Aggregating Judgements by Merging Evidence

JON WILLIAMSON, *University of Kent. E-mail: j.williamson@kent.ac.uk*

Abstract

The theory of belief revision and merging has recently been applied to judgement aggregation. In this article I argue that judgements are best aggregated by merging the evidence on which they are based, rather than by directly merging the judgements themselves. This leads to a three-step strategy for judgement aggregation. First, merge the evidence bases of the various agents using some method of belief merging. Second, determine which degrees of belief one should adopt on the basis of this merged evidence base, by applying objective Bayesian theory. Third, determine which judgements are appropriate given these degrees of belief by applying a decision-theoretic account of rational judgement formation.

Keywords: Judgement aggregation, belief merging, belief revision, objective Bayesianism, decision theory, maximum entropy

1 Introduction

In recent years formal methods originally developed by computer scientists and logicians to deal with inconsistencies in databases have been increasingly applied to problems in social epistemology. Thus, the formal theory of belief revision and merging, developed since the 1980s to help maintain consistency when revising and merging sets of propositions, has been applied by Meyer *et al*. [18] and Choi [3] to the problem of aggregating preferences, applied by Gauwin *et al*. [7] to the problem of judgement deliberation and conciliation, and applied by Pigozzi [22] to the problem of judgement aggregation.

It is this latter application, to judgement aggregation, that is the focus of this article. In Section 2, I shall introduce the problem of judgement aggregation and some of the difficulties encountered in trying to solve this problem. Section 3 will introduce the theory of belief revision and merging and Pigozzi's application of this theory to judgement aggregation. I argue, in Section 4, that judgements should not be merged directly; rather, one should merge the evidence on which the judgements are based. Given this merged epistemic background, probability theory and decision theory can be used to derive an appropriate set of judgements—the resulting judgements should be viewed as the aggregate of the judgements of the original individuals (Section 5).

2 Judgement aggregation

In many cases it is important for a collection of agents with their own individual judgements to come to some agreed set of judgements as a group. The question of how best to do this is the problem of judgement aggregation. (In discussions of judgement aggregation it is normally assumed that a judgement is a proposition or sentence and that by providing a judgement one *endorses* the relevant proposition.)

Simple majority voting on each of the propositions in an agenda is unsatisfactory as a judgement aggregation procedure for the following reason: while each individual may have a consistent set of judgements, the aggregated set of judgements may be inconsistent. This is known as the *discursive dilemma* or the *doctrinal paradox* [13]. Table 1 displays a simple example, taken from Dietrich and

Vol. 19 No. 3,
Published online 2 May 2008 doi:10.1093/logcom/exn011

TABLE 1. An example of the discursive dilemma

	p	$p \to q$	q
A	true	true	true
B	true	false	false
C	false	true	false
Majority	true	true	false

List [6]. Here three individuals have consistent sets of judgements (e.g. A judges that $p, p \to q$ and q), but majority voting yields an inconsistent set $\{p, p \to q, \neg q\}$ of aggregated judgements.

Problems with majority voting have led to a quest to find a better judgement aggregation procedure. However, a number of impossibility theorems limit the options available—see, e.g. List and Pettit [16, 17]; Dietrich and List [6]; Pauly and van Hees [21]. In particular Dietrich and List [6] show that if the agenda is sufficiently rich and if (formalizations of) the following conditions hold, then the only aggregation functions are *dictatorships*, i.e. those that simply take the judgements of a particular individual as the judgements of the group:

Universal Domain: the domain of the aggregation function is the set of all possible profiles of consistent and complete individual judgement sets,

Collective Rationality: the aggregation function generates consistent and complete collective judgement sets,

Independence: the aggregated judgement on each proposition depends only on individual judgements on that proposition,

Unanimity: if each individual judges a particular proposition true then so will the aggregate.

(As Dietrich and List [6] show, this result generalises Arrow's celebrated impossibility theorem concerning preference aggregation in social choice theory.)

The question thus arises as to whether any reasonable aggregation function remains. Dictatorship aside, is there any reasonable judgement aggregation rule? If so, which of the above conditions does it violate? We shall see next that Pigozzi [22] advocates a judgement aggregation rule that violates Independence. I will argue in Section 5 that judgement aggregation requires a rule that may violate *all* of these conditions. This rule will appeal to a richer epistemology than that invoked by the standard examples of the judgement aggregation literature; in particular it will distinguish an agent's judgements and the evidence on which those judgements are based. I shall argue that only by considering such evidence can one aggregate judgements properly. Thus problems like that of Table 1 turn out to be underspecified.

3 Belief revision and merging

The question of how to revise and merge sets of propositions is addressed by the theory of belief revision and merging. *Belief revision* seeks to say how one should revise a deductively closed set of sentences T in the light of new information, a consistent sentence θ. The theory imposes the following desiderata for a belief revision operator $\star$ (known as the AGM postulates, after Alchourrón *et al.* [1]):

$\star$0: $T \star \theta$ is consistent,

$\star$1: $T \star \theta$ is deductively closed,

⋆2: $\theta \in T \star \theta$,

⋆3: if $T+\theta$, the deductive closure of $T \cup \theta$, is consistent, then $T \star \theta = T + \theta$,

⋆4: if θ and φ are logically equivalent then $T \star \theta = T \star \varphi$,

⋆5: if $T \star \theta + \varphi$ is consistent then $T \star \theta + \varphi = T \star \theta \wedge \varphi$.

