Judgment Aggregation as Maximization of Epistemic and Social Utility

Szymon Klarman

Abstract

We restate the problem of judgment aggregation and approach it using the decision-theoretic framework applied by I. Levi to modeling acts of rational acceptance in science. We propose a method of aggregation built on the concepts of epistemic and social utility of accepting a collective judgment, which accounts for such parameters as the factual truth of the propositions, reliability of agents, information content (completeness) of possible collective judgments and the level of agreement between the agents. We argue that the expected utility of accepting a judgment depends on the degree to which all those objectives are satisfied and that groups of rational agents aim at maximizing it while solving judgment aggregation problems.

1 Introduction

The problem of judgment aggregation, concerning the issue of building up collective judgments by groups of agents, lies at the intersection of social choice theory and epistemology. On the one hand it deals with the question of what is a good procedure by which individual viewpoints should contribute to the collective one — a central matter of concern of social choice theory. On the other, since the objects between which the choice is made are *judgments*, any proposed method of aggregation has to be verified against logical, or even broader, epistemological criteria, guaranteeing soundness of the outcomes. For that reason the problem falls also in the scope of some essentially philosophical considerations, of which the most important regards the rationality of acceptance of propositions in general.

Due to the twofold character of the problem an aggregation procedure is expected to be socially fair and epistemologically reliable at the same time. It might be the case, as C. List and P. Pettit show in [12], that these two requirements cannot be satisfied simultaneously and one has to prioritize between them. Nonetheless, even if this state of affairs is inevitable, it is still worth asking where exactly the trade-off takes place and whether it is possible to capture it formally and gain substantial control over it.

An interesting framework for such an analysis has been proposed by I. Levi [9, 10], who applied the notion and a simple measure of *epistemic utility* of accepting a proposition in order to deal with the problem of underdetermination of inductive inferences in science.[1] As a result, induction has been formally reinterpreted in terms of the trade-off between rival epistemic goals that drive scientific inquiry. This approach has been recently recalled in the context of judgment aggregation by Levi himself [11] and also by D. Fallis in [3]. In this paper we propose and discuss a possible aggregation procedure based on those grounds, which explicitly parameterizes the trade-offs underlying the problem of aggregating judgments, accounting for the relevant epistemological and social-theoretical aspects.

The reminder of the paper is organized as follows. First, we outline the formal frames for the judgment aggregation problem together with the discursive dilemma. Then we point out a correspondence between the dilemma and the lottery paradox and recall the work of I. Levi used for circumventing the latter. In Section 3 we introduce the utilitarian judgment aggregation model, followed by sample aggregation results and the discussion of the method.

[1] The idea was first brought about by R. Jeffrey [6] and C. G. Hempel [5], however the approach was not formally elaborated and introduced as a self-standing proposal until Levi's work.

2 Discursive Dilemma and Lottery Paradox

The *judgment aggregation problem* can be shortly characterized as follows (cf. [12, 1]). Let $\mathcal{A} = \{1, \ldots, n\}$ be a set of n agents and Φ an *agenda*, i.e. a set of well-formed propositional formulas. It is assumed that for every $\varphi \in \Phi$ there is also $\neg\varphi \in \Phi$. Each agent has an *individual set of judgments* with respect to Φ. A judgment regarding a proposition is understood here as an unequivocal act of *acceptance* or *rejection* of that proposition. An individual set of judgments of agent i can be represented as a subset $\Phi_i \subseteq \Phi$ of exactly those propositions that are accepted by i. All the propositions that are not in Φ_i are the ones that i rejects. Further, we require every individual set of judgments Φ_i to satisfy three rationality constraints: 1) *completeness*: for every $\varphi \in \Phi$, either $\varphi \in \Phi_i$ or $\neg\varphi \in \Phi_i$; 2) *consistency*: there is no φ such that $\varphi \in \Phi_i$ and $\neg\varphi \in \Phi_i$; 3) *deductive closure* with respect to the agenda: $\mathrm{Cn}(\Phi_i) \cap \Phi \subseteq \Phi_i$.

A *collective judgment* is a subset $\Psi \subseteq \Phi$ such that Ψ also satisfies the rationality constraints and (in some sense) it is *a response* to the individual judgments of all and only the agents from $\mathcal{A}$. The desired properties of responsiveness are characterized by three requirements imposed on the judgment aggregation function (JAF), i.e. a function that, given a profile of all individual judgments $\{\Phi_i\}_{i\in\mathcal{A}}$, should uniquely determine the collective judgment. These are: 1) *universal domain*: a JAF should yield a collective judgment for any possible profile of individual judgments; 2) *anonymity*: the individual judgments of agents should have equal importance in determining the outcome; 3) *independence*:[2] for every proposition φ and any two profiles of individual judgments, if for every $i \in \mathcal{A}$ it holds that $\varphi \in \Phi_i$ iff $\varphi \in \Phi'_i$ then $\varphi \in \mathrm{JAF}(\{\Phi_i\}_{i\in\mathcal{A}})$ iff $\varphi \in \mathrm{JAF}(\{\Phi'_i\}_{i\in\mathcal{A}})$.

