{\{A,E\}}, u_\varphi^{\{A,E\}}, c_\varphi^{\{A,E\}}).$$

Note, that by the map tr_{ns}^4, the non-standard probabilities $p(\varphi)$ and $p(\neg\varphi)$ can be calculated from the four-valued probability $\hat{p}(\varphi)$. Hence, if $\hat{p}_{\{A,E\}}(\varphi)$ is defined, a corresponding two-valued aggregation mechanism for $p_{\{A,E\}}(\varphi)$ and $p_{\{A,E\}}(\neg\varphi)$ follows immediately. However, the opposite does not hold. $p_{\{A,E\}}(\varphi)$ and $p_{\{A,E\}}(\neg\varphi)$ do not fully determine $\hat{p}_{\{A,E\}}(\varphi)$ and hence the various policies defined in the last section do not readily translate into four-valued aggregation procedures. In fact, when employing the map tr_4^{ns}, the three values $p(\varphi)$, $p(\neg\varphi)$ and $p(\varphi \wedge \neg\varphi)$ are required to determine $\hat{p}(\varphi)$. In the case of weighted averaging, this is not a problem. By Lemma 7, setting

$$p_{\{A,E\}}^k(\varphi \wedge \neg\varphi) = k p_A^k(\varphi \wedge \neg\varphi) + (1-k) p_E^k(\varphi \wedge \neg\varphi)$$

yields a consistent set of requirements and the corresponding four-valued aggregation rule is exactly $\hat{p}_{\{A,E\}}^k(\varphi) = k\hat{p}_A(\varphi) + (1-k)\hat{p}_E(\varphi)$.

However, the situation is different in the case of credulous or cautious updating. As shown in Fact 9, requiring that $p_{\{A,E\}}(\varphi) = max(p_A(\varphi), p_E(\varphi))$, $p_{\{A,E\}}(\neg\varphi) = max(p_A(\neg\varphi), p_E(\neg\varphi))$ and $p_{\{A,E\}}(\varphi \wedge \neg\varphi) = max(p_A(\varphi \wedge \neg\varphi), p_E(\varphi \wedge \neg\varphi))$ may yield an *inconsistent* set of requirements. Hence, other choices are needed.

The vector $p_{\{A,E\}}(\varphi)$ is determined by four choices. With two of them given by $p_{\{A,E\}}(\varphi) = max(p_A(\varphi), p_E(\varphi))$ and $p_{\{A,E\}}(\neg\varphi) = max(p_A(\neg\varphi), p_E(\neg\varphi))$, and a third by axiom (D2), one last condition is missing. In the case of credulous update, we would arguably expect that $c_\varphi^{\{A,E\}} \geqslant max(c_\varphi^A, c_\varphi^E)$: If an agent opts to be credulous about both φ and $\neg\varphi$, she could not expect her conflict to fall below any of the input conflicts. Within this restriction, the below definition of credulous update, assumes $c_\varphi^{\{A,E\}}$ to be as close to $max(c_\varphi^A, c_\varphi^E)$ as possible while maintaining consistency.

Likewise, in the case of cautious update, we would arguably expect overall uncertainty to grow, or, at least, not to shrink through aggregation. That is, we would expect that $u_\varphi^{\{A,E\}} \geqslant max(u_\varphi^A, u_\varphi^E)$. Again, We will demand that $u_\varphi^{\{A,E\}}$ is the maximal possible consistent value with this property.

Definition 18. Assume agents A and E provide four-valued probabilities $\hat{p}_A(\varphi) = (b_\varphi^A, d_\varphi^A, u_\varphi^A, c_\varphi^A)$ and $\hat{p}_E(\varphi) = (b_\varphi^E, d_\varphi^E, u_\varphi^E, c^A E_\varphi)$. Then the **credulously aggregated four-valued probability** $\hat{p}_{\{A,E\}}(\varphi) = (b_\varphi^{\{A,E\}}, d_\varphi^{\{A,E\}}, u_\varphi^{\{A,E\}}, c_\varphi^{\{A,E\}})$ is given by the following four conditions

$$b_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = max(b_\varphi^A + c_\varphi^A, b_\varphi^E + c_\varphi^E)$$
$$d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = max(d_\varphi^A + c_\varphi^A, d_\varphi^E + c_\varphi^E)$$
$$b_\varphi^{\{A,E\}} + d_\varphi^{\{A,E\}} + u_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = 1$$
$$c_\varphi^{\{A,E\}} = max\left(c_\varphi^E,\ c_\varphi^A,\ (b_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}}) + (d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}}) - 1\right)$$