Alchourrón *et al*. [1] also put forward postulates for a *belief contraction* operator $-$, where $T-\theta$ is the result of taking θ away from deductively closed T. Revision and contraction turn out to be related by the Levi identity $T \star \theta = (T-\theta)+\theta$ and by the Harper identity $T-\theta = (T \star \theta) \cap T$ (for θ not a tautology). The theory has been further extended to give an account of *belief update*: while for belief revision θ and T are understood as referring to the same situation, for belief update T is taken to refer to the past and θ to the present. See Katsuno and Mendelzon [10]; del Val and Shoham [4].

A third extension of the theory is of special interest here. *Belief merging* seeks to provide an account of how one should combine sets $T_1, \ldots, T_{n_T}$ of sentences. Konieczny and Pino Pérez [11] put forward the following postulates for a merging operator Δ acting on a multiset $T = \{T_1, \ldots, T_{n_T}\}$, where each T_i is assumed consistent, the notation T_i may be used to refer to the set of sentences or the conjunction of its members, and where $T \sqcup U$ is the multiset $\{T_1, \ldots, T_{n_T}, U_1, \ldots, U_{n_U}\}$:

Δ0: ΔT is consistent,

Δ1: if $\bigwedge T$ is consistent then $\Delta T = \bigwedge T$,

Δ2: if T_i and U_i are logically equivalent for $i = 1, \ldots, n_T = n_U$, then ΔT and ΔU are logically equivalent,

Δ3: if $\bigwedge T \wedge \bigwedge U$ is inconsistent then $\Delta(T \sqcup U)$ does not logically imply T,

Δ4: $\Delta T \wedge \Delta U$ logically implies $\Delta(T \sqcup U)$,

Δ5: if $\Delta T \wedge \Delta U$ is consistent then $\Delta(T \sqcup U)$ logically implies $\Delta T \wedge \Delta U$.

Konieczny and Pino Pérez [12] generalize this framework to provide postulates for a merging operator Δ_ι with *integrity constraints* ι that the merged set has to satisfy. Belief merging is related to belief revision by the identity $T \star \iota = \Delta_\iota\{T\}$, where T is a set of sentences and where merging is subject to a deductive closure condition ([15], Section 5; [12], Section 5).

Pigozzi [22] applies belief merging to the problem of judgement aggregation. The idea is this: if n agents have propositional judgements $T = \{T_1, \ldots, T_n\}$, then one should take $\Delta_\iota T$ as the aggregate of these judgements, where Δ_ι is a particular belief merging operator motivated by majority voting considerations. In the example of the discursive dilemma of Table 1, no integrity constraints ι are required, since postulate Δ0 ensures that the aggregate set is consistent. Pigozzi [22] shows that a judicious choice of integrity constraints avoids any paradoxical outcome in another example of the discursive dilemma. However, in all such cases the aggregation procedure results in several equally optimal merged sets—the selection of one of these sets as the aggregate is hence rather arbitrary. Arbitrariness is a fundamental problem when using belief merging techniques to directly aggregate judgement sets, since there is rarely enough information in the judgement sets themselves to allow one to decide which way to resolve an inconsistency in the majority view. However, I shall suggest next that arbitrariness can be controlled by merging at the level of evidence rather than at the level of judgement.

TABLE 2. The epistemic states of A and B

	Grants	Believes	Judges
A	$\neg l$	$\neg r$	$\neg r, \neg c$
B	h	$\emptyset$	r, c

4 Merging evidence

When devising a method of aggregating judgements, one might have one of two goals in mind. One natural goal is to try to find a *fair* procedure—a procedure that treats all agents equally, where all agents' judgements contribute in the same way in determining the aggregate. Since the recent literature on judgement aggregation has stemmed from work in social choice theory—e.g. work on trying to devise fair voting systems—this goal has hitherto been paramount. Unfortunately, the discursive dilemma and the impossibility results suggest that this goal may be unattainable.[1]

Another natural goal is to try to find a judgement aggregation procedure that yields the *right judgements*. Arguably, if this is the primary goal then it matters little whether different agents' judgements play an equal role in determining the aggregate. Thus the impossibility results, whose assumptions are motivated by fairness considerations, need not apply. This latter goal will be our goal here.[2]

Consider the following judgement aggregation problem. A patient has received some treatment for breast cancer and two consultants, A and B, need to make the following judgements. First, they need to judge whether or not the patient's cancer will recur, r, in order to inform the patient of her prospects. Second, they need to judge whether or not chemotherapy is required, c (if recurrence is unlikely then aggressive treatments such as chemotherapy, which have harsh side effects, may not be justified). In this kind of example, when aggregating judgements it is far more important that the collective judgements be the right judgements than that they be fair to the individual agents' opinions: the patient's health is more important than the egos of the consultants.

