Proceedings of Machine Learning Research 103:38–48, 2019 ISIPTA 2019

Aggregating Belief Models

Seamus Bradley S.C.BRADLEY@LEEDS.AC.UK
University of Leeds

Abstract

This paper has two goals. The first goal is to say something about how one might combine different agents' imprecise probabilities to generate an aggregate imprecise probability. The second goal is to champion the very general theory of "belief models" (de Cooman "Belief models: an order theoretic investigation" Annals of Mathematics and AI 2005) which, I think, deserves more attention. The belief models framework is interesting partly because many other formal models of reasoning appear as special cases of belief models (for example, propositional logic, ranking functions, imprecise probability).

Keywords: Belief Models, Aggregation, Lower Previsions, Lattice Theory

1. Belief Structures

The basic idea of this paper is to start with the idea that the most important structural features of epistemic states are informativeness and coherence. This gives rise to the "belief models" framework first put forward by Gert de Cooman [5]. We then show that this very general model is sufficient to generate a rich and interesting theory of aggregation.

Consider representing agents' epistemic states using sets of sentences in some propositional language. One state is more informative than another just in case the latter is a subset of the former. There is a least informative epistemic state: the empty set. This relation of informativeness partially orders the possible epistemic states, and further gives the epistemic states the structure of a bounded lattice. That is, the least upper bound of a set of epistemic states is itself an epistemic state: it is the intersection of the sets of sentences. Likewise for greatest lower bound. There are two ways a set of sentences can be defective as an epistemic state: it can be self-contradictory or incoherent, or it can be incomplete in the sense of not including the logical consequences of elements of the set. Sets of sentences that have neither of these flaws are privileged: they are the good states to be in. Call these the coherent states. States that don't include all of their logical consequences can be "completed", that is, mapped to a coherent state by a closure operator. The top element of the lattice – the set of all sentences – is not coherent: it is contradictory.

Informativeness and coherence are also characteristic of Imprecise Probability models of epistemic states. We'll be working with lower previsions [14, 2, 25]. For example, for lower previsions, $\underline{P}$ and $\underline{P}'$, say that $\underline{P}$ *dominates* $\underline{P}'$ when $\underline{P}(g) \geq \underline{P}'(g)$ for all g. $\underline{P}$ is more informative than $\underline{P}'$ just in case $\underline{P}$ dominates $\underline{P}'$. This relation of informativeness gives the set of lower previsions the structure of a bounded lattice. The bottom element is the vacuous prevision. Infima and suprema of lower previsions (pointwise minima and maxima) are lower previsions. A lower prevision can be defective in two ways: it can suffer sure loss, or it can be partial. Lower previsions with neither of these defects are privileged, and as before we can call them "coherent". Partial lower previsions can be "completed" through the procedure of natural extension. The top element of this lattice – the lower prevision that assigns value ∞ to every gamble – is obviously incoherent.

In what follows, it will be conceptually useful to be able to identify a lower prevision with its associated closed convex set of linear previsions and thus with a closed convex set of probabilities (a credal set). While there are some subtle distinctions that such loose talk will fail to make, nothing will hang on that in my discussion of lower previsions as a belief structure. Note that thinking in terms of credal sets makes the idea that the incoherent lower prevision is the top element of the lattice more intuitive: an incoherent lower prevision dominates no linear previsions, and thus the associated credal set is empty. Since a more informative credal set is a subset of another, the incoherent credal set – the empty set – is a subset of all others and is thus the top element.

So both propositional logic and lower previsions have this structure of informativeness and coherence. Abstracting from the details of the two cases gives us the general theory of belief models. Let $\mathbf{S}$ be a set of *belief models*, partially ordered by $\preceq$ (read as "is less informative than"), such that $\langle \mathbf{S}, \preceq \rangle$ is a complete bounded lattice.[1] Let $\mathbf{C} \subseteq \mathbf{S}$ be the subset of *coherent* belief models, and stipulate that $\mathbf{C}$ is closed under arbitrary non-empty infima. To be clear, we are saying that there is some, exogenously specified, subset of $\mathbf{S}$ that we are taking to be the coherent belief

1. That is, the relation $\preceq$ is reflexive, transitive and antisymmetric; the partially ordered set is such that the operations of taking a least upper bound or greatest lower bound of a set of elements results in an element of the set; the poset has elements 1 and 0 such that for all elements of the lattice x, $0 \preceq x \preceq 1$;

models. In most particular instances of belief models, we will have some internal grasp on what the coherent models are, for example, the logically closed and consistent sets of sentences, but as far as the abstract formal theory of belief models is concerned, $\mathbf{C}$ is just exogenously given. Stipulate further that $1_{\mathbf{S}} \notin \mathbf{C}$. $\langle \mathbf{S}, \mathbf{C}, \preceq \rangle$ is called a *belief structure*. The elements of $\mathbf{S}$ are the *belief models*: the representations of potential epistemic states. The elements of $\mathbf{C}$ are the coherent belief models. Let $\overline{\mathbf{C}} = \mathbf{C} \cup \{1_{\mathbf{S}}\}$, and define a closure operator:

$$Cl_{\mathbf{S}}(\varphi) = \inf\{\kappa \in \overline{\mathbf{C}}, \varphi \preceq \kappa\}$$

A further property shared by propositional logic and lower previsions is that the maximally informative coherent elements play a special role: the full structure of the space of coherent belief models can be recovered from the maximal coherent elements.

In the propositional logic case, the maximally informative coherent sets of sentences are, in effect, the sets of consequences of a state description or possible world. This means that the logically closed and logically consistent sets of sentences are in a one-to-one correspondence with sets of state descriptions. Likewise, the maximal elements of the belief structure of lower previsions are the linear previsions. Indeed, the set of linear previsions that dominate a given coherent $\underline{P}$ will be a closed and convex set of linear previsions, and any closed and convex set of linear previsions gives rise to a lower prevision by taking the pointwise minimum [25, p. 71].

Let $\mathbf{M} = \{m \in \mathbf{C} : \text{For all } \varphi \in \mathbf{C}, m \preceq \varphi \Rightarrow m = \varphi\}$. The $m \in \mathbf{M}$ are the belief models that are "just under" the maximal element. They are the atoms of the dual structure. Let $M(\varphi) = \{m \in \mathbf{M}, \varphi \preceq m\}$. Call a belief structure a *strong* belief structure, when, for all $\varphi \in \mathbf{C}$, $\varphi = \inf M(\varphi)$. That is, a belief structure is strong just when you can recover all the information about the coherent belief models from looking at the maximally informative elements of the structure.[2]

The formal framework of belief models was introduced in [5], and it provides a very general theory of rational belief. In that paper de Cooman went on to show that one can recover a significant portion of AGM belief revision theory [1, 9] in the framework of belief models. In particular, belief models admit of an analogue of the representation theorem of expansion, and strong belief models also admit of a representation of revision. Contraction appears to be more recalcitrant to a belief models treatment.[3] In essence what de Cooman did was point out that the proofs of those representation theorems relied only on the lattice structure, and not really on any distinctively logical structure. My goal is to point out that the same is true for a large body of literature on "logic based merging" aggregation operators [10, 11]. This gives us a new suite of tools to bring to bear on the question of how to aggregate various kinds of epistemic states.

Lower previsions and propositional logic are not the only theories that have this informativeness and coherence structure. Possibility theory [4] and ranking functions [21, 22] are belief structures, with pointwise dominance as the relation of informativeness, although this sort of structure appears not to be strong (but see footnote 2). Various representations of preference are a belief model. Informativeness corresponds to refinement of the relation. This idea will not be pursued in this paper, but preferences are important in the literature on aggregation more generally, so followup work on the impossibility results (e.g. Arrow's theorem) will pay much more attention to preferences. Strongness of these sorts of structures depends on the details. Other examples of belief models more akin to the lower previsions approach I am about to discuss include sets of desirable gambles [19], and choice functions [20, 26].