A seemingly natural choice for a JAF is the *propositionwise majority voting rule*, which advises accepting collectively all and only those propositions from Φ that are accepted individually by the majority of agents. As it turns out, however, if Φ contains at least two different propositions and their conjunction, the majority procedure may lead to obtaining inconsistent collective judgments. This effect is known as the *discursive dilemma* or *doctrinal paradox* and has recently attracted much attention in the field of computational social choice, e.g. [12, 1, 2, 14]. The dilemma, as List and Pettit [12] have proved, is unavoidable under the six previously listed constraints. Nevertheless, there is a number of ways of escaping it by relaxing some of the requirements. In the same paper the authors present a comprehensive discussion of different solutions, of which we shall mention two.

The first one avoids the paradox by dropping the completeness requirement on the collective judgment. Under specific provisions a group might simply suspended its judgment on particular propositions. This way the outcome is deductively weaker and does not provide a full solution to the given aggregation task, but inconsistency does not occur. The other method resolves the paradox by relaxing the independence assumption through conditioning acceptance of certain propositions on the judgment on some others. Two typically invoked strategies include the premise- and the conclusion-driven aggregation, according to which the priority is given to those propositions that serve respectively as the premises or the conclusion in the agenda. The judgment on the remaining ones is then suitably adjusted in order to avoid inconsistency.

The basic shortcoming of both approaches is their inability of resolving the paradox in an unambiguous and nonarbitrary manner. For instance, for the same input the premise- and the conclusion-driven procedures can easily yield incompatible outcomes, while none of them is more intuitive than the other.

Recently, an argument-driven approach, inspired by an operation of merging belief bases in AI [7, 8], has been also investigated as a strategy of relaxing independence [14, 15]. The

[2] The *independence* condition is weaker than the original *systematicity* requirement used in [12] and so, as slightly less controversial, tends to replace the former constraint in more recent literature, e.g. [1].

method is considerably better justified and well-behaved than the aforementioned procedures. It employs a simple distance measure of individual judgments from possible consistent collective judgments, and chooses the one (or those) that minimize it. Thus the preference is given not to the premises or the conclusion but to the argument as a whole. Our proposal rests upon similar principles, but instead builds on certain results from philosophy of science and significantly generalizes the approach.

The problem of judgment aggregation, interpreted as a specific case of propositional acceptance, can be related to a similar question of how to aggregate logically connected hypotheses about the world into a coherent body of scientific knowledge. Interestingly, a direct analogue of the discursive dilemma occurs also in that context [11, 2] and has been a recurrent subject of debates and analyses in the XXth century philosophy of science, e.g. [5, 9]. Essentially, the *lottery paradox* shows that propositionwise acceptance over logically connected statements fails in general in yielding a consistent set of formulas. According to I. Levi [9], the problem stems from a too narrow perspective on acceptance in science. Scientific inquiry is a goal-oriented activity, driven by (at least) two rival goals of a purely epistemic nature: obtaining *true* and highly *informative* statements. Any plausible conclusion of an inference may satisfy them to different extent, and so be relatively better or worse with respect to other candidate answers. If the scientist is able to assess the degree to which the goals are met by particular conclusions and possesses some probabilistic knowledge about the possible states of the world, then he should evaluate the *expected epistemic utility* of the conclusions and simply accept the one that *maximizes* it. The *cognitive decision model*, proposed by Levi, aims to give a formal account of such a mechanism of acceptance.

The aggregation method presented in the following section is an extension of Levi's model. It borrows its all basic assumptions and the chief part of its formal apparatus. The novel share involves defining a measure of the social agreement, a method of generating probability distributions from profiles of individual judgments, and restating the judgment aggregation problem as a task of satisfying (often rival) epistemic and social goals. Some limited experimental results are presented in Section 4.

3 The Utilitarian Model of Judgment Aggregation

A group of agents striving to construct a collective judgment on some issue wants the judgment to have certain good properties. Namely, it has to reflect individual judgments of the group's members and moreover it has to be a rational statement by itself. The former requirement, to which discussions on judgment aggregation have been mainly confined, involves applying some measure of responsiveness of the collective judgment to individual beliefs of agents. The latter rests upon the assumption that a collective judgment quite often conveys a particular claim about the world, which can be evaluated with respect to the epistemic objectives mentioned above.

To start with, the utilitarian model of aggregation requires recognition of the set of all possible states of the world associated, for instance, with the logical models of the agenda. Consider $\Phi = \{p, \neg p, q, \neg q, r, \neg r\}$ and the background knowledge[3] $b = \{p \wedge q \leftrightarrow r\}$. Under the given constraints there are four distinct truth valuation functions over the formulas from Φ, corresponding to four complete, consistent and deductively closed sets of judgments on Φ: $\{p, q, r\}, \{\neg p, q, \neg r\}, \{p, \neg q, \neg r\}, \{\neg p, \neg q, \neg r\}$. We shall interpret the collection of these functions $\mathcal{M}_{\Phi,b} = \{v_1, ..., v_4\}$ as the set of all possible and mutually exclusive ways the world might be with respect to Φ and b. By convention we posit that judgment Ψ is satisfied by state v_i, i.e. $v_i \vDash \Psi$, whenever v_i makes all formulas in Ψ true.

[3]The assumption of background knowledge is used only to abbreviate the notation, and can be dropped at any time by replacing atomic formulas in Φ by respective compound ones, defined as in b.

