# Probability and non-classical logic

J. Robert G. Williams

April 15, 2012

Classical tautologies have probability 1. Classical contradictions have probability 0. These familiar features reflect a connection between standard probability theory and classical logic. In contexts in which classical logic is questioned—to deal with the paradoxes of self-reference, or vague propositions, for the purposes of scientific theory or metaphysical anti-realism—we must equally question standard probability theory.[1]

Section 1 reviews the connection between classical logic and classical probability. Section 2 briefly reviews salient aspects of nonclassical logic, laying out a couple of simple examples to fix ideas. In section 3 I explore modifications of probability theory, adopting an experimental methodology: the variations are motivated simply by formal analogies to the classical setting. In section 4, we look at two foundational *justifications* for the presentations of 'nonclassical probabilities' that are arrived at. In section 5-8, we look at extensions of the nonclassical framework: to conditionalization and decision theory in particular. Section 9 evaluates our progress.

I will assume that probabilities attach to *sentences* in a propositional language, which are also the locus for logical properties and relations (logical truth, falsity, consistency, consequence). And I will work initially with a *subjective* interpretation of probability.[2] On this picture, believe-true sentences to various degrees. The *probabilist* then maintains that to be ideally rational, the distribution of these degrees of belief must be *probabilistic*—i.e. satisfy the probability axioms.

## 1 Classical logic and semantics and classical probability

Consider a colour swatch, Patchy. Patchy is borderline between red and orange. The classical law of excluded middle requires the following be true:

(LEM) Either Patchy is red, or Patchy is not red.

[1]For the paradoxes of self-reference, (Field, 2008) provides a recent survey of nonclassical approaches. For vagueness, see inter alia (Williamson, 1994; Keefe, 2000; Smith, 2008). (Hughes, 1992) is a relatively accessible approach to the issues surrounding quantum logic, with chs.7 and 8 particularly pertinent to our concerns. For metaphysical anti-realism and logic, a locus classicus is Dummett (1991).

[2]For more, see (Hájek, Summer 2003).

Many regard (LEM) as implausible for borderline cases like Patchy—intuitively there is no fact of the matter about whether Patchy is red or not, and endorsing (LEM) suggests otherwise. This motivates the development of non-classical logic and semantics on (LEM) is no longer a logical truth.[3] But if one doubts (LEM) for these reasons, one surely cannot regard it is a constraint of rationality that one be certain—credence 1—in it, as classical probabilism would insist. One does not have to be a convinced revisionist to feel this pressure—even one who is (rationally) agnostic over whether or not logic should be revised in these situations, and so has at least some inclination to doubt LEM, is in tension with the principle that belief states that are not probablistic are irrational.[4]

We can view the distinctively classical assumptions embedded in standard probability theory from at least two perspectives. First, the standard *axiomatization* of probability (over sentences) makes explicit appeal to (classical) logical properties. Second, probabilities can be identified with expectations of *truth values* of sentences, where those 'truth values' are assumed to work in classical ways. We briefly review these two perspectives below in the classical setting, before outlining in the next section how they may be adapted to a nonclassical backdrop.

The following is a standard Kolmogorov-style set of axioms for probability over sentences in the propositional language $L$:

P1c. (Non-negativity) $\forall S \in L, P(S) \in \mathbb{R}_{\geq 0}$

P2c. (Normalization) For $T$ a (classical) logical truth, $P(T) = 1$

P3c. (Additivity) $\forall R, S \in L$ with $R$ and $S$ (classically) inconsistent, $P(R \vee S) = P(R) + P(S)$.

Various theorems of this system articulate further relations between logic and probability:

P4c. (Zero) For $F$ a (classical) logical falsehood, $P(F) = 0$;

P5c. (Monotonicity) If $S$ is a (classical) logical consequence of $R$, $P(S) \geq P(R)$;

(Normalization) is problematic for the logical revisionist who seeks to deny the law of excluded middle: under our interpretation of probability, it says that rational agents must be fully confident in instances of excluded middle. But it is not the only problematic principle. Some popular nonclassical settings say that (LEM) is true, but assert the following:

[3]The literature on this topic is vast. Two representatives of the contemporary debate are (Wright, 2001) and (Smith, 2008). Williamson (1994) is the most influential critic of nonclassical approaches in this area.

[4]The connection between logic and probability in these contexts is a major theme of Hartry Field's work in recent times. See Field (2000, 2003b,a, 2009).

(TVG) It's not true that Patchy is red, and it's not true that Patchy is not red.

On this—supervaluation-style—nonclassical setting, a disjunction can be true, even though each disjunct is untrue.[5] This motivates allowing high confidence in 'either Patchy is red or Patchy isn't red', and yet ultra low confidence in each disjunct. But this violates (Additivity).[6]

Still other, dialethic, nonclassical settings allow contradictions to be true. Let $L$ be the liar sentence ("this sentence is not true"). Some argue that the following holds:

(TC) $L \wedge \neg L$

Advocates of this view presumably have reasonably high confidence in (TC). But (Zero) rules this out.[7]

(Monotonicity) inherits problems both from (Zero) and (Normalization). Since $A \vee \neg A$ classically follows from anything, (Monotonicity) tells us that rational confidence in excluded middle is bounded below by our highest degree of confidence in anything. And since $L \wedge \neg L$ classically entails anything, (Monotonicity) tells us that rational confidence in the conjunction of the liar and its negation is bounded above by our lowest degree of confidence in anything. But revisionists object: one such revisionist thinks we should have higher confidence that hands exist (for example), than in sentence (LEM). Another thinks we should have lower confidence in the moon being made of green cheese, than we do in the conjunction of the liar and its negation.

Finally, what of (Non-negativity)? Many revisionists will find this unproblematic; notice that it doesn't appeal to (classical) logical relations explicitly at all. But the assumption that (subjective) probabilities are non-negative reals builds in, inter alia, that rational degrees of belief are *linearly ordered*. It's not crazy for a nonclassicst to question this assumption.[8]

The obvious moral from this brief review is that it would be madness for a logical revisionist to endorse *as articulating rationality constraints on belief* a probability theory that is based on the 'wrong' logic.

[5] Some further details are given later. For supervaluations, see inter alia (van Fraassen, 1966; Fine, 1975; Keefe, 2000).

[6] Compare (Field, 2000).

[7] See (Priest, 2006), who explicitly discusses the modifications of standard probability theory required to accommodate paraconsistent logics.

(Zero) of course follows from (P1c-P3c)—the issue is traceable to Additivity. Consider the pair $L \wedge \neg L$ and $\neg(L \wedge \neg L)$. The latter is a classical tautology so by P2c gets probability 1. Their disjunction is also a tautology (since it has a tautology as a disjunct) so also gets probability 1. The dialethist need not demur to this point (leading systems agree that these are tautologies). But because the pair are classically inconsistent (in virtue of having a classical logical falsehood as an element), P3c applies, giving $1 = P(L \wedge \neg L) + 1$, and hence forcing $P(L \wedge \neg L) = 0$.

[8] To take one example: Macfarlane (2010) argues for that certain graphs, rather than point-like credences, capture our doxastic states in the nonclassical setting he considers. Of course, many rivals to probabilism exist within the nonclassical setting (the indeterminate probabiities of (Levi, 1974) for example) and these may also have analogues here—but I'm setting aside such considerations for now.

