

ACCEPTANCE, AGGREGATION AND SCORING RULES

JAKE CHANDLER*

(Forthcoming in *Erkenntnis*)

Abstract This article provides a novel perspective on the vexed issue of the relation between probability and rational acceptability, exploiting a recently-noted structural parallel with the problem of judgment aggregation. After offering a number of general desiderata on the relation between finite probability models and sets of accepted sentences in a Boolean sentential language, it is noted that a number of these constraints will be satisfied if and only if acceptable sentences are true under all valuations in a distinguished non-empty set W . Drawing inspiration from distance-based aggregation procedures, various scoring rule based membership conditions for W are discussed and a possible point of contact with ranking theory is considered. The paper closes with various suggestions for further research.

Keywords acceptance - distance-based aggregation - model coarsening / refinement - probability - ranking functions - scoring rules

1 Introduction

The formal modeling of doxastic states appears to operate at different levels of granularity. At the finer end of the scale, we encounter for instance—in the so-called Bayesian tradition—‘graded’ representations in terms of sets of real-valued functions over some formal language. At the coarser

*Address for correspondence: Center for Logic and Analytic Philosophy, HIW, KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium. Email: jacob.chandler[at]hiw.kuleuven.be

end, we find—most notably in some areas of the belief-revision literature—‘all-or-nothing’ representations in terms of sets of sentences or valuations. These differences in modeling choices seemingly map onto a corresponding heterogeneity in folk psychological practice. Belief reports do indeed come both in the form of attributions of degrees of confidence, aka ‘credences’, (e.g. ‘I am pretty sure that he needs help.’) and in the form of unqualified attributions of ‘full’ belief (e.g. ‘He believes in fairies.’).

Whilst a unification of these various types of models, and of the phenomena they purport to represent, would be an extremely desirable achievement, it unfortunately seems fair to say that the history of attempts to provide a single, overarching framework—in particular a reduction of the coarser level of description to the finer one—encourages a certain amount of pessimism. In what follows, however, I hope to show that this pessimism is premature. The current track record of the unificationist camp is indeed poor; there is no disputing that. But, as Chandler (2010) has argued, the root cause of this mediocre performance is easy to pinpoint: the blame lies in the uncritical endorsement of a rather dubious would-be constraint on the relation between graded and all-or-nothing belief. This putative constraint, which I shall call INDEPENDENCE, turns out to be a close cousin of a homonymous troublemaker discussed in the judgment aggregation literature. LOSS of INDEPENDENCE, one might say, paves the way to unification. But now since existing proposals invariably satisfy the constraint, it is of interest to see what kinds of alternatives may be on offer. The aim of this article is to outline and briefly discuss a family of such alternatives, inspired by recent work on the problem of judgment aggregation.

The paper will proceed as follows. Section 2 presents the basic framework and notation employed throughout, discuss some desiderata on the mapping between degrees of confidence and full belief and present some baseline results. Section 3 spells out the relationship between the task at hand and the problem of judgment aggregation, offering a brief overview of a family of distance-based aggregation methods that have enjoyed a certain degree of popularity in the recent computer science literature. Section 4 then outlines a corresponding family of scoring rule -based mappings from degrees of confidence onto sets of accepted sentences. To illustrate this general approach, a sample of noteworthy members of this family are presented and some of their respective properties flagged out. Section 5 notes that these distance-based mappings suggest yet a *further* potential unificatory strategy, this time pertaining to the vexed issue of the relation

between probabilistic and ranking-theoretic models of belief. Section 6 briefly concludes with some suggestions for future research.

2 Formal preliminaries

For the purposes of the current paper, the credal state of mind of a rational agent will be assumed to be representable by a probability model \mathcal{M} . Whilst more expressive types of representations, such as sets of probability models, are of course available, this simplification will help keep the discussion focused. \mathcal{M} will be defined as a pair $\langle \mathcal{L}, \text{Pr} \rangle$, where: (i) \mathcal{L} is a sentential language constructed from a *finite* set \mathcal{P} of m atomic sentences by means of the usual Boolean connectives \wedge , \vee and \neg , and (ii) Pr is a probability function with domain \mathcal{L} , obeying the following standard axioms, where \vdash denotes the relation of classical consequence and \top and \perp respectively denote an arbitrary classical tautology and an arbitrary classical contradiction: for any $\varphi, \psi \in \mathcal{L}$, (i) if $\varphi \vdash \perp$, then $\text{Pr}(\varphi) = 0$, (ii) if $\top \vdash \varphi$, then $\text{Pr}(\varphi) = 1$, (iii) if $\varphi \vdash \psi$, then $\text{Pr}(\varphi) \leq \text{Pr}(\psi)$ and (iv) if $\varphi \wedge \psi \vdash \perp$, then $\text{Pr}(\varphi \vee \psi) = \text{Pr}(\varphi) + \text{Pr}(\psi)$.

V will denote the set of valuations $\{v_1, v_2, \dots, v_{2^m}\}$ of \mathcal{L} , which are total functions from \mathcal{L} to the set of classical truth values $\{0, 1\}$. Furthermore, $\llbracket \varphi \rrbracket$ will denote the set of valuations that validate φ (i.e. $\{v \in V : v(\varphi) = 1\}$) and conversely, where W is some set of interpretations, φ_W will be used to denote an arbitrary sentence such that $W = \{v \in V : v(\varphi_W) = 1\}$. By abuse of notation, we shall sometimes write v for $\{v\}$ and vice versa. Finally, a rational acceptance function Acc is a particular kind of function that maps pairs of probability models $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentences $\varphi \in \mathcal{L}$ onto $\{0, 1\}$. The intended interpretation is that $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff it is rationally permissible to fully believe φ given that one's credences are representable by \mathcal{M} .^{1,2} $\text{Acc}(\mathcal{M}, \cdot)$ is, in other terms, the 'characteristic function' of the belief set that is rationally permissible given \mathcal{M} .

Clearly, in order to ensure that the set of accepted sentences is 'well-behaved', one should require the following:

ZERO-NORMALISATION: For any probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentence $\varphi \in \mathcal{L}$, if $\varphi \vdash \perp$ then $\text{Acc}(\mathcal{M}, \varphi) = 0$.

UNIT-NORMALISATION: For any probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentence $\varphi \in \mathcal{L}$, if $\top \vdash \varphi$ then $\text{Acc}(\mathcal{M}, \varphi) = 1$.

DEDUCTIVE CLOSURE: For any probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$, sentence $\varphi \in \mathcal{L}$ and set of sentences $\Gamma \subseteq \mathcal{L}$, if for any $\psi \in \Gamma$, $\text{Acc}(\mathcal{M}, \psi) = 1$ and $\Gamma \vdash \varphi$, then $\text{Acc}(\mathcal{M}, \varphi) = 1$.