Consultant A has clinical evidence: the tumour has not spread to lymph nodes, $\neg l$. This leads her to believe that the cancer will not recur, $\neg r$. She is not so sure about whether chemotherapy is required, but on balance she judges that it is not. Thus A judges $\neg r, \neg c$.

Consultant B has molecular evidence—the presence of certain hormone receptors in the patient. This evidence indicates a less favourable prognosis. B is not convinced enough to say that he *believes* r and c, but since a judgement is required, he judges r, c.

Table 2 represents the epistemic states of the two consultants. The epistemological picture here is that the agents have three grades of propositions: evidence, beliefs and judgements. The agent's *evidence base* includes all that she *takes for granted* in the context of the task in hand. As well as the results of observation, it includes any theoretical and background assumptions that are not currently open to question. The evidence base may equally be called her *epistemic background* or *data*. (As to whether the agent is rational to take these propositions for granted will depend on the goals of her

[1] Some of those working on judgement aggregation argue that they are trying to *model* judgement aggregation scenarios such as expert panels, rather than directly trying to find a fair aggregation procedure. This may be so; however, the situations being modelled are themselves typically constructed with fairness in mind. Thus fairness remains an indirect goal.

[2] Bovens and Rabinowicz [2]; Hartmann *et al*. [8] pursue this goal for instance, assessing standard judgement aggregation procedures from the point of view of correctness. This is a normative goal: the question is how one *should* aggregate judgements to yield the right judgements, not how people actually *do* aggregate judgements. Thus there is no need for an answer to this question to be psychologically realistic. There is a need for any answer to be computationally feasible, but it will be beyond the scope of this paper to address computational concerns in any detail.

current enquiry as well as the provenance of the propositions themselves.) The agent's beliefs are not taken for granted, but are credible enough to be construed as *accepted*, and hopefully rationally so. As can be seen from the above example, judgements often need to be made when there may be little to decide one way or the other: they are *speculated* propositions—again, one hopes, rationally so. Thus, the items of evidence are practically certain, beliefs have high credence, while a judgement may be much more weakly believed.

If we are to apply belief revision and merging to the problem of judgement aggregation, the question arises as to what exactly we merge. It is natural to try to merge the judgements themselves—this is the strategy of Pigozzi [22], for instance. But then we come up against the problem of arbitrariness. In our example, this amounts to merging $\{\neg r, \neg c\}$ with $\{r, c\}$. This can be done by arbitrarily choosing one agent's judgements over the other, or by taking the disjunction of the judgements, $(\neg r \wedge \neg c) \vee (r \wedge c)$, as the aggregate. This last strategy does not avoid arbitrariness: the aggregated judgements do not contain judgements on the propositions we want, so to make judgements about recurrence and chemotherapy on the basis of the disjunctive aggregate requires arbitrary choice.

Perhaps the theory of belief revision and merging should be applied at the level of belief—this is ostensibly what the theory concerns. I think that this is a mistake, for two reasons. First, the theory of belief revision and merging is more suited to revising and merging evidence than belief. This can be seen as follows. Belief revision is related to nonmonotonic logic: there is a correspondence between revision operators $\star$ and consistency-preserving *rational consequence relations* $\mid\!\sim$ (i.e. $\mid\!\sim$ satisfying the Gabbay-Makinson conditions for nonmonotonic consequence together with the rule $\theta \not\models \neg\tau \Rightarrow \theta \not\mid\!\sim \neg\tau$ for tautology τ) via the equivalence $\theta \mid\!\sim \varphi \Leftrightarrow$ either θ is inconsistent or else $\varphi \in T \star \theta$. In turn, nonmonotonic logic is related to probability: there is a correspondence between non-trivial rational consequence relations $\mid\!\sim$ and ε-probability functions p (probability functions whose range is $[0,1]$ augmented with infinitesimal numbers) via the equivalence $\theta \mid\!\sim \varphi \Leftrightarrow$ either $p(\theta) = 0$ or else $p(\neg\varphi|\theta)$ is infinitesimal or zero.[3] Thus there is a correspondence between revision operators $\star$ and the ε-probability functions p that only award probability zero to contradictions: $T \star \theta$ consists of those sentences whose probability conditional on θ is infinitesimally close to 1. Similarly, there is a correspondence between merging operators Δ_ι (under a deductive closure condition) and such ε-probability functions: $\Delta_\iota\{T\}$ consists of those sentences whose probability conditional on ι is infinitesimally close to 1. Now under a Bayesian account, ε-probability is construed as a measure of degree of belief. Thus $T \star \theta$ is the set of sentences that are believed to degree infinitesimally close to 1, given θ; $\Delta_\iota\{T\}$ is the set of sentences that are believed to degree infinitesimally close to 1, given ι. Under the Bayesian account, degrees of belief are understood as indicative of betting intentions; an infinitesimal difference between two degrees of belief has no consequence in terms of betting; hence 'infinitesimally close to 1' corresponds to practical certainty. Thus $\Delta_\iota\{T\}$ is the set of sentences that are practically certain given ι. But then $\Delta_\iota\{T\}$ cannot be interpreted as an agent's qualitative beliefs, since practical certainty is an absurdly high standard for qualitative belief. A sentence might be deemed *believed* simpliciter if it is strongly believed—say to degree 0.8 or 0.9 depending on circumstances—but it is far too stringent to insist that one's beliefs must be practically certain. Rather, it is the agent's evidence that is practically certain. Therefore belief revision and merging should be used to revise and merge evidence rather than belief.[4]

