Axiomatising Judgement Aggregation Procedures in a Minimal Logical Language

Marc Pauly*

March 30, 2005

Abstract

We introduce a language similar to modal logic for reasoning about judgement aggregation procedures. In this language, the formula $\Box\varphi$ expresses that φ is collectively accepted, or that φ is a group judgement based on voting. Different judgement aggregation procedures may be underlying the group decision making. Here we investigate majority voting, where $\Box\varphi$ holds if a majority of individuals accepts φ, consensus voting, where $\Box\varphi$ holds if all individuals accept φ, and dictatorship. We provide complete axiomatisations for all three aggregation procedures. The results obtained are sensitive to the axiomatic language used for expressing properties of judgement aggregation. In order to clarify this point, we compare the results obtained to standard axiomatisation results of social choice theory.

1 Introduction

Social choice theory has traditionally been concerned with the aggregation of individual preferences. More recently some researchers have shifted the focus from preference aggregation to judgement aggregation. The difficulties arising in judgement aggregation are illustrated by the *discursive dilemma* (also known as *doctrinal paradox*): Consider a situation with three individuals each of which has an associated judgement set, the proposition he considers to be true. If the judgement sets for the three individuals are given by $\{p, q, p \land q\}$, $\{p, \neg q, \neg(p \land q)\}$ and $\{\neg p, q, \neg(p \land q)\}$, then each individual is logically consistent, but judgement aggregation based on proposition-wise majority voting will produce an inconsistent set of group judgements, namely $\{p, q, \neg(p \land q)\}$. This simple observation has been generalised in [4] and subsequent articles to impossibility results concerning judgement aggregation which go beyond majority voting.

From a different perspective, the discursive dilemma shows the importance of procedural effects in social decision making. If a group needs to decide whether or not to accept a conjunction $p \land q$, one can imagine at least two different procedures: The standard conclusion-driven procedure simply asks for a majority vote on $p \land q$. An alternative premise-driven procedure would accept $p \land q$ precisely when both p and q separately are accepted by majority vote. The discursive dilemma shows that these two procedures may produce different

*CNRS, Institut de Recherche en Informatique de Toulouse; e-mail: pauly@irit.fr

results, essentially because a majority for p and a majority for q together do not imply a majority for $p \wedge q$.

In this paper, a logical model for judgement aggregation procedures is developed. The premise- and conclusion-driven procedures mentioned are only two examples in a wide range of possible procedures for deciding complex issues. What is the logical relationship between these different procedures? For majority-voting, for instance, accepting $p \wedge q$ on the conclusion-driven procedure implies acceptance on the premise-driven procedure, but not vice versa. If consensus voting is applied instead of majority voting, however, both procedures are equivalent. Similar logical questions we might ask: What are the views that can be consistently held by different majorities? What can we conclude from the fact that there is a majority for φ or a majority for ψ?

Our aim in this paper is to develop a logical framework that allows us to characterise precisely the logical relationships that exist between different procedures involving not only majority voting, but also consensus voting and dictatorship. On the one hand, this will yield an axiomatisation of these voting mechanisms in the context of judgement aggregation. On the other hand, it will allow us to compare what logical properties distinguish, e.g., consensus voting from majority voting.

The axiomatisations obtained in this paper differ from those normally presented in social choice theory, like May's characterisation of majority voting [5], in that they depend on the language used for describing judgement aggregation. We aim for finding an axiomatisation which is syntactically minimal, i.e., in a logical language which extends propositional logic in a minimal way, where we can only talk about the propositions accepted by the group or society and nothing else. Thus, we may not, for instance, refer to propositions accepted by particular individuals, the fact that at least two individuals accept a certain proposition, and so on. As a consequence, these axiomatisations will completely characterise an aggregation procedure only w.r.t. the properties expressible in our language.

The paper is structured as follows: In section 2, we begin by introducing the syntax and semantics of a minimal language for describing aggregation procedures. Since the axiomatisation of majority voting makes essential use of a result about simple games, some time is spent introducing these in section 3, also because one can actually view an axiomatisation as establishing a relation between two kinds of simple games. Section 4 presents the axiomatisation results for consensus voting, dictatorship and majority voting, and section 5 discusses these results form a more methodological perspective, relating them to previous axiomatisation results in judgement aggregation as well as to the methodology of axiomatisation in social choice theory more generally. The conclusion (section 6) discusses some implications of the results obtained as well as future research.

2 A Formal Model of Collective Judgements

Given a finite nonempty set of propositional atoms Φ_0, we define the set of *individual formulas* Φ_I as the set of all formulas α generated by the following grammar, where $p \in \Phi_0$:

$$\alpha := p \mid \neg\alpha \mid \alpha_1 \wedge \alpha_2$$

An *individual valuation* is a function $v : \Phi_0 \to \{0, 1\}$, and we let V_I be the set of all individual valuations. In the standard way, we extend an individual valuation v by induction to a function $\hat{v} : \Phi_I \to \{0,1\}$ which assigns truth values also to complex propositions:

$$\begin{aligned} \hat{v}(p) &= v(p) \text{ for } p \in \Phi_0 \\ \hat{v}(\neg\alpha) &= 1 - \hat{v}(\alpha) \\ \hat{v}(\alpha \wedge \beta) &= min(\hat{v}(\alpha), \hat{v}(\beta)) \end{aligned}$$

In general, we shall usually identify v and $\hat{v}$, simply writing $v(\alpha)$ instead of $\hat{v}(\alpha)$. Note that since Φ_0 is finite, valuations can be characterised completely by individual formulas. Given an individual valuation $v \in V_I$, we define $[v] = \bigwedge_{\{p \in \Phi_0 | v(p)=1\}} p \wedge \bigwedge_{\{p \in \Phi_0 | v(p)=0\}} \neg p$. For a set of valuations $V \subseteq V_I$, we define $[V] = \bigvee_{\{v \in V\}} [v]$. Finally, we shall let $V_I(\alpha) = \{v \in V_I | v(\alpha) = 1\}$ denote all the individual valuations satisfying $\alpha \in \Phi_I$. Analogously for a set of formulas $\Sigma \subseteq \Phi_I$, we let $V_I(\Sigma) = \{v \in V_I | \forall \sigma \in \Sigma : \; v(\sigma) = 1\}$.