Further, it is necessary to specify the *answer set*, i.e. the set of all collective judgments whose acceptance could present certain value to the voting group. Typically, the judgments satisfying the rationality constraints should be permitted in the first order as possible outcomes of the aggregation procedure. Also, in many cases, a group might want to relax some of the requirements — predominantly completeness — to extend the range of possible outcomes. The following partial, though still consistent and deductively closed judgments on Φ could be often seen as interesting in a variety of contexts: $\{\{p\}, \{q\}, \{\neg p, \neg r\}, \{\neg q, \neg r\}, \{\neg r\}\}$. For instance, if r is a legal verdict based upon two premises p and q, it could be enough for the jury to agree on the negation of only one of the premises, since this alone allows for determining the conclusion $\neg r$. Even if the jury cannot decide on the truth value of the second premise, it can still make a final judgment and justify it.

Partial judgments can be evaluated with respect to the amount of information they convey. By analogy to the cognitive decision model, this can be defined in terms of the proportion of possible states of the world that are excluded by the given judgment:[4]

$$\text{cont}(\Psi) = \frac{|v_i \in \mathcal{M}_{\Phi,b} : v_i \nvDash \Psi|}{|\mathcal{M}_{\Phi,b}|}$$

In the example under discussion we obtain:

$$\text{cont}(\{p, q, r\}) = 0.75 \qquad \text{cont}(\{p\}) = 0.5 \qquad \text{cont}(\{\neg r\}) = 0.25$$

Measure cont, based on purely structural properties of the problem, can be suitably parameterized to account also for additional pragmatic value that a judgment offers to the group. Figure 1 shows three sample ways of evaluating information relative to the proportion of possible states that are excluded by the judgment. Notice, that whereas the linear plot represents the standard measure, the two others may be adopted by a group revealing a weaker (top) or a stronger (bottom) bias for completeness of the outcome.

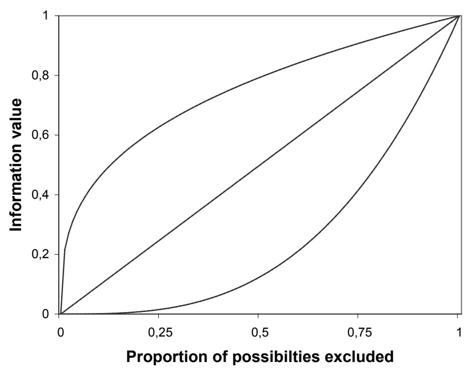

Figure 1: The value of information with respect to information content.

Another possibility of extending the range of permissible judgments regards dropping the consistency requirement. According to D. Fallis [3] inconsistent judgments do not have to be always worthless. The information content of an inconsistent judgment should by definition of cont be equal 1, which neatly harmonizes with the property of inferential explosion. However, a group might want to assign other values, according to its particular understanding of informativeness of an inconsistent statement. Also judgments that are not deductively closed can be formally incorporated into the model, if only the group can assess

[4]Other interesting measures of information content are discussed in e.g. D. Fallis [4] and P. Maher [13].

their information value in a meaningful sense. Finally, $\emptyset$ with $\text{cont}(\emptyset) = 0$ is worth including in the answer set as an option for suspending the judgment, which often might be the only reasonable decision.

Let then $\mathcal{CJ} = \{\Psi_1, \ldots, \Psi_m\}$ be a set of possible collective judgments, preselected and evaluated with respect to information content by the group. Following the model of cognitive decision we will assess the value of truth — the second epistemic objective traded off against the information content — relatively to the state of the world, using a simple binary measure:

$$\text{T}(\Psi, v_i) = \begin{cases} 1 \text{ iff } v_i \models \Psi \\ 0 \text{ iff } v_i \not\models \Psi \end{cases}$$

The measure takes value 1 in state v_i, whenever a judgment is true in it, and 0 otherwise. The overall utility of accepting a collective judgment, provided that state v_i holds, is given by the following function, which assigns numerical values to potential judgments according to the degree to which they satisfy the two epistemic goals:

$$u_\varepsilon(\Psi, v_i) = \alpha \ \text{cont}(\Psi) + (1 - \alpha) \ \text{T}(\Psi, v_i)$$

Coefficient $\alpha \in [0, 1]$ serves here as an epistemic preference indicator, with 0 standing for the full preference for truth, and 1 for the information content of a judgment.

In order to complete the framework one has to employ knowledge about the probability of the possible states. Although in a typical judgment aggregation problem no such information is explicitly provided, there is a way of inducing a probability distribution using the profile of individual judgments. Notice, that every individual judgment, as satisfying the rationality constraints, corresponds to exactly one model from $\mathcal{M}_{\Phi,b}$. Assuming that agents are characterized by a certain degree of reliability, it is justified to regard their judgments as good indicators of the truth of the states (cf. [15]). Formally, this information can determine the probability distribution over $\mathcal{M}_{\Phi,b}$ by means of Bayes' Theorem.

