Collective Decision-Making without Paradoxes: An Argument-Based Account

Gabriella Pigozzi*
King's College London
Department of Computer Science
26/29 Drury Lane
WC2B 5RL London, UK
email: pigozzi@dcs.kcl.ac.uk

Abstract

The combination of individual judgments on logically interconnected propositions into a collective decision on the same propositions is called judgment aggregation. Literature in social choice and political theory has claimed that judgment aggregation raises serious concerns. For example, consider a set of premises and a conclusion in which the latter is logically equivalent to the former. When majority voting is applied to some propositions (the premises) it may give a different outcome than majority voting applied to another set of propositions (the conclusion). This problem is known as the *doctrinal paradox*. The doctrinal paradox is a serious problem since it is not clear whether a collective outcome exists in these cases, and if it does, what it is like. Moreover, the two suggested escape-routes from the paradox - the so-called premise-based procedure and the conclusion-based procedure - are not, as I will show, satisfactory methods for group decision-making. In this paper I introduce a new aggregation procedure inspired by an operator defined in artificial intelligence in order to merge knowledge bases. The result is that we do not need to worry about paradoxical outcomes, since these arise only when inconsistent collective judgments are not ruled out from the set of possible solutions.

1 Introduction

Consider a group of teachers of the Elite High School contemplating the purchase of an espresso machine for their common room. After a careful discussion, they agree

*This research was supported by the Economic and Social Research Council, UK, Grant: PTA-026-27-0046.

to buy the machine only if they deem that they need an espresso machine *and* that the price is good. Let C be the proposition "to buy the espresso machine", A be "the machine is needed", and B be the proposition "the price is good". Each teacher casts her vote on C (the conclusion) depending on her judgments on A and B (the premises). Suppose that there are three teachers and that they vote as shown in the table below.

	A=Needed?	B=Good price?	C=Buy?	$(A \wedge B) \leftrightarrow C$
Voter 1	Yes	Yes	Yes	Yes
Voter 2	Yes	No	No	Yes
Voter 3	No	Yes	No	Yes
Majority	Yes	Yes	No	Yes

A majority voting on the propositions produces an inconsistent result, as $((A \wedge B) \leftrightarrow C)$ is unanimously accepted, and yet a majority accepts A, a majority accepts B (therefore, a majority accepts $(A \wedge B)$), but a majority rejects C. The above is an example of the so-called *doctrinal paradox* or discursive dilemma (Brennan 2001, Kornhauser 1992, Kornhauser and Sager 1986, 1993). In order to escape the paradox and obtain a final decision, two procedures have been suggested. One procedure is to let each member publicly vote on each premise and proceed to the purchase only if a majority of teachers believes that they need an espresso machine *and* that the price is good (this is called the premise-based procedure). The second procedure is that each member decides about A and B and then publicly casts her vote on the conclusion C only if she believes that an espresso machine is needed and the price is good (this is called the conclusion-based procedure). If the latter procedure is followed, the espresso machine will be purchased if and only if a majority of teachers voted for C.

Consider the following scenario. Voter 1 (the mathematics teacher, who is aware of the paradox) persuades her colleagues to vote on the premises, A and B, and consequently they buy the espresso machine. Voter 2 (the literature teacher, also known as Ms Stingy) immediately realizes that if they applied the majority rule to the conclusion, they would avoid spending money since a majority of 2/3 is against the purchase of the espresso machine. Voter 3 cannot understand why the two procedures give different results.

The moral of the story is that a collective outcome can only be defined if either the premise-based procedure or the conclusion-based procedure is adopted. But the two procedures may lead to contradictory results (one saying yes and the other saying no to C), depending on whether the majority is taken on the individual judgments of A and B, or whether the majority is calculated on the individual votes of C. This is obviously a serious problem, not only because it is unclear which of these two is the correct method (to let the individuals vote on A and B, or to let them vote on C), but also because it makes the collective outcome open to manipulation through the stipulation of a specific procedure.

Furthermore, the doctrinal paradox illustrates that collective inconsistencies can be obtained when consistent sets of propositions are aggregated. The idea is that individuals vote (in the form of yes/no) on *logically connected* propositions, and that different (and equally sensible) judgment aggregation procedures give contradictory collective outcomes, though each individual behaves perfectly rationally. As in the High School example, each person says yes to C only if she says yes both to A and to B (like voter 1). Voting against A (voter 3) or against B (voter 2) forces the voter to reject C. All three voters accept $((A \wedge B) \leftrightarrow C)$ and conform their judgment to the logical constraints between the premises (A and B) and the conclusion C. Since the conjunction of the premises is equivalent to the conclusion, one would expect that an aggregation procedure applied on the individual judgments on the premises would give the same result as the aggregation on the individual conclusions. But this is not the case, and it is precisely where the paradox arises. Moreover, List and Pettit (2002) show that the contradiction does not depend on the specific choice of aggregation procedure. Rather, they prove a general impossibility theorem such that there exists no aggregation function that satisfies a minimal set of conditions.[1]

The doctrinal paradox is a new problem in social choice theory. The classic theorem by Arrow proves that there is no procedure, given some minimal conditions, to aggregate individual *preferences* into a collective preference. The doctrinal paradox is a separate, but related[2] problem arising when individual *judgments* (in the form of propositional logic) are aggregated to form a collective judgment.[3]

I will argue in this paper that we need not to worry about the doctrinal paradox, provided that we fully recognize the logical relations between premises and conclusion, and that we understand the logic of aggregating sets of judgments. The framework I use to define and analyze the aggregation function profits from the collaboration of two (so far separate) research areas: knowledge fusion and collective decision. Knowledge fusion and group decision-making share a similar objective, viz. the definition of operators that produce collective knowledge from individual (and possibly conflicting) knowledge bases, and operators that produce a collective decision from individual decisions. Some of the fusion operators proposed in the literature were inspired by some of the voting procedures studied in social choice theory. I believe that also methods from knowledge fusion can be fruitfully imported into group decision-making where, like in judgment aggregation, the individual's decision does not necessarily come as a preference ordering, but has a propositional form.