In order to talk about collective judgements, we shall use the modal $\Box$-operator. The formula $\Box\alpha$ will refer to the collective or aggregate judgement on formula α. Since in principle we want to allow for arbitrary collective judgements to be made, boxed formulas will be treated like atoms which can be assigned arbitrary truth values. Formally, we define the set of *collective formulas* Φ_C as the set of all formulas φ generated by the following grammar, where $\alpha \in \Phi_I$:

$$\varphi := \Box\alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2$$

For both individual and collective formulas, we use the standard abbreviations for the remaining connectives: $\top := p \vee \neg p$, $\bot := \neg\top$, $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $\varphi \to \psi := \neg\varphi \vee \psi$ and $\varphi \leftrightarrow \psi := (\varphi \to \psi) \wedge (\psi \to \varphi)$.

To give some examples of aggregation procedures expressible in this language, note that the premise-driven procedure mentioned in the introduction can be written as $\Box p \wedge \Box q$ whereas the conclusion-driven procedure can be written as $\Box(p \wedge q)$. The fact that a positive result in the conclusion-driven procedure implies a positive result in the premise-driven procedure is accordingly expressed by the formula $\Box(p \wedge q) \to (\Box p \wedge \Box q)$ which itself can be viewed as the aggregation procedure returning no only in the case where the premise-base procedure returns no but the conclusion-driven procedure returns yes.

Let $\Phi_{\Box I} = \{\Box\alpha | \alpha \in \Phi_I\}$. Then we can also define a *collective valuation* or *model* as a function $v : \Phi_{\Box I} \to \{0,1\}$, and we can analogously extend a collective valuation to a function $\hat{v} : \Phi_C \to \{0,1\}$. We shall let V_C denote the set of collective valuations. A formula $\varphi \in \Phi_C$ is a *collective (individual) tautology* iff $v(\alpha) = 1$ for all collective (individual) valuations v. A collective (individual) formula γ is *satisfiable* if there is a collective (individual) valuation v such that $v(\gamma) = 1$. Analogous to individual formulas, we let $V_C(\varphi) = \{v \in V_C | v(\varphi) = 1\}$ for $\varphi \in \Phi_C$ and $V_C(\Sigma) = \{v \in V_C | \forall \sigma \in \Sigma : \; v(\sigma) = 1\}$ for $\Sigma \subseteq \Phi_C$.

The notion of logical consequence is defined in the standard way: φ is a logical consequence of Γ, denoted as $\Gamma \models \varphi$, provided $V_C(\Gamma) \subseteq V_C(\varphi)$. We say that a set of collective formulas $\Delta \subseteq \Phi_C$ is sound for a class of models $\mathcal{C}$ iff $\mathcal{C} \subseteq V_C(\Delta)$. Δ is *complete* for $\mathcal{C}$ iff $V_C(\Delta) \subseteq \mathcal{C}$. An *axiomatisation* of a class of models is a set of formulas which is both sound and complete for that class.

We are interested in collective valuations which arise by means of certain decision methods. For a finite set of individuals $N = \{1, \ldots, n\}$, a *decision method* $D : \{0,1\}^N \to \{0,1\}$ maps n individual yes/no-decisions into a collective yes/no-decision. We shall be interested in axiomatising models based on a few specific decision methods. An n-ary decision method D is a *dictatorship* if there is some individual $d \in N$ such that $D(x_1, \ldots, x_n) = x_d$ for all

$x_1, \ldots, x_n \in \{0,1\}$. *Majority voting* is the decision method D where $D(x_1, \ldots, x_n) = 1$ iff $|\{x_i | x_i = 1\}| > \frac{1}{2}|N|$. Note that we have chosen here for strict majority, but the axiomatisation results to be presented later have their analogues for weak majority, where for an even number of voters exactly half of the voters will also suffice. Finally, we call D *consensus voting* provided that $D(x_1, \ldots, x_n) = 1$ iff all $x_i = 1$.

The terminology we applied to decision methods can also be lifted to collective valuations. A model $v \in V_C$ is *n-systematic* iff there is some n-ary decision method D and there are n individual valuations $v_1, \ldots, v_n \in V_I$ such that for all $\alpha \in \Phi_I$, $v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$. We denote the class of n-systematic models as $\mathcal{SYS}_n$. In systematic models, the group judgement on some formula will only depend on the individual judgements concerning that formula. Furthermore, the way in which the the group judgement depends on the individual judgements is uniform for all formulas. The class of all majority models, i.e., n-systematic models where the decision method is majority voting, will be denoted as $\mathcal{MAJ}_n$, and the class of all consensus models based on n individuals will be denoted as $\mathcal{CON}_n$. Finally, we let $\mathcal{DIC}$ denote the class of all dictatorial models. Note that the arity of the decision method does not matter in the case of dictatorships, so there is no need to specify the number of individuals for dictatorships on the level of collective valuations.

Figure 1 summarises the definitions of the different model classes. We occasionally want to abstract over the number of individuals, using $\mathcal{CON}$ as an abbreviation for $\bigcup_n \mathcal{CON}_n$, and similarly for the other model classes. The relationships between the different model classes are given in theorem 1.

$\mathcal{SYS}_n$	systematic	$v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$
$\mathcal{DIC}$	dictatorial	$v(\Box\alpha) = v_0(\alpha)$, for some $v_0 \in V_I$
$\mathcal{MAJ}_n$	majority	$v(\Box\alpha) = 1$ iff $|\{i \leq n | v_i(\alpha) = 1\}| > \frac{1}{2}n$
$\mathcal{CON}_n$	consensus	$v(\Box\alpha) = 1$ iff $\forall i \leq n : v_i(\alpha) = 1$

Figure 1: Model Classes

Theorem 1

1. *$\mathcal{DIC} \subseteq \mathcal{MAJ}_n, \mathcal{CON}_n \subseteq \mathcal{SYS}_n$, for all n.*

2. *$\mathcal{DIC} = \mathcal{MAJ}_1 = \mathcal{CON}_1$ and $\mathcal{MAJ}_2 = \mathcal{CON}_2$.*

3. *$\mathcal{MAJ}_n \not\subseteq \mathcal{CON}$, for $|\Phi_0| \geq 2$ and $n \geq 3$.*

4. *$\mathcal{CON}_n \not\subseteq \mathcal{MAJ}$, for $|\Phi_0| \geq 2$ and $n \geq 3$.*

Proof. Properties 1 and 2 are easy to verify, we shall only demonstrate properties 3 and 4. For property 3, consider the situation of the discursive dilemma generalised to n individuals, where there are individual valuations $v_1, \ldots, v_n$ defined as follows: $v_i(p) = 1$ iff $i \leq \frac{1}{2}n + 1$, and $v_i(q) = 1$ iff $i \geq \frac{1}{2}n$. If $v \in V_C$ is the collective valuation associated to majority voting, we have $v(\Box p) = v(\Box q) = 1$ while $v(\Box(p \wedge q)) = 0$ which shows that $v \notin \mathcal{CON}$, since $\Box$ distributes over conjunction for consensus voting (see axiom C in section 4.2).