First, a uniform prior distribution is posed over the states, i.e. $P(v_i) = \frac{1}{|\mathcal{M}_{\Phi,b}|}$ for every $v_i \in \mathcal{M}_{\Phi,b}$. Let r, such that $0.5 > r > 1$, be the degree of reliability of agents, meaning that given state v_i is true, the probability $P(\Phi|v_i)$ that an agent makes a correct judgment ($v_i \models \Phi$) is r, whereas the probability of the opposite is $1 - r$. Starting with the prior probabilities every individual judgment is used to update the distribution over $\mathcal{M}_{\Phi,b}$, so that posterior probability of a state, given an individual judgment Φ, equals to:

$$P(v_i|\Phi) = \frac{P(\Phi|v_i)P(v_i)}{\sum_j P(\Phi|v_j)P(v_j)}$$

It can be shown that the resulting distribution after n consecutive updates, i.e. for n individual judgments, is given by the equation:

$$P^\star(v_i) = \frac{r^{n_i}(1-r)^{n-n_i}}{\sum_i r^{n_i}(1-r)^{n-n_i}}$$

where $\sum_i n_i = n$ and each n_i is the number of supporters of state v_i. A distribution generated in this way has two interesting properties. First, it reflects the beliefs of agents to the degree that these beliefs can be deemed correct, thus complying to the principles of the Bayesian epistemology. Judgments of unreliable agents (for r approaching 0.5) do not have a strong influence on the distribution, as opposed to those of highly reliable judges (for r being close to 1), which are given maximal weight. Second, for a finite number of agents none of the possible states is ever completely excluded, i.e. $P^\star(v_i) > 0$ for all $v_i \in \mathcal{M}_{\Phi,b}$, hence the fundamental uncertainty, which often motivates the need for judgment aggregation, is not eliminated totally.

Up to this point the model roughly mimics the design of Levi's framework, giving account of the epistemic aspects of acceptance that, as we argue, play essential role in the problem of aggregating judgments. The central claim of this paper, however, is that an aggregation procedure is driven by the endeavor to maximize both epistemic and social benefits following from acceptance of a collective judgment. Epistemically good judgments are not fully satisfactory unless they also justly reflect opinions of the agents involved. Intuitively, socially the fairest collective judgment is the one selected by propositionwise majoritarian rule. Nevertheless, as already pointed out, this strategy cannot be reliably employed. Instead, we should allow the choice to be made only from the answer set $\mathcal{CJ}$. For this purpose we will use a measure of social agreement interpreted as a form of distance of a judgment from the majoritarian choice. Let SA be a function whose values are assigned as follows:

- for any $\varphi \in \Phi$: $\mathrm{SA}(\varphi) = \frac{|\mathcal{A}_\varphi|}{|\mathcal{A}|}$, i.e. the social agreement on a proposition is the ratio of the number of agents who accept it to the total number of agents;
- for any $\Psi \in \mathcal{CJ}$: $\mathrm{SA}(\Psi) = \frac{1}{|\Psi|}\sum_{\varphi\in\Psi}\mathrm{SA}(\varphi)$, i.e. the social agreement on a collective judgment is equal to the arithmetic mean of the degrees of social agreement on propositions accepted in the judgment.

SA expresses what proportion of propositions from a judgment is on average accepted by an agent. We thus stay very close to the notion of Hamming distance proposed as the basis for the argument-driven approach to judgment aggregation [14].[5] The suspension of judgment can be assigned value 0, as it is in no way responsive to individual judgments.

SA, as considered independently from the rest of the model, shares interesting similarities with the afformentioned aggregation procedures that rest on the relaxation of the independence constraint. If we restrict $\mathcal{CJ}$ to contain all and only complete sets of judgments, including the inconsistent ones, then the judgment Ψ such that $\mathrm{SA}(\Psi) = \max_i\ \mathrm{SA}(\Psi_i)$ is equivalent to the choice of the propositionwise majority voting rule. This is due to the fact, that out of all complete judgments, the one with maximum value of SA is the one whose every accepted proposition has $\mathrm{SA}(\varphi) \geq 0.5$. If inconsistent judgments are left out, then the one that maximizes the degree of social agreement can be seen as socially the closest in $\mathcal{CJ}$ to the majoritarian choice. This resolves the dilemma of arbitrary selection between the premise- and the conclusion-driven strategy. The measure SA alone points at the judgment that violates the majoritarian vote to the smallest degree, no matter whether violation occurs on the side of premises or conclusions. Consider again the example with $\Phi = \{p, \neg p, q, \neg q, r, \neg r\}$ and $b = \{p \wedge q \leftrightarrow r\}$. If the profile of individual judgments is such that the following degrees of social agreement are assigned to the propositions:

$$\mathrm{SA}(p) = 0.8 \qquad \mathrm{SA}(q) = 0.7 \qquad \mathrm{SA}(r) = 0.4$$

then the socially best, complete and consistent judgment is $\{p, q, r\}$, since the value of r is the closest to 0.5, and so accepting r, even though it is against the majoritarian vote, brings the least disagreement amongst agents. The choice is therefore equivalent to the outcome of the premise-driven majoritarian procedure. In case of another distribution of agreement, for instance:

$$\mathrm{SA}(p) = 0.7 \qquad \mathrm{SA}(q) = 0.6 \qquad \mathrm{SA}(r) = 0.3$$

the judgment $\{p, \neg q, \neg r\}$ would be the most preferable, which is compatible with the conclusion-driven strategy.

Finally, the epistemic and social factors can be embraced within a single utility function defined for any $\Psi \in \mathcal{CJ}$ and $v_i \in \mathcal{M}_\Phi$ as:

$$\begin{aligned} u(\Psi, v_i) &= \beta \quad \underbrace{(\alpha\ \mathrm{cont}(\Psi) + (1-\alpha)\ \mathrm{T}(\Psi, v_i))} + (1-\beta)\ \mathrm{SA}(\Psi) \\ &= \beta \qquad\qquad\quad u_\varepsilon(\Psi, v_i) \qquad\qquad\quad + (1-\beta)\ \mathrm{SA}(\Psi) \end{aligned}$$

[5] In fact SA is a normalized form of the measure adopted by G. Pigozzi in [14].