The main goal of this paper is to reveal that the doctrinal paradox disappears as soon as we recognize that the propositions voted by the group majority do not necessarily define a unique, consistent and collective outcome. It is indeed often the case that a tie occurs on several consistent and collective outcomes. Hence, to

[1]We will go back to these conditions in section 5, to show that the new aggregation procedure proposed here relaxes one of the conditions of List and Pettit.

[2]The relations between preference aggregation and judgment aggregation have been investigated by List (2003) and by List and Pettit (2004).

[3]Beside the work of List and Pettit, other impossibility theorems on judgment aggregation have been proved by Dietrich (2004) and by Pauly and van Hees (2004).

make complex collective decisions without running into a paradox we have to allow multiple outcomes. An outcome in the new aggregation procedure is an argument, i.e. a consistent assignment to a conclusion *and* to the premises supporting that conclusion. The premise-based procedure and the conclusion-based procedure are therefore included in a unitary approach, which I will call *argument-based procedure*.

The structure of the paper is as follows. In section 2 it will be argued that the premise-based procedure and the conclusion-based procedure cannot be rationally justified and therefore do not solve the doctrinal paradox. In section 3 the general framework of the new model will be introduced, while the formal framework will be discussed in section 4. In section 5 it will be shown that the doctrinal paradox can be avoided when the argument-based procedure is used. This procedure focuses on the argument structure that connects (and logically constraints) the premises with the conclusion. Compared to the premise-based and the conclusion-based procedures, the new approach displays some additional interesting properties which allow, for example, the relaxation of various typically made conditions in judgment aggregation, like the completeness of the sets of judgments, leading to a more realistic account of group decision-making.

2 Neither Premise nor Conclusion-Based Procedure

As I sketched in the previous section, the premise-based and conclusion-based procedures have been used as escape-routes from the doctrinal paradox. In this section I will criticize these two approaches and claim that we need an aggregation method that provides a conclusion *together* with proper reasons to support it.

The premise-based and the conclusion-based procedures were both discussed by Pettit (Pettit 2001) within the context of *deliberative democracy* (for an introduction see Elster 1998).[4] The ideal of deliberative democracy requires that the whole

[4]For a discussion of these two procedures in an epistemic perspective, see Bovens and Rabinowicz

community is involved in the decision-making process: people have the right to ask for the reasons that supported a decision and to question them. In this respect, the premise-based procedure seems to be more appropriate. Yet there are at least two serious worries about this method, both of which concern whether it can guarantee citizen control of the decision.

The *first* problem is that when we apply the majority rule to each premise separately, we forget how the premises are related. The logical connectives are reintroduced only between the propositions that received the highest degree of support. In the example of the High School there is a majority for A and a majority for B and this is enough to infer that there is a majority for the conjunction $(A \wedge B)$. Nonetheless, only one voter (voter 1) cast her vote for both A and B.

The *second* problem is that it is not clear how one identifies the premise in a complex compound. Suppose that a group of 7 voters votes on $(P \wedge (Q \vee R) \leftrightarrow T)$ according to the following table:

	P	Q	R	$Q \vee R$	$P \wedge (Q \vee R)$	T
Voter 1	1	0	0	0	0	0
Voter 2	0	1	1	1	0	0
Voter 3	0	1	0	1	0	0
Voter 4	0	0	1	1	0	0
Voter 5	1	0	1	1	1	1
Voter 6	1	0	1	1	1	1
Voter 7	1	0	1	1	1	1
Majority	1	0	1	1	0	0

Majority voting gives two divergent results depending on what we take to be the premises. If the premise-based procedure is applied to the atomic propositions

2004 and List 2004.

P, Q, and R, then the conclusion T should be accepted. If the majority voting is applied to the conjuncts P and $(Q \vee R)$, then the conclusion T is rejected.[5] If we now compare these results with the outcome obtained by the majority rule applied on the conclusion T, we see that only the aggregation on the conjuncts P and $(Q \vee R)$ agrees with it and rejects T. Needless to say it is problematic to rationally justify whether we should take a premise to be an atomic proposition or not, and this renders the whole premise-based procedure open to manipulation.

Let us now turn to the conclusion-based procedure. Kornhauser and Sager (Kornhauser 1992, Kornhauser and Sager 1986, 1993) were the first to discover the doctrinal paradox and in their juridical example a three-member court has to decide whether a defendant is liable under a charge of breach of a contract. According to the legal doctrine, the defendant is liable (R) if and only if the contract was valid (P) and there was a breach (Q). This case is logically equivalent with our espresso machine example, and thus a possible doctrinal paradox.

But unlike the decision to buy or not to buy an espresso machine, a verdict in court is a public act. Not only is it a defendant's right to know the reasons for which she could be convicted, but also these reasons will guide future decisions - they are patterns for future verdicts. Here the final decision must be supported and justified by reasons. For cases similar to this a method like the conclusion-based procedure that defines the collective outcome without providing support for it, is not satisfactory. We need an approach that provides a collective decision *and* the reasons for that collective decision.

In section 5 I will show that the paradox disappears as soon as exclusively consistent sets of judgments are accepted as candidates for group decisions and, at the same time, we dispose of an aggregation procedure that authorizes multiple collective outcomes. The new method will select the most popular argument (and eventually more than one). By argument I mean a consistent judgment on a conclusion together

[5]Bovens and Rabinowicz 2004 referred to this problem as the instability of the premise-based procedure.

with the reasons to support that conclusion. The premise-based procedure and the conclusion-based procedure will be reconciled in a unique approach with no fear of paradoxical results.

3 The Framework

The aggregation procedure that will be used in this paper is inspired by a fusion operator defined in artificial intelligence. A fusion operator combines[6] various and possibly conflicting knowledge bases. A knowledge (or belief) base K_i is a finite set of propositional formulas representing the explicit beliefs of the individual i. The merged base is a set which *consistently* integrates parts of the knowledge from all the initial bases, and satisfies some additional conditions, such as *integrity constraints*.[7] The items of the resulting global knowledge base are identified by rules like the majority rule. It should therefore be clear that belief fusion and social choice theory share a similar objective, i.e. the definition of operators that produce collective knowledge from individual knowledge bases, and operators that produce a collective decision from individual preferences. Some researchers have found inspiration in belief fusion operators by examining voting procedures from social choice theory (Konieczny 1999, Lin and Mendelzon 1999). Here I want to show how methods from belief fusion can be fruitfully imported into social choice theory to tackle specific problems.