For property 4, we will show in the proof of theorem 8 that every collective valuation in $\mathcal{MAJ}$ will satisfy axiom T introduced in section 4.4. Hence, it suffices to show that there are collective valuations based on consensus voting which fail to satisfy the axiom. Consider $n \geq 3$ individual valuations $v_1, \ldots, v_n$ such that we have $v_1 \neq v_2$, $v_1 \neq v_3$ and $v_2 \neq v_3$. Let $v \in V_C$ be the collective valuation associated with these individual valuations under consensus voting. Then the following formula

$$\neg\Box\bot \wedge \neg\Box[\{v_1, v_2\}] \wedge \bigwedge_{i \leq n} \neg\Box[v_i] \wedge \neg\Box\neg[v_i]$$

is true in v. If v satisfied axiom T, this would imply that $\neg\Box\top$ is also true in v, a contradiction. Hence, v cannot satisfy axiom T, and so $v \notin \mathcal{MAJ}$. ⊣

A final word about logical consistency. Models require logical consistency only on the nonatomic level. On the atomic level, no logical consistency is required for the collective judgement, since the atoms of a collective valuation can be assigned arbitrary values. This also points to the raison d'être of the $\Box$-operator, for it is needed to separate the logically consistent part of a collective formula from the possibly inconsistent part. Within the scope of the $\Box$-operator, no logical consistency is required. This means that group judgements are not required to be logically consistent, since $\Box(p \wedge \neg p)$ may be true. Similarly, it is not necessarily the case that $v(\Box(p \wedge q)) = 1$ if and only if $v(\Box p) = v(\Box q) = 1$. Outside of the $\Box$-operator, however, we do require logical consistency: $\Box p \wedge \neg\Box p$ can never be true.

3 Simple Games

In order to obtain a complete axiomatisation of $\mathcal{MAJ}$ we will make use of a characterisation result for simple games. As it turns out, simple games also provide for an interesting alternative perspective on the models we have used and the axiomatisation results we are going to establish.

Simple games, introduced in [6], are a very simple model of coalitional power. Formally, a *simple game* is a pair (N, W) where N is a finite nonempty set of individuals, and $W \subseteq \mathcal{P}(N)$ is the collection of winning coalitions of individuals. The UN security council as well as the original European Economic Community provide examples of simple games where different coalitions of nations may be winning, i.e., they are able to pass any motion. Simple games have been studied extensively in game theory and social choice theory, see [10] for a recent survey.

Some of the notions introduced earlier can also be rendered using the terminology of simple games. With every systematic model $v \in V_C$ we can associate two simple games:

Decision Game A decision method $D : \{0,1\}^N \to \{0,1\}$ is nothing but the simple game $(N, \{X \subseteq N | D(x_1, \ldots, x_n) = 1, \text{ where } \forall i \in N : x_i = 1 \text{ iff } i \in X\}$. Consequently, a systematic model $v \in V_C$ is associated with n not necessarily distinct individual valuations and a simple game (N, W). When viewing a decision method as a simple game, we shall refer to it as a *decision game*.

Truth Game Every systematic model $v \in V_C$ induces a simple game (V_I, T), where the winning coalitions are all the sets of individual valuations corresponding to true atomic formulas in v, i.e., $X \in T$ iff $v(\Box[X]) = 1$. Note that the fact that v is systematic implies that for all $\alpha \in \Phi_I$, $V_I(\alpha) \in T$ iff $v(\Box\alpha) = 1$. We shall call this second simple game the *truth game* associated with model v and denote it as $Tr(v)$.

Hence, simple models can be used both to define a systematic model and to define the collectively accepted formulas of the model. The link between these two simple games is provided by our central truth definition for systematic models. In terms of simple games, we say that a decision game (N, W) and truth game (V_I, T) *correspond*, notation $(N, W) \sim (V_I, T)$, if there exist individual valuations $v_1, \ldots, v_{|N|}$ such that for all $\alpha \in \Phi_I$ we have

$$V_I(\alpha) \in T \text{ iff } \{i \in N | v_i(\alpha) = 1\} \in W.$$

Hence, for a systematic model v based on decision game (N, W), we have $(N, W) \sim Tr(v)$ simply by definition. The notion of correspondence can also be lifted to classes of simple games. A class of decision games $\mathcal{D}$ and a class of truth games $\mathcal{T}$ *correspond*, denoted as $\mathcal{D} \sim \mathcal{T}$, iff for all truth games (V_I, T),

$$(V_I, T) \in \mathcal{T} \text{ iff } \exists (N, W) \in \mathcal{D} : (V_I, T) \sim (N, W).$$

The axiomatisation results we will be demonstrating can now be expressed in terms of simple games as follows. Let $\mathcal{D}$ be a class of decision games and let $v(\mathcal{D})$ be the associated class of systematic models, i.e. $v(\mathcal{D}) = \{v \in V_C | \exists D \in \mathcal{D} \exists v_1, \ldots, v_n \in V_I \forall \alpha \in \Phi_I : v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))\}$. If Δ is a set of collective formulas, an axiomatisation result shows that $V_C(\Delta) = v(\mathcal{D})$. If $Tr(\Delta) = \{Tr(v) | v(\Delta) = 1\}$ denotes all the truth games associated with models satisfying Δ, then

$$V_C(\Delta) = v(\mathcal{D}) \text{ iff } Tr(\Delta) \sim \mathcal{D}.$$

Hence, our axiomatisation results can also be expressed as correspondences between two types of simple games associated with models, decision games and truth games.

Simple games not only provide an interesting alternative perspective on the axiomatisations we are after. We will also make essential use of them in our axiomatisation of majority voting. For this purpose, it is useful to slightly generalise the notion to incorporate voting ties. A *simple game with ties* (SGT, also known as a prehypergraph with ties in [10]) $G = (N, W, T, L)$ consists of a finite nonempty set of individuals N and sets of winning (W), tied (T) and losing (L) coalitions, where $W, T, L \subseteq \mathcal{P}(N)$ and these sets are pairwise disjoint. We call an SGT (N, W, T, L) *weighted* iff there exists a weight function $w : N \to \mathbb{R}$ and a threshold or quota $q \in \mathbb{R}$ such that for all $X \subseteq N$,

1. if $X \in W$ then $\sum_{x \in X} w(x) > q$,
2. if $X \in T$ then $\sum_{x \in X} w(x) = q$, and
3. if $X \in L$ then $\sum_{x \in X} w(x) < q$.