One property that de Cooman does not appeal to, but that will be necessary in the current setting is the following.

$$\text{For distinct } a, b, c \in \mathbf{M} \text{ we have } c \npreceq (a \wedge b) \qquad (*)$$

This property is a consequence of distributivity,[4] and since both propositional logic and lower previsions give rise to distributive lattices, it is something that holds of both cases above.

2. Merging Operators for Belief Structures

The main topic of the remainder of this paper is aggregation. Consider a group of agents each of whom has some opinion represented by a belief model. The question naturally arises: how might we generate a new belief model that is, in some sense, an aggregate of the belief models of the individuals? As [24] discuss, there are various understandings of how an aggregate might be interpreted, and so there is not a univocal answer as to what is the right way to aggregate. For example, one might be aggregating so as to make a decision, in which case, an uninformative aggregate will be unhelpful, whereas if the goal of the aggregation is to generate a "common ground" position as the basis for future collaborative enquiry, perhaps the uninformative aggregate is more appealing. In short, pluralism about aggregation seems a natural position.

2. I conjecture that this "strongness" condition can be weakened to the condition that says that every element of $\mathbf{C}$ is such that it is the supremum of the set of maximal ideals that contain it. This would allow some versions of ranking functions to be "strong".

3. What is missing is the representation of a contraction operator by a function determined by a relation on the states. One could *define* a contraction operator using the Harper identity $K_A^- = K \cap (K_{\neg A}^*)$, so the sense in which the belief models theory of AGM is incomplete is minimal.

4. For distinct co-atoms a, b, c we have $c \vee (a \wedge b) = (c \vee a) \wedge (c \vee b) = 1 \wedge 1 = 1 \neq c$, so $c \npreceq (a \wedge b)$. And thus for any c distinct from a, b we have $c \notin M(a \wedge b)$.

So, let Ψ stand for the collection of individual agents attitudes: Ψ is a multiset of belief models. Consider further the possibility that there might be some independent constraints that we would like our merging operator to satisfy. Let the belief model μ represent these constraints. In the propositional case, we can represent a constraint by a set of possible worlds – the worlds where the constraint holds – and thus by a proposition, and thus by a set of sentences. The idea is that whatever else the merging operator does, it must output a belief model that is more informative than the constraint.

Recall that here and throughout, $\wedge$ and $\vee$ are order-theoretic "meet" and "join", not logical "and" and "or".

As de Cooman did for belief revision, we will start our study of aggregation for belief models by taking inspiration from the literature on propositional logic approaches to aggregation, in particular from [10, 11]. Call $\Delta_\mu(\Psi)$ a *merging operator* if Ψ is a multiset of belief models, and μ is a belief model representing the constraints the aggregate belief must satisfy, and Δ satisfies:

IC0 $\mu \preceq \Delta_\mu(\Psi)$

IC1 If μ is consistent then $\Delta_\mu(\Psi)$ is consistent

IC2 If $\bigvee\Psi\vee\mu$ is consistent then $\Delta_\mu(\Psi) = \bigvee\Psi\vee\mu$

IC4 If $\mu \preceq \varphi_1$ and $\mu \preceq \varphi_2$ then $\Delta_\mu(\varphi_1 \sqcup \varphi_2) \vee \varphi_1$ is consistent if and only if $\Delta_\mu(\varphi_1 \sqcup \varphi_2) \vee \varphi_2$ is consistent

IC5 $\Delta_\mu(\Psi_1 \sqcup \Psi_2) \preceq \Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2)$

IC6 If $\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2)$ is consistent then, $\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2) \preceq \Delta_\mu(\Psi_1 \sqcup \Psi_2)$

IC7 $\Delta_{\mu_1\vee\mu_2}(\psi) \preceq \Delta_{\mu_1}(\Psi) \vee \mu_2$

IC8 If $\Delta_{\mu_1}(\Psi) \vee \mu_2$ is consistent then $\Delta_{\mu_1}(\Psi) \vee \mu_2 \preceq \Delta_{\mu_1\vee\mu_2}(\psi)$

These conditions can be motivated in the following way (following [10]). IC0 just says that the output of the merging operator had better satisfy the constraints that we want it to. IC1 says that we want the output of the merging operator to be consistent, unless we require it to be inconsistent by plugging in inconsistent constraints. IC2 tells us that in the ideal case where we are merging a collection of belief models (and constraints) that are such that they are all consistent with each other, then the merging operator should output the supremum of the set of belief models (and constraints). The equivalent claim for sets of sentences says that if there is a non-empty intersection of the belief sets, that should be the output of the merging operator. We might want to split IC2 into two sub-conditions:

IC2a $\Delta_\mu(\Psi) \preceq \bigvee\Psi\vee\mu$

IC2b If $\bigvee\Psi\vee\mu$ is consistent then $\bigvee\Psi\vee\mu \preceq \Delta_\mu(\Psi)$

There is no IC3 for belief models, since the condition is trivial. The condition would have said something like "logically equivalent multisets of belief models should be treated equivalently", but since we have effectively quotiented out logical equivalence, there's nothing for the condition to do. We skip number 3 in order to keep our numbering of the conditions correspond to that of [10]. IC4 states a kind of "fairness" condition. Other things being equal, the output of merging φ_1 with φ_2 should not be consistent with one of them but not the other. IC5 and IC6 together, following on from the intuition driving IC2, say that if there's something common to the merging of two subgroups, then that should be the output of merging the groups together. IC7 and IC8 govern how merging should respond to making the constraints stronger (more informative). These conditions are fairly plausible constraints on aggregation, but finding an operator that satisfies them is non-trivial.

Note that, holding Δ and Ψ fixed, we can define a relation over maximal belief models by looking at pairwise comparisons as follows:

$$m \trianglelefteq_\Psi m' \text{ iff } \Delta_{m\wedge m'}(\Psi) \preceq m \tag{1}$$

Indeed, each Ψ gives rise to such an ordering,[5] and thus Δ gives rise to a *syncretic assignment*: a syncretic assignment is an assignment of a relation over $\mathbf{M}$, $\trianglelefteq_\Psi$ to each multiset Ψ, such that:

S0 For each Ψ, $\trianglelefteq_\Psi$ is a total order on $\mathbf{M}$

S1 If $a \in M(\bigvee\Psi)$ and $b \in M(\bigvee\Psi)$ then $a \trianglelefteq_\Psi b$

S2 If $a \in M(\bigvee\Psi)$ but $b \notin M(\bigvee\Psi)$ then $a \triangleleft_\Psi b$

S4 For all $a \in M(\varphi)$ there is some $b \in M(\varphi')$ such that $b \trianglelefteq_{\varphi\sqcup\varphi'} a$

S5 If $a \trianglelefteq_{\Psi_1} b$ and $a \trianglelefteq_{\Psi_2} b$ then $a \trianglelefteq_{\Psi_1\sqcup\Psi_2} b$

S6 If $a \triangleleft_{\Psi_1} b$ and $a \trianglelefteq_{\Psi_2} b$ then $a \triangleleft_{\Psi_1\sqcup\Psi_2} b$

S7 $\trianglelefteq_\Psi$ is *smooth*, meaning for all μ, for all $m \in M(\mu)$, if m is not minimal with respect to $\trianglelefteq_\Psi$ then there is an $m' \in M(\mu)$ such that m' is minimal and $m' \triangleleft_\Psi m$.

And, given a syncretic assignment we can define a merging operator as follows:

$$\Delta_\mu(\Psi) = \inf_{\preceq} \min_{\trianglelefteq_\Psi}\{M(\mu)\} \tag{2}$$

Indeed, the two concepts are interdefinable.