We turn now from axiomatics to the second perspective: probabilities as expectations of truth value. To articulate this, we presuppose an 'underlying' credence function on a maximally fine-grained partition of possibilities ("worlds") (for simplicity, we take this to be finite). The only constraints imposed on this underlying credence $c$ is that the total credence invested across all possibilities sums to 1, and that the credence in a given possibility lies between 0 and 1 inclusive.[9]

Sentences are true or false relative to these worlds. Let $|S|_w$ be a function that takes value 1 iff $S$ is true at $w$, and value 0 iff $S$ is false at $w$—this we call the *truth value* of the sentence at $w$. Each underlying division of credence then induces a function from sentences to real numbers, as follows:

$$f(S) = \sum_w c(w)|S|_w$$

It turns out that such $f$s are exactly the probabilities over sentences. Classical probabilities are *convex combinations* of classical truth values; if we think of the truth values as a random variable over the space of worlds, they are *expectations* of truth values relative to the underlying $c$.[10]

In characterizing this connection, the association of 1s and 0s with truth statuses truth and falsity is crucial. *The True* and *the False* can't themselves be arithmetically manipulated; the arithmetical manipulations of 1 and 0 above make perfect sense. So why call these 'truth values'? The answer I will explore—and extend to the nonclassical case—is that the representations are justified only because they are the degree of belief that omniscient agents should invest in $S$, in situations where $S$ has that truth status. They reflect the omniscience-appropriate cognitive states; the 'cognitive loading' of the classical truth statuses.[11]

Since expectations of (classical) truth values lead to our familiar probability functions, satisfying the familiar axioms, all the problematic consequences for the logical revisionist arise once again. The revisionist faced with the expectational characterization of probabilities will pinpoint the appeal to *classical* truth value distributions as what causes the trouble. Faced with classical axiomatics, the natural strategy is to consider revised principles appealing to a nonclassical consequence

[9]It's worth noting that by calling this a 'credence' and the derived quantity 'expectations', I am building in an interpretation of the formalism. From a formal perspective, the characterization of probability being appealed to is simply as convex combinations of truth value assignments. Compare (Paris, 2001; **?**).

[10]A quick word on the use of 'expectations'. One can view the above as setting out a relation between two probability spaces—one taking a sample space of worlds (and an event space on which probabilities are defined, of sets of those worlds); the other ranging over sentences. The truth values of each sentence defines a random variable over the sample space of the world-probabilities. Where there are only finitely many worlds, the full technology of a worldly probability space isn't needed. But if there are infinitely many (so that we can no longer rely on summations to converge) the appropriate generalization will be to appeal to expectations (assuming that the set of worlds where $S$ takes a given truth value is a member of the worldly event space).

[11]Of course, It is ultimately conventional that we represent full belief via the number 1—what we're really pointing to here is a match between the representation of truth values and the representation of maximal degree of belief.

relation. Faced with the expectational characterization, the natural strategy is to explore variations where nonclassical *truth value distributions* are appealed to.

## 2 Non-classical logic and semantics

Non-classical logics come in wild and wonderful variety.[12] To fix ideas, I set out a three-valued setting that allows us to characterize a handful of sample logics. A *Kleene truth status assignment* involves, not a scattering of two statuses (Truth, Falsity) over sentences, but a scattering of three—call them for now $T$, $F$ and $O$. The distribution over compound sentences must accord with the (strong) Kleene truth-tables for negation, conjunction and disjunction:

| $A$ | $\neg A$ |
|---|---|
| $T$ | $F$ |
| $O$ | $O$ |
| $F$ | $T$ |

| $A \wedge B$ | $T$ | $O$ | $F$ |
|---|---|---|---|
| $T$ | $T$ | $O$ | $F$ |
| $O$ | $O$ | $O$ | $F$ |
| $F$ | $F$ | $F$ | $F$ |

| $A \vee B$ | $T$ | $O$ | $F$ |
|---|---|---|---|
| $T$ | $T$ | $T$ | $T$ |
| $O$ | $T$ | $O$ | $O$ |
| $F$ | $T$ | $O$ | $F$ |

(In the last two tables, the horizontal headers represent the truth status of $A$, and the vertical headers the truth status of $B$, and corresponding entry the resultant truth status of the complex sentence). We have various options for characterizing logical consequence on this basis:

| | |
|---|---|
| **Kleene logic:** | $A \vdash_K B$ iff on every Kleene truth status assignment, if $A$ is $T$, then $B$ is $T$ too. |
| **LP:** | $A \vdash_L B$ iff on every Kleene truth status assignment, if $A$ is $T$ or $O$, then $B$ is $T$ or $O$ too. |
| **Symmetric logic:** | $A \vdash_S B$ iff on every Kleene truth status assignment, if $A$ is $T$, then $B$ is $T$; and if $A$ is $T$ or $O$, then $B$ is $T$ or $O$. |

However we characterize consequence, logical truths (tautologies) are those that are logical consequences of everything; logical falsehoods are those sentences of which everything is a logical consequence; an inconsistent set is a set of sentences of which everything is a logical consequence.

The strong Kleene logic is a simple example of a non-classical logic where excluded middle is no tautology: if $A$ has status $O$, then so will $\neg A$, and looking up the truth table above, so will $A \vee \neg A$. If $B$ has value $T$, this provides a Kleene-logic countermodel to the claim that $A \vee \neg A$ follows from everything. By contrast, excluded middle will be a tautology on the LP understanding of consequence. $A \vee \neg A$ can never have the status $F$; and that suffices to ensure it follows from everything on the LP definition. But LP provides us with a simple example of a paraconsistent logic—one on which explicit

[12]For a general introduction to nonclassical logics, including the Kleene logic and LP discussed below, see (Priest, 2001) and for further philosophical discussion, see(Haack, 1978).

'contradictions' $L\wedge\neg L$ do not 'explode'—they do not entail everything. The symmetric characterization has both features—contradictions are not inconsistent (/explosive) and excluded middle is no tautology.

What shall we make of these $\boldsymbol{T}$s, $\boldsymbol{F}$s and $\boldsymbol{O}$s? In the classical setting, we ordinarily assume that (in context) sentences are true or false *simpliciter*—that these are monadic properties that sentences (in context) either possess or fail to possess. Truth status distributions represent possible ways in which such properties can be distributed. We could regard the Kleene distributions in the same way: rather than two properties, there are three; but we can ask about what the *actual* distribution is, and about the nature of the properties so distributed. And perhaps such information could then be used to motivate one choice of logic over another. A package of non-classical logic and semantics meeting this description we call *semantically driven.*

But one needn't buy into this picture, to use the abstract three valued 'distributions' to characterize the relations $\vdash_K$, $\vdash_L$ and $\vdash_S$. Hartry Field has urged this approach to logic in recent times. Semantics for Field does not involve *representing* real alethic statuses that sentences possess. It is rather an *instrumental device* that allow us to characterize the relation that is of real interest: logical consequence.[13] The $\boldsymbol{T}$s, $\boldsymbol{F}$s and $\boldsymbol{O}$s could remain uninterpreted, since they're merely an algebraic tool used to describe the consequence relation. And the question of which category a sentence (like (LEM)) falls into would simply be nonsense.

Let's suppose that we do not go Field's way, but take our nonclassical logic to be semantically driven, so that sentences have properties corresponding to (one of) $\boldsymbol{T}$, $\boldsymbol{F}$ and $\boldsymbol{O}$. What information would we like about these statuses, in order to further understand the view being put forward? Consider the classical case. Here the statuses were Truth and Falsity; and these statuses were each 'cognitively loaded': we could pinpoint the ideally appropriate attitude to adopt to each. In the case of a true sentence this was full belief (credence 1); and in the case of a false sentence, utter rejection (credence 0). We'd like to know something similar about the nonclassical statuses $\boldsymbol{T}$, $\boldsymbol{F}$ and $\boldsymbol{O}$. If $S$ has status $\boldsymbol{O}$, should an omnisicient agent invest confidence in $S$? If so, to what level? Would they instead suspend judgement? Or feel conflicted? Or groundlessly guess?

Call a semantics *cognitively loaded* when each alethic status that it uses is associated with an 'ideal' cognitive state. Nonclassicists endorsing a semantically-driven conception of logic may still not regard it as cognitively loaded. For example: Tim Maudlin (2004) advocates a nonclassical three valued logic (the Kleene logic, in fact) in the face of semantic paradoxes, but explicitly *denies* that there is any cognitive loading *at all* to the middle status $\boldsymbol{O}$. Indeed, he thinks that the distinctive characteristic of $\boldsymbol{O}$ that makes it a 'truth value gap' rather than a 'third truth value', is that it gives no guidance for belief or assertion.