For ease of readability, we shall usually write $w(X)$ instead of $\sum_{x \in X} w(x)$. The simple games mentioned earlier of the UN security council and the original European Economic Community (EEC) provide examples of strongly weighted voting games. In the EEC, for instance, France, West Germany, and Italy each had a weight of 4 votes, Belgium and the Netherlands had 2 votes and Luxembourg 1 vote. Every motion which got at least 12 of the 17 possible votes passed.

We call an SGT $G = (N, W, T, L)$ *k-trade robust* iff the following condition holds: For all sequences of coalitions $\langle X_1, \ldots, X_k \rangle$ and $\langle Y_1, \ldots, Y_k \rangle$ such that (1) for every $p \in N$, $|\{i : p \in X_i\}| = |\{i : p \in Y_i\}|$, (2) for every $i \leq k$ we have $X_i \in W \cup T$, and (3) for every $i \leq k$ we have $Y_i \in L \cup T$, we have $X_i \in T$ and $Y_i \in T$ for every $i \leq k$. G is *trade robust* iff it is k-trade robust for all k. Intuitively, suppose we have k non-losing coalitions and k non-winning coalitions. The non-losing coalitions can be obtained from the non-winning coalitions by trading, i.e., if an individual occurs x times in the non-losing coalitions, he will appear also x times in the non-winning coalitions. In this situation, all the coalitions involved must be tied.

The following combinatorial result due to [11] generalises a result of Taylor and Zwicker [9] for simple games (without ties). Since the completeness theorem for majority voting relies essentially on this highly non-trivial result, we present its proof in the appendix in order to make the exposition self-contained, and also because the proof has not been published before.

Theorem 2 ([11]) *For any simple game with ties $G = (N, W, T, L)$, the following are equivalent:*
(i) G is weighted.
(ii) G is trade robust.
(iii) G is 2^k-trade robust, with $k = 2^{|N|}$.

4 Axiomatisation Results

We now continue to axiomatise a number of classes of models. Due to the link with non-normal modal logic, our terminology shall be close to the one used in [2].

4.1 Systematicity

Systematicity is the minimal requirement which all our model classes share, so it is instructive to see what axioms can enforce systematicity. Let $\mathtt{E} = \{\Box\alpha \leftrightarrow \Box\beta \in \Phi_C \mid \alpha \leftrightarrow \beta \in \Phi_I \text{ is a tautology}\}$.

Theorem 3 $V_C(\mathtt{E}) = \mathcal{SYS}_n$*, provided $n \geq 2^{|\Phi_0|}$. Hence,* $\mathtt{E}$ *also axiomatises* $\mathcal{SYS}$*.*

Proof. For soundness, all axioms of $\mathtt{E}$ are easily seen to be true in any systematic model. For completeness, suppose that $v_c(\mathtt{E}) = 1$ for some collective valuation $v_c \in V_C$ and that $n \geq 2^{|\Phi_0|}$. We need to find individual valuations $v_1, \ldots, v_n \in V_I$ and a decision method D such that $v_c(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$.

As for the individual valuations, we will assign every possible valuation $v \in V_i$ to some individual, i.e., we will simply order all the valuations in V_i in some way, assigning the first valuation to individual 1, the second to individual 2, etc. and the last possible valuation to all the remaining players. Note that since $n \geq 2^{|\Phi_0|}$, there are sufficiently many players so that every world view (i.e. valuation) will be represented. As for D, we define $D(x_1, \ldots, x_n) = v_c(\Box\alpha)$, in case there is some $\alpha \in \Phi_I$ such that for all $i \leq n$ we have $v_i(\alpha) = x_i$, and we let $D(x_1, \ldots, x_n) = 0$ otherwise. Given that $v_c(\mathtt{E}) = 1$, D is well-defined: if there is also a second formula $\beta \in \Phi_I$ such that for all $i \leq n$, $v_i(\alpha) = x_i$, we know (since all valuations are present among the individuals) that α and β are logically equivalent, so $v_c(\Box\alpha) = v_c(\Box\beta)$. By definition, we have $v_c(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$. ⊣

The result implies that if we are interested in systematic judgement aggregation, group decision making is invariant under logical equivalence, i.e., within a decision procedure φ which contains $\Box\alpha$ as a subformula, we can substitute $\Box\beta$ for $\Box\alpha$ provided that α is logically equivalent to β.

Two further remarks concerning the result. First, the restriction that there are sufficiently many individuals to express all the possible valuations ($n \geq 2^{|\Phi_0|}$) is only needed to show completeness; soundness holds for an arbitrary number of individuals. Second, completeness does fail if there are not sufficiently many individuals. For $n = 1$, the formula $\Box p \wedge \Box\neg p \rightarrow \Box q$ holds in all systematic models, while it does not hold for $n > 1$. More generally, the formula $\bigvee_{v \in V_I}(\Box[v] \leftrightarrow \Box\bot)$ holds in all systematic models if $n < 2^{|\Phi_0|}$, whereas it is easy to falsify the formula in a systematic model with $n \geq 2^{|\Phi_0|}$. Hence the bound on the number of individuals given in theorem 3 is tight.

4.2 Consensus Voting

Let `EMCN` or `K` denote the set of collective formulas containing `E` and all instances of the following three axiom schemes:

$$
\begin{array}{ll}
\mathtt{M.} & \Box(\alpha \wedge \beta) \rightarrow (\Box\alpha \wedge \Box\beta) \\
\mathtt{C.} & (\Box\alpha \wedge \Box\beta) \rightarrow \Box(\alpha \wedge \beta) \\
\mathtt{N.} & \Box\top
\end{array}
$$

If, in addition, the axiom `D` $\neg\Box\bot$ is added, we obtain the set of formulas `KD`.

Lemma 4 (Monotonicity) `KD` $\models \Box\alpha \rightarrow \Box\beta$, *provided* $\alpha \rightarrow \beta$ *is a tautology.*

Proof. If $\alpha \rightarrow \beta$ is a tautology, then so is $\alpha \leftrightarrow (\alpha \wedge \beta)$. Consequently, any model v satisfying `E` will satisfy $\Box\alpha \leftrightarrow \Box(\alpha \wedge \beta)$, and applying the `M` axiom v must also satisfy $\Box\alpha \rightarrow \Box\beta$. ⊣

Lemma 5 *For any set* $\Gamma \subseteq \Phi_I$, *if every finite subset of* Γ *is satisfiable, then so is* Γ.