These correspond, respectively, to the claims that it is never permissible to accept a contradiction, always permissible to accept a tautology and always permissible to accept the logical consequences of what one is permitted to accept. It seems, however that one should *not* require

OPINIONATION: For any probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentences $\varphi, \psi \in \mathcal{L}$, either $\text{Acc}(\mathcal{M}, \varphi) = 1$ or $\text{Acc}(\mathcal{M}, \neg\varphi) = 1$.

which states that, for any sentence φ that a rational agent S can entertain, S always accepts either φ or its negation. Another important and fairly uncontroversial constraint is the following:

NON-UNANIMITY: For some probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and some sentence $\varphi \in \mathcal{L}$, $\text{Pr}(\varphi) < 1$ and $\text{Acc}(\mathcal{M}, \varphi) = 1$.

This reflects the commonly-held intuition that one can accept sentences whose truth one is not one hundred percent certain of. Perhaps more controversial, but in my view *prima facie* plausible, is

STRUCTURALITY: For any model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$, any automorphism π of \mathcal{M} and any sentence $\varphi \in \mathcal{L}$, $\text{Acc}(\mathcal{M}, \varphi) = \text{Acc}(\mathcal{M}, \pi(\varphi))$, where an *automorphism* π of a probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ is a 1 : 1 function $\mathcal{L} \mapsto \mathcal{L}$, such that, for any $\varphi, \psi \in \mathcal{L}$, (i) $\pi(\varphi \wedge \psi) = \pi(\varphi) \wedge \pi(\psi)$, (ii) $\pi(\neg\varphi) = \neg\pi(\varphi)$, (iii) $\pi(\varphi \vee \psi) = \pi(\varphi) \vee \pi(\psi)$ and (iv) $\text{Pr}(\varphi) = \text{Pr}(\pi(\varphi))$.

This principle can be understood as stating that the acceptability of a sentence with respect to a model \mathcal{M} supervenes on the logical and probabilistic properties of \mathcal{M} . As Douven and Williamson (2006) point out, STRUCTURALITY, DEDUCTIVE CLOSURE and ZERO-NORMALISATION jointly entail

LOTTERY-PROOFNESS: For any finite probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ such that for any $v, v^* \in V$, $\text{Pr}(\varphi_v) = \text{Pr}(\varphi_{v^*})$ and any sentence $\psi \in \mathcal{L}$, if $\text{Pr}(\psi) < 1$, then $\text{Acc}(\mathcal{M}, \psi) = 0$.³

which states that no contingent sentence is rationally acceptable with respect to a finite uniform probability model.⁴ Douven and Williamson appeared to reject this principle. In fact, however, as argued by Chandler

([2010, pp.8–9]), this property is intuitively *correct*, if one grants one of the working assumptions of this paper, namely that acceptability is determined by a credal state modeled as a single probability model. The grounds for Douven and Williamson’s objection may have been rooted in a tacit commitment to the following principle:

MONOTONICITY: For any pair of probability models $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and $\mathcal{M}^* = \langle \mathcal{L}^*, \text{Pr}^* \rangle$ and sentences $\varphi \in \mathcal{L}$ and $\psi \in \mathcal{L}^*$, if $\text{Pr}(\varphi) \leq \text{Pr}^*(\psi)$ then $\text{Acc}(\mathcal{M}, \varphi) \leq \text{Acc}(\mathcal{M}^*, \psi)$.

Indeed, it is easy to show the following:

Theorem 2.1. LOTTERY-PROOFNESS, MONOTONICITY and NON-UNANIMITY are jointly inconsistent.

In response to this, it was argued that no clear justification for MONOTONICITY had been given and that, in view of the *prima facie* plausibility of STRUCTURALITY, DEDUCTIVE CLOSURE and ZERO-NORMALISATION, as well as the independent plausibility of LOTTERY-PROOFNESS, MONOTONICITY ought to be rejected. The same line can be taken with respect to the somewhat weaker

INDEPENDENCE: For any pair of probability models $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and $\mathcal{M}^* = \langle \mathcal{L}^*, \text{Pr}^* \rangle$ and sentence $\varphi \in \mathcal{L}, \mathcal{L}^*$, if $\text{Pr}(\varphi) = \text{Pr}^*(\varphi)$ then $\text{Acc}(\mathcal{M}, \varphi) = \text{Acc}(\mathcal{M}^*, \varphi)$.

By a similar chain of reasoning to the one above, INDEPENDENCE can be shown to be incompatible with the conjunction of LOTTERY-PROOFNESS and a seemingly unobjectionable strengthening of NON-UNANIMITY to the claim that there exists a *rational-valued* sub-unit probability sentence that is rationally acceptable (calls this NON-UNANIMITY⁺). But it would seem entirely arbitrary to insist that no rational-valued sub-unit probability sentence is rationally acceptable, but maintain that some *irrational-valued* sub-unit probability sentence is so. So denying the stronger version of NON-UNANIMITY, presumably leads to denying the weaker one, which, I submit, is unacceptable. So INDEPENDENCE must surely be given up too. Finally, to the aforementioned desiderata, it is clear that one should furthermore add

RESPONSIVENESS: For some pair of probability models $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and $\mathcal{M}^* = \langle \mathcal{L}, \text{Pr}^* \rangle$ and some sentence $\varphi \in \mathcal{L}$, $\text{Acc}(\mathcal{M}, \varphi) \neq \text{Acc}(\mathcal{M}^*, \varphi)$.

This states that the underlying probability distribution makes a difference as to whether or not a sentence is rationally acceptable.

Let us briefly take stock. We have argued that the following are to be categorically endorsed: UNIT-NORMALISATION, ZERO-NORMALISATION, DEDUCTIVE CLOSURE, NON-UNANIMITY⁺, LOTTERY-PROOFNESS and RESPONSIVENESS. We have more tentatively endorsed STRUCTURALITY and have categorically rejected OPINIONATION. In the light of Theorem 2.1, INDEPENDENCE, and hence MONOTONICITY, has also been categorically rejected, due the endorsement of NON-UNANIMITY⁺ and LOTTERY-PROOFNESS.

Can we find a function that simultaneously satisfies all of these constraints? It turns out that the range of possible avenues is limited by the following elementary result:

Theorem 2.2. *UNIT-NORMALISATION, ZERO-NORMALISATION and DEDUCTIVE CLOSURE hold iff there exists a non-empty set $W \subseteq V$ of valuations such that, for any $\varphi \in \mathcal{L}$, $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff $W \subseteq \llbracket \varphi \rrbracket$.*

So we need a procedure to select this non-empty set W of valuations, and one that will yield an acceptance function that fares well on the remaining desiderata. How should this be done?

3 Taking cue from judgment aggregation

In a recent article, Douven and Romeijn (2007) hint at the existence of a striking parallel between the issue of individual-level rational acceptability or sentences relative to a probability model and the issue of group-level rational acceptability of sentences relative to the opinions of a set of agents, aka the problem of rational judgment aggregation.