The nonclassical logics we will focus on will be semantically driven, cognitively loaded, and further, will be loaded with cognitive states of a particular kind: with standard degrees of belief, represented by real numbers between 1 (full certainty) and 0 (full

[13](Field, 2009, passim). Field doesn't propose eliminate truth-*talk* from our language—he favours a deflationarist approach to the notion—but he does not regard truth-statuses as something in terms of which logic can be characterized.

rejection, anti-certainty). This last qualification is yet another restriction. There's no a priori reason why the cognitive load appropriate to non-classical statuses shouldn't take some other form—calling for some non-linear structure of degrees of belief, or suspension rather than positive partial belief, or somesuch. Such views motivate more radical departures from classical probabilities than the ones to be explored below.[14]

Consider the following three loadings of Kleene distributions (numerical values represent the degree of belief that an omniscient agent should adopt to a sentence having that status):

| **Status:** | **T** | **O** | **F** |
|---|---|---|---|
| **Kleene loading:** | 1 | 0 | 0 |
| **LP loading:** | 1 | 1 | 0 |
| **Symmetric loading:** | 1 | $\frac{1}{2}$ | 0 |

The loadings differ on the attitude they prescribe for $O$ under conditions of omniscience: utter rejection, certainty, or half-confidence respectively. They motivate informal glosses on this truth status, respectively: neither true nor false; half-true; or both true and false. Furthermore, the loadings correspond systematically to the logics mentioned earlier: in each case, logical consequence requires that there be no possibility of a drop in truth *value*, where the truth value is identified with the cognitive load of the truth status.[15]

We continue to use the three Kleene-based logics as worked examples. But there are many, many ways of setting up non-classical logics. So long as the logics are semantically driven, and cognitively loaded with real values, then our discussion will cover them.

# 3 Truth value expectations and the logic in probability

The significance of cognitive loadings for us is that they give a natural way to extend the expectational characterization.[16] Recall the classical case: for an appropriate $c$, the probability of each $S$ must satisfy:

$$P(S) = \sum_w c(w)|S|_w$$

[14]Three potential examples of this are Wright's notion of a *quandary* (Wright, 2001); Macfarlane's credence profiles (Macfarlane, 2010) and whatever corresponds to the partially ordered statuses for which (Weatherson, 2005) argues.

[15]If the loadings are simply 1 and 0, this corresponds to the familiar distinction between 'designated' and 'undesignated' truth statuses, and the characterization of consequence as preservation of designated status (cf. Dummett, 1959, e.g.). The 'no drop' characterization generalizes this somewhat.

[16]See in particular (Paris, 2001) for this strategy. Compare also (Zadeh, 1968) and (Smith, 2010).

Consider the limiting case where $c$ is zero everywhere but the actual world (i.e. conditions of omniscience). The above equation then simplifies to $P(S) = |S|_w$. That is: under conditions of omniscience, the subjective probability matches the numerical value assigned as $S$'s truth value; hence, that number will be the cognitive load of the truth status.[17] In this way, the Kleene, LP and Symmetric loadings induce three kinds of 'nonclassical probabilities', as expectations of the respective truth values.

A nice feature of this approach is that the axiomatic perspective generalizes in tandem with the expectational one.[18] Consider the following principles, for parameterized consequence relation $\vdash_x$:

P1x. (Non-negativity) $\forall S \in L, P(S) \in \mathbb{R}_{\geq 0}$

P2x. (Normalization) If $\vdash_x T$, then $P(T) = 1$

P3x. (Additivity) $\forall R, S \in L$ such that $R, S \vdash_x$, $P(R \vee S) = P(R) + P(S)$

P4x. (Zero) If $F \vdash_x$, then $P(F) = 0$;

P5x. (Monotonicity) If $R \vdash_x S$, then $P(S) \geq P(R)$;

If we pick the Kleene loadings, then these five principles are satisfied by any 'nonclassical probability' (expectation of truth value), so long as we use the Kleene logic (set $x = K$). Mutatis mutandis for the LP and Symmetric loadings and logics.

It's useful to add two further principles—extensions and variations on (Additivity)—which are also satisfied by expectations of truth values, for the appropriate consequence relation:

P3x+. (IncExc) $\forall R, S \in L, P(R) + P(S) = P(R \vee S) + P(R \wedge S)$

P3x$^d$. (Dual additivity) $\forall R, S \in L$, if $R \vee S$ is a tautology, then $P(R) + P(S) - 1 = P(R \wedge S)$

In the presence of (Zero) and (Normalization) respectively, (IncExc) will entail the original (Additivity) and (Dual Additivity).[19] (Additivity) itself is weak in logics with few or no inconsistencies, such as $LP$ (if there are no inconsistent pairs of sentences, then the antecedent is never satisfied, and the principle becomes vacuously true); dual

[17] If the cognitive loads are something *other* than particular degrees of belief—if they require 'suspension' or some other more complex attitude; or if the degrees of belief are no longer representable by reals—then we cannot generalize the 'expected truth value' characterization of probability so easily. At minimum, we'd need to think about what we mean by expectations of things that do not take real number values. This can be done; it is interesting to explore; but I will not pursue it here. A similar situation arises if the logical revisionists refuses to acknowledge or answer the question about the cognitive load of their putative nonclassical truth statuses; or more radically, if like Field they do not take an alethically-driven approach to logic in the first place.

[18] For instances of this observation in specific settings, see e.g. (Weatherson, 2003; Field, 2003b; Priest, 2006). As we shall see, (Paris, 2001) gives a particularly elegant treatment of many cases.

[19] In addition to (P4x) and (P3x*), we need to assume that the conjunction of an inconsistent pair is a logical falsehood. But this is not often targetted for revision.

Additivity is correspondingly weak in logics with few or no tautologies, such as the Kleene logic. The weaknesses are combined in the Symmetry logic. But their generalization, (IncExc), makes no mention of the logical system in play, and so retains its strength throughout.

The connection between expectational and logical characterizations illustrated above is very general. Scatter truth statuses over sentences howsoever one wishes, with whatever constraints on permissible distributions you like. Make sure you associate them with real-valued 'cognitive loads'—degrees of belief within $[1,0]$, so that we can straightforwardly define the notion of possible expected truth value, by letting $|S|_w$ be equal to the cognitive load of the status that $S$ has at $w$. We consider the following logic:

**No drop:**

$A \vdash B$ iff on every truth status assignment $w$, $|A|_w \leq |B|_w$.

It's straightforward to check that (P1x), (P2x), (P4x) and (P5x) will then hold of all the expected truth values.

The status of (Additivity) and its variants is more subtle. These principles make explicit mention of a particular connective, so its no surprise that whether or not they hold depends how those connectives behave. (IncExc) will hold iff we have the following:[20]

$$|A|_w + |B|_w = |A \vee B|_w + |A \wedge B|_w$$

Classical logic, and many non-classical logics, satisfy this principle. But others do not.[21]

In the classical setting, we had more than just a grabbag of principles satisfied by probabilities: we had a complete axiomatization. An obvious question is whether some subset of nonclassical variants is complete with respect to nonclassical expected truth values in a similar way.

Paris (2001) delivers an elegant result on this front. Among much else of interest to us, he shows that the non-classical verisons of (Normalization), (Zero), (Monotonicity) and (IncExc) deliver complete axiomatizations of a wide range of nonclassical probabilities. The conditions for this result to hold are that: (i) truth values (in our terminology: the cognitive loadings of truth statuses) are taken from $\{0,1\}$; (ii) $A \vdash_k B$ is given the 'no drop' characterization mentioned earlier;[22] and (iii) the following pair is satisfied:

[20] The 'if' direction follows by the linearity of expectations. The 'only if' direction holds by considering the special case of probability where the underlying credence all lies on a single world, $c$, and hence the probability coincidens with truth values.

[21] The supervaluational setting mentioned earlier is one example. Suppose supervaluational truth value gaps are value 0 (as in the Kleene loading). Then if $A$ is gappy, $A$ and $\neg A$ are value 0, $A \vee \neg A$ is value 1, and $A \wedge \neg A$ is value 0. So the LHS is smaller than the RHS.