Proof. While this result is simply the compactness theorem for propositional logic treated in most logic textbooks, in our case, the result can be obtained in a simpler manner. Suppose that Γ is not satisfiable. Due to the fact that Φ_0 is finite, there must be a finite $\Gamma_0 \subseteq \Gamma$ such that for every $\alpha \in \Gamma$ there is some $\beta \in \Gamma_0$ such that α and β are logically equivalent. Hence, Γ_0 cannot be satisfiable. ⊣

Theorem 6 $V_C(\mathtt{KD}) = \mathcal{CON}_n$, *provided* $n \geq 2^{|\Phi_0|}$. *Hence,* `KD` *also axiomatises* $\mathcal{CON}$.

Proof. The proof is analogous to the proof of theorem 3. Soundness is easy to check, and for completeness, suppose that $v_c(\mathtt{KD}) = 1$ for some collective valuation $v_c \in V_C$. We will construct individual valuations $v_1, \ldots, v_n \in V_I$ such that $v_c(\Box\alpha) = 1$ iff for all $i \leq n$ we have $v_i(\alpha) = 1$.

Consider the sets $W = \{v \in V_I | \forall \alpha : v_c(\Box\alpha) = 1 \Rightarrow v(\alpha) = 1\}$ and $\Gamma = \{\alpha \in \Phi_I | v_c(\Box\alpha) = 1\}$. Since $v_c(\Box\top) = 1$ by axiom N, we know that Γ is nonempty. Also W must be nonempty. For suppose to the contrary that W were empty. Then there is no individual valuation satisfying all of Γ, and so by lemma 5, there is some finite $\Gamma_0 = \{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$ such that $\bigwedge \Gamma_0 \to \bot$ is an individual tautology. By lemma 4, this implies that $v_c(\Box \bigwedge \Gamma_0 \to \Box\bot) = 1$, and by axiom C, $v_c((\Box\gamma_1 \wedge \ldots \wedge \Box\gamma_m) \to \Box\bot) = 1$. Consequently, $v_c(\Box\bot) = 1$, a contradiction with axiom D. Hence, W is indeed nonempty.

Now we will assign every valuation $v \in W$ to some individual, and since $n \geq 2^{|\Phi_0|}$, there are sufficiently many players to cover all valuations in W.

It remains to show that $v_c(\Box\alpha) = 1$ iff $\forall i \leq n$ we have $v_i(\alpha) = 1$. First, if $v_c(\Box\alpha) = 1$, then we have $v_i(\alpha) = 1$ by definition of W for all $i \leq n$. Conversely, suppose that $v_c(\Box\alpha) = 0$. In this case it suffices to find a single individual valuation $v_i \in W$ such that $v_i(\alpha) = 0$. For this, it suffices to show that $\Gamma \cup \{\neg\alpha\}$ is satisfiable. Suppose again to the contrary that it is not, then by lemma 5 there must be a finite $\Gamma_0 = \{\gamma_1, \ldots, \gamma_m\} \subseteq \Gamma$ such that $\bigwedge \Gamma_0 \to \alpha$ is an individual tautology. Then by lemma 4 and axiom C, $v_c((\Box\gamma_1 \wedge \ldots \wedge \Box\gamma_m) \to \Box\alpha) = 1$, and hence $v_c(\Box\alpha) = 1$, a contradiction. Hence, $\Gamma \cup \{\neg\alpha\}$ is indeed satisfiable, and so there must be some individual valuation $v_i \in W$ such that $v_i(\alpha) = 0$. $\dashv$

As in the case of systematicity, completeness may fail if there are not enough individuals present. The formula $\bigvee_{v \in V_I} \Box[V_I - \{v\}]$ holds in all consensus models if $n < 2^{|\Phi_0|}$, but not necessarily if $n \geq 2^{|\Phi_0|}$. Hence, also the bound in theorem 6 is tight.

4.3 Dictatorship

Let MCY denote the set of collective formulas containing M, C and all instances of the following axiom

$$\mathtt{Y.} \quad \neg\Box\alpha \leftrightarrow \Box\neg\alpha.$$

Note first that MCY $\models$ E: Suppose that $v(\mathtt{MCY}) = 1$. If $\alpha \leftrightarrow \beta \in \Phi_I$ is an individual tautology, then the collective formula φ, obtained by replacing every propositional atom $p \in \Phi_0$ by $\Box p$, must be a collective tautology, and hence $v(\varphi) = 1$. But given the axioms v satisfies, $v(\varphi) = v(\Box\alpha \leftrightarrow \Box\beta)$, and hence axiom E is indeed a logical consequence of MCY.

Furthermore, MCY $\models$ N: In the presence of axiom Y, $\neg(\Box p \wedge \Box\neg p)$ must be true and using axiom C, $\neg\Box\bot$ must be true which yields N using again axiom Y. Note finally that we also have MCY $\models$ D.

Theorem 7 $V_C(\mathtt{MCY}) = \mathcal{DIC}$, *i.e.,* MCY *axiomatises* $\mathcal{DIC}$.

Proof. Soundness is easy to check. For completeness, given a model $v_c \in V_C$ satisfying MCY, we simply need to find an individual valuation $v_1 \in V_I$ such that $v_c(\Box\alpha) = v_1(\alpha)$ for all $\alpha \in \Phi_I$. Simply define $v_1(p) = v_c(\Box p)$ for all $p \in \Phi_0$. We then verify by induction on α that $v_1(\alpha) = v_c(\Box\alpha)$. The base case holds by definition, the inductive step for conjunction follows from axioms M and C and the inductive step for negation from axiom Y. ⊣

Intuitively, axioms M, C and Y together require the group decision to be logically consistent. A dictatorship, being based on a logically consistent individual decision maker, clearly satisfies these requirements. Conversely, any logically consistent group decision can be represented by a dictatorship based on an individual who holds precisely the group's views. This reasoning is formalised in theorem 7. Since in a distributive logic the $\Box$ operator distributes over all connectives, there are no procedural effects for dictatorship, i.e., no matter how a decision problem for a complex formula is proceduralised, the outcome will always be the same.

4.4 Majority Voting

For the purposes of formulating our axiom system, it will be useful to define a number of abbreviations referring to non-strict majority, ties, etc. We will use $[>]\varphi$ for $\Box\varphi$ simply to remind the reader that our basic modality refers to strict majority. Furthermore, we define $[=]\varphi$ as $\neg\Box\varphi \wedge \neg\Box\neg\varphi$, $[\geq]\varphi$ as $[>]\varphi \vee [=]\varphi$, $[\leq]\varphi$ as $\neg[>]\varphi$ and $[<]\varphi$ as $\neg[\geq]\varphi$.