In formal terms, the problem of rational judgment aggregation involves the characterisation of a two-place aggregation function Agg , this time mapping pairs of *opinion models* \mathcal{M} and sentences in \mathcal{L} onto $\{0, 1\}$, such that $\text{Agg}(\mathcal{M}, \varphi) = 1$ iff φ is rationally acceptable, at the group level, with respect to \mathcal{M} . An opinion model is a pair $\langle \mathcal{L}, \mathcal{O} \rangle$, where \mathcal{O} , known as an ‘opinion profile’, is an n -tuple $\langle \psi_1, \dots, \psi_n \rangle$ of classically consistent sentences in \mathcal{L} . These sentences are known as ‘opinions’, and represent the respective full beliefs of the members of a group of n rational agents.⁵

In cases in which the range of Pr is a subset of the rational numbers \mathbb{Q} , we could think of the problem of rational acceptability as a special case of the problem of judgment aggregation in which (i) the opinions aggregated

are maximally strong consistent sentences in \mathcal{L} and (ii) the aggregation function Agg is subject to the following constraints:

ANONYMITY: For any opinion models $\mathcal{M} = \langle \mathcal{L}, \langle \psi_1, \dots, \psi_n \rangle \rangle$ and $\mathcal{M}^* = \langle \mathcal{L}, \langle \pi(\psi_1), \dots, \pi(\psi_n) \rangle \rangle$, with $\psi_i \in \mathcal{L} (1 \leq i \leq n)$, where π is a permutation of $\langle \psi_1, \dots, \psi_n \rangle$, and any $\varphi \in \mathcal{L}$, $\text{Agg}(\mathcal{M}, \varphi) = \text{Agg}(\mathcal{M}^*, \varphi)$.

DUPLICATION: For any opinion model $\mathcal{M} = \langle \mathcal{L}, \langle \psi_1, \dots, \psi_n \rangle \rangle$ and $\mathcal{M}^* = \langle \mathcal{L}, \langle \psi_1, \dots, \psi_n, \psi_1, \dots, \psi_n \rangle \rangle$, with $\psi_i \in \mathcal{L} (1 \leq i \leq n)$, and any $\varphi \in \mathcal{L}$, $\text{Agg}(\mathcal{M}, \varphi) = \text{Agg}(\mathcal{M}^*, \varphi)$.

Indeed, we could view an assignment of rational-valued probabilities to maximally strong consistent sentences as a function returning the relative frequencies of corresponding opinions in the opinion set of a maximally opinionated group.

The parallel deepens when we turn to the kinds of constraints that are typically imposed on aggregation functions. We find, for instance, endorsements of precise aggregation-theoretic analogues of UNIT-NORMALISATION, ZERO-NORMALISATION, DEDUCTIVE CLOSURE, NON-UNANIMITY⁺ and RESPONSIVENESS. Furthermore, although the analogue of OPINIONATION is an admittedly widespread constraint, this has also occasionally been waived. Finally, concerns with respect to the obvious analogue of INDEPENDENCE, which is crucially involved in a number of aggregation-theoretic impossibility results⁶, have also been raised.

In view of all this, it will come as little surprise that recent developments in the aggregation literature suggest an attractive answer to the question posed at the end of the previous section. Konieczny, Lang and Marquis (2004) provide an overview of a large family of aggregation functions based on the notion of distance between valuations and opinion profiles.⁷ The idea behind this class of proposals is strikingly simple: first (i) provide a measure of the distance between the various valuations of \mathcal{L} and the tuple of individual-level opinions, then (ii) select, as adopted at the group level, all and only those sentences that are validated by all the valuations that are close enough. As we shall see, Konieczny *et al.* take ‘close enough’ to mean ‘closest’. It is however worth noting that alternative options could be explored. One could also for instance select all and only those valuations that are situated within some suitably chosen distance t . I shall return to this kind of variant in the next section.⁸

With some relevant simplifications, the calculation of the distance between a valuation v and an opinion profile \mathcal{O} proceeds as follows.

Step 1: For every $v^* \in V$, provide a measure of the distance between v and v^* .

Step 2: For every ψ in \mathcal{O} , construct a measure of the distance between v and ψ , by aggregating the distances between v and the valuations that validate ψ .

Step 3: Construct a measure of the distance between v and \mathcal{O} , by aggregating, in turn, the distances between v and each ψ in \mathcal{O} .

So we start off with a ‘distance’ function d mapping pairs of valuations onto \mathbb{R}_+ . The requirements imposed by Konieczny and his colleagues are very minimal:

Definition 3.1. *d is a distance between valuations iff it is a total function from $V \times V$ to \mathbb{R} such that, for every $v, v^* \in V$, (i) $d(v, v^*) \geq 0$ (Positivity), (ii) $d(v, v^*) = d(v^*, v)$ (Symmetry), and (iii) $d(v, v^*) = 0$ iff $v = v^*$ (Minimality).⁹*

Of course, one could just take d as primitive and simply add a further argument to the aggregation function. Primitivism regarding distance between valuations is not unheard of in philosophical circles (e.g. Hilpinen 1976), but it is somewhat mysterious and unparsimonious nevertheless. Indeed, Konieczny *et al* do not even discuss the option and focus instead on two particular alternatives, both of which are, incidentally, bounded above (by 1), which does not seem to be an undesirable feature. One of these alternatives is what they call the ‘drastic distance’ (d_D):

Definition 3.2. $d_D(v, v^*) := 0$ if $v = v^*$ and equal to 1 otherwise.

The other is the normalised weighted Hamming distance (d_{H_q}):

Definition 3.3. $d_{H_q}(v, v^*) := \sum_{p \in S} q(p)$, where $S := \{p \in \mathcal{P} : v(p) \neq v^*(p)\}$ and q is a total function from \mathcal{P} to \mathbb{R}_+ such that $\sum_{p \in \mathcal{P}} q(p) = 1$.¹⁰

If we take q to be a constant function, treating all mismatches symmetrically, we obtain the familiar normalised Hamming distance (d_H).

Definition 3.4. $d_H(v, v^*) := \frac{|S|}{|\mathcal{P}|}$, where S is defined as in Definition 3.3.

Once a measure of the distance between valuations has been settled on, we obtain distances between valuations and opinion profiles in steps 2 and 3 by means of two successive distance aggregation procedures. Konieczny *et al.* offer three constraints on such procedures:

Definition 3.5. g is a distance aggregation function iff it is a function from \mathbb{R}^n to \mathbb{R}^+ , such that (i) g is non-decreasing in every argument, (ii) $g(r_1, r_2, \dots, r_n) = 0$ iff $r_1 = r_2 = \dots = r_n = 0$ and (iii) $g(r) = r$.