[22] Paris formulates this in terms of the preservation of value 1 across the sequent, but given (i), these are equivalent.

$$
\begin{array}{lllll}
(\mathcal{T}2) & V(A)=1 \wedge V(B)=1 & \iff & V(A \wedge B)=1 \\
(\mathcal{T}3) & V(A)=0 \wedge V(B)=0 & \iff & V(A \vee B)=0.
\end{array}
$$

This applies, for example, to the Kleene and LP loadings mentioned above, as well as the original classical case. It's application goes well beyond this: for example, to appropriate formulations of intuitionistic logic.[23]

Beyond this, it is a matter of hard graft to see whether similar completeness results can be derived for settings that fail the Parisian conditions (one representative of which is our Symmetric logic) Drawing on the work of Gerla (2000) and Di Nola *et al.* (1999), Paris shows that a similar result holds for finitely valued (Łukasiewicz) setting and Mundici (2006) later extended this to the continuum valued fuzzy setting.[24]

We have already mentioned supervaluational logics several times. These are widely appealed to in the philosophical literature. They arise as a generalization of classical truth values, via the assumption that the world and our linguistic conventions settles, not a *single* intended classical truth value assignment over sentences, but a set of co-intended ones. Sentences are supertrue if they are true on all the cointended assignments, and superfalse if they are false on all of them. This allows supertruth gaps: cases where the assignments for $S$ differ, and so it is neither supertrue nor superfalse. We shall assume that supertruth has a cognitive loading of 1, and other statuses have a loading of 0 (compare the Kleene loading earlier). This delivers the results already flagged. For example, as a classical tautology, (LEM) is true on each classical assignment, and a fortiori true on the set of co-designated ones, so it will always be supertrue (value 1). But this is compatible with each disjunct being a supertruth gap (value 0).

This case is covered by a theorem that Paris gives, drawing on the work of Shafer (1976) and Jaffray (1989). For the propositional language under consideration, the results show that convex combinations of such truth values are exactly the *Dempster-Shafer belief functions*. These may be axiomatized thus:

$$
\begin{array}{llll}
(\mathcal{DS}1) & \vdash A & \Rightarrow & P(A)=1 \\
 & A \vdash & \Rightarrow & P(A)=0 \\
(\mathcal{DS}2) & A \vdash B & \Rightarrow & P(A) \leq P(B) \\
(\mathcal{DS}3) & \multicolumn{3}{l}{P(\bigvee_{i=1}^{m} A_i) \geq \sum_S (-1)^{|S|-1} P(\bigwedge_{i \in S} A_i)}
\end{array}
$$

(where $S$ ranges over non-empty subset of $\{1,\ldots,m\}$).[25] Note how the third condition here replaces the earlier (IncExc), which fails in the supervaluational setting.

[23]For the intuitionistic case, compare Weatherson (2003). Paris reports the general result as a corollary of a theorem of Choquet (1953)

[24]The major difference between the 3-valued Kleene based setting and the Łukasiewicz settings is the addition of a stronger conditional—and this is crucial to the proofs mentioned. Its notable that Paris provides axiomatizations not in terms of a 'no drop' logic, but in terms of the logic of 'preserving value 1'. This is possible because the 'no drop' consequence is effectively encoded in the 1-preservation setting via tautological Łukasiewicz conditionals.

[25]Paris's initial formulation is slightly different, and uses classical logic (p.7), but as he notes this is extensionally equivalent to current version using the 'no drop' logic over the 'supervaluational' truth

It's important to note that these kind of completeness results are often sensitive to the exact details of the sentences we are considering. We do *not* have a guarantee that the completeness result will generalize when we add expressive resources to the language. This is one reason why the earlier Choquet-Paris result, which applies to *all* languages equipped with a semantics meeting the stated conditions, are so attractive.

While the completeness proofs are interesting and elegant, from a philosophical perspective, the identification of a reasonably rich body of principles that are *hold good* of nonclassical probabilities is of independent philosophical interest. Only the most radical Bayesians think that satisfying probabilistic coherence is *all* that there is to rationality; and so even if satisfying the axioms sufficed for probabilistic coherence, it would be contentious to conclude that it sufficed for rationality. On the other hand, so long as probabilistic coherence *is* a constraint on rational belief in the nonclassical setting, then what we learn from the above is that *violating* the principles suffices for *irrationality*. A natural next question, therefore, is whether the 'nonclassical probabilities' that we have identified so far, have the same claim as classical probabilities in the classical setting, to provide constraints on rational belief.

# 4 Foundational considerations: dutch books and accuracy domination

In well-behaved nonclassical settings, we have seen a nice generalization of probability theory in prospect. But is this just coincidence? Or can we argue that this is the *right* way to theorize about subjective probabilities in such a setting?

I will focus on two arguments for 'probabilism' familiar from the classical case: the dutch-bookability of credences that violate the axioms of probability, and the accuracy-domination arguments advocated by Jim Joyce.[26] Jeff Paris has shown how the first can be generalized, showing that credences that are not *expectations of truth value* in the relevant sense are dutch-bookable.[27] And for similar formal reasons, such credences are also 'accuracy dominated'.[28]

In the classical case, accuracy domination arguments consist taking a belief state $b$, and assessing it at each world $w$ for its degree of 'accuracy'. How accuracy to be measured is the leading issue for this approach; but in all cases, the starting point is to compare a degree of belief (within $[1,0]$) to the 'truth value' of the sentence in question. But comparing a number with *Truth* or *Falsity* is not terribly tractable. So one standardly compares a given degree of belief with the cognitive loading of the truth status—how close, overall, the degrees of belief are to the 1s and 0s that an omniscient (perfectly accurate) agent would have in that world.

values (p.10).

[26](Joyce, 1998, 2009).

[27](Paris, 2001).

[28]See (De Finetti, 1974) for the formal background to both results (in the latter case, with a quite different interpretation of its significance). (Williams, forthcominga) examines the relation between the two results.

The result (relative to many plausible ways of measuring accuracy) is the following: if one's beliefs $b$ are not (classical) probabilities), then it is always possible to construct a rival belief state $c$ such that $c$ is more accurate than $b$ *no matter which world is actual.* If such accuracy-domination is an epistemic flaw, then only probabilistic belief states can be flawless. This is offered as a rationale for why subjective *probabilities* in particular, should be constraints on ideally rational belief.

What's important for our purposes is that the argument generalizes.[29] As earlier, suppose the cognitive loading of some nonclassical truth statuses are real numbers between $[1,0]$. We use the very same accuracy measures as previously, to measure closeness of beliefs to these nonclassical truth values. And it turns out that if the belief state is not representable as an expectation of truth values (relative to some underlying $c$ or other) then it will be accuracy-dominated.[30] The accuracy-based arguments for probabilism thus offer a justification for the claim that nonclassical probabilities, as characterized in the previous section, should indeed play the role of constraints on rational partial belief. (Of course, whether it's a good justification in either setting is contested).[31]

Perhaps the most familiar foundational justification for the claim that rational partial beliefs must be probabilistic comes from dutch book arguments. The key claim is that one's degrees of belief are fair betting odds in the following sense: if offered a bet that pays £1 if $A$, and £0 if $\neg A$, then if you believe $A$ to degree $k$, then you should be prepared to buy or sell the bet for $k$. Suppose that degrees of belief do play this role. Then if $b$ is an improbabilistic belief state, there is a set of bets—a 'dutch book'—such that you are prepared to buy each bet within it, but which ends up giving you a loss *no matter what.* Pragmatically viewed, a set of beliefs that open you up to sure-losses may seem flawed. Or more generally, one might think that the belief state is flawed because it commits you to viewing the book as both to be accepted (since consisting of bets that your belief state makes you prepared to accept) and also obviously to be rejected (since it leads to sure loss). So you are committed to inconsistency.