Let STEM denote the set of collective formulas containing M, E and all instances of the following two axiom schemes

S. $[>]\alpha \rightarrow \neg[>]\neg\alpha$
T. $([\geq]\alpha_1 \wedge \ldots \wedge [\geq]\alpha_k \wedge [\leq]\beta_1 \wedge \ldots \wedge [\leq]\beta_k) \rightarrow \bigwedge_{1\leq i\leq k}([=]\alpha_i \wedge [=]\beta_i)$
where $\forall v \in V_I : |\{i : v(\alpha_i) = 1\}| = |\{i : v(\beta_i) = 1\}|$

Axiom T is easily seen to express trade-robustness, and axiom S states that there can be no strict majority for φ and its negation at the same time.

Theorem 8 STEM *axiomatises* $\mathcal{MAJ}$, *i.e.,* $V_C(\mathtt{STEM}) = \mathcal{MAJ}$.

Proof. As for soundness, all axioms except T are straight forward to verify. For trade-robustness, consider a model v_c based on individual valuations $v_1, \ldots, v_n$ and majority voting. In order to obtain a contradiction, suppose that T is false in v_c, i.e., without loss of generality, one of the majorities for some α_i is strict. Then

$$\sum_{i\leq k}\sum_{p\leq n} v_p(\alpha_i) > k \cdot \frac{1}{2}n \geq \sum_{i\leq k}\sum_{p\leq n} v_p(\beta_i).$$

But since the trading property $\forall v \in V_I : |\{i : v(\alpha_i) = 1\}| = |\{i : v(\beta_i) = 1\}|$ implies that $\sum_{p\leq n}\sum_{i\leq k} v_p(\alpha_i) = \sum_{p\leq n}\sum_{i\leq k} v_p(\beta_i)$, we have a contradiction.

For completeness, consider any model $v_c \in V_C$ satisfying STEM, and consider the simple game with ties $G = (V_I, W, T, L)$ where $(V_I, W) = Tr(v_c)$ is simply the truth game of v_c, i.e., $W = \{V_I(\alpha)|v_c(\Box\alpha) = 1\}$. For the tied and losing coalitions, we define $L = \{V_I(\alpha)|v_c(\Box\neg\alpha) = 1\}$ and $T = \{V_I(\alpha)|v_c(\Box\alpha) = v_c(\Box\neg\alpha) = 0\}$. Note that due to axiom S, sets W and L are disjoint, and hence G is a well defined SGT. Axiom T guarantees that G is trade robust, and hence by applying theorem 2 we know that G is weighted, with some weight function $w : V_I \rightarrow \mathbb{R}$ and some threshold $q \in \mathbb{R}$. We now proceed to show a few properties of the weight function w and the threshold q.

First, we show that we can assume that $q = \frac{1}{2}\sum_{v\in V_I} w(v)$. If $v_c(\Box\alpha) = 1$, by axiom S, $v_c(\Box\neg\alpha) = 0$, and hence $\sum_{v\in V_I:v(\alpha)=1} w(v) > q > \sum_{v\in V_I:v(\alpha)=0} w(v)$, showing that $\sum_{v\in V_I:v(\alpha)=1} w(v) > \frac{1}{2}\sum_{v\in V_I} w(v)$. The case where $v_c(\Box\neg\alpha) = 1$ is analogous. Finally, if $v_c(\Box\alpha) = v_c(\Box\neg\alpha) = 0$, $V_I(\alpha)$ is a tie, and hence $\sum_{v\in V_I:v(\alpha)=1} w(v) = q = \sum_{v\in V_I:v(\alpha)=0} w(v)$, showing that $\sum_{v\in V_I:v(\alpha)=1} w(v) = \frac{1}{2}\sum_{v\in V_I} w(v)$.

Second, we show that we can assume the weight function to be nonnegative, i.e., for all $v \in V_I$, $w(v) \geq 0$. We distinguish two cases. (i) Suppose that v is irrelevant, i.e., for all $\alpha \in \Phi_I$, $v_c(\Box\alpha) = v_c(\Box(\alpha \vee [v]))$. In that case, the presence of v never matters, and we can take $w(v) = 0$. (ii) If v is relevant, there is some $\alpha \in \Phi_I$ such that $v_c(\Box\alpha) \neq v_c(\Box(\alpha \vee [v]))$. Using the monotonicity axiom M we know that $v_c(\Box\alpha) < v_c(\Box(\alpha \vee [v]))$, hence $w(v) > 0$.

Third, note that we can assume all weights and the threshold q to be integers. Due to the discreteness of the domain of the weight function w, all weights and the threshold can be assumed to be rational numbers, and hence we only need to multiply these by a sufficiently high integer to obtain integer weights and threshold.

Hence, taking these three observations together, we know that there exists a weight function $w : V_I \to \mathbb{N}$ such that

$$v_c(\Box\alpha) = 1 \text{ iff } \sum_{v\in V_I:v(\alpha)=1} w(v) > \frac{1}{2}\sum_{v\in V_I} w(v).$$

Now we can consider as individual valuations $v_1, \ldots, v_n$ precisely all those valuations v for which $w(v) > 0$, and we let the number of individuals with worldview v equal $w(v)$. Hence, in total, we have $\sum_{v\in V_I} w(v)$ individual valuations. If D is the decision method of majority voting, then we have for all $\alpha \in \Phi_I$

$$D(v_1(\alpha), \ldots, v_n(\alpha)) = 1 \text{ iff } \sum_{v\in V_I:v(\alpha)=1} w(v) > \sum_{v\in V_I:v(\alpha)=0} w(v) \text{ iff } v_c(\Box\alpha) = 1,$$

thereby showing that v_c is a majority model based on individual valuations $v_1, \ldots, v_n$. $\dashv$

Note that in contrast to the previous axiomatisation results, theorem 8 does not provide a characterisation of $\mathcal{MAJ}_n$. Hence, the result is weaker than the others in the sense that we do not know beforehand how many individuals there need to be. The result can be thought of as a result *in the limit*. The model constructed in the proof will usually have a number of individuals exponential in $2^{|\Phi_0|}$, since the weights and quota obtained after transformation into integers may be exponential in $2^{|\Phi_0|}$ (cf. the results in [1] on linear threshold units).

5 Varieties of Axiomatisation

The axiomatisation results obtained rely on a particular conceptualisation of judgement aggregation and on a particular description language. The results differ in a number of respects from standard axiomatisation results in social choice theory. For this reason, we will spend some time reflecting on the nature of the axiomatisation results obtained. More specifically, we shall compare our axiomatisation of majority voting (theorem 8) with that of May [5], and our axiomatisation of dictatorship with a previous result obtained in the context of judgement aggregation [8].