To these, one would presumably want to add the following internality requirement:

$$(iv) \min\{r_1, r_2, \dots, r_n\} \leq g(r_1, r_2, \dots, r_n) \leq \max\{r_1, r_2, \dots, r_n\}.$$

Step 2 establishes the distance d' between valuations and sentences, by aggregating the distances between these valuations and the valuations that validate the sentences. Although there is a fair amount that could be said here, given (iv), which is satisfied by Konieczny *et al.*'s own proposal for d' , the specifics of this step are of no great relevance to the present paper. Indeed, in the cases that we are interested in, the ψ_i are maximally strong consistent sentences, and the $\llbracket \psi_i \rrbracket$ are singletons, so we will wind up with the same value for $d'(v, \varphi)$ whatever the particular d' we settle on. Of more central concern to the present case is Step 3, in which we aggregate the distances $d'(v, \psi_i)$ between v and the ψ_i in \mathcal{O} to obtain a distance $D(v, \mathcal{O})$ between v and \mathcal{O} . There is a large number of available options here, such as the max and min functions, as well as various weighted and non-weighted means:

As mentioned earlier, Konieczny *et al.* simply make use of distances between valuations and opinion profiles to establish a total preorder in V , taking as acceptable, at the group level, all and only those sentences that are validated by the valuations that are in the minimal set. This leaves us with:

For any opinion models $\mathcal{M} = \langle \mathcal{L}, \mathcal{O} \rangle$ and $\varphi \in \mathcal{L}$, $\text{Agg}(\mathcal{M}, \varphi) = 1$ iff $v(\varphi) = 1$, for any $v \in V$ such that, for any $v^* \in V$, $D(v, \mathcal{O}) \leq D(v^*, \mathcal{O})$.

It is worth taking a brief look at some of the general properties of the resulting proposal, in the case of interest in which the opinions are maximally strong. By the aggregation-theoretic analogue of Theorem 2.2, we already know that Agg will satisfy the analogues of $\text{UNIT-NORMALISATION}$, $\text{ZERO-NORMALISATION}$ and DEDUCTIVE CLOSURE . RESPONSIVENESS straightforwardly holds by virtue of definitions 3.1 and 3.5. ANONYMITY , which we are keen to preserve in the present context, obviously corresponds to symmetry of D , i.e. its invariance under argument permutation. This fails for weighted means (with non-uniform weighting function). In the presence of ANONYMITY , the analogue of STRUCTURALITY would be secured by the symmetry of the distance matrices of d_D and d_H along their main diagonal. The analogue of INDEPENDENCE , however, needn't hold:

Example 3.3. *Let $\mathcal{P} = \{\varphi, \psi\}$, $\mathcal{O}_1 = \langle \varphi \wedge \psi, \varphi \wedge \neg\psi, \neg\varphi \wedge \psi, \neg\varphi \wedge \neg\psi \rangle$ and $\mathcal{O}_2 = \langle \varphi \wedge \psi, \varphi \wedge \neg\psi, \neg\varphi \wedge \psi, \neg\varphi \wedge \neg\psi \rangle$. Let $d = d_D$ and D be the arithmetic mean function. It is easily verified that all valuations are equidistant from \mathcal{O}_1 but that the sole member of $\llbracket \varphi \wedge \psi \rrbracket$ is uniquely closest to \mathcal{O}_2 . Hence $\text{Agg}(\mathcal{M}_1, \varphi \wedge \psi) = 0$ but $\text{Agg}(\mathcal{M}_2, \varphi \wedge \psi) = 1$.*

The same example demonstrates that OPINIONATION can also fail, since we also have $\text{Agg}(\mathcal{M}_1, \neg(\varphi \wedge \psi)) = 0$. Finally, the analogue of NON-UNANIMITY fails in some cases, for instance if we set $d = d_H$ and $D = \max$.

4 Acceptance and scoring rules

We can adapt and generalise the kind of strategy outlined in the previous section in the following manner:

- (a) Provide a measure D of the distance between the $v \in V$ and Pr , which, we shall assume, satisfies constraints of both positivity and minimality (see Definition 3.1).
- (b) Use this to select a set W of ‘close enough’ valuations, such that, for any $v, v^* \in V$, if $v \in W$ and $D(v^*, \text{Pr}) \leq D(v, \text{Pr})$, then $v^* \in W$.
- (c) Define Acc as follows: For any $\varphi \in \mathcal{L}$, (i) if $W = \emptyset$, then $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff $\top \vdash \varphi$, (ii) otherwise $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff $W \subseteq \llbracket \varphi \rrbracket$.¹¹

Regarding membership conditions for W , we have two salient options. The first of these, the analogue of which was endorsed by Konieczny *et al.*, we shall call ‘MIN’, for ‘minimising’. The second, whose analogue has, to the best of my knowledge, yet to be discussed in the aggregation-theoretic literature, we shall call ‘SAT’, for ‘satisficing’.

MIN: $v \in W$ iff $v \in V$ and $\forall v^* \in V (D(v^*, \text{Pr}) \geq D(v, \text{Pr}))$.

SAT: $v \in W$ iff $v \in V$ and $D(v, \text{Pr}) \leq t$, for some appropriate $t \in \mathbb{R}^+$.

Regarding D , the literature is already replete with various suggestions under the name of so-called ‘scoring rules’. It turns out however that most proposals in usage satisfy

\geq_{Pr} -**REVERSAL:** $D(v^*, \text{Pr}) \geq D(v, \text{Pr})$ iff $\text{Pr}(\varphi_v) \geq \text{Pr}(\varphi_{v^*})$.

These include the family of exponential scores, of which the well-known Brier score is a member, as well as the logarithmic score (D_{\log}), which we will return to in the next section:

Definition 4.1. $D_{\log}(v, \text{Pr}) := -\log(\text{Pr}(\varphi_v))$.

Given \geq_{Pr} -REVERSAL, whatever the particular choice of D , the upshot of MIN would be that it is permissible to accept all and only those sentences that are true in all the most probable worlds. Call the resulting function ‘Acc₁’. This does have the arguable drawback of allowing for acceptance of sentences in whose truth one has an arbitrarily small degree of confidence. Opting for SAT avoids this consequence, since it permits acceptance of all and only those sentences that are true in all the those worlds whose probability exceeds a certain threshold. Call the resulting function ‘Acc₂’. It is easily checked that both Acc₁ and Acc₂ satisfy both STRUCTURALITY and CONSENSUS PRESERVATION, and violate OPINIONATION. Furthermore NON-UNANIMITY fails for Acc₁ and does so for Acc₂ as well, unless we set t to 0.¹²

Interestingly enough, some relevant functions *do* violate \geq_{Pr} -REVERSAL. One particularly interesting example, which isn’t found in the scoring rule literature but whose analogue appears to be taken quite seriously in the aggregation literature, is the following:

$$D_H(v, \text{Pr}) = \sum_{v^* \in V} \text{Pr}(\varphi_{v^*}) d_H(v, v^*)$$

There are a number of potential issues worth noting here, however. To begin with, as has been famously noted in the verisimilitude debate, the Hamming distance is not invariant under translation (see Miller’s (1974) well-known criticism of Tichý (1974, 1976)).