Dutch book justifications for probabilism, like accuracy domination arguments, are contentious.[32] But independently of whether they persuade, are they adaptable to our case? Suppose one has bought a bet that pays out 1 if $A$ and 0 otherwise. If one is in a nonclassical setting, one can be faced with a situation where $A$ takes some nonclassical truth status. The returns on such a bet then depends on how the bookie reacts. Call a real number $k \in [1,0]$ the *pragmatic loading* of a truth status $X$, just in case the right way for the bookie to resolve such a bet, given that $A$ has status $X$, is to give the gambler £$k$.[33] Clearly the pragmatic loading of classical truth should be 1, and the pragmatic loading of classical falsehood is 0. Just as with cognitive loadings of nonclassical truth statuses, there are many many ways one might consider assigning pragmatic loadings

[29] At least, the results in Joyce (1998) generalize. As discussed in (Williams, forthcomingb) the situation is much more complex for the main argument in (Joyce, 2009). The variation in proofs is significant, since different assumptions about the accuracy measure are involved in each case.

[30] For further details, see (Williams, forthcomingb).

[31] For a recent review of some criticisms of the accuracy-based argument for probabilism, see (Hájek, 2008).

[32] Again, (Hájek, 2008) reviews and extends the critiques of such arguments.

[33] Compare the stipulations in (Paris, 2001) on the returns of bets in a nonclassical setting.

(and just as with cognitive loadings, there are pragmatic loadings for truth statuses that don't fit the above description—the option of 'cancelling the bet' for example—as well as the option to deny that truth statuses *have* any identifiable pragmatic loading).

Suppose we have real-valued pragmatic loadings for truth statuses, however. Then we can make sense of resolving bets in a nonclassical setting, and can consider what kinds of belief states are immune from dutch books. Happily, the answer is just as you would expect: immunity from dutch books is secured when (and only when) the belief state is a 'nonclassical probability'—an expectation of the relevant truth values.[34]

It's worth noting that in this last result, the 'truth value' of a sentence refers to the *pragmatic loading* of the relevant truth status, whereas in the previous results it referred to the *cognitive loading* of the truth statuses. If they differed, then we might have inconsistent demands—for example, if the cognitive loading of the 'other' status was 0.5 (omniscient agents are half-confident in $A$, when $A$ is $O$), but its pragmatic loading was zero (one doesn't receive any reward for a bet on $A$, given that it is half-true) then being 0.5 confident in $A \wedge \neg A$ might be entirely permissible from the an accuracy-domination point of view, but one that makes you dutch-bookable. The way to avoid this, of course, is to have cognitive and pragmatic loadings coincide. It is interesting to speculate on *whether* they should coincide, and if so *why* (I can imagine philosophers taking cognitive value as primary, and arguing on this basis that the *right* way to resolve bets accords with the pragmatic loading; but I can equally envisage philosophers arguing that pragmatic loadings are primary, and that these *give the reasons* why a particular cognitive loading attaches to a truth status. I can also imagine someone who takes both as coprimitive, but argued ('transcendentally') that they must coincide, otherwise rationality would place inconsistent demands on agents).

Both dutch book and accuracy arguments—and much of the debate between their advocates and critics—can be replayed in nonclassical setting. This should bolster our confidence that we have the *right generalization* of probability theory for the cases under study. And none of the results just mentioned make *any* assumptions about the particular kind of truth value distributions or logical behaviour of the connectives in question—other that the truth values, in the relevant sense, lie within $[1,0]$. These are extremely general results.

# 5 Conditional probabilities and updating

Subjective probability without a notion of conditional probability would be hamstrung. If we are convinced (at least pro tem) that we have a nonclassical generalization of probability, then the immediate question is how to develop the theory of conditional probability within this setting. Three approaches suggest themselves. The first is simply to carry over standard characterizations of conditional probabilities, for example, the ratio formula (restricted to cases where $P(A) \neq 0$; I often leave such constraints tacit in what follows):

[34] The result follows from dutch book arguments for expectations in (De Finetti, 1974) and is interpreted in the way just mentioned in (Paris, 2001). For more discussion, see (Williams, forthcominga).

$$P(B|A) := P(B \wedge A)/P(A)$$

The second is to investigate axiomatizations of probability in which conditional probability is the basic notion (of course, if left unchanged, these lead to classical probabilities). One experimentally or foundationally investigates variations of these axioms, much as we did for the unconditional case above. Third, we can look to the work we want conditional probability to do, and try to figure out what quantity is suited, within the non-classical setting, to play that role. It is this third approach we adopt here, with a focus initially on the role of conditional probability in *updating* credences.

Conditional probabilities will be two-place functions from pairs of propositions to real numbers, written $P(\cdot|\cdot)$. The key idea will be that this should characterize an update policy: when one receives total information $A$, one's updated unconditional beliefs, should match the old beliefs conditional on $A$: $P_{\text{new}}(\cdot) = P_{\text{old}}(\cdot|A)$. If updating on information isn't to lead us into irrationality, then a minimal constraint on conditional probabilities fit to play this role is that the result of 'conditioning on $A$' as above, should be a probability. (It turns out, incidentally, that straightforwardly transferring the 'ratio formula' treatment of conditional probabilities can violate this constraint).[35]

Classical conditionalization on $A$ can be thought of as the following operation: one first sets the credence in all $\neg A$ worlds to zero; leaving the credence in $A$-worlds untouched. This, however, won't give you something that's genuinely a probability (for example, the 'base credences' no longer sum to 1). So one renormalizes the credences to ensure we do have a probability, by dividing each by the total remaining credence $P(A)$.

We could generalize this in several ways, but here is the one we will consider. Take the first step in the classical case: wiping out credence in worlds where the proposition is false (truth value 0) and leave alone credence in worlds where the proposition is true (truth value 1). Another way to put this is that the updated credence in $w$, $c_A(w)$ (prior to renormalization) is given by $c(w)|A|_w$: the result of multiplying the prior credence in $w$ by the truth value of $A$ at that possibility. Since we have real-valued truth values in our nonclassical settings, we can simply transfer this across. The credence is scaled in proportion to *how true* $A$ is at a given possibility. Renormalizing is achieved just as in the classical setting, by dividing by the prior credence in $A$.[36] Notice that by focusing on how the underlying credence $c$ is altered under conditionalization, we have guaranteed that the function $P_A(X)$ defined by this procedure will be an expectation of truth values, and so a nonclassical probability in our sense. We set $P(X|A) := P_A(X)$.

The characterization of the update procedure can be set down as follows:

[35] Suppose that we are working within the 'symmetric/half truth' nonclassical setting, suppose $P(A) = P(\neg A) = P(A \wedge \neg A) = 0.5$—which is certainly permitted by the relavant nonclassical probabiltiies. Now consider $P(A \wedge \neg A|A)$. By the ratio formula, this would be $P(A \wedge \neg A \wedge A)/P(A) = P(A \wedge \neg A)/P(A) = 0.5/0.5 = 1$. So $P_{\text{new}}(A \wedge \neg A) = 1$. But no probability (expectation of truth value) in this setting can have this exceed 0.5.

[36] To see this, note that the result of the process is to give a 'base credence' over worlds which may add to less than 1. The sum total is given by $\sum_{w \in W} c_A(w) = \sum_{w \in W} c(w)|A|_w$. But by construction this is exactly $P(A)$. Hence dividing by $P(A)$ will renormalize the base credence, making it sum to 1, after the procedure described above.

$$P_A(X) = \sum_{w \in W} \frac{c_A(w)}{P(A)} |X|_w$$

Rearranging the right hand side, this gives the following fix on conditional probability:

$$P(X|A) = \frac{\sum_{w \in W} c(w)|A|_w|X|_w}{P(A)}$$

Now, if we have a connective $\circ$ such that for arbitrary $A$ and $B$, at any $w$:

$$|A|_w|B|_w = |A \circ B|_w$$

then it follows:

$$P(X|A) = \frac{\sum_{u \in W} c(u)|A \circ X|_u}{P(A)} = \frac{P(X \circ A)}{P(A)}$$

Certain nonclassical settings already have a connective $\circ$—in the classical, Kleene and LP settings, $\wedge$ plays the role, and so the familiar ratio formula with relatively familiar conjunctive connectives is derived. A more exciting example is the main conjunctive connective of the product fuzzy logic.[37] In other settings, it is well defined truth function, but would require an extension of the language to introduce. But it is *not automatic* that the conditions for $\circ$ can be met by a truth function, in arbitrary nonclassical systems. Consider, for example, the Symmetric loading of the Kleene assignments. A sentence that has the truth status $O$ gets the truth value 0.5. So $A \circ A$ would by construction have to take the truth value 0.25; $(A \circ A) \circ A$ would have to take the value 0.125, and so on. But in the symmetric setting, there are no truth statuses that have these loads. (A fortiori, $\circ$ is clearly *not* the Symmetric connective $\wedge$). The process of conditionalization that was described works perfectly well in the Symmetric setting as a way to shift from one nonclassical probability to another on receipt of the information that $A$. It's just that it doesn't have a neat formulation that mirrors the ratio formula.