A *judgement aggregation function* $A : (V_I)^N \to V_C$ maps n individual valuations into a collective valuation. Judgement aggregation functions have been considered in [4] and subsequent papers, but in contrast to our work, most of the results concerning judgement aggregation functions assume that the aggregated judgement is logically consistent (see [3] for a web site covering judgement aggregation and the discursive dilemma). We will say that a judgement aggregation function A satisfies a formula $\varphi \in \Phi_C$, denoted as $A \models \varphi$, provided that for all $v_1, \ldots, v_n \in V_I$ we have $A(v_1, \ldots, v_n)(\varphi) = 1$. A satisfies a set of formulas Σ (denoted as $A \models \Sigma$) if $A \models \sigma$ for all $\sigma \in \Sigma$. Let $\mathcal{A}$ be the class of all judgement aggregation functions. Then given a class of aggregation functions $\mathcal{C}$, we say that $\Delta \subseteq \Phi_C$ *axiomatises* $\mathcal{C}$ iff $\{A \in \mathcal{A} | A \models \Delta\} = \mathcal{C}$.

The judgement aggregation functions defined and the associated notion of axiomatisation present a standard example of axiomatisation in social choice theory, with the exception that we use a formal language for expressing our axioms. In order to relate the axiomatisation results we obtained in the preceding section to axiomatisations of judgement aggregation functions, we still need to relate classes of judgement aggregation functions to classes of collective valuations. For $\mathcal{C} \subseteq \mathcal{A}$, we define $\mathcal{C}\downarrow = \{v \in V_C | \exists A \in \mathcal{C}\ \exists v_1, \ldots, v_n :\ A(v_1, \ldots, v_n) = v\}$. We can now consider analogues of the classes of collective valuations defined in section 2. Let $\mathcal{SYS}^* \subseteq \mathcal{A}$ consist of all the systematic aggregation functions, i.e., the aggregation functions A for which there is a decision method D such that for all $v_1, \ldots, v_n \in V_I$ and $\alpha \in \Phi_I$ we have $A(v_1, \ldots, v_n)(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$. Let $\mathcal{CON}^*$ consist of all the systematic aggregation functions based on consensus voting, and similarly for $\mathcal{MAJ}^*$ and $\mathcal{DIC}^*$.

Theorem 9 *(1)* $\mathcal{SYS}^*\downarrow = \mathcal{SYS}$, $\mathcal{CON}^*\downarrow = \mathcal{CON}$, $\mathcal{DIC}^*\downarrow = \mathcal{DIC}$ *and* $\mathcal{MAJ}^*\downarrow = \mathcal{MAJ}$. *(2) If* $\Delta \subseteq \Phi_C$ *axiomatises a class of aggregation functions* $\mathcal{C} \subseteq \mathcal{A}$, *then* Δ *also axiomatises* $\mathcal{C}\downarrow$. *(3) If* Δ *axiomatises* $\mathcal{C}\downarrow$ *but not* $\mathcal{C} \subseteq \mathcal{A}$, *then* $\mathcal{C}$ *cannot be axiomatised.*

Proof. (1) We shall only prove the second equality, the others are completely analogous. If $v \in \mathcal{CON}$, then there are some $v_1, \ldots, v_n \in V_I$ such that for all $\alpha \in \Phi_I$ we have $v(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$, where D is consensus voting. Now let A be the n-ary systematic aggregation function based on consensus voting. Since $A \in \mathcal{CON}^*$ and $A(v_1, \ldots, v_n)(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$, $A(v_1, \ldots, v_n) = v$ and hence $v \in \mathcal{CON}^*\downarrow$. Conversely, if $v \in \mathcal{CON}^*\downarrow$, there is some $A \in \mathcal{CON}^*$ and some $v_1, \ldots, v_n \in V_I$ such that $A(v_1, \ldots, v_n) = v$. Hence, for the consensus decision method D we have $A(v_1, \ldots, v_n)(\Box\alpha) = D(v_1(\alpha), \ldots, v_n(\alpha))$ for all $\alpha \in \Phi_I$ which shows that $v \in \mathcal{CON}$.

(2) If $v \in \mathcal{C}\downarrow$, there is some $A \in \mathcal{C}$ and some $v_1, \ldots, v_n \in V_I$ such that $A(v_1, \ldots, v_n) = v$. Since Δ axiomatises $\mathcal{C}$, $A \models \Delta$, and so in particular $A(v_1, \ldots, v_n)(\Delta) = v(\Delta) = 1$. Conversely, if $v(\Delta) = 1$, consider the aggregation function A defined as $A(v_1, \ldots, v_n) = v$ for all $v_1, \ldots, v_n \in V_I$. Then by definition, $A \models \Delta$, and so $A \in \mathcal{C}$ which suffices to show that $v \in \mathcal{C}\downarrow$.

(3) Suppose that Δ axiomatises $\mathcal{C}\downarrow$ but not $\mathcal{C}$, and suppose that Γ axiomatises $\mathcal{C}$. Then

$$\begin{aligned} v(\Delta) = 1 \quad &\Leftrightarrow \exists A \in \mathcal{C}\ \exists v_1, \ldots, v_n \in V_I : A(v_1, \ldots, v_n) = v \\ &\Leftrightarrow \exists A \models \Gamma\ \exists v_1, \ldots, v_n \in V_I : A(v_1, \ldots, v_n) = v \\ &\Leftrightarrow v(\Gamma) = 1 \end{aligned}$$

and hence Γ and Δ must be logically equivalent. As a consequence, Δ also axiomatises $\mathcal{C}$, a contradiction. $\dashv$

The result leaves open the possibility that the two notions of axiomatisation are actually equivalent. This, however, turns out not to be the case. The following result not only shows that an axiomatisation of aggregation functions is stronger than an axiomatisation of collective valuations, it also shows that we cannot hope to generalise the axiomatisation results obtained for majority voting, etc. to aggregation functions. By theorem 9 and theorems 3, 6 and 8, we have axiomatisations of $\mathcal{SYS}^*\!\downarrow$, $\mathcal{CON}^*\!\downarrow$ and $\mathcal{MAJ}^*\!\downarrow$. The following result shows that, in contrast, these sets of formulas do not axiomatise the corresponding aggregation functions.

Theorem 10 *The classes of aggregation functions $\mathcal{SYS}^*$, $\mathcal{CON}^*$, $\mathcal{DIC}^*$ and $\mathcal{MAJ}^*$ cannot be axiomatised.*

Proof. By theorem 9, it suffices to show that the axiomatisations already obtained in theorems 3, 6, 7 and 8 for the respective model classes do not axiomatise the respective classes of aggregation functions.