---

[3] See Paris [20] for a nice introduction to these connections.

[4] The distinction between evidence and belief tends to be passed over in the literature; Pigozzi [22] and many of those working in the area of belief revision and merging use 'knowledge' and 'belief' interchangeably. Note that Kyburg Jr. *et al.* ([14], Section 6) suggest some modifications to the belief revision framework if it is to used to handle propositions that are *rationally accepted*, rather than known. The view of rational acceptance that they consider is that provided by the theory of *evidential probability*.

There is a second reason why merging agents' beliefs is mistaken. In our example, such a merging operation would yield $\neg r$ as the merged belief set. Presumably, then, judging $\neg r$ ought to be reasonable given these merged beliefs. But in fact such a judgement may be very unreasonable: it may be that the presence of both the patient's symptoms together makes recurrence quite likely. Merging beliefs will thus ignore any interactions between the pieces of evidence that give rise to the beliefs.

These two considerations motivate merging the agents' *evidence* instead of their beliefs or judgements. On the one hand, this is the correct domain of application of the theory of belief revision and merging. On the other, one must merge the *reasons* for the agents' beliefs and judgements, rather than the beliefs or judgements themselves, if one is to achieve the goal of making the right judgements. The right judgement is that which, considering all available evidence, is most appropriate given the uses to which the judgement will be put. Only by merging the evidence itself can one consider all available evidence; merging beliefs or judgements directly will ignore interactions amongst items of evidence.

Merging at the level of evidence has a further advantage: it reduces arbitrariness. We often have to make judgements in the face of extensive uncertainty—there may be little to decide between a proposition and its negation, yet we must make a call one way or the other. Different agents are likely to make different calls, yielding mutually inconsistent judgements. If these judgements are to be merged, this leads to arbitrariness on the part of the merging operator. Beliefs are of a higher epistemological grade than judgements: to count as a belief in the qualitative sense, a proposition must be strongly believed in a quantitative sense. Fewer propositions are likely to make this grade and hence there are likely to be fewer inconsistencies between different agents' beliefs than between different agents' judgements. Thus merging beliefs involves less arbitrariness than merging judgements. But evidence is of a higher grade yet: an item of evidence is taken for granted and so practically certain. Of course, items of evidence may be false; hence inconsistencies may arise between different agents' data bases. But such inconsistencies will be considerably rarer than those that arise between belief bases, let alone judgements. Therefore, merging evidence avoids much of the arbitrariness that besets attempts to merge judgements directly.

I have argued thus far that a merging operator should operate on agents' evidence bases rather than their beliefs or judgements. If merging is to be applied to judgement aggregation, we then require some way of determining a set of judgements from a merged evidence base. These judgements can be viewed as the aggregate. Is it this task to which we shall turn next.

5 From merged evidence to judgements

Judgement is essentially a decision problem. For each proposition in an agenda, one must decide whether to endorse that proposition or its negation. One makes the *right* judgements to the extent that the judgements are most appropriate considering the uses that will be made of them. Typically, as in the case of our example, one would like one's judgements to be true, but it is sufficient that they are determined in the right way from one's epistemic background: a false judgement may be the right judgement if it is most plausible on the basis of the limited evidence available; conversely a true judgement may be the wrong judgement if it is unlikely given what is known.

Since judgement is a decision problem, decision theory can be applied to judgement formation. A simple decision-theoretic account might proceed like this. Let $u(x|y)$ represent the utility of judging x given that y is the case. For example, Table 3 gives a table of utilities for judging cancer recurrence: the judgement will be used to inform the patient of her prospects, so judging the true outcome has

TABLE 3. A utility matrix for judging recurrence

		Judgement	
		r	$\neg r$
Case	r	1	−1
	$\neg r$	−3	1

TABLE 4. A utility matrix for judging chemotherapy

		Judgement	
		c	$\neg c$
Case	r	5	−10
	$\neg r$	−4	1

positive utility while judging the false outcome has negative utility and is particularly bad when falsely judging recurrence because it leads to needless anxiety. Then decide on the judgement that maximises expected utility, $EU(x) = \sum_y u(x|y)p(y)$ where $p(y)$ is the probability of y. In our example, judge recurrence if $EU(r) > EU(\neg r)$, i.e. if $p(r) > 2/3$. Judgements may depend on propositions other than the judged proposition. Table 4 gives a utility matrix for judging chemotherapy: judging in favour of chemotherapy is a good thing if the patient's cancer would otherwise recur but quite bad if it would not recur; judging against chemotherapy is very bad given recurrence but quite good given non-recurrence. Decide in favour of chemotherapy if $EU(c) > EU(\neg c)$, i.e. if $p(r) > 1/4$.