Much more on these nonclassical conditional probabilities (focusing on fuzzy logic settings) is available in (Milne, 2007, 2008)—who cites (Zadeh, 1968) as the source for the conception. Milne shows how to provide a synchronic dutch book argument for this characterization of conditional probability, relative to the assumption (i) that conditional probabilities give fair betting odds for conditional bets; (ii) that nonclassical conditional bets are to be resolved a certain way (in particular, that they are 'progressively more and more called off' as the truth value of the condition gets lower and lower).[38] As Milne emphasizes, the assumption (ii) is crucial; in principle, there are many ways one might consider handling conditional bets in this setting, which would vindicate different conceptions of conditional probability. (Williams, forthcominga) gives

[37]For the product logic, see (Hájek, 1998).

[38]If we want to run these arguments in a *general* nonclassical setting, we need to be careful about how such

a nonclassical generalization of the Teller-Lewis diachronic dutch book argument, but (although it gives us non-trivial information) it is even worse at giving leverage on the crucial case of conditionalizing on nonclassical propositions.

In the light of this, the real test for a proposed generalization of conditional probabilities lies in its applications—as an update procedure and elsewhere. To give a flavour of some important ways it generalizes classical conditional probability, we show how some key results generalize.

The analogue of Bayes' theorem is immediate:

$$P(A|B) = \frac{P(A \circ B)}{P(B)} = \frac{P(A \circ B)}{P(A)} \frac{P(A)}{P(B)} = P(B|A) \frac{P(A)}{P(B)}$$

Further key classical results also carry over:

1. **Lemma.** Assume that $\forall w, |\neg A|_w = |1 - A|_w$. Then $P(C) = P(C \circ A) + P(C \circ \neg A)$.

   Proof. First note that relative to arbitrary $w$,

   $$|C| = |C|(|A| + 1 - |A|) = |C|(|A| + |\neg A|) = |C||A| + |C||\neg A| = |C \circ A| + |C \circ \neg A|$$

   For arbitrary nonclassical probability $P$, there's an underlying credence-over-worlds $c$ such that $P(A) = \sum_w c(w)|A|_w$. So in particular

   $$P(C) = \sum_w c(w)|C|_w = \sum_w c(w)(|C \circ A| + |C \circ \neg A|)$$

   But this in turn is equal to:

   $$\sum_w c(w)|C \circ A| + \sum_w c(w)|C \circ \neg A| = P(C \circ A) + P(C \circ \neg A)$$

   as required.

2. **Corollary.** $P(C) = P(C|A)P(A) + P(C|\neg A)P(\neg A)$. Follows immediately from the above by the $\circ$-ratio formula for conditional probability.

More generally, call $\Gamma$ a *nonclassical partition* if in each world, the sum of the truth values of the propositions in $\Gamma$ is 1 (thus our assumption that $|\neg A| = 1 - |A|$ ensured that $A, \neg A$ was a partition). Then replicating the above proof delivers:

1. **Generalized Lemma.** $P(C) = \sum_{\gamma \in \Gamma} P(C \circ \gamma)$, so long as $\Gamma$ is a partition.

2. **Generalized Corollary.** $P(C) = \sum_{\gamma \in \Gamma} P(C|\gamma)P(\gamma)$, so long as $\Gamma$ is a partition.

It's nice to have this general form since there are some settings (supervaluational semantics for example) where the truth values of $A$ and $\neg A$ don't sum to 1; the partition-form is still applicable even though the first result is not.

Another useful result is that, if $P_C$ is the probability that arises from $P$ by updating on $C$, then $P_C(A|B) = P(A|B \circ C)$. This follows straightforwardly from the ratio formula. For:

$$P_C(A|B) = P_C(A \circ B)/P_C(B) = \frac{\frac{P(A \circ B \circ C)}{P(C)}}{\frac{P(B \circ C)}{P(C)}} = \frac{P(A \circ B \circ C)}{P(B \circ C)} = P(A|B \circ C)$$

This all looks promising. On the other hand, there are some surprising divergences. Conditional probability so generalized does not guarantee that $P(A|A) = 1$. Consider the Symmetric loaded Kleene setting. Suppose all credence is invested in a world in which $A$ has truth value 0.5. It turns out by the recipe above that $P(A|A) = 0.5$.

# 6 Jeffrey-style Decision theory

An important application of probability is within the theory of rational decision making. We want to say something about a decision situation taking the following form: there are a range of actions $A$. There are factors $S \in \Gamma$, which fix the consequences of the action. $\Gamma$ form a partition, and we are uncertain which element of that partition obtains. We are in a position to judge the desirability of the total course of events, representable by $A \wedge S$. But our uncertainty over which $S$ obtains means that we have work to do to in order to figure out the desirability of $A$ itself.

Jeffrey's decision theory (Jeffrey, 1965) provides a way to calculate the desirability $D$ of a course of action, from the desirability of the outcomes, plus one's subjective probabilities. The desirability of the action is a weighted average of the desirability of the outcomes, with the weights provided by how likely the outcome is to obtain, given you take the action. The recipe is:

$$D(A) = \sum_{S \in \Gamma} P(S|A) D(A \wedge S)$$

Notice the crucial role given to conditional probabilities.

The application of probabilities within the theory of decision making is important, and if we couldn't recover a sensible account, this would render the whole enterprise of nonclassical (subjective) probability less interesting. As a proof of principle, I'll show that Jeffrey's recipe can indeed be generalized. I do not claim here to *justify* this as the *right* theory of decision in the nonclassical setting, but just to show that such theories are *available*.

Here's one way in which the Jeffrey decision rule can arise. Start by introducing a valuation function from worlds to reals, $v$—intuitively, a measure of how much we'd like the world in question to obtain. Then the desirability of an arbitrary *sentence* $A$, is defined as follows:

$$D(A) := \sum_{w} P(w|A)v(w)$$

Now, one might wonder if this is the *right* way to define desirability; but there is no question that it is *well-defined* in terms of the specific underlying valuation $v$. Now take any partition $\Gamma$ of sentences (in the generalized sense of partition of the previous section). By the corollary in that section, applied to the nonclassical probability $P_A$ that arises from conditioning on $A$, we have:

$$P(w|A) = P_A(w) = \sum_{S\in\Gamma} P_A(S)P_A(w|S)$$

Using another fact noted there, $P_A(W|S) = P(w|S\circ A)$, and putting these two together and substituting for $P(w|A)$ in the definition above, we obtain:

$$D(A) = \sum_{w}[\sum_{S\in\Gamma} P(S|A)P(w|S\circ A)]v(w)$$

Rearranging gives:

$$D(A) = \sum_{S\in\Gamma} P(S|A)[\sum_{w} P(w|S\circ A)v(w)]$$

But the embedded sum here is by construction equal to $D(S\circ A)$. Thus we have:

$$D(A) = \sum_{S\in\Gamma} P(S|A)D(S\circ A)$$

This is the exact analogue of the Jeffrey rule. So valuations over worlds allow us to define a notion of desirability that satisfies the generalized form of Jeffrey's equation.