Example 4.1. *Let \mathcal{L}_1 be built up from the set of atomic sentences $\mathcal{P}_1 = \{\varphi, \psi, \chi\}$ and \mathcal{L}_2 from $\mathcal{P}_2 = \{\varphi, \beta, \gamma\}$. Let the binary relation R pair up sentences in \mathcal{L}_1 with their synonymous counterparts in \mathcal{L}_2 and let $\langle \varphi \leftrightarrow \psi, \beta \rangle, \langle \varphi \leftrightarrow \chi, \gamma \rangle \in R$. We have, in R , $\langle \varphi \wedge \psi \wedge \chi, \varphi \wedge \beta \wedge \gamma \rangle, \langle \neg\varphi \wedge \psi \wedge \chi, \neg\varphi \wedge \neg\beta \wedge \neg\gamma \rangle$ and $\langle \neg\varphi \wedge \neg\psi \wedge \neg\chi, \neg\varphi \wedge \beta \wedge \gamma \rangle$. Now $d_H(\llbracket \varphi \wedge \psi \wedge \chi \rrbracket, \llbracket \neg\varphi \wedge \psi \wedge \chi \rrbracket) = 1/3 < d_H(\llbracket \varphi \wedge \psi \wedge \chi \rrbracket, \llbracket \neg\varphi \wedge \neg\psi \wedge \neg\chi \rrbracket) = 1$. However, the ordering is reversed when we substitute the following \mathcal{L}_2 counterparts: $d_H(\llbracket \varphi \wedge \beta \wedge \gamma \rrbracket, \llbracket \neg\varphi \wedge \neg\beta \wedge \neg\gamma \rrbracket) = 1 > d_H(\llbracket \varphi \wedge \beta \wedge \gamma \rrbracket, \llbracket \neg\varphi \wedge \beta \wedge \gamma \rrbracket) = 1/3$.*

But this, so the worry might go, is bad news, since it is easy to show that it will result, given either MIN (call the resulting function ‘Acc₃’) or SAT (call the resulting function ‘Acc₄’), in an undesirable translation-sensitivity of acceptability.^{13, 14}

Setting this issue aside, both Acc₃ and Acc₄ violate the following property of CONSENSUS PRESERVATION, which could be argued to be intuitively compelling:

CONSENSUS PRESERVATION: For any *finite* probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentence $\varphi \in \mathcal{L}$, if $\text{Pr}(\varphi) = 1$ then $\text{Acc}(\mathcal{M}, \varphi) = 1$.¹⁵

To illustrate the result for Acc₃:

Example 4.2. *Let $\mathcal{P} = \{\varphi, \psi\}$ and $\text{Pr}(\neg\varphi \wedge \psi) = \text{Pr}(\varphi \wedge \neg\psi) = 1/2$. Let $D = D_H$. It is easily verified that all valuations are equidistant from Pr and hence, although $\text{Pr}(\neg(\varphi \leftrightarrow \psi)) = 1$, $\text{Acc}(\mathcal{M}, \neg(\varphi \leftrightarrow \psi)) = 0$.*

For these reasons, and pending further discussion of the normative status of the constraints flouted, it seems fair to treat both Acc₃ and Acc₄ with a certain degree of caution.

To wrap up this section, it is worth briefly considering an issue that has been somewhat overlooked in the literature on probability and acceptability, namely the behaviour of acceptability under probability model refinement or coarsening. Indeed, it has been put to me that the following property may be desirable:

PRESERVATION UNDER REFINEMENT: Where $\mathcal{M}^+ = \langle \mathcal{L}^+, \text{Pr}^+ \rangle$ is a refinement of $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$, for any $\varphi \in \mathcal{L}$, if $\text{Acc}(\mathcal{M}, \varphi) = 1$, then $\text{Acc}(\mathcal{M}^+, \varphi) = 1$.

PRESERVATION UNDER COARSENING: Where $\mathcal{M}^+ = \langle \mathcal{L}^+, \text{Pr}^+ \rangle$ is a refinement of $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$, for any $\varphi \in \mathcal{L}$, if $\text{Acc}(\mathcal{M}^+, \varphi) = 1$, then $\text{Acc}(\mathcal{M}, \varphi) = 1$.

Where a refinement is defined as follows:

Definition 4.2. $\mathcal{M}^+ = \langle \mathcal{L}^+, \text{Pr}^+ \rangle$ is a refinement of $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and \mathcal{M} a coarsening of \mathcal{M}^+ iff $\mathcal{L} \subset \mathcal{L}^+$ and, for any $\varphi \in \mathcal{L}$, $\text{Pr}(\varphi) = \text{Pr}^+(\varphi)$.

It is worth noting that Smith (forth.) appears to have recently—albeit cautiously—endorsed the second of the above principles, which plays a crucial role in the impossibility result that is central to his paper.

NOW INDEPENDENCE obviously guarantees the satisfaction of both constraints. However, both are violated—and hence so to is INDEPENDENCE—by both Acc_1 and Acc_2 ¹⁶. The following simple example illustrates the point for Acc_1 :

Example 4.3. Let $\mathcal{P} = \{\varphi\}$ and $\mathcal{P}^+ = \{\varphi, \psi\}$. Furthermore, let $\text{Pr}(\varphi) = 0.6$, $\text{Pr}^+(\varphi \wedge \psi) = 0.3$, $\text{Pr}^+(\varphi \wedge \neg\psi) = 0.3$, $\text{Pr}^+(\neg\varphi \wedge \psi) = 0.4$, $\text{Pr}^+(\neg\varphi \wedge \neg\psi) = 0$. It is easily verified that we have $\text{Acc}_1(\mathcal{M}, \varphi) = 1$ and $\text{Acc}_1(\mathcal{M}, \neg\varphi) = 0$, but $\text{Acc}_1(\mathcal{M}^+, \varphi) = 0$ and $\text{Acc}_1(\mathcal{M}^+, \neg\varphi) = 1$.