By tr_{ns}^4, the first two of these equations correspond to the two conditions of credulous non-standard updates, i.e. $p_{\{A,E\}}(\varphi) = max(p_A(\varphi), p_E(\varphi))$ and $p_{\{A,E\}}(\neg\varphi) = max(p_A(\neg\varphi), p_E(\neg\varphi))$. The third equation is axiom (D2). The last equation, finally expresses that $c_\varphi^{\{A,E\}}$ is the minimal consistent choice such that $c_\varphi^{\{A,E\}} \geqslant max(c_\varphi^A, c_\varphi^E)$. To see this, note that by (D2), we have $b_\varphi^{\{A,E\}} + d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} \leqslant 1$ and hence

$$b_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} + d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} - 1 \leqslant c_\varphi^{\{A,E\}}.$$

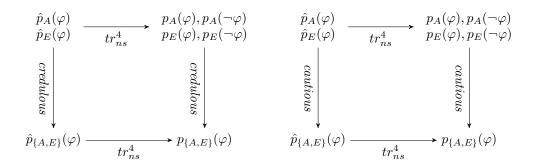Likewise we can define a cautious aggregation of four-valued probabilities:

Definition 19. For $\hat{p}_A$ and $\hat{p}_E$ as above, the **cautiously aggregated four-valued probability** $\hat{p}_{\{A,E\}}(\varphi) = (b_\varphi^{\{A,E\}}, d_\varphi^{\{A,E\}}, u_\varphi^{\{A,E\}}, c_\varphi^{\{A,E\}})$ is given by the following four equations

$$b_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = min(b_\varphi^A + c_\varphi^A, b_\varphi^E + c_\varphi^E)$$
$$d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = min(d_\varphi^A + c_\varphi^A, d_\varphi^E + c_\varphi^E)$$
$$b_\varphi^{\{A,E\}} + d_\varphi^{\{A,E\}} + u_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}} = 1$$
$$u_\varphi^{\{A,E\}} = max\left(u_\varphi^E,\ u_\varphi^A,\ 1 - (b_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}}) - (d_\varphi^{\{A,E\}} + c_\varphi^{\{A,E\}})\right)$$

Credulous and cautious aggregation as defined here cohere with their definition for non-standard probabilities.

Lemma 9. *Assume that agents A and E provide four-valued vectors $\hat{p}_A$ and $\hat{p}_E$ respectively. Then the following diagrams commute, where the application of tr_{ns}^4 makes use of the fact that $p(\varphi)$ and $p(\neg\varphi)$ can be calculated from $\hat{p}(\varphi)$.*

$$\begin{array}{ccc}
\hat{p}_A(\varphi) & \xrightarrow{\ tr^4_{ns}\ } & p_A(\varphi), p_A(\neg\varphi) \\
\hat{p}_E(\varphi) & & p_E(\varphi), p_E(\neg\varphi) \\
\Big\downarrow{\scriptstyle credulous} & & \Big\downarrow{\scriptstyle credulous} \\
\hat{p}_{\{A,E\}}(\varphi) & \xrightarrow{\ tr^4_{ns}\ } & p_{\{A,E\}}(\varphi)
\end{array}
\qquad
\begin{array}{ccc}
\hat{p}_A(\varphi) & \xrightarrow{\ tr^4_{ns}\ } & p_A(\varphi), p_A(\neg\varphi) \\
\hat{p}_E(\varphi) & & p_E(\varphi), p_E(\neg\varphi) \\
\Big\downarrow{\scriptstyle cautious} & & \Big\downarrow{\scriptstyle cautious} \\
\hat{p}_{\{A,E\}}(\varphi) & \xrightarrow{\ tr^4_{ns}\ } & p_{\{A,E\}}(\varphi)
\end{array}$$

The algebraic structure of credulous and cautious aggregation.