# 7 Conclusions

To recap:

1. Nonclassical probabilities can be viewed as expectations of truth values, and standard principles of probability can be carried over to the nonclassical case, if we

substitute an appropriate nonclassical logic for appeals to the classical one. The appropriate nonclassical logic can be generally characterized as one guaranteeing no drop in truth value.

2. The truth values concerned should be thought of as the cognitive loadings of the non-classical truth status. But the general recipe may break down if (a) one's attitude to nonclassical logic is not semantically driven (one will not have 'truth statuses' to play with); (b) one does not regard the statuses as cognitively loaded; or (c) the cognitively loadings are not representable as real-valued degrees of belief.

3. The nonclassical probabilities so defined can be justified as constraints on rational belief via analogues of dutch book and accuracy-domination arguments.

4. A notion of conditional probability can be defined, that preserves important features of classical conditional probability. It satisfies a ratio formula, though only for an appropriately chosen conjunctive connective that is not available in all nonclassical settings.

5. An analogue of Jeffrey's recipe for calculating desirability of actions with respect to an arbitrary partition of states can be provided.

This provides a rich field for further investigation:

1. Studying *axiomatizations* of nonclassical probabilities is an open-ended task. Can we extend the results of Paris, Mundici et al, and get more general sense of what set of axioms are generally sufficient to characterize expectations of truth value? A major obstacle here is the appeal throughout to the additivity principle (P3) and its variants, which is the only one to turn essentially on the behaviour of particular connectives. Is there a way of capturing its content in terms of logical relations between sentences rather such hardwired constraints?[39]

2. We have focused on cases where 'truth values' (the cognitive loading of nonclassical truth statuses take a particularly tractable form: represented by reals in $[1, 0]$. Can we get a notion of nonclassical probability in more general setting, where the cognitive loads are not linear ordered, or where some truth statuses are missing such loading altogether? Perhaps the notion of expectations of non-real valued random variables may provide a lead here.

3. What is the relation between the cognitive loading of a nonclassical truth status (appealed to directly in the accuracy-domination argument) and the pragmatic loading (relevant to the generalized dutch book argument). Must they coincide? If so, why?

4. How much of the theory of conditional probability transfers to the nonclassical treatment? In the classical case, it is possible to start with an axiomatization of conditional probability, and develop unconditional probability on this basis. Is an analogue available here?

[39](Williams, forthcomingb) sketches one potential approach, appealing to constraints derived from a (strange) multi-conclusion logic.

5. Many foundational and formal questions about nonclassical decision theory deserve exploration. Is the characterization of desirability sketched above a cogent one? Are analogues of classical representation theorems available, relative to a set of rational constraints on qualitative preference?

It's worth emphasizing a remark we made right at the start. It is not only convinced revisionists that need to be concerned about these issues. It is anyone not dogmatically opposed to logical revision. For, prima facie, if one is open to the possibility of the failure of excluded middle, for example, then one shouldn't invest full confidence in these principles, as classical probabilism requires. And yet, it does not seem that one is irrational in harbouring such doubts, as the interpretation of classical probabilities as constraints on rational belief would suggest.[40] Now, perhaps one could argue that in the end, such doubts manifest a lack of perfect, ideal rationality. But this seems a strong assumption. Why think that our total evidence, and superhuman processing power, would convey *total* conviction in the correctness of classical logic?

The revisionist shouldn't gloat too soon, however, since the same point seem to carry across to any one of the *nonclassical* settings. It would be bold to for revisionists to proclaim utter certainty the Kleene logic as the One True logical framework. They're not in quite as embarressing a position, in that investing some confidence in excluded middle, for example, is entirely compatible with Kleene probabilities (the Kleene constraints are strictly weaker than the classical ones). But the nonclassicist always faces the possibility of challenges from yet weaker systems.

One way of dealing with this is to drop the assumption that there is a space of truth value distributions (classical or otherwise) over which to define probabilities, independent of one's doxastic state. Perhaps the theory of subjective probability should be developed relative to a set of truth value distributions *that the agent regards as open possibilities*, $\boldsymbol{Z}$. The arguments above can be used to characterize expectations of truth value over the possibilites in $\boldsymbol{Z}$, and a consequence relation defined via the no drop method used before. If the open possibilities include as many varieties of truth value distributions as has been suggested, then the $\boldsymbol{Z}$-logic will be weak indeed, and the constraints on rational degrees of belief also weak. However, perhaps the majority of a sensible person's credence will be devoted to some $\boldsymbol{C} \subset \boldsymbol{Z}$ which contains—say—classical truth value distributions. And if rational degrees of belief have to be expectations of truth value, then we do get the non-trivial result that the degrees of belief *conditional on* $\boldsymbol{C}$ have to meet the classical constraints. Mutatis mutandis for other interesting regions of the open possibilities $\boldsymbol{Z}$—for example, the probabilities conditional on the Kleene distributions $\boldsymbol{K}$ should be Kleene probabilities. So even though the statable constraints on $\boldsymbol{P}$ itself are rather minimal, it inherits as constraints of rationality the much more demanding requirements on its updates $\boldsymbol{P_C}(\cdot)$ (that they be classical probabilities); $\boldsymbol{P_K}(\cdot)$ (that they be Kleene probabilities) and so forth.

[40]'Logical omnscience' is a familiar worry for probabilism. But often this is presented as the worry that probabilism requires certainty even in very complex tautologies, the proof of which is too long to convey certainty. But in this case (i) one might think the right attitude is simply to suspend judgment on the question; and (ii) the lack of *ideality* seems built into the case, in that it is lack of *processing power* that seems to be the bar on being certain in the tautologies. The case at hand is one where one is perfectly well aware that LEM is a classical tautology, but one takes oneself to have principled reasons nevertheless to not be fully confident of it.

Throughout this discussion, I have been operating under two assumptions. The first is that the objects of probability are sentences; the second is that we are interested in *subjective* probabilities. I finish by considering what happens when the assumptions are relaxed.

The focus on *sentences* is inessential to the above. I would quite happily substitute Fregean thoughts, or Russellian propositions, or other fine-grained truth-bearers, in their place. What is important is that they stand in logical relations, and have contingent truth statuses.

But a very common approach to probability starts not from truth-bearers like sentences, thoughts or structured propositions, but entities assumed to be relevantly *coarse-grained*. For example, one finds a probability defined in terms of a triple $(\Omega, F, P)$, where $\Omega$ is a set (the 'sample space'), the event space $F$ is an algebra of subsets of $\Omega$, and $P$ is a function from $F$ to reals. We then have the familiar Kolmogorov axioms, analogous to the ones cited earlier:

P1. (Non-negativity) $\forall E \in F, P(E) \in \mathbb{R}_{\geq 0}$

P2. (Normalization) $P(\Omega) = 1$

P3. (Additivity) $\forall D, E \in F$ with $D$ and $E$ disjoint, $P(D \cup E) = P(D) + P(E)$

On one reading, $\Omega$ could be the set of possible worlds, and then $F$ would be coarse grained *propositions* in the sense of Lewis (1986) and Stalnaker (1984): sets of possible worlds. The interesting thing about this setting from our perspective is that *logic* seems to have disappeared from view. To be sure, we have analogues of 'inconsistency' (disjointness) and so forth—but the classicality is not explicit. What are we to make of this?

One diagnosis is that classical logic is still tacitly built into this framework—by the assumption that the event space $F$ forms a Boolean algebra. And indeed, *Quantum* probabilities are developed exactly on this diagnosis; quantum events are held to form a non-distributive lattice, and quantum probabilities are built on a framework appropriate to the structure of subspaces of Hilbert space, rather than to the structure of a subsets of a classical sample space.[41] This is a way of deriving a 'nonclassical probability' very different in inspiration and articulation from those sketched earlier (note, for example, the appeal in the expectations of truth value, to an underlying classical sample space). This illustrates a more general strategy for revising the above characterize: replace the Boolean algebraic structure associated with classical logic for some alternative, and study analogues of the standard probability axioms in that setting.