First, some terminology is in order. Following the literature on belief fusion, the words 'knowledge' and 'belief' will be used interchangeably in this paper.[8] This

[6]Konieczny (2000) refers the term 'combination' to the syntax-based fusion operators. We will instead use the verbs 'to combine', 'to merge' and 'to fuse' as synonyms.

[7]An integrity constraint (Kowalski 1978, Reiter 1988) is a sentence that has to be satisfied by the merged base. It is usually not required that the individual bases satisfy the integrity constraints, though I will assume this. Doing so, I will maintain the model as close as possible to the original formulation of the doctrinal paradox, where each voter consistently casts her vote on a set of propositions. We will see that integrity constraints play a crucial role in avoiding irrational sets of collective judgments.

[8]Also in the belief revision literature (a separate but related area to belief fusion, see Konieczny 1999) the use of the word 'knowledge' in a broader sense than in the epistemological literature - such that 'knowledge' also covers what is traditionally meant by 'belief' - is commonly accepted

is because a fusion operator can be interpreted as an operator merging knowledge or belief bases. Similarly, the aggregation of judgments can be interpreted as the aggregation of beliefs or as the aggregation of desires. In the first case, a member of the group votes yes for a proposition if she believes that that proposition is true. In the second case, she casts her vote for a proposition if she desires that proposition to be true. Furthermore, the elements of a knowledge base are not required to be true. This would simply be too big an idealization for practical applications in which knowledge is taken to be defeasible - for example when an expert system is defined by merging the knowledge of a group of human experts.

When merging individual belief bases, two cases can occur. If all individual bases are mutually consistent, then the collective outcome can easily be constructed: it is the union of all the individual bases. More interesting, however, is the case when the individual belief bases are in conflict with each other. There are two main approaches to this problem depending on whether all the individuals are treated equally or not.[9] For the purpose of the paper, fusion operators for individuals who have the same power to influence the final decision will be considered. I will, in fact, restrict myself to the classical doctrinal paradox where all voters have equal power.[10]

Some merging operators for knowledge bases of equally reliable sources have been proposed in Borgida and Imielinski (1984), Konieczny and Pino-Pérez (1998), Lin and Mendelzon (1999). These operators come in two types. On the one hand, there are majoritarian operators that minimize the level of total dissatisfaction (Konieczny 1999, Lin and Mendelzon 1996). On the other hand, there are egalitarian operators, which define rules to equally distribute the level of individual dissatisfaction among the group members (Konieczny 1999). For the sake of generality, the model presented here will make use of a utilitarian operator studied in Konieczny (1999). This fusion operator identifies the (possibly more than one) collective outcome with the model

(Gärdenfors 1988).

[9]In belief fusion, the latter corresponds to belief bases provided with different reliability values (see Benferhat *et al.* 1999, Cholvy 1994, Lin 1996, Maynard-Reid and Shoham 1998).

[10]For a study of expert rights in judgment aggregation see Dietrich and List 2004.

that is closest to the individual bases. The result in many cases is equivalent to the outcome we obtain by propositionwise majority voting. Yet, the paradox is avoided by adding domain-specific restrictions (integrity constraints) to the collective base. One of the key points in the literature of knowledge fusion is that merging consistent knowledge bases does not guarantee a consistent collective outcome. To overcome this problem, the integrity constraints are imposed on the final base as well as on the individual ones. This ensures that inconsistent models for the collective are ruled out from the set of possible group decisions. Finally, unlike the existing models for judgment aggregation, all possible consistent results for the collective outcome are explored and a ranking of them is defined. It will turn out that the discursive dilemma hides a tie among the possible outcomes.

4 The Formal Model

Let $N = \{1, 2, ..., n\}$ ($n \geq 2$) be a set of individuals. Let $\mathcal{L}$ be a finitary propositional language with the usual connectives of propositional logic ($\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$). $A, B, C, ...$ are propositions expressing knowledge, beliefs, or desires. The belief base K_i of an agent i is a consistent and complete[11] finite set of atomic propositions like A and B, and compound propositions like $\neg A$, $(A \wedge B)$, $((A \vee B) \leftrightarrow C)$, and so on.[12] K_i can also be represented as the conjunction of its propositional formulas.

IC is the belief base whose elements are the integrity constraints. Given a belief set $E = \{K_1, K_2, \ldots, K_n\}$ and IC, a fusion operator $\mathcal{F}$ is a function that assigns a belief base to E and IC. Let $\mathcal{F}_{IC}(E)$ be the resulting collective belief base from the IC fusion on E. Fusion operators come in two types: model-based and syntax-based. The latter (Baral *et al.* 1992, Konieczny 2000) are usually based on the selection of some consistent subsets of E. The bases K_i in E can be inconsistent and

[11] In the new argument-based procedure the completeness requirement can be relaxed. However, here K_i is assumed to be complete to keep the model as close as possible to the original formulation of the discursive dilemma.

[12] K_i is the correspondent in knowledge fusion of what, in judgment aggregation, is called a personal profile.

the result does not depend on the distribution of the wffs over the members of the group. Since in judgment aggregation we assume that each person forms a consistent set of judgments, and the distribution of wffs over the members is crucial for the definition of the collective outcome, here I will use the model-based merging operator introduced by Konieczny and Pino-Pérez (1999). The idea of a model-based fusion operator is that models of $\mathcal{F}_{IC}(E)$ are models of IC, which are preferred according to some criterion depending on E. Usually the preference information takes the form of a total pre-order on the interpretations induced by a notion of distance $d(w, E)$ between an interpretation w and the collective belief base E. Model-based merging operators have been discussed in Konieczny and Pino-Pérez (1999), Liberatore and Schaerf (2000), Lin and Mendelzon (1999), Revesz (1997).