We see then that if we are to apply this decision-theoretic approach to judgement, we need to determine probabilities, e.g. $p(r)$.

Objective Bayesianism can be used to determine the required probabilities. According to objective Bayesian epistemology, probabilities are to be interpreted as an agent's degrees of belief, and it is the agent's evidence which determines the degrees of belief that she should adopt [27]. Evidence determines degrees of belief in two ways. First, an agent's degrees of belief should be *calibrated* with her evidence: if she grants a, she should set $p(a) = 1$; if her evidence base determines a suitable frequency for a, she should set $p(a)$ to that frequency. (e.g. if she knows just that the frequency of recurrence in patients with breast cancer is 0.4, then she should set $p(r) = 0.4$). Second, her degrees of belief should otherwise be as *equivocal* as possible: if empirical calibration does not fully determine $p(a)$, e.g. $p(a) \in [0.3, 0.4]$, then she should set $p(a)$ to the value that most equivocates between a and $\neg a$, $p(a) = 0.4$ in this case. (She should equivocate because if she were to believe propositions to a greater or lesser amount than warranted by evidence, she would open herself up to unjustified risks: a middling value for $p(r)$ might lead the agent to collect further evidence, while a more extreme value might lead to unjustified chemotherapy or an unjustified failure to administer chemotherapy—see [23, 26] on this point.)

Assuming for simplicity that an agent's language is a finite propositional language $\mathcal{L} = \{a_1, \ldots, a_n\}$, objective Bayesian epistemology can be explicated by the following principles:

Probability: The agent's degrees of belief can be represented by a probability function p over $\mathcal{L}$.

Calibration: If evidence determines that empirical probability (frequency) lies in some set $\mathbb{P}$ of probability functions then $p \in [\mathbb{P}]$, where $[\mathbb{P}]$ is the smallest closed convex set of probability functions containing $\mathbb{P}$.

Equivocation: p is the most equivocal probability function in $[\mathbb{P}]$, where the degree to which a probability function equivocates is measured by its entropy $-\sum_{\omega} p(\omega) \log p(\omega)$ where ω ranges over the atomic states $\pm a_1 \wedge \cdots \wedge \pm a_n$ of $\mathcal{L}$.[5]

It turns out that p is uniquely determined by the agent's evidence.[6]

We have, now, all the tools we need for a normative account of judgement aggregation. First, use a merging operator to merge the agents' evidence bases. This merged evidence base can be thought of as the evidence base of a hypothetical agent M. Objective Bayesianism can then be used to determine the degrees of belief that M should adopt, given this merged evidence. Finally, decision theory can be used to determine the judgements that M should make given these degrees of belief. These judgements can be viewed as the aggregate of the individual agents' judgements, although they are a function of the individual agents' evidence bases rather than of their judgement sets. (Note that it is *not* assumed that the original agents determine their own judgements via objective Bayesianism and decision theory. Indeed, the procedure outlined here does not depend on the original agents' own judgement procedures at all. All that is assumed is that these agents have some evidence and that this evidence can be made explicit in order for merging to take place. After merging, it is the hypothetical agent who conforms to the norms of objective Bayesian epistemology and decision theory.)

In our example the items of evidence $\neg l$ and h are consistent. Hence, by merging postulate $\Delta 1$, their merger is $\{\neg l, h\}$. To make it more interesting, we shall assume that the two consultants have some common evidence gleaned from hospital data, namely that the frequency of recurrence given no lymph cancer is 0.2, and that the frequency of recurrence given the presence of the relevant hormone receptor is 0.6. Thus, calibration yields the constraints $p(r|\neg l) = 0.2, p(r|h) = 0.6$. We shall also suppose that the consultants know that r is not a cause of l or h. (As explained in Williamson [24], in the presence of such causal knowledge entropy must be maximised sequentially: first maximise entropy to find the probability distribution over l and h, then maximise entropy again to find the probability distribution over r, given the previously determined probabilities of l and h.) Then maximising entropy yields $p(r) = p(r|\neg l \wedge h) = 2/5$. Now $p(r) < 2/3$, so M should judge $\neg r$. Further, $p(r) > 1/4$ so M should judge c. The epistemic state of the merged agent is depicted in Table 5; rather than M having qualitative beliefs she has a quantitative belief—she believes r to degree 0.4. Clearly, the aggregate judgement set $\{\neg r, c\}$ could not be produced by merging the agents' judgements, nor could it be produced by merging their beliefs. To find the right judgements we need to merge evidence.