Definition 20. For an aggregation strategy S, we call φ a neutral element if for all ψ we have

$$S(\varphi, \psi) = S(\psi, \varphi) = \psi,$$

and we call φ an anihilator if for all ψ

$$S(\varphi, \psi) = S(\psi, \varphi) = \varphi.$$

Proposition 1. *The subjective four-valued probability assignment $(0,0,0,1)$, i.e. the element of maximal conflict, is an anihilator with respect to credulous updating. Likewise, the subjective four-valued probability assignment $(0,0,1,0)$, representing maximal uncertainty, is an anihilator for the cautious strategy.*

Proof. Let $\hat{p}_A(\varphi) = (0,0,0,1)$, let $\hat{p}_E(\varphi)$ be arbitrary and denote the result of credulous updating by (B,D,U,C). Then by definition $B + C = D + C = 1$ which, together with (D1) and (D2) implies $C = 1$, $B = D = 0$, and hence $U = 0$. In a similar manner, let $\hat{p}_A(\psi) = (0,0,1,0)$ and let $\hat{p}_E(\psi)$ be arbitrary, and denote the result of cautious updating by (B',D',C',U'). then by definition $B' + C' = D' + C' = 0$, which implies $B' = D' = C' = 0$ and hence $U' = 1$. $\square$

Proposition 2. *The subjective four-valued probability assignment $(0,0,0,1)$, i.e. the element of maximal conflict, is a neutral element with respect to cautious updating. Likewise, the subjective four-valued probability assignment $(0,0,1,0)$, representing maximal uncertainty, is neutral with respect to credulous updating.*

Proof. Let $\hat{p}_A(\varphi) = (0,0,0,1)$, let $\hat{p}_E(\varphi) = (b,d,u,c)$ be arbitrary and denote the result of cautious updating by (B,D,U,C). Then by definition $B+C = b+c$ and $D + C = d + c$ which implies

$$B + D + 2C = b + d + 2c. \tag{4}$$

Using this, the last condition of cautious updating yields $U = max(0,\ u,\ 1 - b - d - 2c)$. Since $u = 1 - b - d - c$, this implies $U = u$. Together with $1 = U + B + C + D = u + b + c + d$, it follows that $B + C + D = b + c + d$. In

30

combination with equation (4), this implies $C = c$. With this, $B + C = b + c$ and $D + C = d + c$ imply that $B = b$ and $D = d$. The proof for the second claim follows from a similar argument.

$\square$

9 Conclusions

Many classical approaches to reasoning address idealized situations, where the agents' information is consistent, closed under logical implication, and possibly even complete. These assumptions, of course, are at odds with many realistic reasoning scenarios, where the available evidence may be scarce and memory or observation faulty. In short, there is no guarantee for our available information to be consistent, nor complete. Yet, we would arguably hold that *some* valid inferences can be drawn from such imperfect information, as partial incompleteness or local contradictions may not preclude us from drawing conclusions about other parts of the data. As automated reasoning systems are becoming increasingly important, there is a need for a rigorous formal treatment of inferences from non-ideal information. To this end, a wealth of non-classical logical systems for dealing with uncertainty or conflict has been put forward, with Belnap-Dunn logic (BD) arguably the most prominent such framework.

However, the reasons for moving to non-normal, BD like frameworks apply equally well to probabilistic settings. Agents may, for instance, have inconclusive, probabilistic evidence for the truth or falsity of various statements. Just as in the classic case, if such information comes from different sources or different experiments, it needs not add up to 1, nor be mutually exclusive. It hence seems natural to investigate probabilistic extensions of BD. This was the focus of the current paper.

Paralleling recent work by Dunn (cf. Dunn, 2010; Dunn and Kiefer, 2019), we have investigated four-valued probability assignments that permit agents to have probabilistic beliefs about the truth and falsity of a statement, and about its gaps and gluts. More specifically, we have provided a theory of four-valued probabilities that slightly departs from Dunn's in its treatment of conjunctions. Yet, both are generalizations of Belnap-Dunn logic in that they coincide with BD whenever all probabilities are extremal, i.e. only assume the values of 0 and 1.