Theorem 1 *If Δ satisfies IC0–8, then the syncretic assignment defined by equation (1) satisfies S0–7.*

Theorem 2 *If $\trianglelefteq_\Psi$ satisfies S0–7, then the merging operator defined by equation (2) satisfies IC0–8*

Remark 3 *Using equation (1) to define a syncretic assignment, and then using equation (2) to define a merging operator returns the original merging operator.*

Before continuing, let us note a further interesting connection between this work and that of [5]. Letting K^*_A stand for revising knowledge base K by proposition A, [11] show that $K^*_A = \Delta_A(K)$. That is, if Δ is a merging operator, then one can recover an AGM revision operator through the above equation.[6]

5. One can think of this along the lines of an entrenchment ordering á la AGM theory [1, 9].

6. The converse is also sort of true, but much more complicated.

So we have seen that we can generate a merging operator by defining a syncretic assignment: a class of relations over the maximal elements of the belief structure. One natural thought is that these relations are determined by some "distance" over elements of **M**. The following discussion builds on [12]. For convenience, I am going to call maximal coherent belief models "worlds", in analogy to the propositional case. Start from a distance between worlds: $D(m,m')$ satisfying:

D0 D maps pairs of worlds to real numbers
D1 $D(m,m') = D(m',m)$
D2 $D(m,m') = 0$ iff $m = m'$

Define a distance between worlds and belief models:

$$D(m,\varphi) = \min_{\varphi \preceq m'} \{D(m,m')\} \tag{3}$$

Define a distance aggregation function F satisfying:

F0 F takes a sequence of real numbers and outputs a real number
F1 If $x \leq y$ then $F(x_1,\dots,x,\dots,x_n) \leq F(x_1,\dots,y,\dots,x_n)$
F2 $F(x_1,\dots,x_n) = 0$ iff $x_1 = \dots = x_n = 0$
F3 For all $x \in \mathbb{R}$, $F(x) = x$

Let $\Psi = \{\{\varphi_1,\varphi_2,\dots,\varphi_n\}\}$ Define a distance between worlds and multisets of belief models:

$$D(m,\Psi) = F(D(m,\varphi_1), D(m,\varphi_2),\dots,D(m,\varphi_n))$$

Define a relation between worlds depending on Ψ:

$$m \trianglelefteq_\Psi m' \text{ iff } D(m,\Psi) \leq D(m',\Psi) \tag{4}$$

And use equation (2) to define a merging operator. Call this $\Delta^{D,F}$.

Theorem 4 *If D satisfies D0–2, F satisfies F0–3, and the relation defined by equation (4) is smooth (satisfies S7) then $\Delta^{D,F}$ satisfies IC0–2 and IC7,8.*

Consider the further properties that F might satisfy.

F4 For a permutation σ, $F(x_1,\dots,x_n) = F(\sigma(x_1),\dots,\sigma(x_n))$
F5 $F(x_1,\dots,x_n) \leq F(y_1,\dots,y_n) \Rightarrow F(x_1,\dots,x_n,z) \leq F(y_1,\dots,y_n,z)$
F6 $F(x_1,\dots,x_n) \leq F(y_1,\dots,y_n) \Leftarrow F(x_1,\dots,x_n,z) \leq F(y_1,\dots,y_n,z)$

Theorem 5 *If D satisfies D0–2, F satisfies F0–3 and the relation defined by equation (4) is smooth (satisfies S7) then: F satisfies F4–6 iff $\Delta^{D,F}$ satisfies IC0–8.*

Among the functions that satisfy F0–6 are summation and a lexmax.[7]

The only slight wrinkle in the belief models version of this theorem is that the relation defined by the distance must be smooth (i.e. satisfy S7). This can be achieved in a couple of different ways: have the codomain of the distance function be well-ordered ([12] have "distances" into $\mathbb{N}$); or consider cases where D is a metric on **M** and **C** is a collection of closed (in the induced topology) subsets. This latter case is what happens with lower previsions with a well-behaved distance such as Euclidean distance.

For more on distance based merging in general, see [12, 11]. It is interesting to note that there are, in essence, two moving parts here: D and F. However, not much has been written about alternatives to min in equation (3). Exploring this possibility must be left for a later paper.

The upshot of the above theorems is that we have a shortcut to generating aggregation rules for belief models. That is, whenever we have a formal theory that satisfies the pretty minimal conditions of belief structures, we can generate aggregation rules for such a structure simply by thinking about "distances" between the maximal elements of the structure.[8] In the next section we will do this for credal sets. Since the maximal elements of the belief structure of credal sets are the probability functions, we can use the extensive literature on distances between probabilities as a springboard to new aggregation procedures.

Dubois et al. also discuss a formal theory of aggregation with an extremely wide scope [7]. In their framework, they take the triple of core (most plausible elements), support (minimally plausible elements) and plausibility ordering to be primitive, and they show that many formal models fit this mould including the ones discussed here. They then discuss an "information ordering" as derived from the above. What the belief models approach does is shows how far you can get with *just* the structure imposed by the informativeness relation. This is interesting for two reasons. First, informativeness and coherence seem like more natural first principles for a general representation of attitudes than do core, support and plausibility ordering. We can see this from the fact that this structure applies in a kind of trivial way to, for example, propositional logic belief sets. (the core is identical to the support, and the plausibility ordering is the equivalence relation for a two-element partition). Second, informativeness and coherence seem applicable even in the context of non-epistemic attitudes such as preference, whereas the notions of support and core seem inescapably epistemic. We also go beyond existing accounts of generalised merging by making room for "independent constraints".

7. Lexmax returns a value such that the ordering on **M** coincides with the following procedure: Associate each m with a list of distances between m and each φ, with the distances in ordered from biggest to smallest, order the m according to the lexicographic ordering of the lists of ordered distances. See [12, p. 60] for details.

8. "Distances" in inverted commas here because D0–2 are weaker than the typical properties of distance (e.g. of a metric).

3. Merging Operators for Lower Previsions and Probabilities

We now focus on the question of how to define interesting merging operators for belief models, and in particular what this can tell us about aggregating lower previsions. We will pay particular attention to drawing connections to the literature on aggregating probabilities, which are a special case. We will, for the most part, be suppressing the independent constraints μ: we can presume for this section that $\mu = 0_{\mathbf{S}}$.

3.1. A Crude Merging Operator

As a warm up, let's consider a crude form of aggregation for credal sets, as discussed by several authors [7, 6].[9] Let's presume for the moment that the independent constraints are trivial. Consider the following simple aggregation rule: if the credal sets have a non-empty intersection then the aggregate should be that intersection, and if not, then the aggregate should be the (closure of the convex hull of) the union of the credal sets. Does this aggregation rule satisfy IC0–8? It might not be immediately obvious that it does, but we can show it does by constructing a syncretic assignment that satisfies S0-7. Divide $\mathbf{M}$ into three parts: A the possibly empty set that contains all probabilities in the intersection of the credal sets in Ψ; B the set containing all probability functions that are in some but not all members of Ψ; C the set containing probability functions in no member of Ψ. Recalling that $\bigvee\Psi$ is the intersection of the credal sets in Ψ, and $\bigwedge\Psi$ is the union, we have: $A = M(\bigvee\Psi)$, $B = M(\bigwedge\Psi)\setminus A$, and $C = \mathbf{M}\setminus B\cup A$. Recall that the point of a syncretic assignment is that the things that are minimal in the ordering are included in the aggregation. So we want members of A (if there are any) to be lowest, and members of B to be higher in the ordering than those in A, but lower than those in C. So, define $\trianglelefteq_\Psi$ as follows: for all $a, a' \in A, b, b' \in B, c, c' \in C$,

- $a \trianglelefteq_\Psi a'$ and $a' \trianglelefteq_\Psi a$ and likewise for b, b' and c, c'.
- $a \triangleleft_\Psi b \triangleleft_\Psi c$

Alternatively, the same ordering is induced by this function via Equation (4):

$$D_c(m,\Psi) = \begin{cases} 0 \text{ if } \bigvee\Psi \preceq m \\ 1 \text{ if } \bigwedge\Psi \preceq m \\ 2 \text{ otherwise} \end{cases}$$

It's pretty clear that such a class of relations is indeed a syncretic relation in the sense of satisfying S0–7.