According to Konieczny and Pino-Pérez, $\mathcal{F}$ is an IC merging operator if and only if it satisfies the following postulates:[13]

Definition 1 (IC merging operator $\mathcal{F}$)

1. $\mathcal{F}_{IC}(E) \vdash IC$

2. *If IC is consistent, then* $\mathcal{F}_{IC}(E)$ *is consistent*

3. *If* $\bigwedge E$ *is consistent with IC, then* $\mathcal{F}_{IC}(E) = \bigwedge E \wedge IC$

4. *If* $E_1 \leftrightarrow E_2$ *and* $IC_1 \leftrightarrow IC_2$, *then* $\mathcal{F}_{IC1}(E_1) \leftrightarrow \mathcal{F}_{IC2}(E_2)$

5. *If* $K \vdash IC$ *and* $K' \vdash IC$, *then* $\mathcal{F}_{IC}(K \cup K') \wedge K \nvdash \bot \Rightarrow \mathcal{F}_{IC}(K \cup K') \wedge K' \nvdash \bot$

6. $\mathcal{F}_{IC}(E_1) \wedge \mathcal{F}_{IC}(E_2) \vdash \mathcal{F}_{IC}(E_1 \cup E_2)$

[13] For a discussion on these postulates, see Konieczny and Pino-Pérez 1999. These postulates are inspired by the AGM postulates for belief revision (Alchourrón *et al.* 1985, Gärdenfors and Rott 1995). It has indeed been shown (Konieczny 1999) that belief fusion is a generalization of belief revision, and that an IC belief fusion operator can be defined from a belief revision operator.

7. *If* $\mathcal{F}_{IC}(E_1) \wedge \mathcal{F}_{IC}(E_2)$ *is consistent, then* $\mathcal{F}_{IC}(E_1 \cup E_2) \vdash \mathcal{F}_{IC}(E_1) \wedge \mathcal{F}_{IC}(E_2)$

8. $\mathcal{F}_{IC1}(E) \wedge IC_2 \vdash \mathcal{F}_{IC1 \wedge IC2}(E)$

9. *If* $\mathcal{F}_{IC1}(E) \wedge IC_2$ *is consistent, then* $\mathcal{F}_{IC1 \wedge IC2}(E) \vdash \mathcal{F}_{IC1}(E)$

Majority fusion operators satisfy the additional majority postulate (Konieczny and Pino-Pérez 1999):

(Maj) $\exists n \mathcal{F}_{IC}(E_1 \cup E_2^n) \vdash \mathcal{F}_{IC}(E_2)$

This postulate states that the collective outcome endorses a set of opinions if this is supported by a large part of the group.

Konieczny and Pino-Pérez prove that a family of pre-orders on models corresponds to the IC majority merging operator. To do so, they first define a syncretic assignment as follows:

Definition 2 (Syncretic assignment) *A syncretic assignment is a function mapping each belief set* E *to a total pre-order* $\leq_E$ *over interpretations such that for any belief sets* E_1, E_2 *and for any belief bases* K_1, K_2*:*

1. *If* $I \models E$ *and* $J \models E$*, then* $I \simeq_E J$
2. *If* $I \models E$ *and* $J \not\models E$*, then* $I <_E J$
3. *If* $E_1 \leftrightarrow E_2$*, then* $\leq_{E1} = \leq_{E2}$
4. $\forall I \models K_1 \; \exists J \models K_2$ *such that* $J \leq_{K1 \cup K2} I$
5. *If* $I \leq_{E1} J$ *and* $I \leq_{E2} J$*, then* $I \leq_{E1 \cup E2} J$
6. *If* $I <_{E1} J$ *and* $I \leq_{E2} J$*, then* $I <_{E1 \cup E2} J$

In particular, a majority syncretic assignment is a syncretic assignment that satisfies the following condition:

(SynMaj) If $I <_{E2} J$, then $I <_{E1 \cup E2^n} J$

With this the following theorem can be proved.

Theorem 1 *(Konieczny and Pino-Pérez 1999) An operator is a majority IC fusion operator if and only if there exists a majority syncretic assignment that maps each belief set E to a total pre-order $\leq_E$ such that:*

$$mod(\mathcal{F}_{IC}(E)) = min(mod(IC), \leq_E)$$

In order to define $\leq_E$ we need to determine a distance between interpretations (possible worlds), and an aggregation function that assigns a natural number to a finite sequence of natural numbers. Intuitively, the distance between interpretations measures the level of dissatisfaction of each person in the group given a possible collective outcome. If a belief state K_i is assigned a distance n from a possible outcome, and another belief state K_j is at distance m ($n<m$) from the same outcome, this means that the individual i will be 'happier' than the individual j if that outcome is selected as the collective decision. The selection is made after an aggregation function allocates a number to the sequence of the individual distances. As a matter of fact, this number measures the distance between an interpretation and a belief set. The aggregation function takes the sequence of dissatisfaction levels of the group members, given a certain belief set E, and associates with them a non-negative number (in our case, it will be the sum of the individual distances). This number is then used to define a total pre-order of all possible collective outcomes. The minimal distance identifies the final collective outcome, i.e. the belief base with the lowest total level of dissatisfaction among all possible models satisfying IC.

The distance between interpretations is a function that assigns a natural number to each pair of interpretations I and J such that $d(I,J) = d(J,I)$ and $d(I,J) = 0$ iff $I = J$. Since a belief base can have more than one interpretation that makes it true, the distance between an interpretation I and a belief base K is:

$$d(I,K) = min_{J \models K} d(I,J)$$

In the following, the Dalal's distance (Dalal 1988a, 1988b) will be used as a distance between interpretations. The Dalal's distance is based on the numbers of propositional letters on which two interpretations I and J differ. For example, the distance between $I = (1, 0, 0, 1)$ and $J = (0, 1, 0, 1)$ is $d(I, J) = 2$. There are several reasons to choose the Dalal's distance. From a logical perspective, it is sensible to choose a distance that was originally proposed in the framework of belief revision (see footnote 12). But what is more interesting, at least for the purpose of the present paper, is that the results obtained from the fusion on individual bases using the Dalal's distance, display a strong similarity to the results obtained via propositionwise majority. This shows that the fusion method used to tackle the doctrinal paradox is not an *ad hoc* method, but really a more fine-grained procedure than propositionwise majority. This is for two reasons. First, because it allows us to take into account the fact (well-known in knowledge fusion) that the aggregation of consistent belief bases is not necessarily a consistent belief base, and therefore, the notion of integrity constraints must play a role in the merging process. Interestingly, it often happens that the belief base that would be selected in the first place as the collective outcome violates the integrity constraints - this is exactly when the doctrinal paradox appears. Second, there are cases in which there is no unique outcome as a result of a fusion of belief bases, but rather several options that dissatisfy the group members at the same minimum total level.