6 Concluding remarks

If one is concerned with finding the right judgements, one should not apply belief merging directly to agents' judgement sets. Instead, one should apply belief merging to the agents' evidence bases and

[5] It should be emphasized that Equivocation is entirely subsidiary to Calibration and that while Calibration is motivated largely by evidential concerns, Equivocation is motivated by pragmatic considerations. It is not evidence that motivates the choice of the most equivocal function in $[\mathbb{P}]$ since all the evidence is 'used up' in determining $[\mathbb{P}]$ in the first place. All the functions in $[\mathbb{P}]$ are compatible with the evidence, and the maximum entropy function is chosen because it is most cautious on average with respect to risky decisions ([26], Section 8). It might be objected that an appeal to risk smuggles decision theoretic considerations into probability, but in response it should be pointed out that (i) the whole notion of degree of belief is intertwined with pragmatic concerns—the justification of Probability typically appeals to betting considerations (the Dutch book argument), and Calibration is only desirable inasmuch as one wants one's bets to be successful, and (ii) these appeals to pragmatic considerations are independent of any particular decision theory—whatever decision theory one advocates, one's degrees of belief ought to satisfy the above three principles.

[6] Note that p may vary as the agent's language changes, and this fact has been cited as grounds for concern. In Williamson [24, Chapter 12], I argue that this relativity to language is unobjectionable: an agent's language expresses implicit evidence about her domain, and an agent's degrees of belief *should* change as her evidence base changes.

TABLE 5. The epistemic states of A, B, and the merged agent M

	Grants	Believes	Judges
A	$\neg l$	$\neg r$	$\neg r, \neg c$
B	h	$\emptyset$	r, c
M	$h, \neg l$	$r^{0.4}$	$\neg r, c$

consider a single hypothetical agent with this merged evidence base. One can determine appropriate degrees of belief for this hypothetical agent using objective Bayesian theory and then determine appropriate judgements using decision theory. These judgements can be viewed as the aggregate of the original agents' judgement sets.

This judgement aggregation framework appeals to an epistemology of evidence, rational degree of belief and rational judgement, rather than the epistemology of evidence, qualitative belief and judgement outlined in Section 4. It integrates logic (merging evidence), probability theory (determining degrees of belief) and decision theory (determining judgements). While objective Bayesianism maps evidence to degrees of belief and decision theory maps degrees of belief and utilities to judgements, the logical stage—merging—is thus far underspecified: no particular merging operator has been advocated for merging evidence. Clearly, a lot will depend on choice of merging operator, and some remarks are in order. Merging operators are often motivated by majority voting considerations: if two agents advocate a and only one agent advocates $\neg a$ then a is normally taken to be the result of the merger. This strategy is rather dubious when it comes to merging *evidence*, since given that the agents disagree about a it seems unreasonable to take a for granted. On the other hand, we do have some information about a, and it does seem reasonable that the agent should confer higher degree of belief on a than on $\neg a$.[7] One can model this kind of situation as follows. Let variable A take possible assignments $a, \neg a$. For each agent i who has either a or $\neg a$ in her evidence base, construct a variable A_i with possible assignments a_i (signifying that i grants a), and $\neg a_i$ (signifying that i grants $\neg a$). Take A to be a cause of each A_i, as in Figure 1. Suppose we have some *reliability threshold* $\tau \in [1/2, 1]$ such that for all $i = 1, \dots, k, p(a_i|a) \geq \tau$ and $p(\neg a_i|\neg a) \geq \tau$. Let $a_i^{\varepsilon_i}$ signify the assignment that A_i actually takes (a_i if i grants a, $\neg a_i$ if i grants $\neg a$) and let m be the *majority* granting a, i.e. the number of agents granting a minus the number granting $\neg a$. Then we can find the objective Bayesian recommendation by maximising entropy:[8]

$$p(a|a_1^{\varepsilon_1} \cdots a_k^{\varepsilon_k}) = \frac{\tau^m}{\tau^m + (1-\tau)^m}.$$

This function is portrayed in Figure 2. We see then that if we have a reliability threshold τ, we have what we need to define a merging operator. The characterization and properties of the ensuing merging operator are questions for further research.

[7]Some merging operators in the literature are not based on majority voting considerations and in this example would not include a in the merger. In particular, *consistency-based merging operators* look for maximal consistent subsets of the union of the evidence bases—see, e.g. Delgrande and Schaub [5]. But these merging operators just ignore the evidence in favour of a, which itself seems unreasonable.

[8]The maximum entropy function is represented by the Bayesian network with the graph of Figure 1 and conditional distributions $p(a) = 1/2, p(a_i|a) = \tau, p(a_i|\neg a) = 1 - \tau$. Such a network is called an *objective Bayesian net*—see, e.g. Williamson [25] for further discussion of the use of objective Bayesian nets to handle threshold probabilities.

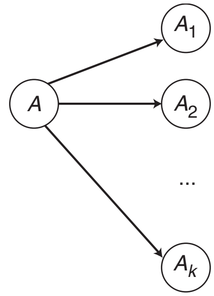

FIGURE 1. A causes $A_1, \ldots, A_k$.