Note that we are defining $D_c(m,\Psi)$ directly, rather than have it determined by a D, F pair as above, but we define the merging operator on the basis of this function through equations (4) and (2). This may seem a trivial, and unhelpful sort of aggregation, but it commutes with aggregating by taking products [6], and with conditionalisation [24, 23]. This sort of approach also seems in line with what [8] say about resolving disagreement through IP.

Consider now another aggregation procedure. Start with the so-called "drastic distance":

$$D_d(m,m') = \begin{cases} 1 \text{ if } m = m' \\ 0 \text{ otherwise} \end{cases}$$

Pair this with summation as the F function and we know that this determines an aggregation function that satisfies IC0–8. We need to consider which $m \in \mathbf{M}$ will have the smallest sum of minimum distances with respect to D_d. It is easy to see that those m will be the ones that are in the most of the $M(\varphi)$s for $\varphi \in \Psi$. In essence, what this approach does is find the biggest number n such that a subset of Ψ of size n has a coherent supremum, collect together all probability functions that appear in a subset of that size, and take the closure of the convex hull of that set to be the aggregate. This is motivated by the same intuitions that support the above intersection-or-union merging operator, but it is a little more fine-grained. It is, in essence, a maximal consistent subset approach, where maximality is determined by the cardinality of the subset of Ψ, rather than in terms of set inclusion, as is typical. This kind of MCS has been discussed in the logic case by, for example, [11]. This approach is still somewhat unintuitive as can be seen in Figure 1:[10] since the intersection of credal sets is empty, the merging operator yields the closure of the convex hull of the union of the credal sets. If one member of the collective grows slightly such that the intersection is now non-empty, the aggregate shrinks discontinuously to become the intersection of the credal sets.

3.2. Distance Based Merging

Let's look at what distance-based merging actually gives us with more discriminating distance functions.

$$\Delta_\mu^{D,F}(\Psi) = \inf_{\preceq} \arg\min_{\mu\preceq m} \left\{ F\Big(\min_{\varphi_1\preceq m_1} D(m,m_1),\ldots,\min_{\varphi_n\preceq m_n} D(m,m_n)\Big)\right\} \tag{5}$$

If we let F be summation, this bears some surface resemblance to the (unweighted) "coherent approximation principle" [16, 18, 17], which, translated into our current framework[11] looks like this:

$$\Delta_\mu^{CAP}(\Psi) =$$

9. The discussion in this paper is, as they say, inspired mostly by Peter Walley's 1982 Technical Report "The elicitation and aggregation of imprecise probabilities". I have been unable to find a copy of this report, so I base this discussion on what Dubois et al. say about it.

10. A quick note on how to read these diagrams. Each point in the space represents a probability function over three mutually exclusive and exhaustive events, a, b, c. The closer the point is to the vertex labelled a, the higher the probability of a and so on. The purple shape marks the agggregate credal state.

11. Nobody in this literature is thinking about independent constraints, but the translation here seems plausible in they were.

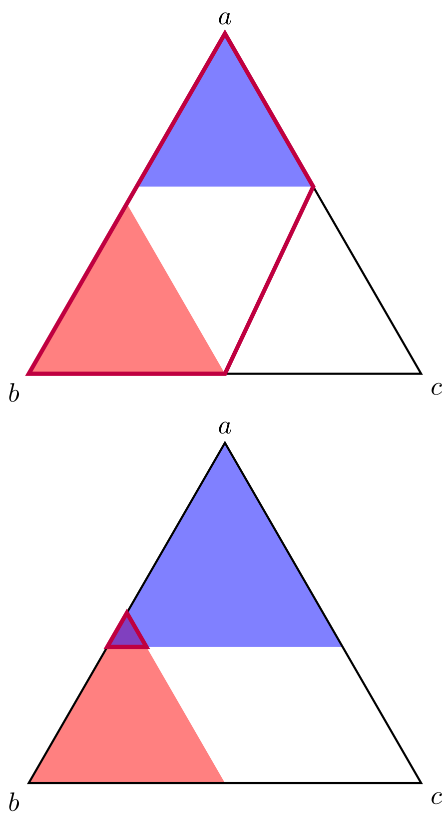

Figure 1: Discontinuous aggregation

$$\inf_{\preceq} \arg\min_{\mu \preceq m} \left\{ \sum_{\varphi \in \Psi} D(m, \varphi) \right\}$$

The difference between the two is that the D in the CAP has to be defined over all belief models, whereas $\Delta^{D,F}$ needs D over $\mathbf{M}$ only. We know that these are distinct aggregation procedures, however, since the CAP will typically output a precise probability, whereas, since (5) satisfies IC2, it will sometimes output a larger credal set. So this is a new aggregation rule discovered through the use of the belief models approach. It's not necessarily a rule I would endorse for reasons we will come to shortly, but note that because we arrived at the rule through the belief models approach we immediately know that it satisfies IC0–8.

Note that if every $\varphi \in \Psi$ is in $\mathbf{M}$ then the two expressions coincide and for that special case, results in [17] tell us that if D is squared Euclidean distance, the aggregate is equivalent to linear pooling and if D is Generalised Kullback-Leibler divergence, merging is by geometric pooling. That is, if you input precise probabilities, the aggregate output is also precise. And in fact, this sort of distance based merging will often yield a precise probability. Essentially, if D is some conventional distance function and F is continuous, you're looking for members of $\mathbf{M}$ that minimise a continuous real-valued function: it's rare that there'll more than one such m. Thinking in terms of credal sets, Δ will output a coherent (i.e. closed and convex) credal set, such that each extremal element of that set minimises a sum of minimum distances. Only in specific circumstances will such a set not be a singleton. One such non-singleton case is where the multiset of credal sets has a non-empty intersection. To give the flavour of another possible kind of

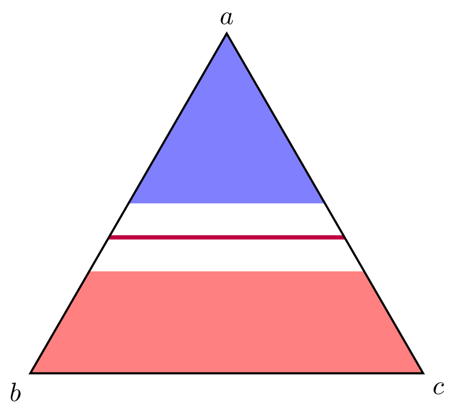

Figure 2: Non-singleton aggregation

case, consider Figure 2. But in most other circumstances, it seems like the set of points minimising the above formula will be a singleton. This seems a somewhat unsatisfactory result for IP aggregation.

A further issue with this approach is that it doesn't seem to "respect" how imprecise each agent is. For example, consider Figure 3. In both cases, the aggregate is the singleton that assigns $\frac{1}{3}$ to each of a, b, c. And yet, in the lower case, it seems like the green agent's confidence that c's probability is about $\frac{1}{2}$ ought to affect the aggregation in some way.[12] This is, of course, merely an impressionistic sketch of what seems unsatisfying about this approach, and more work could certainly be done on merging along the lines of Equation (5).

To give another example of an aggregation rule, consider using Euclidean distance for D and the *lexmax* rule for F. This amounts to aggregating by determining, for each probability function, a sequence of minimum distances, one for each member of Ψ. We then order these from greatest to lowest, and then pick the elements that are minimal with respect to the lexicographic order. So, essentially, (lexi-)minimise maximum minimum distances. We can contrast this with taking F to be summation by considering a case like the top image of Figure 3 except that there are, say, ten "copies" of the blue credal set and one each of the red and green. For the lexicographic rule, the fact that there are

12. How the precision of the agents being aggregated ought to affect the aggregation is a question I have not explored deeply, but it certainly seems like a method that precludes that precision having any effect seems to have a flaw. For one take on aggregation and precision, see [13].