The second, and last, component we need to introduce in order to define $\leq_E$ is the aggregation function that takes the sequence of distances between each $\{K_i\} \in E$ and each possible collective outcome, and associates it with a natural number (in our case, the sum of all the distances).

Definition 3 ($\mathcal{F}^{\Sigma}$) *The distance between an interpretation I and a belief set E is:*

$$d_{\Sigma}(I, E) = \Sigma_{K \in E} d(I, K)$$

The pre-order between two interpretations I and J can now be defined as:

$$I \leq_E^{\Sigma} J \text{ if and only if } d_{\Sigma}(I, E) \leq d_{\Sigma}(J, E)$$

Finally the IC majority fusion operator $\mathcal{F}^{\Sigma}$ is:

$$mod(\mathcal{F}_{IC}^{\Sigma}(E)) = min(mod(IC), \leq_E^{\Sigma})$$

We are now ready to apply the method of belief fusion to judgment aggregation.

5 The Argument-Based Procedure

The lesson of the doctrinal paradox is that a majoritarian procedure will fail to assign a consistent collective judgment to certain individual judgments. List and Pettit (2002) prove a general impossibility theorem to show that the occurrence of an inconsistency does not depend on a specific aggregation procedure. Their impossibility result holds for all aggregation functions that satisfy some minimal conditions, such as universal domain, anonymity and systematicity. *Universal domain* states that any logically possible individual judgment is accepted as an input by the aggregation procedure. Moreover, *anonymity* ensures that all the voters are equally treated. *Systematicity* is the condition requiring that "the collective judgment on each proposition should depend exclusively on the pattern of individual judgments on that proposition. In particular, the collective judgment on no propositions should be given special weight in determining the collective judgments on others" (List and Pettit 2002, p.98). List and Pettit acknowledge that the condition of systematicity is controversial and that, when coupled with anonymity, it implies that if two propositions receive the same degree of support (not necessarily from the same individuals) the collective view on the two propositions should be the same.[14] As in the High School example, where a majority of teachers supports A and *another* majority supports B, systematicity requires that the collective outcome endorses $(A \wedge B)$ in contradiction to the majority voting on the conclusion against C. Systematicity is relaxed in the argument-based procedure that I propose here and instead it uses the fusion operator introduced in section 4. As the next example will make clear, the argument-based procedure gives priority to the integrity constraints in IC over the other propositions. The view of the group is selected from the models of IC according to a distance measure that takes into account each K_i.

Let us consider again the High School example. The argument-based procedure takes each individual judgment as a belief base. A set of possible interpretations is

[14] See Chapman 2002 for a critique on systematicity.

assigned to each belief base. Suppose that a belief base is $K_i = \{((A \vee B) \wedge C)\}$. The set I of interpretations for K_i respectively on the proposition variables A, B, C is $Mod(K_i) = \{(1,0,1),(0,1,1),(1,1,1)\}$. If, for example, an integrity constraint requires A to be true, $Mod(K_i)$ reduces to the set $\{(1,0,1),(1,1,1)\}$.

In the High School example all three teachers agree that the espresso machine will be bought (C) only if the machine is needed (A) and the price is convenient (B); that is they accept $((A \wedge B) \leftrightarrow C)$. Therefore $E = \{K_1, K_2, K_3\}$ and $IC = \{(A \wedge B) \leftrightarrow C\}$. Each individual makes a judgment on A, B and C that satisfies the integrity constraint. We can therefore write:[15]

$K_1 = \{A, B, C\}$

$K_2 = \{A, \neg B, \neg C\}$

$K_3 = \{\neg A, B, \neg C\}$

The interpretations for each belief base are the following:

$Mod(K_1) = \{(1,1,1)\}$

$Mod(K_2) = \{(1,0,0)\}$

$Mod(K_3) = \{(0,1,0)\}$

The table below shows the result of the IC majority fusion operator on $E = \{K_1, K_2, K_3\}$. The first three columns are all the possible interpretations for the propositional variables A, B and C. The rows with a shaded background correspond to the interpretations excluded by IC. The numbers in the columns of K_1, K_2 and K_3 are the Dalal's distances of each K_i from the correspondent interpretation. Finally, in the last column is $d_\Sigma(I, E)$.

[15] Given that I assumed that each K_i satisfies IC, to avoid redundancies I will not write $((A \wedge B) \leftrightarrow C)$ in K_i.

A	B	C	K_1	K_2	K_3	$d_\Sigma(I,E)$
1	1	1	0	2	2	4
1	1	0	1	1	1	3
1	0	1	1	3	1	5
1	0	0	2	2	0	4
0	1	1	1	1	3	5
0	1	0	2	0	2	4
0	0	1	2	2	2	6
0	0	0	3	1	1	5

Because $\mathcal{F}^{\Sigma}_{IC}(E)$ is an IC fusion operator, the collective outcomes are chosen among the interpretations that are not excluded by IC. Only $(1,1,1), (1,0,0), (0,1,0)$ and $(0,0,0)$ are the available candidates for the collective judgments. Moreover, $\mathcal{F}^{\Sigma}_{IC}(E)$ is a majority operator, and so the interpretations associated with the minimum distance value are selected. Thus, no paradox arises using the fusion operator. $Mod(\mathcal{F}^{\Sigma}_{IC}(E)) = \{(1,1,1), (1,0,0), (0,1,0)\}$, which is equivalent to saying that the collective outcome is a tie: $\{K_1 \vee K_2 \vee K_3\}$; We do not have enough information to select a unique collective judgment. Two of the three possible outcomes endorse the same conclusion $\neg C$ and the group may agree that $\neg C$ is the outcome to which they should come. Furthermore, unlike the conclusion-based procedure, $\mathcal{F}^{\Sigma}_{IC}(E)$ also indicates the precise reasons that properly support that conclusion.