But *do* the principles constitute intuitive desiderata in the first place? This far from clear. If constraints on the preservation of acceptability under model coarsening or refinement are indeed in order, these should at the very least be restricted to changes in granularity that are not accompanied by the loss or acquisition of further information. The above conditions are simply too strong. A more promising, weaker constraint would be invariance of acceptability in the presence of a relation of *conservative* refinement between the two models:

Definition 4.3. $\mathcal{M}^+ = \langle \mathcal{L}^+, \text{Pr}^+ \rangle$ is a conservative refinement of $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ iff (i) $\mathcal{L} \subset \mathcal{L}^+$ (ii) for any maximally strong consistent $\varphi \in \mathcal{L}$, and maximally strong consistent $\varphi_1^+, \varphi_2^+ \in \mathcal{L}^+$, such that $\varphi_i^+ \not\vdash \neg\varphi$, with $1 \leq i \leq 2$, $\text{Pr}^+(\varphi_1^+) = \text{Pr}^+(\varphi_2^+)$.¹⁷

It is easy to show that the resulting weakened principles are both satisfied by Acc_1 . This follows pretty immediately from the fact that, for any maximally strong consistent sentences $\varphi_1, \varphi_2 \in \mathcal{L}$, the sets of \mathcal{L}^+ -valuations $\llbracket \varphi_1 \rrbracket$ and $\llbracket \varphi_2 \rrbracket$ will have the same cardinality. Matters, however, are somewhat different for Acc_2 , when $t > 0$. The weakening of PRESERVATION UNDER REFINEMENT clearly fails. The weakening of PRESERVATION UNDER COARSENING, however, holds.

5 Probability and ranking theory

For a number of years, Wolfgang Spohn has been floating a fairly influential model of rational degrees of belief that appears, on the face of it, to be a genuine alternative to the probabilistic view. This well worked-out model, known as ‘ranking theory’ ships with a number of attractive features, including various credence update procedures, as well as a story regarding the relation between graded and full belief. Ranking functions are defined as follows:

Definition 5.1. *A ranking function κ is a function from \mathcal{L} to $\mathbb{R}^+ \cup \{\infty\}$ such that, for any $\varphi, \psi \in \mathcal{L}$, (i) if $\varphi \vdash \perp$, then $\kappa(\varphi) = \infty$, (ii) if $\top \vdash \varphi$, then $\kappa(\varphi) = 0$, (iii) if $\varphi \vdash \psi$, then $\kappa(\psi) \leq \kappa(\varphi)$ and (iv) $\kappa(\varphi \vee \psi) = \min\{\kappa(\varphi), \kappa(\psi)\}$.*

A rank of ∞ plays a role that is somewhat analogous to that of a credence of 0 in probabilistic frameworks. In particular, the rank of ∞ is invariant under strict ranking-theoretic conditionalisation, just as a credence of 0 is invariant under strict probabilistic conditionalisation, as the following definition of a conditional ranking functions makes clear:

Definition 5.2. *Where κ is a ranking function with domain \mathcal{L} , $\varphi, \psi \in \mathcal{L}$ and $\kappa(\varphi) < \infty$, the quantity $\kappa(\psi \mid \varphi) = \kappa(\psi \wedge \varphi) - \kappa(\varphi)$ is the conditional rank of ψ given φ .¹⁸*

Adapting the notation to harmonise with the present paper, the account of acceptability on offer is the following, where Acc denotes this time a function from pairs of ranking models $\mathcal{M} = \langle \mathcal{L}, \kappa \rangle$ and sentences $\varphi \in \mathcal{L}$:

MIN RANK: For any ranking models $\mathcal{M} = \langle \mathcal{L}, \kappa \rangle$ and all $\varphi \in \mathcal{L}$, $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff $\kappa(\varphi) < \kappa(\neg\varphi)$.

In a recent article, Spohn (2009) ponders over the ‘suprisingly complex and fascinating’ relation between the ranking-theoretic and probabilistic pictures. He notes some superficial connections and suggests a potential explanation:

...translate the sum of probabilities into the minimum of ranks, and the quotient of probabilities into the difference of ranks. Thereby, the probabilistic law of additivity turns into the law of disjunction, the probabilistic law of multiplication into the law of conjunction (for negative ranks), and the definition of probabilities into the definition of conditional ranks... [T]ake any probabilistic theorem, apply the above translation to it, and you are almost guaranteed to get a ranking theorem... The translation of products and quotients of probabilities suggests that negative ranks simply are the logarithm of probabilities. [*ibid*, p. 209]¹⁹

He does then however note that the suggested picture may not be so clear, mentioning discrepancies regarding conditional independence and ‘positive and non-negative instancial relevance’, as well as the translation of sums of probabilities.

However, the translation is not fool-proof... The issue is not completely cleared up... [The view of ranks as logarithms of probabilities] does not seem to fit with the translation of sums of probabilities. But it does fit when the logarithmic base is taken to be some infinitesimal i ...

But the discussion in the previous section suggests a somewhat more straightforward picture of the relationship between ranks and probabilities. Indeed, it is worth noting the following alternative formulation of MIN RANK:

For any ranking models $\mathcal{M} = \langle \mathcal{L}, \kappa \rangle$ and all $\varphi \in \mathcal{L}$, $\text{Acc}(\mathcal{M}, \varphi) = 1$ iff, for any $v \in W$, $v(\varphi) = 1$, where $W = \{v \in V : \kappa(\varphi_v) = 0\}$.

It is easily verified that, in virtue of the constraints on ranking functions, W is guaranteed to be non-empty.

But now we can see what is a really quite remarkable similarity with the kind of proposal made in the previous section. In both cases, we start

with some positive real -valued function f on the elements of V (previous section) or on the set of strongest consistent sentences in \mathcal{L} (current section). This enables us to select a lower set W of valuations, such that if $v \in W$ and $f(v^*) \leq f(v)$, or $f(\varphi_{v^*}) \leq f(\varphi_v)$, then $v^* \in W$. Acceptability is then identified with truth in all elements of W .

And the similarity is all the more striking if we recall the logarithmic distance D_{\log} , briefly mentioned earlier on. Indeed, the range of the function is $\mathbb{R}^+ \cup \{\infty\}$, just as is the case with ranking functions, with $D_{\log}(v, \text{Pr}) = -\log(\text{Pr}(\varphi_v)) = \infty$ iff $\text{Pr}(\varphi_v) = 0$. This suggests an interpretation of ranks of maximally strong consistent sentences as renormalised negations of logarithms of their probabilities. Given **MIN RANK**, the details of the renormalisation then have to depend on whether we opt for **MIN** or opt for **SAT**, to ensure consistency. Here are two simple correspondences that would do the trick:

For any $v \in V$, $\kappa(\varphi_v) = \min_{v^* \in V}(-\log(\text{Pr}(\varphi_{v^*}))) + \log(\text{Pr}(\varphi_v))$ (for **MIN**)

For any $v \in V$, $\kappa(\varphi_v) = t - \max\{t, -\log(\text{Pr}(\varphi_v))\}$ (for **SAT**)

We then straightforwardly recover the ranks for the remainder of the sentences in the language using Definition 5.1, obtaining a many-to-one mapping from probability function onto ranking functions. Interestingly, this would somewhat vindicate Spohn's intuition about a logarithmic connection between ranks and probabilities. The vindication, however, would only be partial: the connection would not quite be the one anticipated.