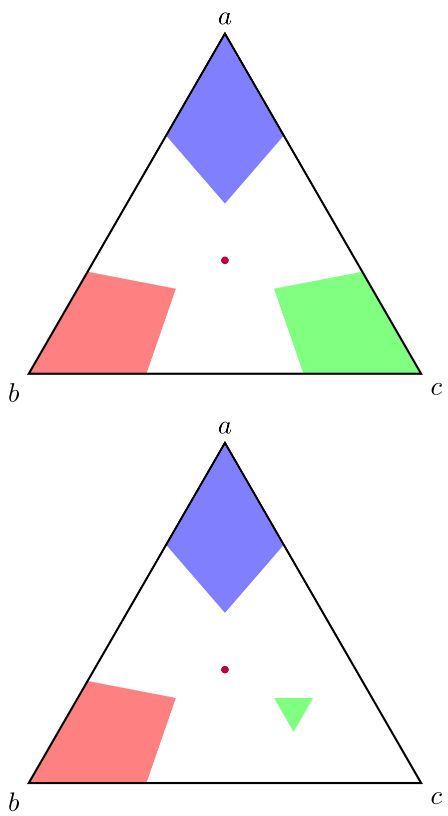

Figure 3: Aggregation and precision

more a-favouring members makes no difference, and the purple point in the figure is where the aggregation would be. For the sum rule, the aggregate would be closer to the bottom tip of the blue shape, since the copies would make a difference to the sum of minimum distances. This illustrates the difference between majoritarian (sum) and arbitration (lexicographic) aggregation rules [10, 11].

3.3. Other Work on Merging Credal Sets

An important work on the aggregation of IP is an unpublished technical report by Peter Walley. I haven't found a copy of this report, but [7] discusses Walley's work, and I discuss it here, based on my second hand knowledge of it. In the interests of space I suppress discussion of those principles Dubois et al. say Walley considers dubious. The remaining criteria, translated into the above formalism look like this:[13]

Coherence $\Delta(\Psi)$ should be a coherent belief model (i.e. a convex set of probabilities).

Unanimity $\bigwedge\Psi \preceq \Delta(\Psi$

Reconciliation If $\bigvee\Psi$ is consistent then $\Delta(\Psi) \vee \varphi$ is consistent for all $\varphi \in \Psi$.

Indeterminacy $\Delta(\Psi) \preceq \bigvee\Psi$

Strong reconciliation If $\bigvee\Psi_1$ is consistent then for any Ψ_2, $\Delta(\Psi_1 \sqcup \Psi_2) \vee (\bigvee\Psi_1)$ is consistent

Conjunction If $\bigvee\Psi$ is consistent then $\bigvee\Psi \preceq \Delta(\Psi)$

Symmetry Permuting the elements of Ψ does not influence the aggregate

Complete ignorance $\Delta(\Psi \sqcup 0_{\mathbf{S}}) = \Delta(\Psi)$ (i.e. an agent who provides no information can be ignored)

Updating Merging should commute with rational update

Coherence is, essentially, IC1. Unanimity seems natural, but in a context where arbitrary independent constraints are possible, we can't rule out the possibility that the independent constraints would force us to output an aggregate incompatible with all the probabilities in the union of the credal sets. Reconciliation is a bit like IC4, but IC4 is weaker in being a biconditional rather than an absolute requirement. Strong reconciliation is also not a consequence of IC0–8. Indeterminacy is, in effect, IC2a and Conjunction is IC2b. Symmetry is a consequence of taking the input to be a multiset, rather than, say, a sequence of belief models. It follows from S1 that $a \triangleq_{0_{\mathbf{S}}} b$ for all $a, b, \in \mathbf{M}$. Together with the other properties S0–7, it appears that every merging operator in this framework will satisfy Complete ignorance.[14] Whether all merging operators satisfy updating is an open question. As mentioned above, it appears that certain operators do.

Moral and del Sagrado also discuss some desirable properties of credal set aggregation functions. They mention Symmetry (which they call P1), Indeterminacy (P3) and Conjunction (P5) all of which are validated in the current framework. They also discuss Complete ignorance (P4) and Unanimity (P6) which are probably validated by most if not all merging operators. Their final property, (P2) exploits the mixture space structure of lower previsions that has, so far, made no appearance in this paper.

Mixture closeness If we have $\varphi_1, \dots, \varphi_n$ and φ_1' is such that for every $p \in \varphi_1, q \in \varphi_j (j \geq 2)$ there is a $p' \in \varphi_1'$ such that p' is a convex combination of p and q, then $\Delta(\varphi_1', \varphi_2, \dots, \varphi_n) \subseteq \Delta(\varphi_1, \varphi_2, \dots, \varphi_n)$

P2 intuitively says that if you replace φ_1 with φ_1' which is uniformly "closer" to each other φ_j, then the aggregation should be more precise (a smaller set of probabilities). I don't yet have a firm opinion on whether this is a sensible thing to require of aggregation.

3.4. Linear Pooling

In the discussion of the CAP above, we mentioned linear pooling. Now, we know that a convex combination of coherent lower previsions is coherent [25, p. 59]. So can we recover linear pooling as an aggregation method in this framework? One of the few papers to discuss aggregating

13. I have suppressed the independent constraints, since they are not something Walley considered.

14. It remains a project for future work to see what this means for majoritarian merging operators [11].

imprecise probabilities [15] proposes this as a method.[15] Define:

$$\Psi^*_{\{\lambda_i\}}(X) = \sum_{\varphi_i \in \Psi} \lambda_i \varphi_i(X) \tag{6}$$

For any non-negative weights that sum to one, $\bigwedge \Psi \preceq \Psi^* \preceq \bigvee \Psi$. And, indeed, the latter inequality can be strict even when $\bigvee \Psi$ is consistent, thus this aggregation procedure does *not* satisfy IC2a.[16] The linear pool can also violate IC2b. Despite this, it is an interesting and natural aggregation worthy of further study.

4. Conclusion

I hope to have shown that the belief models framework is worthy of study. I think this is so both as a general model for theorising about rational belief, and also as a method for importing valuable work from propositional logic approaches into a broadly probabilistic way of doing things.

Appendix A. Proofs

First, we summarise, without proof, some results from [5]. Call a belief model φ *consistent* when its closure $Cl_{\mathbf{S}}(\varphi)$ is in $\mathbf{C}$.

Lemma 6 *The following are equivalent:*

- *φ is consistent*
- *$Cl_{\mathbf{S}}(\varphi) \prec 1_{\mathbf{S}}$*
- *$\varphi \preceq \kappa$ for some $\kappa \in \mathbf{C}$*

If we are talking about a strong belief structure, then each of the above is equivalent to:

- *$M(\varphi) \neq \emptyset$*

Lemma 7

- *$M(\varphi) = M(Cl_{\mathbf{S}}(\varphi))$*
- *$Cl_{\mathbf{S}}(\varphi) = \inf M(\varphi)$*
- *$\varphi \preceq \varphi'$ iff $M(\varphi') \subseteq M(\varphi)$*
- *$M(\varphi \vee \varphi') = M(\varphi) \cap M(\varphi')$*

The first three proofs follow those of [10, 3]. The innovation here is to demonstrate that their proofs only essentially rely on the order-theoretic structure of the propositional logic they took to be their target.

Proof of Theorem 1 S0: $m \wedge m'$ is consistent (since $\mathbf{C}$ is closed under infima), and thus by IC1 $\Delta_{m\wedge m'}(\Psi)$ is consistent. By IC0 $\Delta_{m\wedge m'}(\Psi) \preceq m \wedge m'$, thus $M(\Delta_{m\wedge m'}(\Psi)) \subseteq M(m \wedge m') = \{m, m'\}$. So at least one of m or m' is in $M(\Delta_{m\wedge m'}(\Psi))$, and thus either $m \trianglelefteq_\Psi m'$ or $m' \trianglelefteq_\Psi m$.

15. There is a great deal more in this paper than I have space to discuss here.
16. Note that if every member of Ψ is maximal, then either they all coincide, or their supremum is inconsistent, thus this result does not conflict with what I said earlier about $\Delta^{D,F}$ coinciding with CAP and thus with linear pooling in this circumstance.

$\emptyset \neq M(\Delta_m(\Psi)) \subseteq M(m) = \{m\}$. So $m \preceq \Delta_{m\wedge m}(\Psi)$ and thus $m \trianglelefteq_\Psi m$.