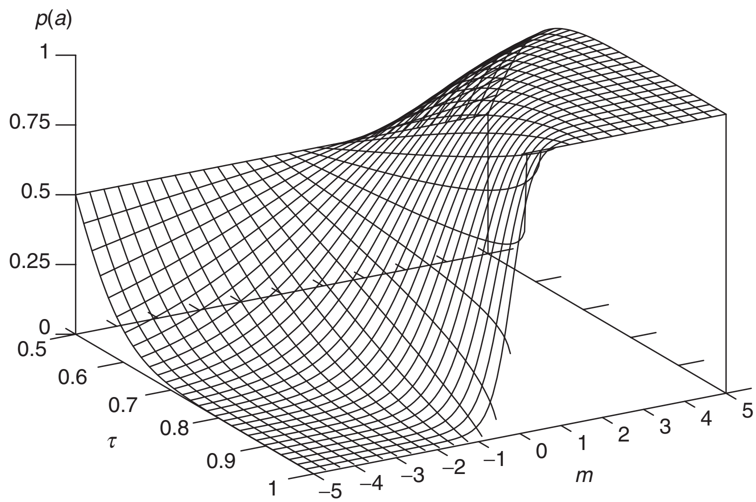

FIGURE 2. The degree to which the merged agent should believe a.

It may seem that the judgement aggregation procedure advocated in this article suffers because it ignores the expertise of the two consultants in our example: their judgements are aggregated by disregarding their judgements and merging their evidence instead, so the agents play no role in the process. But this latter conclusion does not follow. While this procedure does disregard the agents' actual judgements, their expertise in forming those judgements need not be ignored since it can be put to good use in the formulation of decision problem. Deciding what to judge on the basis of merged evidence requires the calculation of utilities and here the consultants' expertise will be crucial. Moreover, the aggregation procedure requires that the agents' evidence bases be merged, and this evidence will go beyond a simple list of the patients' symptoms—it will encapsulate a large part of their expertise. Thus, the agents' expertise feeds in to the first and last steps of the aggregation procedure.

There is an interesting question as to how utilities should be determined for the decision-theoretic component of the judgement aggregation procedure. Perhaps the simplest approach involves aggregating the utilities of the individual agents. Since utilities are numbers rather than propositions, one can simply average the utilities of the respective agents. However, there are several potential problems with this approach. First, the individual agents may have no utilities—they need not have used formal decision theory to come up with their judgements. Second, there may be no guarantee that any requirements of the decision theory (e.g. transitivity of preferences) remain satisfied by the aggregated utilities [9, 19]. Third, just as some agents may be better than others at latching onto the right judgements, some may be better than others at determining appropriate utilities, in which case treating the agents symmetrically may not yield the right utilities. Fourth, the required utilities depend to a certain extent on the uses of the aggregated judgements (e.g. palliative versus curative care of the patient in the cancer example), and these uses may vary considerably from the intentions of the original agents. These considerations motivate an alternative strategy for determining utilities that is analogous to the objective Bayesian procedure for determining degrees of belief. The available evidence, which may include evidence of the original agents' utilities and intentions, imposes constraints on an appropriate utility function for the decision-theoretic component of the aggregation problem (this is the analogue of the Calibration principle of Section 5). The decision theory will also impose certain constraints (the Probability principle). One may then choose a utility function, from all those that satisfy the constraints, that is as equivocal as possible (Equivocation). Exactly how these principles are to be fleshed out remains a question for future research.

The normative goal sets the approach of this article apart from previous work in judgement aggregation, which tends to be motivated by considerations of fairness. In particular, the impossibility results discussed in Section 2 do not get off the ground because the assumptions they make need not hold. Universal Domain fails because the individual judgement sets are not even in the domain of the aggregation function. (One can salvage Universal Domain by construing 'judgements' liberally to include evidence as well as judgements in the literal sense, but there seems little to be gained by this since the whole approach outlined here is predicated upon a distinction between evidence and judgement.) As to whether Collective Rationality holds will depend on the particular decision theory employed; consistency may seem desirable but it is not too hard to envisage very complex agendas where some more limited paraconsistency is all that is required. Independence fails because the aggregate judgements depend on the agents' evidence bases rather than their judgements. Unanimity fails because each agent's evidence may motivate a judgement a while the collective evidence indicates $\neg a$ (as happens with the lottery paradox, for instance).[9]

Returning to the discursive dilemma of Table 1, we see that the problem is under-specified. There is no obvious solution to the paradox because there is no information about the evidence on which the judgements are based. Only by considering the total evidence available can we determine the right judgements.

Acknowledgements

This research was carried out as a part of the academic network *Probabilistic logic and probabilistic networks*, funded by The Leverhulme Trust. I am very grateful to Gabriella Pigozzi, Jan Sprenger, Gregory Wheeler and an anonymous referee for helpful comments.