While α controls the trade-off on the purely epistemic level, coefficient $\beta \in [0, 1]$ is supposed to reflect the upper-level preferences of a group, balancing the trade-off between the epistemic and social perspective. For β approaching 1, the epistemic criteria take over and an act of judgment aggregation becomes an act of rational acceptance, analogous to the typical cases modeled in Levi's framework. When it is close to 0, the social agreement measure becomes a decisive factor, rendering the procedure majoritarian, but still free of paradoxes. If the group is able to correctly diagnose its preferences and set the values of α and β accordingly, then the judgment maximizing the expected utility, i.e.:

$$\mathrm{EU}(\Psi) = \sum_{v_i \in \mathcal{M}_\Phi} P^\star(v_i) u(\Psi, v_i)$$

should be the one to be rationally accepted, as satisfying to the greatest extent the collective preferences of the group. For a complete picture, the procedure requires a tie-breaking rule. The prescription proposed by Levi [9], namely to accept the disjunction of all the answers with maximum expected utility, could be interpreted in this context as acceptance of the *common information* included in the judgments maximizing the expected utility. If we provisionally accept such a rule,[6] we can conclude the presentation with the utilitarian judgment aggregation function, formulated as follows:

$$\text{(UJAM)} \qquad \mathrm{JAF}(\{\Phi_i\}_{i \in \mathcal{A}}) = \bigcap \Psi \quad \text{such that} \quad \Psi \in \arg\max_{\Psi \in \mathcal{CJ}} \mathrm{EU}(\Psi)$$

⋆

Before concluding the presentation we will summarize the model's ingredients introduced in this section. Given a judgment aggregation problem specified by a set of agents $\mathcal{A}$, a set of propositional formulas Φ and a profile of individual judgments $\{\Phi_i\}_{i \in \mathcal{A}}$, where each judgment has to be a complete, consistent, and deductively closed subset of Φ, the group can define utilitarian aggregation model consisting of the following elements:

$\mathcal{CJ} = \{\Psi_1, ..., \Psi_m\} \subseteq 2^\Phi$	the answer set: a set of potential collective judgments preselected by the group as presenting certain value,
$\mathcal{M}_\Phi = \{v_1, ..., v_l\}$	the set of all possible, mutually exclusive states of the world (models) with respect to Φ,
$v_i \in \mathcal{M}_\Phi : \Phi \to \{0, 1\}$	a truth valuation function on the propositions from Φ,
$P^\star : \mathcal{M}_\Phi \to [0, 1]$	a probability function over possible states, whose values are derived from the profile of individual judgments,
$r \in (0.5, 1)$	the degree of reliability of individual judgments / agents,
$\alpha, \beta \in [0, 1]$	the coefficients controlling the information-truth and epistemic-social trade-offs,
$\mathrm{cont} : \mathcal{CJ} \to [0, 1]$	the valuation measure of information contained in the collective judgments,
$\mathrm{T} : \mathcal{CJ} \times \mathcal{M}_\Phi \to \{0, 1\}$	the valuation measure of truth with respect to the collective judgments and possible states of the world,
$\mathrm{SA} : \mathcal{CJ} \cup \Phi \to [0, 1]$	the measure of social agreement on the propositions and collective judgments.

[6] Clearly, the outcome of such a rule does not have to necessarily maximize the expected utility. Note also, that alternative interpretations of Levi's proposal, though less appealing, are possible. Yet other options of establishing a tie-breaking rule include typical decision-theoretic escape routes, for instance: 1) providing an a priori preference ranking over the goals, so that the judgment which satisfies the most preferred goal better is picked; 2) employing another decision criterion (e.g. Hurwicz, Laplace), which have chances to yield a unique outcome. Both solutions suffer, however, from the same arbitrariness as was previously pointed out in the standard aggregation methods.

4 Experimentation

In this section we will present sample aggregation results obtained with the utilitarian judgment aggregation model for set of propositions $\Phi = \{p, \neg p, q, \neg q, r, \neg r\}$ and background knowledge $b = \{p \wedge q \leftrightarrow r\}$. In the scenario there are four possible states of the world (models): $\mathcal{M}_{\Phi,b} = \{v_1, v_2, v_3, v_4\}$, defined by the following truth valuations:

$$\begin{array}{ll} v_1: & v_1(p) = 1,\ v_1(q) = 1,\ v_1(r) = 1 \\ v_2: & v_2(p) = 0,\ v_2(q) = 1,\ v_2(r) = 0 \\ v_3: & v_3(p) = 1,\ v_3(q) = 0,\ v_3(r) = 0 \\ v_4: & v_4(p) = 0,\ v_4(q) = 0,\ v_4(r) = 0 \end{array}$$

The value of information content for all collective judgments Ψ is assigned according to the standard measure cont(Ψ). The shaded rows in the tables below mark the collective judgments selected by the model.

Information content vs. truth: Tables 1 - 2. For a constant profile of individual judgments and a fixed value of the epistemic-social coefficient $\beta = 0.8$, we change the value of the information-truth coefficient α. For $\alpha = 0.7$ (Table 1) $\{p, \neg q, \neg r\}$ is chosen as the most informative judgment among those which are still sufficiently probable and agreeable. Decreasing the value of α results gradually in selection of less informative (more incomplete) collective judgments. For $\alpha = 0.5$ we get $\{\neg q, \neg r\}$, whereas for $\alpha = 0.1$, $\{\neg r\}$ (Table 2). When fixed to $\alpha = 0.01$ the aggregation becomes so truth-oriented that only the suspension of judgment is a plausible choice.