Let $a, b, c \in \mathbf{M}$ and assume for contradiction that $a \trianglelefteq_\Psi b$, $b \trianglelefteq_\Psi c$ but $a \not\trianglelefteq_\Psi c$. Since $a \not\trianglelefteq_\Psi c$, we have $\Delta_{a\wedge c}(\Psi) = c$. Now consider $\Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge c)$. Take the case where this is consistent first. By IC7,8 we know that this expression is equal to $\Delta_{\inf\{a,b,c\} \vee (a\wedge c)}(\Psi) = \Delta_{a\wedge c}(\Psi) = c$. This means that, since $M(\Delta_{\inf\{a,b,c\}}(\Psi))$ is non-empty, it must be equal to $\{c\}$ or $\{b, c\}$.

Take these cases in turn. Assume $M(\Delta_{\inf\{a,b,c\}}(\Psi)) = \{b, c\}$. Consider $M(\Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge b)) = M(\Delta_{\inf\{a,b,c\}}(\Psi)) \cap M(a \wedge b) = b$. Thus, since it is consistent, by IC7 and IC8, $b = \Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge b) = \Delta_{a\wedge b}(\psi)$, which contradicts the fact that $a \trianglelefteq_\Psi b$.

Now assume $M(\Delta_{\inf\{a,b,c\}}(\Psi)) = \{c\}$. For similar reasons as above, $c = \Delta_{\inf\{a,b,c\}}(\Psi) \vee (b \wedge c) = \Delta_{b\wedge c}(\Psi)$, which, again, contradicts our original statement.

The other case to consider is when $\Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge c)$ is not consistent, meaning $M(\Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge c)) = M(\Delta_{\inf\{a,b,c\}}(\Psi)) \cap M(a \wedge c) = \emptyset$. However, $M(\Delta_{\inf\{a,b,c\}}(\Psi) \subseteq \{a, b, c\}$ and is non-empty, thus $\Delta_{\inf\{a,b,c\}}(\Psi) = b$. By considering $\Delta_{\inf\{a,b,c\}}(\Psi) \vee (a \wedge b)$, we conclude that $\Delta_{a\wedge b}(\Psi) = b$ meaning $b \triangleleft_\Psi a$, which contradicts our original statement. So $\trianglelefteq_\Psi$ is transitive.

S1: By IC2, $\Delta_{a\wedge b}(\Psi) = a \wedge b$ so $a \trianglelefteq_\Psi b$ and $b \trianglelefteq_\Psi a$.

S2: $\Delta_{a\wedge b}(\Psi) = a$ (by IC2) so $a \triangleleft_\Psi b$.

S4: Let $a \in M(\varphi)$, and thus $\varphi \wedge \varphi'$ is consistent. Therefore, by IC1, $\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi')$ is consistent. Assume for contradiction that $M(\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi') \vee \varphi) = \emptyset$. Thus, since $\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi')$ is not empty, and $M(\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi')) \subseteq M(\varphi \wedge \varphi)$, it must be that $M(\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi')) \subseteq (M(\varphi') \setminus M(\varphi))$. This contradicts our assumption, and thus $\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi') \vee \varphi$ is consistent. By S4, $\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi') \vee \varphi'$ is consistent. Let $b \in M(\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi') \vee \varphi')$. So $b \in M(\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi'))$, and thus $\Delta_{\varphi\wedge\varphi'}(\varphi \sqcup \varphi') \vee (a \wedge b)$ is consistent, and thus by IC7,8 $b \in M(\Delta_{a\wedge b}(\varphi \sqcup \varphi')$. Therefore, $b \trianglelefteq_{\varphi \sqcup \varphi'} a$ as required.

S5: By hypothesis, $\Delta_{a\wedge b}(\Psi_1) \preceq a$ and $\Delta_{a\wedge b}(\Psi_2) \preceq a$. So by IC5, $\Delta_{a\wedge b}(\Psi_1 \sqcup \Psi_2) \preceq \Delta_{a\wedge b}(\Psi_1) \vee \Delta_{a\wedge b}(\Psi_2) \preceq a$ and thus $a \trianglelefteq_{\Psi_1 \sqcup \Psi_2} b$.

S6. From S5, we know that $a \trianglelefteq_{\Psi_1 \sqcup \Psi_2} b$, so we need to show that $b \not\trianglelefteq_{\Psi_1 \sqcup \Psi_2} a$, or in other words, that $\Delta_{a\wedge b}(\Psi_1 \sqcup \Psi_2) \not\preceq b$. $\Delta_{a\wedge b}(\Psi_1) \not\preceq b$ and thus, $\Delta_{a\wedge b}(\Psi_1) \vee \Delta_{a\wedge b}(\Psi_2) \not\preceq b$. Since $\Delta_{a\wedge b}(\Psi_1) \vee \Delta_{a\wedge b}(\Psi_2)$ is consistent (a is in both halves), by IC5,6, $\Delta_{a\wedge b}(\Psi_1) \wedge \Delta_{a\wedge b}(\Psi_2) = \Delta_{a\wedge b}(\Psi_1 \sqcup \Psi_2)$. Thus, $\Delta_{a\wedge b}(\Psi_1 \sqcup \Psi_2) \not\preceq b$ so $b \not\trianglelefteq_{\Psi_1 \sqcup \Psi_2} a$.

S7: The proof that $\trianglelefteq_\Psi$ is smooth follows [3] proposition 40 (vi). Consider $m \in M(\mu)$ such that m is not minimal, meaning $m \notin M(\Delta_\mu(\Psi))$. Now take some $n \in M(\Delta_\mu(\Psi))$; we can be sure there is one because μ is consistent (and IC1). Such an n is minimal in $M(\mu)$ with respect to $\trianglelefteq_\Psi$, by the above proof. So it just remains to show that $n \triangleleft_\Psi m$. $m, n \in M(\mu)$, but $m \notin M(\Delta_\mu(\Psi)) = \min_\trianglelefteq\{M(\mu)\}$, so it is not the case that $m \trianglelefteq_\Psi n$. Since $\trianglelefteq_\Psi$ is a complete relation, it must be that $n \triangleleft_\Psi m$. ■

Proof of Remark 3 Assume that $a \in M(\Delta_\mu(\Psi))$ but $a \notin \min_{\trianglelefteq_\Psi} M(\mu)$. So there is some $b \in M(\mu)$ such that $b \triangleleft_\Psi a$. $\Delta_\mu(\Psi) \vee (a \wedge b)$ is consistent (since $a \in M(\Delta_\mu(\Psi))$, so $\Delta_\mu(\Psi) \vee (a \wedge b) = \Delta_{a\wedge b}(\Psi)$. This is so because, since $\mu \preceq a, b$ and thus $\mu \vee (a \wedge b) = a \wedge b$. Now, $a \notin \Delta_{a\wedge b}(\Psi)$, because $b \triangleleft_\Psi a$. Consider $a \notin M(\Delta_{a\wedge b}(\Psi)) = M(\Delta_\mu(\Psi) \vee (a \wedge b)) = M(\Delta_\mu(\Psi)) \cap M(a \wedge b)$. Since $a \in M(a \wedge b)$, it must be that $a \notin M(\Delta_\mu(\Psi))$. This contradicts our original assumption. Thus, $M(\Delta_\mu(\Psi) \subseteq \min_{\trianglelefteq_\Psi} M(\mu)$.