Before moving to other examples where the result of the fusion operator is not just the disjunction of the belief bases, there are three points that we should note. The *first* is that had it not been discarded by IC, $\mathcal{F}^{\Sigma}_{IC}(E)$ would have selected $(1,1,0)$ as the model for the collective outcome. This is precisely the result associated with the doctrinal paradox. We can then say that the argument-based procedure reveals the nature of the doctrinal paradox. Propositionwise majority finds the closest interpretation to all the K_i in E. Unfortunately, it may happen that the closest model violates the integrity constraints. This, however, is not a danger with $\mathcal{F}^{\Sigma}_{IC}(E)$ because the inconsistent options are rejected by IC. The *second* point is that while majority rule on propositions can find only one model, the fusion operator looks at all possible outcomes and orders them according to some criteria (for the majoritarian fusion operator the criterion is the minimal distance). This is important especially in those cases where the fusion operator is used to aggregate several knowledge bases, since it guarantees that no relevant information is discarded. *Third*, note that normally a fusion operator does not require the belief bases to satisfy IC. Here I assume that each K_i satisfies IC for ease of comparison with the classical doctrinal paradox. When this assumption is relaxed, the interpretations of the belief bases that violate the integrity constraints are ruled out by the IC imposed on the final outcome. The collectivity is ensured to be rational, despite the fact that some of its members (or all) made irrational judgments.

Let us now turn to another problem. As we have seen the premise-based procedure may lead to contradictory results. The example of a group voting on $((P \wedge (Q \vee R)) \leftrightarrow T)$ in section 2 illustrated that different outcomes can be obtained by the premise-based procedure depending on whether the majority is calculated on the atomic propositions P, Q and R, or on the conjuncts P and $(Q \vee R)$. Transforming the judgments of the example in section 2 into interpretations for each belief base, we obtain:

$Mod(K_1) = \{(1,0,0,0)\}$
$Mod(K_2) = \{(0,1,1,0)\}$
$Mod(K_3) = \{(0,1,0,0)\}$
$Mod(K_4) = \{(0,0,1,0)\}$
$Mod(K_5) = Mod(K_6) = Mod(K_7) = \{(1,0,1,1)\}$

The table below displays the results of IC majority fusion operator on $E = \{K_1, K_2, ..., K_7\}$. Interpretations that violate IC do not appear.

P	Q	R	T	K_1	K_2	K_3	K_4	K_5	K_6	K_7	$d_\Sigma(I,E)$
1	1	1	1	3	2	3	3	1	1	1	14
1	1	0	1	2	3	2	4	2	2	2	17
1	0	1	1	2	3	4	2	0	0	0	11
1	0	0	0	0	3	2	2	2	2	2	13
0	1	1	0	3	0	1	1	3	3	3	14
0	1	0	0	2	1	0	2	4	4	4	17
0	0	1	0	2	1	2	0	2	2	2	11
0	0	0	0	1	2	1	1	3	3	3	14

The result $Mod(\mathcal{F}^{\Sigma}_{IC}(E)) = \{(1,0,1,1),(0,0,1,0)\}$ says that the two closest models to the group's view support opposite conclusions. The interpretation $(1,0,1,1)$ accepts the conclusion T and is the same as the premise-based procedure on atomic propositions. The interpretation $(0,0,1,0)$ supports $\neg T$, like the conclusion-based procedure, but, unlike the conclusion-based procedure, it specifies the most popular reasons to justify $\neg T$. Again, the argument-based method reveals that the essence of the doctrinal paradox rests on a tie between different possible outcomes.

In section 2 I criticized the conclusion-based procedure by claiming that there are many scenarios in which the result of collective deliberation needs to be supported by the proper reasons, and I gave the example of the court case. Providing justifications for a group decision can be particularly troublesome when the premises satisfy some logical constraints. This is illustrated by the following story.

Suppose that the Not-so-Posh council in London has to make a complex decision about the reduction of local taxes. The three parties in the council agree that the taxes can be reduced (T) if and only if no new medical equipment is authorized for the local hospital (H), and the local public transport is not improved (P). In order to cast their vote on these issues, they cautiously realize that they also need to consider that $((H \wedge P) \leftrightarrow U)$, where U stands for "being ready to face unpopularity". The last condition for the taxes to be reduced is that the budget of the library has to be reduced for the next year (L). Since the parties assess that this last issue will not affect their popularity, they agree to vote on $((((H \wedge P) \leftrightarrow U) \wedge L) \leftrightarrow T)$. The three parties vote in the following way:

	H	P	U	L	T
Party 1	Yes	Yes	Yes	No	No
Party 2	Yes	No	No	Yes	No
Party 3	No	Yes	No	Yes	No
Majority	Yes	Yes	No	Yes	No

All three parties in the council agree not to reduce the taxes, but they face a serious dilemma when they try to provide the local newspaper with the reasons for this decision. The majority of the council voted against the authorization of the purchase of new equipment for the hospital, a majority opposed the improvement of the local public transport and, again, a majority agreed to reduce the budget of the public library. So far this is the so-called doctrinal paradox. We understand that no party is willing to share with the local newspaper the role that considerations about their popularity played in their judgments. When the three parties meet to discuss how to explain to the public that the taxes have not been reduced, they also have to admit to themselves that they have a serious internal bug. The outcome of the majoritarian voting on the premises is inconsistent, since it assigns yes to H and to P but no to U (recall that $(H \wedge P) \leftrightarrow U$).