In this paper, we have clarified the connection between our four-valued probabilities and single valued non-standard probabilities as introduced by Childers, Majer and Milne (2019). By providing a translation function between the two approaches, we have shown these to be equivalent. Moreover, we have introduced probabilistic models as semantics for four-valued probabilities, and have provided a sound and complete axiomatization with respect to the class of all such models. Lastly, we have enriched our frameworks with dynamical operations for updating and aggregation. As for the former, we have provided versions of Jeffrey and Bayes' conditioning that work in non-standard and four-valued settings and have clarified the relation between these. For aggregation, finally,

we have studied a host of different aggregation policies, some of which go beyond what is available in classic probabilistic settings.

Of course, there are other approaches to weakening classic probability theory, not all of which have a corresponding logic as starting point. Many such approaches take probability or weights as central notion, but consider various cases where no exact probabilistic information is available. A typical example are inner measures intended to approximate probability from below (Fagin and Halpern, 1991). Their underlying idea, briefly, is that an agent might lack probabilistic evidence about some proposition φ, for instance when φ is not in the algebra of (possible) observations. The agent may, though, estimate a lower bound for the probability of φ by building on her available information about other propositions. Formally, this gives rise to an inner measures that only satisfy super-additivity instead of the classic additivity, i.e. $\mu_*(\varphi \vee \psi) \geqslant \mu_*(\varphi) + \mu_*(\psi)$, where $\varphi \wedge \psi$ a classical contradiction.

A related weakening of classic probability theory is Dempster-Shafer (DS) theory of belief (Shafer, 1976; Halpern, 2017). The starting point of this theory is an agent's evidence about some state of affairs, usually represented as a normalized measure on a boolean algebra of possible observations. This evidence then gives rise to a belief function, where $Bel(\varphi)$, the belief in some φ, is derived from all pieces of evidence that entail φ. As the agent might have strong evidence for a compound event, say $\psi \vee \varphi$, without having much evidence that entails either of its compounds alone, this belief function is super-additive in the sense defined above. More specifically, the degree of support for some A needs not be complementary to the support of $\neg A$. That is, $Bel(A)$ may be less than $1 - Bel(\neg A)$, just as in our framework. While $Bel(A)$ can be seen as a lower bound for the classical probability for A, the term $1 - Bel(\neg A)$, sometimes denoted the plausibility of A, is it's upper bound. The interval between both is then interpreted as the agent's uncertainty about A. As our presentation suggests, there is a tight connection between DS theory and inner measures approaches: both are equivalent, at least on a syntactic level where probabilities are associated to formulas, rather than states (Fagin and Halpern, 1991; Zhou, 2013).

Both, inner probabilitiy approaches and DS theory differ in two ways from our framework. In one dimension, our framework is more general than DS belief functions or inner probabilities, as it admits not only for uncertainty but also for conflict in probability assignments. By allowing for gluts, non-standard and four-valued probability assignments can represent contradictory information in ways that DS theory and inner measure frameworks cannot.

For a second difference consider a classic tautology such as $p \vee \neg p$. Working on a classical meta-theory, DS theory associates a probability of 1 to this tautology. Yet, when evidence is scarce, the belief values assigned to p and $\neg p$ need not add up to one, exemplifying the above super-additivity. In fact, it is compatibly with DS theory that both p and $\neg p$ are even assigned a belief of zero. In our framework, in contrast, uncertainty or conflict derive straight from the information available about p and $\neg p$, rather than from evidence about some larger proposition. Working with an non-classic, BD-metatheory, non-

classic information about literals extends to complex formulas such as $p \vee \neg p$, as witnessed in the inclusion-exclusion axiom (A3). This axiom, in fact, can be seen to stand in direct opposition to the theory of inner measures. Our axioms (A3) implies a subadditivity property (i.e. $\mu_*(\varphi \vee \psi) \leqslant \mu_*(\varphi) + \mu_*(\psi)$ when $\varphi \wedge \psi$ is a classical contradiction), in contrast to the superadditivity of DS theory and inner measures. A detailed comparison beyond DS belief functions and our approach would require a more careful analysis that exceeds the scope of this article. We leave this for future work.