But under some interpretations of the above, there may be no motivation for the revisionist to question the standard setting. Perhaps *language* in general has nonclassical truth values; but the structure of the space of *possibilities and propositions* is entirely classical. If so, a nonclassical probability across sentences could live happily

[41](Hughes, 1992, ch.8,9)

with a classical probabilities across possibilities. The bridge would be one we have been presupposing throughout the discussion: *sentences* take nonclassical *truth values* at *worlds*—and the sentential probability of $S$ is the expectation of its truth value relative to the worldly probability distribution. Whether one adopts this conciliatory stance depends, of course, on one's motivations for nonclassicism. For many motivated by failures of bivalence generated by vagueness, it may seem an attractive perspective. For those motivated by the structure of physical theory, as in the case of quantum probabilities, it is not likely to appeal.

What of our focus on subjective probabilities, rather than evidential or objective probabilities? The *justifications* offered for nonclassical probability might be particular to the subjective setting (the dutch book and accuracy domination arguments are both usually presented that way). But the *formulation* of non-classical probabilities is certainly not so limited. One who rejects (LEM) will surely not be happy with the suggestion that this has evidential probability 1, or that there is a chance 1 that it will come about. The nonclassical probabilities we've been looking at are an obvious tool to reach for. Furthermore, if evidential probabilities or objective chance functions themselves are linked to subjective probabilities via principles of rationality—the Principle Principal being a leading contender for such a role[42]—then we may be able to argue from a nonclassical formulation of subjective probability, to nonclassical formulations of objective probability.

# References

Choquet, G. 1953. 'Theory of capacities'. *Annales d'institute fourier*, **V**, 131–295.

De Finetti, Bruno. 1974. *Theory of probability: vol.1.* New York: Wiley.

Di Nola, A., Georgescu, G., & Lettieri, A. 1999. 'Conditional states in finite-valued logic'. *Pages 161–174 of:* Dubois, D, H., Prade, & Klement, E.P. (eds), *Fuzzy sets, logics, and reasoning about knowledge.* Kluwer.

Dummett, Michael A. E. 1959. 'Truth'. *Pages 1–24 of: Truth and other enigmas.* London: Duckworth. First published in *Proceedings of the Aristotelian Society* (59):141–62.

Dummett, Michael A. E. 1991. *The logical basis of metaphysics.* The William James lectures; 1976. Cambridge, Mass.: Harvard University Press.

Field, Hartry. 2009. 'What is the normative role of logic?'. *Proceedings of the aristotelian society*, **83**(3), 251–268.

Field, Hartry H. 2000. 'Indeterminacy, degree of belief, and excluded middle'. *Nous*, **34**, 1–30. Reprinted in Field, *Truth and the Absence of Fact* (Oxford University Press, 2001) pp. 278-311.

Field, Hartry H. 2003a. 'No fact of the matter'. *Australasian journal of philosophy*, **81**, 457–480.

[42](Lewis, 1980, cf)

Field, Hartry H. 2003b. 'Semantic paradoxes and the paradoxes of vagueness'. *Pages 262–311 of:* Beall, J. C. (ed), *Liars and heaps*. Oxford: Oxford University Press.

Field, Hartry H. 2008. *Saving truth from paradox*. Oxford: Oxford University Press.

Fine, Kit. 1975. 'Vagueness, truth and logic'. *Synthese*, **30**, 265–300. Reprinted with corrections in Keefe and Smith (eds) *Vagueness: A reader* (MIT Press, Cambridge MA: 1997) pp.119-150.

Gerla, B. 2000. 'MV-algebras, multiple bets and subjective states'. *International journal of approximate reasoning*, **25**(1), 1–13.

Haack, Susan. 1978. *Philosophy of logics*. Cambridge: Cambridge University Press.

Hájek, Alan. 2008. 'Arguments for—or against—probablism?'. *British journal for the philosophy of science*, **59**.

Hájek, Alan. Summer 2003. 'Interpretations of probability'. *In:* Zalta, Edward N. (ed), *The stanford encyclopedia of philosophy*. `http://plato.stanford.edu/archives/sum2003/entries/probability-interpret/`.

Hájek, Petr. 1998. *Metamathematics of fuzzy logic*. Springer.

Hughes, R.I.G. 1992. *The structure and interpretation of quantum mechanics*. Harvard University Press.

Jaffray, J-Y. 1989. 'Coherent bets under partially resolving uncertainty and belief functions'. *Theory and decision*, **26**, 90–105.

Jeffrey, Richard. 1965. *The logic of decision*. 2nd edn. Chicago and London: University of Chicago Press. Second edition published 1983.

Joyce, James M. 1998. 'A non-pragmatic vindication of probabilism'. *Philosophy of science*, **65**, 575–603.

Joyce, James M. 2009. 'Accuracy and coherence: prospects for an alethic epistemology of partial belief'. *Pages 263–297 of:* Huber, Franz, & Schmidt-Petri, Christoph (eds), *Degrees of belief*. Springer.

Keefe, Rosanna. 2000. *Theories of vagueness*. Cambridge: Cambridge University Press.

Levi, I. 1974. On indeterminate probabilities. *The journal of philosophy*, 391–418.

Lewis, David K. 1980. 'A subjectivist's guide to objective chance'. *Pages 263–93 of:* Jeffrey, Richard (ed), *Studies in inductive logic and probability*, vol. II. University of California Press. Reprinted with postscript in Lewis, *Philosophical Papers II* (Oxford University Press, 1986) 83–113.

Lewis, David K. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Macfarlane, John. 2010. 'fuzzy epistemicism'. *In:* Moruzzi, Sebastiano, & Dietz, Richard (eds), *Cuts and clouds*. Oxford: OUP.

Maudlin, Tim. 2004. *Truth and paradox: Solving the riddles*. Oxford: Oxford University Press.

Milne, Peter. 2007. 'Betting on fuzzy and many-valued propositions (long version)'. *In: ms.*

Milne, Peter. 2008. 'Betting on fuzzy and many-valued propositions'. *In: The logica yearbook.*

Mundici, D. 2006. 'Bookmaking over infinite-valued events'. *International journal of approximate reasoning*, **43**(3), 223–240.

Paris, J.B. 2001. 'A note on the Dutch Book method'. *Pages 301–306 of: Proceedings of the second international symposium on imprecise probabilities and their applications, isipta.* Ithaca, NY: Shaker.

Priest, G. 2006. *In contradiction: a study of the transconsistent.* Oxford University Press, USA.

Priest, Graham. 2001. *An introduction to non-classical logics.* Cambridge: Cambridge University Press.

Shafer, G. 1976. *A mathematical theory of evidence.* Princeton: Princeton University Press.

Smith, Nick J. J. 2008. *Vagueness and degrees of truth.* Oxford: Oxford University Press.

Smith, Nick J. J. 2010. 'Degree of belief is expected truth value'. *In:* Moruzzi, Sebastiano, & Dietz, Richard (eds), *Cuts and clouds.* Oxford: OUP.

Stalnaker, Robert. 1984. *Inquiry.* Cambridge, MA: MIT Press.

van Fraassen, Bas. 1966. 'Singular terms, truth-value gaps, and free logic'. *The journal of philosophy*, **63**(17), 481–495.

Weatherson, Brian. 2003. 'From classical to constructive probability'. *Notre dame journal of formal logic*, **44**, 111–123.

Weatherson, Brian. 2005. 'True, truer, truest'. *Philosophical studies*, **123**, 47–70. Available at `http://brian.weatherson.net/ttt.pdf`.

Williams, J. Robert G. forthcominga. 'Generalized probabilism: dutch books and accuracy domination'. *Journal of philosophical logic.*

Williams, J. Robert G. forthcomingb. 'Gradational accuracy and non-classical logic'. *Review of symbolic logic.*

Williamson, Timothy. 1994. *Vagueness.* London: Routledge.

Wright, Crispin. 2001. 'On being in a quandry'. *Mind*, **110**.

Zadeh, L. A. 1968. Probability measures of fuzzy events. *Journal of mathematical analysis and applications.*, 421–427.