Information content vs. social agreement: Tables 3 - 4. For a constant profile of individual judgments we fix the value of the information-truth coefficient at $\alpha = 1$, meaning that truth is completely excluded from considerations. Shifting the value of the epistemic-social coefficient from $\beta = 0.5$ (Table 3) to $\beta = 0.3$ (Table 4), leads to the change in the accepted collective judgment from $\{p, \neg q, \neg r\}$ to more agreeable though incomplete $\{p\}$.

Truth vs. social agreement: Tables 5 - 6. For a constant profile of individual judgments we fix the value of the information-truth coefficient at $\alpha = 0$, thus discarding information content. Setting $\beta = 0.6$ (Table 5) we reveal a higher preference for a conclusion more likely to be true, which in this case is $\{\neg r\}$. Notice, that $\{\neg r\}$ has the same degree of agreement as $\{\neg q\}$ and even lower than $\{p\}$, but $\{\neg r\}$ is true in three out of four possible states with

Total Number of Agents:	10	Individual Judgments:	2	1	4	3	
Reliability of Agents:	0.55	Probability:	0.22	0.18	0.33	0.27	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.367	0.733	0.493	0.493	0.493	0.546
$\{\neg p,\ q, \neg r\}$	0.75	0.500	0.520	0.760	0.520	0.520	0.563
$\{p, \neg q, \neg r\}$	0.75	0.700	0.560	0.560	0.800	0.560	0.639
$\{\neg p, \neg q, \neg r\}$	0.75	0.633	0.547	0.547	0.547	0.787	0.611
$\{\neg p, \neg r\}$	0.50	0.600	0.400	0.640	0.400	0.640	0.508
$\{\neg q, \neg r\}$	0.50	0.750	0.430	0.430	0.670	0.670	0.574
$\{p\}$	0.50	0.600	0.640	0.400	0.640	0.400	0.532
$\{q\}$	0.50	0.300	0.580	0.580	0.340	0.340	0.436
$\{\neg r\}$	0.25	0.800	0.300	0.540	0.540	0.540	0.487
suspend	0	0	0.240	0.240	0.240	0.240	0.240

Degrees of Agreement on Atoms:		
	p	0.6
	q	0.3
	r	0.2

Trade-offs:	
Information vs. Truth	0.7
Epistemic vs. Social	0.8

Table 1.

Total Number of Agents:	10	Individual Judgments:	2	1	4	3	
Reliability of Agents:	0.55	Probability:	0.22	0.18	0.33	0.27	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.367	0.853	0.133	0.133	0.133	0.292
$\{\neg p,\ q, \neg r\}$	0.75	0.500	0.160	0.880	0.160	0.160	0.290
$\{p, \neg q, \neg r\}$	0.75	0.700	0.200	0.200	0.920	0.200	0.437
$\{\neg p, \neg q, \neg r\}$	0.75	0.633	0.187	0.187	0.187	0.907	0.381
$\{\neg p, \neg r\}$	0.50	0.600	0.160	0.880	0.160	0.880	0.484
$\{\neg q, \neg r\}$	0.50	0.750	0.190	0.190	0.910	0.910	0.621
$\{p\}$	0.50	0.600	0.880	0.160	0.880	0.160	0.556
$\{q\}$	0.50	0.300	0.820	0.820	0.100	0.100	0.389
$\{\neg r\}$	0.25	0.800	0.180	0.900	0.900	0.900	0.741
suspend	0	0	0.720	0.720	0.720	0.720	0.720

Degrees of Agreement on Atoms:		
	p	0.6
	q	0.3
	r	0.2

Trade-offs:	
Information vs. Truth	0.1
Epistemic vs. Social	0.8

Table 2.

Total Number of Agents:	10	Individual Judgments:	4	1	4	1	
Reliability of Agents:	0.6	Probability:	0.39	0.11	0.39	0.11	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.567	0.658	0.658	0.658	0.658	0.658
$\{\neg p,\ q, \neg r\}$	0.75	0.433	0.592	0.592	0.592	0.592	0.592
$\{p, \neg q, \neg r\}$	0.75	0.633	0.692	0.692	0.692	0.692	0.692
$\{\neg p, \neg q, \neg r\}$	0.75	0.433	0.592	0.592	0.592	0.592	0.592
$\{\neg p, \neg r\}$	0.50	0.400	0.450	0.450	0.450	0.450	0.450
$\{\neg q, \neg r\}$	0.50	0.550	0.525	0.525	0.525	0.525	0.525
$\{p\}$	0.50	0.800	0.650	0.650	0.650	0.650	0.650
$\{q\}$	0.50	0.500	0.500	0.500	0.500	0.500	0.500
$\{\neg r\}$	0.25	0.600	0.425	0.425	0.425	0.425	0.425
suspend	0	0	0.000	0.000	0.000	0.000	0.000

Degrees of Agreement on Atoms:		
	p	0.8
	q	0.5
	r	0.4

Trade-offs:	
Information vs. Truth	1
Epistemic vs. Social	0.5

Table 3.