Finally, another open line of inquiry concerns practical implications of the present framework. One may, for instance, ask how an ideally rational agent is to act if she has only imperfect information at her disposal. In future work, we hope to sketch the contours of a non-standard decision theory, that rests on four-valued probabilities in the same manner as traditional decision theory employs classic probability. Doing so, we hope, can help to fill a gap between current frameworks for decisions under risk and under uncertainty.

References

Alchourrón, C. E., P. Gärdenfors, and D. Makinson (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic 50*(2), 510 – 530.

Anderson, A. R. and D. Belnap, Nuel (1975). *Entailment: The Logic of Relevance and Necessity, Volume I*. Princeton: Princeton University Press.

Batens, D. (2001). A general characterization of adaptive logics. *Logique et Analyse 44*(173-175), 45–68.

Belnap, N. D. (1977). A useful four-valued logic. In *Modern uses of multiple-valued logic*, pp. 5–37. Springer.

Belnap, N. D. (2019). *How a Computer Should Think*, pp. 35–53. Springer International Publishing.

Childers, T., O. Majer, and P. Milne (2019). The (relevant) logic of scientific discovery. *(under review)*.

Christensen, D. (2007, 04). Epistemology of Disagreement: The Good News. *The Philosophical Review 116*(2), 187–217.

da Costa, N. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic 15*(4), 497–510.

da Costa, N. and V. Subrahmanian (1989). Paraconsistent logic as a formalism for reasoning about inconsistent knowledge bases. *Artificial Intelligence in Medicine 1*, 167–174.

Dunn, J. M. (1976). Intuitive semantics for first degree entailment and 'coupled trees'. *Philosophicl Studies 29*(3), 149–168.

Dunn, J. M. (2010). Contradictory information: Too much of a good thing. *Journal of Philosophical Logic 39*(4), 425–452.

Dunn, J. M. and N. M. Kiefer (2019). Contradictory information: Better than nothing? the paradox of the two firefighters. In *Graham Priest on Dialetheism and Paraconsistency*, pp. 231–247. Springer.

Elga, A. (2007). Reflection and disagreement. *Noûs 41*(3), 478–502.

Fagin, R. and J. Y. Halpern (1991). Uncertainty, belief, and probability. *Computational Intelligence 7*(3), 160–173.

Font, J. M. (1997). Belnap's four-valued logic and de morgan lattices. *Logic Journal of IGPL 5*(3), 1–29.

Halpern, J. (2017). *Reasoning about Uncertainty*. MIT press.

Jaskowski, S. (1948). Propositional calculus for contradictory deductive systems. *Studia Logica 24*, 143–157.

Jøsang, A. (1997). Artificial reasoning with subjective logic. In *Proceedings of the second Australian workshop on commonsense reasoning*, Volume 48, pp. 34. Citeseer.

Kelly, T. (2010). Peer disagreement and higher order evidence. In A. I. Goldman and D. Whitcomb (Eds.), *Social Epistemology: Essential Readings*, pp. 183–217. Oxford University Press.

Klein, D. and A. Marra (2020). From oughts to goals: A logic for enkrasia. *Studia Logica 108*(1), 85–128.

Kolmogorov, A. N. (2018). *Foundations of the theory of probability*. Courier Dover Publications.

Mares, E. D. (1997). Paraconsistent probability theory and paraconsistent bayesianism. *Logique et analyse 40*(160), 375–384.

Přenosil, A. (2018). *Reasoning with Inconsistent Information*. Ph. D. thesis, Charles University, Faculty of Philosophy.

Priest, G. (1979). Logic of paradox. *Journal of Philosophical Logic 8*, 219–241.

Priest, G. (2002). Paraconsistent logic. *Dov M. Gabbay and Franz Guenthner (eds.) Handbook of Philosophical Logic 6*, 287–393.

Priest, G. (2006). *In contradiction*. Oxford University Press.

Priest, G. (2007). Paraconsistency and dialetheism. *D. Gabbay and J. Woods (eds.) Handbook of the History of Logic 8*, 129–204.

Rescher, N. and R. Manor (1970). On inference from inconsistent premisses. *Theory and Decision 1*(2), 179–217.

Shafer, G. (1976). *A mathematical theory of evidence*, Volume 42. Princeton university press.

Zhou, C. (2013). Belief functions on distributive lattices. *Artificial Intelligence 201*, 1–31.