6 Concluding comments

Due to space limitations and to the amount of ground to be covered, we have had to keep the model rather simple. The most obvious shortcoming is perhaps the chronic lack of expressiveness of the language. A move from a sentential to a predicate or modal language would yield a significant gain in realism. Another, admittedly somewhat minor, irritant is the current limitation to finitely generated languages. The infinitary case certainly does raise some further issues. For instance, proposals based on distance-minimising, rather than satisficing, need to address the issue of what to say with respect to the acceptability of a contingent sentence in the case in which there is an infinite sequence of ever-closer valuations. On the current

proposal, we would wind up with mandatory suspension of judgment, since W would be empty. There may however be a case for claiming that this kind of case, highlights what one might call a ‘deontic blindspot’ a point at which rationality fails to yield any recommendation whatsoever. But there are also technical complications prior to that stage, when it comes to computing the distances themselves. It is far from clear, in particular, how to suitably generalise the concept of normalised Hamming distance without running the risk of having $W = V$.

Acknowledgments

I am grateful to the members of the Formal Epistemology Project, KU Leuven, and to the audience of PROGIC 2009, Groningen for useful feedback on earlier versions of this paper. I am also indebted to two anonymous referees for this journal for the time and trouble that they took to provide exceptionally detailed and insightful reports. Part of the research for this article was funded by a Research Foundation – Flanders (FWO) postdoctoral research grant.

Appendix

Proof of Theorem 2.1. Assume NON-UNANIMITY. It follows that there exists a model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentence $\varphi \in \mathcal{L}$ such that $\text{Pr}(\varphi) = p < 1$ but $\text{Acc}(\mathcal{M}, \varphi) = 1$. Since, as is well-known, there exists a rational number between any two distinct real numbers, there exists a rational number $q = m/n$ (with $m, n \in \mathbb{N}$), such that $p < q < 1$. Let $\mathcal{M}^* = \langle \mathcal{L}^*, \text{Pr}^* \rangle$ be a uniform probability model, such that the cardinality of the set of valuations of \mathcal{L}^* is equal to m . Let ψ denote an arbitrary sentence validated by exactly n of the valuations. Now by LOTTERY-PROOFNESS, $\text{Acc}(\mathcal{M}^*, \psi) = 0$, since $\text{Pr}(\varphi_{\{w_1, \dots, w_m\}}) = q < 1$. By MONOTONICITY, however, $\text{Acc}(\mathcal{M}^*, \varphi_{\{w_1, \dots, w_m\}}) \geq \text{Acc}(\mathcal{M}, \varphi) = 1$. Contradiction. \square

Proof of Theorem 2.2. Let $S := \{\varphi_W : W \subseteq V\}$ and $\varphi_W \geq \varphi_{W^*}$ iff $W^* \subseteq W$. Let L be the lattice $\langle S, \geq \rangle$. Let $B \subseteq S$ denote the set of $\varphi \in S$ s.t. $\text{Acc}(\mathcal{M}, \varphi) = 1$. ZERO-NORMALISATION is true iff $\varphi_\emptyset = \perp \notin B$. UNIT-NORMALISATION and DEDUCTIVE CLOSURE are true iff B is a filter of L . So the three conditions are true iff B is a *proper* filter of L . But every proper

filter is the intersection of a set of ultrafilters. For each of these ultrafilters u , there is some $v \in V$ s.t. $u = \{\varphi \in S : v(\varphi) = 1\}$. \square

Notes

¹Note that the assumption made here, that acceptability is a function of an underlying probability model, whilst commonplace in the literature, is not entirely uncontroversial. It rules out, for instance, the views that the acceptability of a sentence depends also depends on the practical payoffs associated with true/false negatives/positives (see Rudner 1953) or again that it is relative to a specific question, modeled as a partition of the language (Levi 1967). A discussion of these issues is however beyond the scope of the present paper.

²In what follows I will be using the expressions ‘full belief in the truth of’ and ‘acceptance of’ interchangeably.

³The name for this constraint originates in (Chandler 2010), and derives from the fact that were it to be violated, STRUCTURALITY and DEDUCTIVE CLOSURE could be marshalled, lottery paradox -style, to yield a violation of ZERO-NORMALISATION. The lottery paradox -proofness involved was dubbed ‘weak’ in the original paper for reasons that do not apply to the current model.

⁴In fact, Douven and Williamson prove a slightly stronger result, making use of a principle that its strictly weaker than DEDUCTIVE CLOSURE, namely:

AGGREGATIVITY: For any probability model $\mathcal{M} = \langle \mathcal{L}, \text{Pr} \rangle$ and sentences $\varphi, \psi \in \mathcal{L}$, if $\text{Acc}(\mathcal{M}, \varphi) = 1$ and $\text{Acc}(\mathcal{M}, \psi) = 1$, then $\text{Acc}(\mathcal{M}, \varphi \wedge \psi) = 1$.

⁵The reader familiar with the judgment aggregation literature will notice a *considerable* amount of simplification going on here. For instance, I am rather severely restricting the class of possible opinion models, which, in the aggregation-theoretic literature, notably involve opinion profiles that are tuples of possibly inconsistent sets of sentences in \mathcal{L} . For reasons that will become clear shortly, these expository niceties can be dispensed with.

⁶See most notably Theorem 2 (a) of (Dietrich and List 2008), in which, in contrast to many other previous results, OPINIONATION plays no role.

⁷Distance based approaches to aggregation are also more recently discussed in Miller and Osherson (2009) and Pigozzi (2006).

⁸On this option, to ensure that the set of selected valuations isn’t empty, one could specify that, in the event that all valuations are further than t , all valuations are selected.

⁹Strictly-speaking, this is not a distance, as the triangle inequality $d(v, v^*) + d(v^*, v^{**}) \geq d(v, v^{**})$ need not be respected.

¹⁰In fact, Konieczny *et al.* use the non-normalised counterpart; the normalisation is introduced here to harmonise with the range of d_D .

¹¹Clause (i) is crucial here. Without it, in the event that stage (b) allows for W to be the empty set, we would have a violation of ZERO-NORMALISATION, since, obviously, $\emptyset \subseteq \llbracket \varphi \rrbracket$, for *all* $\varphi \in \mathcal{L}$, including contradictions.

¹²Both suggestions were very briefly mentioned by Chandler (2010), where the connection with distance-based aggregation functions had not been drawn.

¹³This is of course not a suitable occasion to address the controversial issue of the normative status of a requirement of translation-invariance. The issues at play are complex and, as far as I can see, remain unresolved at this point. See (Zwart 2001), chapter 5 for a detailed overview of the debate.