Total Number of Agents:	10	Individual Judgments:	4	1	4	1	
Reliability of Agents:	0.6	Probability:	0.39	0.11	0.39	0.11	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.567	0.622	0.622	0.622	0.622	0.622
$\{\neg p,\ q, \neg r\}$	0.75	0.433	0.528	0.528	0.528	0.528	0.528
$\{p, \neg q, \neg r\}$	0.75	0.633	0.668	0.668	0.668	0.668	0.668
$\{\neg p, \neg q, \neg r\}$	0.75	0.433	0.528	0.528	0.528	0.528	0.528
$\{\neg p, \neg r\}$	0.50	0.400	0.430	0.430	0.430	0.430	0.430
$\{\neg q, \neg r\}$	0.50	0.550	0.535	0.535	0.535	0.535	0.535
$\{p\}$	0.50	0.800	0.710	0.710	0.710	0.710	0.710
$\{q\}$	0.50	0.500	0.500	0.500	0.500	0.500	0.500
$\{\neg r\}$	0.25	0.600	0.495	0.495	0.495	0.495	0.495
suspend	0	0	0.000	0.000	0.000	0.000	0.000

Degrees of Agreement on Atoms:		
	p	0.8
	q	0.5
	r	0.4

Trade-offs:	
Information vs. Truth	1
Epistemic vs. Social	0.3

Table 4.

almost uniform probability, and so it is more probable than the other two. When we fix $\beta = 0.4$ (Table 6) the accepted judgment is $\{p\}$ as more agreeable. The same outcome would be yielded for $\beta = 0.6$ were the agents more reliable, say for $r = 0.7$. In that case the probability distribution would be strongly influenced by individual judgments and the two states where $\{p\}$ is true would become much more probable than the others.

Total Number of Agents:	10	Individual Judgments:	4	0	4	2	
Reliability of Agents:	0.51	Probability:	0.26	0.23	0.26	0.24	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.533	0.813	0.213	0.213	0.213	0.372
$\{\neg p,\ q, \neg r\}$	0.75	0.400	0.160	0.760	0.160	0.160	0.295
$\{p, \neg q, \neg r\}$	0.75	0.667	0.267	0.267	0.867	0.267	0.426
$\{\neg p, \neg q, \neg r\}$	0.75	0.467	0.187	0.187	0.187	0.787	0.333
$\{\neg p, \neg r\}$	0.50	0.400	0.160	0.760	0.160	0.760	0.442
$\{\neg q, \neg r\}$	0.50	0.600	0.240	0.240	0.840	0.840	0.546
$\{p\}$	0.50	0.800	0.920	0.320	0.920	0.320	0.638
$\{q\}$	0.50	0.400	0.760	0.760	0.160	0.160	0.454
$\{\neg r\}$	0.25	0.600	0.240	0.840	0.840	0.840	0.681
suspend	0	0	0.600	0.600	0.600	0.600	0.600

Degrees of Agreement on Atoms:		
	p	0.8
	q	0.4
	r	0.4

Trade-offs:	
Information vs. Truth	0
Epistemic vs. Social	0.6

Table 5.

Total Number of Agents:	10	Individual Judgments:	4	0	4	2	
Reliability of Agents:	0.51	Probability:	0.26	0.23	0.26	0.24	
Collective Judgments:	Inform. Value:	Degree of Agreement:	v_1	v_2	v_3	v_4	Expected Utility:
$\{p,\ q,\ r\}$	0.75	0.533	0.720	0.320	0.320	0.320	0.426
$\{\neg p,\ q, \neg r\}$	0.75	0.400	0.240	0.640	0.240	0.240	0.330
$\{p, \neg q, \neg r\}$	0.75	0.667	0.400	0.400	0.800	0.400	0.506
$\{\neg p, \neg q, \neg r\}$	0.75	0.467	0.280	0.280	0.280	0.680	0.378
$\{\neg p, \neg r\}$	0.50	0.400	0.240	0.640	0.240	0.640	0.428
$\{\neg q, \neg r\}$	0.50	0.600	0.360	0.360	0.760	0.760	0.564
$\{p\}$	0.50	0.800	0.880	0.480	0.880	0.480	0.692
$\{q\}$	0.50	0.400	0.640	0.640	0.240	0.240	0.436
$\{\neg r\}$	0.25	0.600	0.360	0.760	0.760	0.760	0.654
suspend	0	0	0.400	0.400	0.400	0.400	0.400

Degrees of Agreement on Atoms:		
	p	0.8
	q	0.4
	r	0.4

Trade-offs:	
Information vs. Truth	0
Epistemic vs. Social	0.4

Table 6.

5 Conclusions and Discussion

The judgment aggregation model introduced in this paper aims at capturing the concept of the utility that an act of acceptance of a collective judgment can offer to a group of rational agents. For this purpose we have reinterpreted the problem of aggregation as a goal-oriented task with (possibly rival) goals of an epistemic and social character involved. Following Levi's cognitive decision model [9, 10], we have identified the epistemic goals as truth and information content and adopted the respective measures of the degrees to which these goals are satisfied by potential collective judgments. Further, we have defined a measure of the agreement on a judgment, which reflects the social objective of the procedure. Finally, an utility function defined as a weighted sum of the three measures has been proposed, together with an acceptance rule based on the criterion of maximizing the expected utility.

As a method of aggregation the model imposes two requirements on a group: designating the answer set and expressing numerically its preferences with respect to the goals involved. Assuming that this is done correctly, the collective judgment that is selected by the model is guaranteed to be the best one among all candidate judgments.