[9] Suppose for example that all agents grant the same item of evidence in favour of a while each agent has her own, slightly weaker, piece of evidence against a, and that these latter items of evidence are all independent. Then while in each agent's case the common favourable evidence outweighs the particular agent's unfavourable evidence, the aggregate of the unfavourable evidence may outweigh the aggregate of the favourable evidence (which is just a single item). Note that it is not assumed that each individual agent is rational, so this scenario may also occur when all agents have the same item of evidence against a and each misreads the evidence, taking it to be evidence in favour of a.

References

[1] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: partial meet functions for contraction and revision. *Journal of Symbolic Logic*, **50**, 510–530, 1985.

[2] L. Bovens and W. Rabinowicz. Democratic answers to complex questions-—an epistemic perspective. *Synthese*, **150**, 131–153, 2006.

[3] S. Choi. Review of 'Bayesian nets and causality'. *Mind*, **115**, 502–506, 2006.

[4] A. del Val and Y. Shoham. A unified view of belief revision and update. *Journal of Logic and Computation*, **4**, 797–810, 1994.

[5] J. P. Delgrande and T. Schaub. A consistency-based framework for merging knowledge bases. *Journal of Applied Logic*, **5**, 459–477, 2007.

[6] F. Dietrich and C. List. Arrow's theorem in judgment aggregation. *Social Choice and Welfare*, **29**, 19–33, 2007.

[7] O. Gauwin, S. Konieczny, and P. Marquis. Conciliation and consensus in iterated belief merging. In *Proceedings of the 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pp. 514–526, Springer, Berlin, 2005.

[8] S. Hartmann, G. Pigozzi, and J. Sprenger. Reliable methods of judgment aggregation. *Technical Report* 3593. Philsci Archive, Pittsburgh, 2007.

[9] M. Hild, R. Jeffrey, and M. Risse. Preference aggregation after harsanyi. In *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*, Fleurbaey, M., Weymark, J. A., and Salles, M., eds, Cambridge University Press, Cambridge, 2008.

[10] H. Katsuno and A. O. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pp. 387–394, Morgan Kaufmann, San Mateo, California, 1991.

[11] S. Konieczny and R. Pino Pérez. On the logic of merging. In *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning*, pp. 488–498, Morgan Kaufmann, San Francisco, California, 1998.

[12] S. Konieczny and R. Pino Pérez. Merging information under constraints: a logical framework. *Journal of Logic and Computation*, **12**, 773–808, 2002.

[13] L. Kornhauser and L. G. Sager. Unpacking the court. *Yale Law Journal*, **96**, 82–117, 1986.

[14] H. E. Kyburg Jr., C. M. Teng, and G. Wheeler. Conditionals and consequences. *Journal of Applied Logic*, **5**, 638–650, 2007.

[15] J. Lin and A. Mendelzon. Merging databases under constraints. *International Journal of Cooperative Information Systems*, **7**, 55–76, 1998.

[16] C. List and P. Pettit. Aggregating sets of judgments. An impossibility result. *Economics and Philosophy*, **18**, 89–110, 2002.

[17] C. List, and P. Pettit. Aggregating sets of judgments. Two impossibility results compared. *Synthese*, **140**, 207–235, 2004.

[18] T. Meyer, A. Ghose, and S. Chopra. Social choice, merging and elections. In *Proceedings of the 6th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Benferhat and Besnard, eds, pp. 466–477. Vol. 2143 of *Lecture Notes in Artificial Intelligence*, Springer, 2001.

[19] P. Mongin. The paradox of the Bayesian experts. In *Foundations of Bayesianism*, Corfield D. and Williamson, J., eds, pp. 309–338. Kluwer, Dordrecht, 2001.

[20] J. B. Paris. *Lecture Notes on Nonmonotonic Logic*. Department of Mathematics, University of Manchester. Available at www.maths.man.ac.uk/∼jeff/lecture-notes/mt4181.ps., 2003.

[21] M. Pauly and M. van Hees. Logical constraints on judgement aggregation. *Journal of Philosophical Logic*, **35**, 568–585, 2006.

[22] G. Pigozzi. Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. *Synthese*, **152**, 285–298, 2006.

[23] F. Russo and J. Williamson. Interpreting probability in causal models for cancer. In *Causality and Probability in the Sciences, Texts in Philosophy*, Russo, F. and Williamson, J., eds, pp. 217–241. College Publications, London, 2007.

[24] J. Williamson. *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford University Press, Oxford, 2005.

[25] J. Williamson. Inductive influence. *British Journal for the Philosophy of Science*, **58**, 689–708, 2007a.

[26] J. Williamson. Motivating objective Bayesianism: from empirical constraints to objective probabilities. In *Probability and Inference: Essays in Honour of Henry E. Kyburg Jr.*, Harper, W. L. and Wheeler, G. R., eds, pp. 151–179. College Publications, London, 2007b.

[27] J. Williamson. Philosophies of probability: objective Bayesianism and its challenges. In *Handbook of the philosophy of mathematics*. Irvine, A., ed. Vol. 4, *Handbook of the Philosophy of Science*, Elsevier, Amsterdam. 2008.

Received 1 December 2007