¹⁴In the original draft of this paper, I had suggested that the issue of translation-sensitivity of distance-based methods had been overlooked in the judgment aggregation literature. As an anonymous referee pointed out to me, however, I was wrong. Indeed, the issue is in fact briefly discussed in a recent piece by Cariani *et al* (2008). There, they first offer a theorem (Theorem 5, p. 17) to the effect that translation sensitivity is a property, not only of the Hamming distance, but of any distance measure that is not ‘trivial’, in the following sense:

A measure d of distance between valuations is trivial iff there exists $r \in \mathbb{R}^+$ such that, for all $v, w \in V$, $d(v, w) = r \times d_D(v, w)$.

After presenting this result, they then go on to say:

In short, only trivial distance measures are translation-invariant... Thus any judgment aggregation procedure that depends on a non-trivial distance measure will fail translation-invariance.

Now the inference from translation-sensitivity of the distance measure to translation-sensitivity of the corresponding aggregation procedure seems basically correct, given certain assumptions about the latter. But these assumptions are not provided and neither is the precise derivation: a little more work is required to establish the result, as Cariani *et al* have acknowledged in recent correspondence.

¹⁵The restriction to finite models is important here. Universal quantification over all probability models—finite or otherwise—would yield a principle that is incompatible with the conjunction of AGGREGATIVITY and ZERO-NORMALISATION. Note, furthermore, that CONSENSUS PRESERVATION entails RESPONSIVENESS.

¹⁶When $t > 0$, since for $t = 0$, INDEPENDENCE is satisfied.

¹⁷As an anonymous referee has pointed out to me, there may be grounds to hold that conservative refinements are perhaps not as informationally innocent as I have suggested and hence that a requirement of preservation of acceptability under such refinements may not be in order. Indeed, the standard Bayesian suggestion of modeling a lack of opinionation with respect to a partition P by a uniform probability distribution over P , faces a number of apparent difficulties: Bertrand-style paradoxes, counterintuitive prescriptions in Ellsberg’s urn decision problem, and so on. Whilst I do share the referee’s worries here, a constraint of preservation under conservative refinement remains the best that can be achieved within the orthodox Bayesian framework that was assumed from the outset of this paper.

¹⁸This is the definition given in (Spohn 2009). Just as the standard ratio definition for conditional probability precludes conditionalisation on probability 0 sentences, it precludes conditionalisation on rank ∞ sentences. There are however alternative accounts of both conditional probabilities and conditional ranks that, rightly or wrongly, waive this prohibition. The former will presumably be well-known to the reader. Regarding the latter, we have the following proposal from (Huber 2009):

$$\kappa(\psi \mid \varphi) = \begin{cases} \kappa(\psi \wedge \varphi) - \kappa(\varphi) & \text{if } \psi \not\vdash \perp; \\ \infty & \text{if } \psi \vdash \perp. \end{cases}$$

The reason for setting the conditional rank to ∞ in case $\psi \vdash \perp$ is presumably to prevent ψ from receiving a rank of 0 upon conditionalisation on a sentence φ of rank ∞ (since we would then have $\kappa(\psi \wedge \varphi) - \kappa(\varphi) = \infty - \infty = 0$), and having $\kappa(\cdot \mid \varphi)$ violate clause (i) of Definition 5.1.

Somewhat curiously, however, note that here, contrary to what was the case in Definition 5.2, we no longer have the result that $\kappa(\psi \mid \neg\psi) = \infty$. Indeed, let $\kappa(\neg\psi) = \infty$ and $\psi \not\vdash \perp$. By the above proposal, $\kappa(\psi \mid \neg\psi) = \kappa(\psi \wedge \neg\psi) - \kappa(\neg\psi) = \infty - \infty = 0$. This result could however be avoided by simply swapping $\psi \vdash \perp$ (resp. $\psi \not\vdash \perp$) for $\psi \wedge \varphi \vdash \perp$ (resp. $\psi \wedge \varphi \not\vdash \perp$) in the above definition.

¹⁹In Definition 5.1, we took the range of κ to be $\mathbb{R}^+ \cup \{\infty\}$. In some publications, however, including the one from which this quote was taken, the range is stated to be $\mathbb{N} \cup \{\infty\}$. But this second option obviously doesn't square with the conjecture that ranks are logarithms of potentially real-valued probabilities.

References

- Cariani, F., M. Pauly and J. Snyder (2008). Decision Framing in Judgment Aggregation. *Synthese*, 163:1–24
- Chandler, J. (2010). The Lottery Paradox Generalised? *British Journal for the Philosophy of Science*, 61(3): 667–679.
- Dietrich, F. and C. List (2008). Judgment Aggregation without Full Rationality *Social Choice and Welfare*, 31(1):15–39.
- Douven, I. and Romeijn, J.-W. (2007). The Discursive Dilemma as a Lottery Paradox. *Economics and Philosophy*, 23:301–319.
- Douven, I. and Williamson, T. (2006). Generalizing the Lottery Paradox. *British Journal for the Philosophy of Science*, 57(4):755–779.
- Hilpinen, R. (1976). Approximate Truth and Truthlikeness. in Przelecki, *et al* (eds.) *Formal Methods in the Methodology of the Empirical Sciences*, Dordrecht, Reidel: 19–42.
- Huber, F. (2009). Belief and Degrees of Belief. In F. Huber and C. Schmidt-Petri (eds) *Degrees of belief*, Springer Synthese Library:1–33.
- Konieczny, S.J. Lang and P. Marquis (2004). DA² Merging Operators. *Artificial Intelligence*, 157(1–2):49–79.

- Levi, I. (1967). *Gambling with Truth* New York: Knopf.
- Miller, D. (1974). On the Comparison of False Theories by their Bases. *British Journal for the Philosophy of Science*, 25:178–88.
- Miller, D. (1976). Verisimilitude Redeflated. *British Journal for the Philosophy of Science*, 27(4):363–81.
- Miller, M. and D. Osherson (2009). Methods for Distance-Based Judgment Aggregation. *Social Choice and Welfare*, 32:575–601.
- Pigozzi, G. (2006) Belief Merging and the Discursive Dilemma: an argument-based account to paradoxes of judgment aggregation. *Synthese*, 152:285–298.
- Rudner, R. (1953). The Scientist Qua Scientist Makes Value Judgments. *Philosophy of Science*, 20(1):1–6.
- Smith, M. (forth.). A Generalised Lottery Paradox for Infinite Probability Spaces. Forthcoming in the *British Journal for the Philosophy of Science*.
- Spohn, W. (2009). A Survey of Ranking Theory. In F. Huber and C. Schmidt-Petri (eds) *Degrees of belief*, Springer Synthese Library:185–228.
- Tichý, P. (1974). On Popper's Definitions of Verisimilitude. *British Journal for the Philosophy of Science*, 25:155–60.
- Tichý, P. (1976). Verisimilitude Redefined. *British Journal for the Philosophy of Science*, 27:25–42.
- Zwart, S.D. (2001). *Refined Verisimilitude*. Springer Synthese Library.