Now, suppose $a \in \min_{\trianglelefteq_\Psi} M(\mu)$ but (for contradiction) $a \notin M(\Delta_\mu(\Psi)$. For every $b \in M(\mu)$, $a \trianglelefteq_\Psi b$. $a \in M(\mu)$ and this μ is consistent. Therefore, by IC1, $\Delta_\mu(\Psi)$ is consistent. Let $b \in M(\Delta_\mu(\Psi)$. We know that $a \trianglelefteq_\Psi b$ and thus $a \in M(\Delta_{a\wedge b}(\Psi))$. But, $b \in M(\Delta_\mu(\Psi)) \cap M(a\wedge)$ and so, $\Delta_\mu(\Psi) \vee (a \wedge b)$ is consistent. Also, by assumption, a is not in that intersection. By IC7,8 we know that this is equal to $\Delta_{a\wedge b}(\Psi)$ which is, in turn equal to b. This contradicts our assumption. Thus $\min_{\trianglelefteq_\Psi} M(\mu) \subseteq M(\Delta_\mu(\Psi)$. ■

Proof of Theorem 2 IC0: By definition, $M(\Delta_\mu(\Psi) \subseteq M(\mu)$. Thus $\mu \preceq \Delta_\mu(\Psi)$.

IC1: $M(\mu)$ is not empty. Smoothness of $\trianglelefteq_\Psi$ guarantees that there are no infinite descending chains of members of $M(\mu)$ and thus, the set of minimal elements is non-empty.

IC2:

$m \in M(\bigwedge \Psi \vee \mu) = M(\bigwedge \Psi) \cap M(\mu)$. All and only those m will be minimal according to $\trianglelefteq_\Psi$, because of S1,2.

IC4: Consider $a \in M(\Delta_\mu(\varphi_1 \sqcup \varphi_2) \vee \varphi_1)$. For, $c \in M(\mu)$, $a \trianglelefteq_{\varphi_1\sqcup\varphi_2} c$. From S4 we know that there is a $b \in M(\varphi_2)$ such that $b \trianglelefteq_{\varphi_1\sqcup\varphi_2} a \trianglelefteq_{\varphi_1\sqcup\varphi_2} c$ So $b \in M(\Delta_\mu(\varphi_1 \sqcup \varphi_2))$. Thus $\Delta_\mu(\varphi_1 \sqcup \varphi_2) \vee \varphi_2$ is consistent.

IC5: Consider $a \in M(\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2))$, meaning for all $b \in M(\mu)$, $a \trianglelefteq_{\Psi_1} b$ and likewise for Ψ_2. Thus, by S5 $a \trianglelefteq_{\Psi_1\sqcup\Psi_2} b$, so $a \in M(\Delta_\mu(\Psi_1 \sqcup \Psi_2))$.

IC6: Let $a \in M(\Delta_\mu(\Psi_1 \sqcup \Psi_2))$. Assume for contradiction that $a \notin M(\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2))$. So $a \notin M(\Delta_\mu(\Psi_1))$ or $a \notin M(\Delta_\mu(\Psi_2))$, without loss of generality, assume it's the former. But since $M(\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2))$ is consistent, pick some b in the maximal elements of it. $b \triangleleft_{\Psi_1} a$ and $b \trianglelefteq_{\Psi_2} a$ so by S6 $b \triangleleft_{\Psi_1\sqcup\Psi_2} a$, which contradicts our assumption. So $a \in M(\Delta_\mu(\Psi_1) \vee \Delta_\mu(\Psi_2))$.

IC7: Let $a \in M(\Delta_{\mu_1}(\Psi) \vee \mu_2)$. Since $a \trianglelefteq_\Psi b$ for all $b \in M(\mu_1)$, we have that $a \trianglelefteq_\Psi b$ for all $b \in M(\mu_1 \vee \mu_2)$. Meaning that $a \in \min_{\trianglelefteq_\Psi}\{M(\mu_1 \vee \mu_2)\} = M(\Delta_{\mu_1\vee\mu_2}(\Psi))$.

IC8: Let $b \in M(\Delta_{\mu_1}(\Psi) \vee \mu_2)$. Consider $a \in M(\Delta_{\mu_1\vee\mu_2}(\Psi))$ but suppose, for contradiction, that $a \notin M(\Delta_{\mu_1}(\Psi))$. So $b \triangleleft_\Psi a$ but $\mu_1 \vee \mu_2 \preceq b$ thus $a \notin \min\{M(\mu_1 \vee \mu_2)\}$. this contradicts what we supposed, thus $a \in M(\Delta_{\mu_1}(\Psi))$. ■

It will be helpful to appeal to the shorthand $F(D(m, \varphi_i)) = F(D(m, \varphi_1), \ldots, D(m, \varphi_n))$.

Lemma 8 *If D satisfies D0–2, then $D(m, \varphi) = 0$ entails that $\varphi \preceq m$. And further, if F satisfies F0–3, then $F(D(m, \varphi_i)) = 0$ iff $\bigvee \Psi \preceq m$.*

The final two proofs are similar to the propositional logic case found in [12].

Proof of Theorem 4 Proofs that $\Delta^{D,F}$ satisfies IC0,IC1 very similar to the corresponding parts of Theorem 1.

IC2 follows from Lemma 8.

IC7: Let $m \in M(\Delta^{D,F}_{\mu_1}(\Psi) \vee \mu_2)$. For any $m' \in M(\mu_1)$ we have $D(m, \Psi) \leq D(m', \Psi)$. Also, $m \in M(\mu_1) \cap M(\mu_2) = M(\mu_1 \vee \mu_2)$. Thus m is minimal in $M(\mu_1 \vee \mu_2)$.

IC8: Consider $m' \in M(\Delta^{D,F}_{\mu_1}(\Psi) \vee \mu_2)$ and $m \in M(\Delta^{D,F}_{\mu_1\vee\mu_2}(\Psi)$. Suppose for contradiction that $m \notin M(\Delta^{D,F}_{\mu_1}(\Psi))$. Thus $D(m', \Psi) < D(m, \Psi)$. Since $m' \in M(\mu_1 \vee \mu_2)$ we must have $m \notin \Delta^{D,F}_{\mu_1\vee\mu_2}(\Psi)$. Contradiction. Thus $m \in M(\Delta^{D,F}_{\mu_1}(\Psi) \cap M(\mu_2)$. ■

Proof of Theorem 5 IC0–2, IC7,8 follow from Theorem 4.

IC4: First note that $D(m, \varphi \sqcup \varphi') = F(D(m, \varphi), D(m, \varphi'))$. Let $\mu \preceq \varphi, \varphi'$ and suppose for contradiction that $\Delta^{D,F}_\mu(\varphi \sqcup \varphi') \vee \varphi$ is consistent but $\Delta^{D,F}_\mu(\varphi \sqcup \varphi') \vee \varphi'$ is not. It follows that $\min_{\varphi\preceq m} D(m, \varphi \sqcup \varphi') < \min_{\varphi'\preceq m} D(m, \varphi \sqcup \varphi')$. Thus $\min_{\varphi\preceq m} F(D(m, \varphi), D(m, \varphi')) < \min_{\varphi'\preceq m} F(D(m, \varphi), D(m, \varphi'))$. By Lemma 8 the RHS is equal to $\min_{m\in M(\varphi')} F(D(m, \varphi), 0)$ and the LHS is equal to $\min_{m\in M(\varphi)} F(0, D(m, \varphi'))$. By F4, and the above, we have that:

$$\min_{m\in M(\varphi')} F(D(m, \varphi), 0) < \min_{m\in M(\varphi)} F(D(m, \varphi'), 0)$$

So by F1 and F6

$$\min_{m\in M(\varphi')} D(m, \varphi) < \min_{m\in M(\varphi)} D(m, \varphi')$$

However, we also have the following:

$$\begin{aligned}
\min_{m\in M(\varphi')} D(m, \varphi) &= \min_{m\in M(\varphi')} \min_{m'\in M(\varphi)} D(m, m') \\
&= \min_{m'\in M(\varphi)} \min_{m\in M(\varphi')} D(m, m') \\
&= \min_{m'\in M(\varphi)} \min_{m\in M(\varphi')} D(m', m) \\
&= \min_{m\in M(\varphi)} D(m, \varphi')
\end{aligned}$$

The first equalities follows by definition of $D(m, \varphi)$, the second is a simple fact about minima, the third is by D1, the fourth is again the definition of $D(m, \varphi)$ plus an exchange of variable. These two results contradict, and thus it must be that $\Delta^{D,F}_\mu(\varphi \sqcup \varphi') \vee \varphi'$ is consistent.