Clearly, the model might not satisfy the six requirements normally imposed on judgment aggregation function and the resulting collective judgments. However, any violation of these receives here a strong justification.

Completeness, consistency, deductive closure: As long as only complete, consistent and deductively closed collective judgments are considered, none of the requirements will be violated by the outcome (provided that tie does not occur). If, on the contrary, the answer set contains other judgments as well, then there is no good reason to defend the constraints. The doctrinal paradox obtains, therefore, a straightforward solution. If the group finds an inconsistent collective judgment undesirable, it should not designate it as a potential outcome; otherwise its occurrence is apparently not troublesome.

Universal domain, anonymity: For any profile of judgments the model designates a unique outcome (assuming the tie-breaking rule is employed) and assigns the same weight to every individual judgment in determining it.

Independence: This constraint can often be not satisfied by the model. Still, the central rationale behind the presented framework is a conviction that independence is a too strong requirement. Investigations into the discursive dilemma and the lottery paradox show that propositionwise acceptance rules — the type of rules enforced by the independence constraint — inevitably lead to inconsistencies when applied to sets of logically connected propositions. Following Levi, we argue that only statements considered in the context of the entire logical structure to which they belong can be rationally accepted or rejected. As a consequence a proposition obtaining the support of exactly the same agents can be once accepted, while rejected another time. This, however, happens only because the judgments containing this proposition are evaluated differently in the two cases, and therefore, in a broader perspective it is clearly a desired effect.

As a theoretical framework the model is universal enough to be amenable to many interesting adjustments and extensions, thus providing a large space of possible applications. Its most important advantage is that it offers a good control over the trade-offs involved in the aggregation task. Predominantly, it gives a clear formal account of how the responsiveness of the procedure — the fundamental social choice postulate — can be weighted against other expectations regarding collective judgments.

From the practical perspective a serious deficiency of the method, which should be a subject to further analysis, concerns its computational tractability. Depending on the structure of the problem it might be necessary to consider up to $2^{|\Phi|}$ possible states and the same number of candidate collective judgments. In any case, significant savings on the computational expense can be achieved at the cost of dropping one or two measures involved. Two scenarios seem especially appealing in this respect: 1) *dropping truth* (and consequently probabilities), which dramatically reduces the computational effort, while still allows for tracing the trade-off between social agreement and completeness of collective judgments; 2) *dropping epistemic utility*, which turns the method into a robust majoritarian procedure, selecting the permissible judgment that is closest to the majoritarian choice (an approach investigated by G. Pigozzi in [14]).

Acknowledgments. I would like to thank Ulle Endriss for his helpful feedback on the initial version of this paper, which has been written while studying the *Master of Logic* programme at the Institute for Logic, Language and Computation, University of Amsterdam. I am also grateful to three anonymous reviewers at COMSOC-2008 for valuable comments and suggestions.

References

[1] F. Dietrich, A generalised model of judgment aggregation. In *Social Choice and Welfare* 28, pp. 529-565, 2007.

[2] I. Douven, JW. Romeijn, The Discursive Dilemma as a Lottery Paradox. In U. Endriss, J. Lang, *Proceedings of the 1st International Workshop on Computational Social Choice (COMSOC-2006)*, ILLC University of Amsterdam, pp. 164-177, 2006.

[3] D. Fallis, Epistemic Value Theory and Judgment Aggregation. In *Episteme* 2(1), pp. 39-55, 2005.

[4] D. Fallis, Measures of Epistemic Utility and the Value of Experiments. In *Proceedings of the 17th Biennial Meeting of the Philosophy of Science Association*, Vancouver 2000.

[5] C. G. Hempel, Inductive Inconsistencies. In *Synthese* 12(4), pp. 439-469, 1960.

[6] R. Jeffrey, Valuation and Acceptance of Scientific Hypotheses. In *Philosophy of Science* 23(3), pp. 237-246, 1956.

[7] S. Konieczny, R. P. Pérez, Merging information under constraints: a logical framework, *Journal of Logic and Computation* 12(5), pp. 773-808, 2002.

[8] S. Konieczny, R. P. Pérez, Propositional belief base merging or how to merge beliefs/goals coming from several sources and some links with Social choice theory. In *European Journal of Operational Research* 160(3), pp. 785-802, 2005.

[9] I. Levi, *Gambling with Truth. An Essay on Induction and the Aims of Science*, MIT Press: Cambridge, 1967.

[10] I. Levi, Information and Inference. In *Synthese* 17(1), pp. 369-391, 1967.

[11] I. Levi, List and Pettit. In *Synthese* 140, pp. 237-242, 2004.

[12] C. List, P. Pettit, Aggregating Sets of Judgments: an Impossibility Result. In *Economics and Philosophy* 18, pp. 89-110, 2002.

[13] P. Maher, *Betting on Theories*, Cambridge University Press: Cambridge, 1996.

[14] G. Pigozzi, Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. In *Synthese* 152(2), pp. 285-298, 2006.

[15] G. Pigozzi, S. Hartmann, Merging Judgments and the Problem of Truth-Tracking. In U. Endriss, J. Lang, *Proceedings of the 1st International Workshop on Computational Social Choice (COMSOC-2006)*, ILLC University of Amsterdam, pp. 408-421, 2006.

Szymon Klarman
ILLC, University of Amsterdam
1018 TV Amsterdam, The Netherlands
Email: `sklarman@science.uva.nl`