The argument-based procedure offers the parties two ways to solve both the internal and the public dilemmas. The interpretations assigned to the three parties are:

$Mod(K_1) = \{(1, 1, 1, 0, 0)\}$

$Mod(K_2) = \{(1, 0, 0, 1, 0)\}$

$Mod(K_3) = \{(0, 1, 0, 1, 0)\}$

The table below highlights the two preferred models for $E = \{K_1, K_2, K_3\}$ among the models satisfying $IC = \{(((H \wedge P) \leftrightarrow U) \wedge L) \leftrightarrow T\}$.

H	P	U	L	T	K_1	K_2	K_3	$d_\Sigma(I,E)$
1	1	1	1	1	2	3	3	8
1	1	1	0	0	0	3	3	6
1	0	0	1	0	3	0	2	5
1	0	0	0	0	2	1	3	6
0	1	0	1	0	3	2	0	5
0	1	0	0	0	2	3	1	6
0	0	0	1	0	4	1	1	6
0	0	0	0	0	3	2	2	7

The argument-based procedure recommends to send a communication to the local newspaper stating that the majority in the council opposed a tax reduction. The reasons for this are as follows: because of the financial situation of the Not-so-Posh council, a tax reduction is possible only if the library has its budget reduced (on which issue the majority agreed), no new medical equipment is provided to the hospital and the local public transport is not improved. On these two last issues, the parties could not find an agreement and there was no majority supporting both these conditions.

6 Conclusion

Two lessons can be drawn from the doctrinal paradox. The first one is that the set of propositions on which most group members agree is not guaranteed to be a candidate for the collective decision because the set can fail to satisfy consistency even though each individual consistently expressed her judgments. The second lesson is that, when this happens, we need to look for the second (or third...) best outcome. Complex collective decisions are therefore paradox-free when the logical relations between propositions are treated as integrity constraints on the collective judgment, and when we drop the ideal that an aggregation function must always find a unique possible solution.

The premise-based procedure and the conclusion-based procedure have been presented in the literature on judgment aggregation as the escape-routes from the doctrinal paradox. In section 2 I criticized both these procedures and urged an approach that reconciles the premises and the conclusion as part of the same argument. The argument-based procedure aims to find, when possible, a unique group decision and, at the same time, to provide the reasons in support of that outcome.

The approach introduced in this paper profits from the work done on belief fusion in artificial intelligence. I have claimed that judgment aggregation and belief fusion are similar processes and that a collaboration between these two areas is definitely fruitful. Following the research done by Konieczny and Pino-Pérez, a majoritarian IC fusion operator for belief bases has been defined. The integrity constraints ensure that the fusion operator looks for solutions only among the consistent models. Each belief base is confronted with a group decision candidate, and a number is assigned as a result of the comparison.[16] Intuitively, this number captures how much the two sets disagree with each other. Finally, the fusion function selects the model (or the models) that keeps the disagreement of most belief bases at the lowest level.

The value of the argument-based procedure rests upon the exclusion of inconsistent sets of judgments from the set of the candidates apt to become collective judgments, and in the definition of a preference order $\leq$ on the remaining candidates. I have shown that this new approach, as well as avoiding paradoxical results, reveals that the discursive dilemma hides incomparable outcomes (or ties).

From these results come implications for all those groups of people - like expert panels, boards, councils, societies, etc. - that have to make decisions on the basis of logically connected propositions, and moreover, want or need to be able to justify

[16] This number is determined by the Dalal's distance. The choice of Dalal's distance was justified in section 3. However, I plan to investigate the behaviour of other distance measures in a future research.

their decisions. In order to define a more realistic model of group decision-making, I plan to investigate extensions of the argument-based procedure. For example, as I already mentioned, belief fusion allows us to relax the completeness requirement of the belief bases, on the grounds that we want to conceive of the possibility of an individual to be indifferent (with respect to a preference) or ignorant (of a certain matter). Let us suppose that we have a finite number of belief bases expressing the decisions of some individuals about three propositions A, B and C. $K_i = \{A, B\}$ is a belief base even if it is not complete. The agent is indifferent to, or unable to make a decision on, C. Therefore K_i is satisfied by $\{(1, 1, 1), (1, 1, 0)\}$.

Also, an IC fusion operator can assign an outcome to belief bases that violate IC. Clearly, the models for the bases that violate IC will not be taken into consideration as candidates for the final decision. But this is an important feature because it gives us the possibility to merge bases even if some of them (or all) do not conform to IC. For instance, an individual who does not obey to IC has no chance that her own view will be adopted by the collectivity, and yet, the fusion operator will find a collective belief base that minimizes K_i's disagreement.

Majoritarian fusion operators are not the only operators that can be introduced to merge individuals' judgments where each member has equal power to influence the final decision. An alternative fusion aggregation is the arbitration. The best outcome for an arbitration operation is a belief base that aims at equally distributing the level of dissatisfaction among the individuals. In some cases this allows us to distinguish between two incomparable majoritarian outcomes. For instance, the two preferred solutions for a majoritarian fusion operator over $E = \{K_1, K_2, K_3\}$ can have $d_\Sigma(I, E)$ equal to 6. Still, the arbitration operator will prefer the one that equally distributes the dissatisfaction among members (like (2,2,2)) to (5,1,0). A question open to investigation is whether an arbitration operator can be used in any kind of judgment aggregation process. Is the arbitration operator likely to capture only the aggregation of individual desires or can it also be used for the aggregation of beliefs?

Finally, it would be interesting to study the aggregation judgment where the individuals do not all have the same power to influence the final decision, following the work done by Dietrich and List (2004). This could be especially useful for those scenarios in which a tie is obtained, but nevertheless a decision must be made.

Acknowledgments

This paper was presented at the workshop *Judgment Aggregation and the Discursive Dilemma* at the University of Konstanz, Germany, June 2004. I wish to thank the organizers, Franz Dietrich and Christian List, and the participants of that workshop. Furthermore, I thank Luc Bovens, Stephan Hartmann and Christian List for valuable comments on a draft of this paper.

References

Alchourrón, C., P. Gärdenfors and D. Makinson, 1985, "On the logic of theory change: Partial meet contraction and revision functions", *Journal of Symbolic Logic*, 50, 510-530.