IC5: Let $m \in M(\Delta^{D,F}_{\mu}(\Psi_1) \cap M(\Delta^{D,F}_{\mu}(\Psi_2)$. Thus for $m' \in M(\mu)$ we have that $D(m,\Psi_1) \leq D(m',\Psi_1)$ and $D(m,\Psi_2) \leq D(m',\Psi_2)$. Let $\Psi_1 = \{\{\varphi_{11},\dots \varphi_{1n_1}\}\}$ and $\Psi_2 = \{\{\varphi_{21},\dots \varphi_{2n_2}\}\}$ And thus:

$$F(D(m,\varphi_{11}),\dots D(m,\varphi_{1n_1})) \leq F(D(m',\varphi_{11}),\dots D(m',\varphi_{1n_1})) \tag{7}$$

$$F(D(m,\varphi_{21}),\dots D(m,\varphi_{2n_2})) \leq F(D(m',\varphi_{21}),\dots D(m',\varphi_{2n_2})) \tag{8}$$

Now take inequality (7) and, using F5 repeatedly, "tack on" each $D(m,\varphi_{2i})$ onto the end, and similarly take inequality (8) and repeatedly "tack on" each $D(m',\varphi_{1i})$. This produces:

$$\begin{aligned} &F(D(m,\varphi_{11}),\dots D(m,\varphi_{1n_1}),D(m,\varphi_{21}),\dots D(m,\varphi_{2n_2}) \\ &\leq F(D(m',\varphi_{11}),\dots D(m',\varphi_{1n_1}),D(m,\varphi_{21}),\dots D(m,\varphi_{2n_2}) \end{aligned} \tag{9}$$

and:

$$\begin{aligned} &F(D(m,\varphi_{21}),\dots D(m,\varphi_{2n_2}),D(m',\varphi_{11},\dots D(m',\varphi_{1n_1})) \\ &\leq F(D(m',\varphi_{21}),\dots D(m',\varphi_{2n_2}),D(m',\varphi_{11},\dots D(m',\varphi_{1n_1})) \end{aligned} \tag{10}$$

Using F4, the LHS of inequality (10) is equal to the RHS of inequality (9), and thus we have that $m \in M(\Delta^{D,F}_{\mu}(\Psi_1 \sqcup \Psi_2))$.

IC6: Suppose for contradiction that $\Delta^{D,F}_{\mu}(\Psi_1) \vee \Delta^{D,F}_{\mu}(\Psi_2)$ is consistent but that $\Delta^{D,F}_{\mu}(\Psi_1) \vee \Delta^{D,F}_{\mu}(\Psi_2) \not\preceq \Delta^{D,F}_{\mu}(\Psi_1 \sqcup \Psi_2)$. So there is some m in the RHS that is not in the LHS. Without loss of generality assume that $m \notin M(\Delta^{D,F}_{\mu}(\Psi_1))$. So there is some $m' \in M(\Delta^{D,F}_{\mu}(\Psi_1)$ such that $D(m',\Psi) < D(m,\Psi)$. Note that, since F has a totally ordered codomain, F6 is equivalent to F5 but with the inequalities being strict. Now, using the same technique as for IC5 we can show that $m \notin M(\Delta^{D,F}_{\mu}(\Psi_1 \sqcup \Psi_2))$ which contradicts our assumption. That concludes one direction of the proof.

F4: $\Delta^{D,F}$ is a function from multisets, so order of arguments cannot make a difference.

F5: Let's say we have x_i and y_i such that $F(x_1,\dots,x_n) \leq F(y_1,\dots,y_n)$. Consider some m,m' and $\Psi = \{\{\varphi_1,\dots,\varphi_n\}\}$ such that $D(m,\varphi_i) = x_i$ and $D(m',\varphi_i) = y_i$. By definition, $\Delta^{D,F}_{m\wedge m'}(\Psi) \preceq m$ and $\Delta^{D,F}_{m\wedge m'}(\Psi) \not\preceq m'$. Now we consider a φ' such that $D(m,\varphi') = D(m',\varphi') = z$. $m,m' \in M(\Delta^{D,F}_{m\wedge m'}(\varphi'))$. So, by IC5 $m \in M(\Delta^{D,F}_{m\wedge m'}((\sqcup\varphi_i) \sqcup \varphi'))$ but not so for m', and thus $F(x_1,\dots,x_n,z) \leq F(y_1,\dots,y_n,z)$.

F6: the proof works similarly to F5. ■

References

[1] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic*, 50:510–530, 1985.

[2] Thomas Augustin, Frank P.A. Coolen, Gert de Cooman, and Matthias C.M. Troffaes, editors. *Introduction to Imprecise Probabilities*. John Wiley and Sons, 2014.

[3] José Luis Chacón and Ramón Pino Pérez. Merging operators: beyond the finite case. *Information Fusion*, 7:41–60, 2006.

[4] Gert de Cooman. Possibility theory I: the measure- and integral-theoretic groundwork. *International Journal of General Systems*, 25(4):291–323, 1997.

[5] Gert de Cooman. Belief models: An order-theoretic investigation. *Annals of Mathematics and Artificial Intelligence*, 45:5–34, 2005.

[6] Gert de Cooman and Matthias Troffaes. Coherent lower previsions in systems modelling: products and aggregation rules. *Reliability engineering and system safety*, 85:113–134, 2004.

[7] Didier Dubois, Weiru Liu, Jianbing Ma, and Henri Prade. The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks. *Information Fusion*, 32:12–39, 2016.

[8] Lee Elkin and Gregory Wheeler. Resolving peer disagreements through imprecise probabilities. *Noûs*, forthcoming.

[9] Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, 1988.

[10] Sébastien Konieczny and Ramón Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.

[11] Sébastien Konieczny and Ramón Pino Pérez. Logic based merging. *Journal of Philosophical Logic*, 40: 239–270, 2011.

[12] Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. *DA*^{2} merging operators. *Artifical Intelligence*, 157:49–79, 2004.

[13] Aidan Lyon. Vague credence. *Synthese*, 194(10): 3931–3954, 2017.

[14] Enrique Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approcimate Reasoning*, 48:628–658, 2008.

[15] Serafín Moral and José del Sagrado. Aggregation of imprecise probabilities. In *Aggregation and Fusion of Imperfect Information*, pages 162–188. Springer, 1998.

[16] Daniel Osherson and Moshe Vardi. Aggregating disparate estimates of chance. *Games and Economic Behavior*, 56:148–173, 2006.

[17] Richard Pettigrew. Aggregating incoherent agents who disagree. *Synthese*, forthcoming.

[18] Joel Predd, Daniel Osherson, Sanjeev Kulkarni, and H. Vincent Poor. Aggregating probabilistic forecasts from incoherent and abstaining experts. *Decision Analysis*, 5(4):177–189, 2008.

[19] Erik Quaeghebeur. Desirability. In Augustin et al. [2], pages 1–27.

[20] Teddy Seidenfeld, Mark J. Schervish, and Joseph B. Kadane. Coherent choice functions under uncertainty. *Synthese*, 172:157–176, 2010.

[21] Wolfgang Spohn. A survey of ranking theory. In Franz Huber and Cristoph Schmidt-Petri, editors, *Degrees of Belief*, Synthese Library, pages 185–228. Springer, 2009.

[22] Wolfgang Spohn. *The Laws of Belief: Ranking Theory and its Philosophical Applications*. Oxford University Press, 2012.

[23] Rush Stewart and Ignacio Ojea Quintana. Learning and pooling, pooling and learning. *Erkenntnis*, 2017.

[24] Rush Stewart and Ignacio Ojea Quintana. Probabilistic opinion pooling with imprecise probabilities. *Journal of Philosophical Logic*, 2017.

[25] Matthias Troffaes and Gert de Cooman. *Lower Previsions*. Wiley, 2014.

[26] Arthur van Camp. Choice functions as a tool to model uncertainty, 2017.