Baral, C., S. Kraus, J. Minker and V. S. Subrahmanian, 1992, "Combining knowledge bases consisting of first-order theories", *Computational Intelligence*, 8(1), 45-71.

Benferhat, S., D. Dubois, H. Prade and M. A. Williams, 1999, "A practical approach to fusing prioritized knowledge bases", *EPIA 1999*, 223-236.

Borgida, A., and T. Imielinski, 1984, "Decision making in committees: A framework for dealing with inconsistency and non-monotonocity", in *Proceedings Workshop on Nonmonotonic Reasoning*, 21-32.

Bovens, L., and W. Rabinowicz, 2004, "Democratic answers to complex questions. An epistemic perspective", *Synthese*, forthcoming.

Brennan, G., 2001, “Collective coherence?”, *International Review of Law and Economics*, 21(2), 197-211.

Chapman, B., 2002, “Rational Aggregation”, *Politics, Philosophy and Economics*, 1(3), 337-354.

Cholvy, L., 1994, “A logical approach to multi-sources reasoning”, in *Knowledge Representation and Reasoning under Uncertainty: Logic at Work*, ed. by M. Masuch and L. Polos, Springer, LNAI 808, 183-196.

Dalal, M., 1988a, “Updates in propositional databases”, Technical Report DCS-TR-222, Department of Computer Science, Rutgers University.

Dalal, M., 1988b, “Investigations into a theory of knowledge base revision: Preliminary report”, *Proceedings of the Seventh National Conference on Artificial Intelligence*, 475-479.

Dietrich, F., 2004, “Judgment aggregation: (Im)possibility theorems”, *Journal of Economic Theory*, forthcoming.

Dietrich, F., and C. List, 2004, “A liberal paradox for judgment aggregation”, *Economics Working Paper Archive at WUSTL*, http://ideas.repec.org/p/wpa/wuwppe/0405003.html

Elster, J. (ed.), 1998, *Deliberative Democracy*, Cambridge University Press.

Gärdenfors, P., 1988, *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, The MIT Press.

Gärdenfors, P., and H. Rott, 1995, “Belief revision”, in *Handbook of Logic in Artificial Intelligence and Logic Programming*, Vol. 4, ed. by D. M. Gabbay, C. J. Hogger and J. A. Robinson, 35-132, Oxford University Press.

Konieczny, S., 1999, *Sur la Logique du Changement: Révision et Fusion de Bases de Connaissance*, Ph.D. dissertation, University of Lille I, France.

Konieczny, S., 2000, "On the difference between merging knowledge bases and combinig them", in *Proceedings of KR'00*, Morgan Kaufmann, Breckenridge, Colorado, USA, 135-144.

Konieczny, S., and R. Pino-Pérez, 1998, "On the logic of merging", in *Proceedings of KR'98*, Morgan Kaufmann, 488-498.

Konieczny, S., and R. Pino-Pérez, 1999, "Merging with integrity constraints", in *Proceedings of ECSQARU'99*, LNAI 1638, 233-244.

Konieczny, S., and R. Pino-Pérez, 2002, "Merging information under constraints: a logical framework", *Journal of Logic and Computation*, 12(5), 773–808.

Kornhauser, L. A., 1992, "Modeling collegial courts II. Legal doctrine", *Journal of Law, Economics and Organization*, 8, 441-470.

Kornhauser, L. A., and L. G. Sager, 1986, "Unpacking the court", *Yale Law Journal*, 96, 82-117.

Kornhauser, L. A., and L. G. Sager, 1993, "The one and the many: Adjudication in collegial courts", *California Law Review*, 81, 1-51.

Kowalski, R., 1978, "Logic for Data Description", in *Logic and Data Bases*, ed. by H. G. J. Minker, New York: Plenum, 77-102.

Liberatore, P., and M. Schaerf, 2000, "Brels: A system for the integration of knowledge bases", *Proceedings of KR 2000*, 145-152.

Lin, J., 1996, "Integration of weighted knowledge bases", *Artificial Intelligence*, 83, 363-378.

Lin, J., and A. Mendelzon, 1996, "Merging databases under constraints", *International Journal of Cooperative Information Systems*, 7, 55-76.

Lin, J., and A. Mendelzon, 1999, "Knowledge base merging by majority", in *Dynamic Worlds: From the Frame Problem to Knowledge Management*, ed. by R. Pareschi and B. Fronhoefer, Kluwer.

List, C., 2003, "A possibility theorem on decisions over multiple propositions", *Mathematical Social Sciences*, 45 (1), 1-13.

List, C., 2004, "The probability of inconsistencies in complex collective decisions", *Social Choice and Welfare*, forthcoming.

List, C., and P. Pettit, 2002, "Aggregating sets of judgments. An impossibility result", *Economics and Philosophy*, 18, 89-110.

List, C., and P. Pettit, 2004, "Aggregating sets of judgments. Two impossibility results compared", *Synthese*, 140, 207-235.

Maynard-Reid, P., and Y. Shoham, 1998, "From belief revision to belief fusion", in *Proceedings of the Third Conference on Logic and the Foundations of Game and Decision Theory (LOFT3)*, ICER, Torino, Italy.

Pauly, M., and M. van Hees, 2004, "Logical constraints on judgment aggregation", *Journal of Philosophical Logic*, forthcoming.

Pettit, P., 2001, "Deliberative democracy and the discursive dilemma", *Philosophical Issues* (supplement 1 of *Nous* 35), 11, 268-295.

Pettit, P., and W. Rabinowicz, 2001, "The jury theorem and the discursive dilemma", *Philosophical Issues* (supplement 1 of *Nous* 35), 11; appendix to Pettit (2001), 295-99.

Reiter, R., 1988, "On integrity constraints", in *Proceedings of the Second Conference on the Theoretical Aspects of Reasoning about Knowledge*, ed. by M. Y. Vardi, San Francisco, Calif.: Morgan Kaufmann, 97-111.

Revesz, P., 1997, "On the semantics of arbitration", *International Journal of Algebra and Computation*, 7(2), 133-160.