`E` does not axiomatise $\mathcal{SYS}^*$: Consider the unary aggregation function A such that $A(v_1) = v$ for all $v_1 \in V_I$ and $v(\Box\alpha) = 1$ iff α is logically equivalent to p. Now consider the particular $v_1 \in V_I$ for which $v_1(q) = 1$ for all propositional atoms $q \in \Phi_0$. Then $A \models \mathtt{E}$, but $A \notin \mathcal{SYS}^*$ for $\Box p$ is true and $\Box q$ is false in $A(v_1)$ even though $v_1(p) = v_1(q) = 1$.

`KD` does not axiomatise $\mathcal{CON}^*$: Consider the binary aggregation function A where for all $v_1, v_2 \in V_I$ and $\alpha \in \Phi_I$, $A(v_1, v_2)(\Box\alpha) = v_1(\alpha)$. Hence, A is a dictatorship and so $A \models \mathtt{KD}$, but A is not based on consensus voting.

`STEM` does not axiomatise $\mathcal{MAJ}^*$: Take the same aggregation function A we used for consensus voting. $A \models \mathtt{STEM}$ but $A \notin \mathcal{MAJ}^*$.

`MCY` does not axiomatise $\mathcal{DIC}^*$: Take any $v_0 \in V_I$, and consider the binary aggregation function A such that $A(v_0, v_2)(\Box\alpha) = v_0(\alpha)$ and $A(v_1, v_2)(\Box\alpha) = v_2(\alpha)$ for all $v_1 \neq v_0$. $A \models \mathtt{MCY}$ but $A \notin \mathcal{DIC}^*$. $\dashv$

These results show that our logical language is too poor to axiomatise majority and consensus voting in the context of judgement aggregation functions. In the case of majority voting, we cannot distinguish a situation of a dictatorship where the dictator has valuation v_i and the majority has a different valuation $v_j \neq v_i$ from a situation of majority voting where the majority shares the same valuation v_i. Since our logical language can only express properties of the collective decision, the difference between these two situations cannot be captured in our language. This example also allows us to further pinpoint the difference between collective valuations and aggregation functions when it comes to axiomatisation: if a collective valuation v satisfies the axioms `STEM`, then it suffices that there are *some* individual valuations which produce the collective valuation under majority voting. On the other hand, any judgement aggregation function satisfying `STEM` has to only yield collective valuations satisfying `STEM` under *any* individual valuations, a much stronger requirement.

On the level of aggregation functions, it has been shown in [8] that any logically consistent systematic aggregation function (i.e., any systematic aggregation function which always produces logically consistent collective valuations) must be a dictatorship, provided that $|\Phi_0| > 1$. As a consequence, the class of logically consistent systematic aggregation functions cannot be axiomatised in our language either, given that $\mathcal{DIC}^*$ cannot be axiomatised. The requirement of logical consistency is essential here, since $\mathcal{DIC}^* \neq \mathcal{SYS}^*$ if $|\Phi_0| > 1$.

The following result finally formalises the intuition that the axiomatisations obtained in section 4 characterise the respective classes of aggregation functions *with respect to the language considered.* Given a class of aggregation functions $\mathcal{C}$ and a collective formula φ, let $\mathcal{C} \models \varphi$ denote that for all $A \in \mathcal{C}$, $A \models \varphi$.

Theorem 11

1. *$\mathcal{SYS}^* \models \varphi$ iff $\mathtt{E} \models \varphi$,*
2. *$\mathcal{CON}^* \models \varphi$ iff $\mathtt{KD} \models \varphi$,*
3. *$\mathcal{DIC}^* \models \varphi$ iff $\mathtt{MCY} \models \varphi$, and*
4. *$\mathcal{MAJ}^* \models \varphi$ iff $\mathtt{STEM} \models \varphi$.*

Proof. For $\mathcal{CON}^*$, if $\mathcal{CON}^* \not\models \varphi$, there must be some $v \in \mathcal{CON}^*{\downarrow}$ such that $v(\varphi) = 0$. By theorems 6 and 9, $v(\mathtt{KD}) = 1$. Conversely, if $\mathtt{KD} \not\models \varphi$, then there is some collective valuation $v \in \mathcal{CON}^*{\downarrow}$ such that $v(\varphi) = 0$. So there is an aggregation function $A \in \mathcal{CON}^*$ and individual valuations $v_1, \ldots, v_n$ such that $A(v_1, \ldots, v_n) = v$. Hence, $\mathcal{CON}^* \not\models \varphi$. Analogously for the other classes of aggregation functions. ⊣

The results on axiomatising collective valuations vs. aggregation functions allow us to make a number of methodological points regarding axiomatisation results in social choice theory. Since axiomatisations of aggregation functions imply axiomatisations of collective valuations (theorem 9), an axiomatisation of the former type is preferable to one of the latter type. Still, as theorem 10 demonstrates, there are cases where only a weaker form of axiomatisation can be obtained, i.e., where only the class of models $\mathcal{C} = \mathcal{C}^*{\downarrow}$, but not its associated class of aggregation functions $\mathcal{C}^*$ is axiomatisable. In general, when a class of aggregation functions $\mathcal{C}^*$ is not axiomatisable, one has two methodological escape routes. One option is to axiomatise $\mathcal{C}^*{\downarrow}$ as we have done here. A second option is to enlarge the language used to express properties of the aggregation functions under considerations. Clearly, if we were able to express also properties of the individual valuations and how they relate, we may be able to axiomatise a class of aggregation functions. In fact, at the extreme, with a very expressive language, such an axiomatisation becomes entirely trivial, since we may then simply write down the definition of, for instance, majority voting in our logical language.

The central question thus becomes which language we want to use to talk about judgement aggregation. If we are only interested in expressing properties of collective decision making, of the aggregated judgements, then the simple language we have used is entirely appropriate, as theorem 11 illustrates, also because any other language would have to include it. In that case, the models we used are also appropriate, since they capture precisely the amount of information needed to define a semantics for our language, and so there is no need to turn to aggregation functions.

While the question of language might appear to be somewhat abstract, it often can be answered on rather pragmatic grounds. In actual situations of judgement aggregation, we have access to certain kinds of information but not to others. In one situation, we may only learn for which propositions there is a majority. In others, we may also know the size of the majority, and in a third type of situation, we may also know who voted for and against the propositions. The language we have used here is appropriate for the first kind of situation and will also be a fragment of the language of the other two. The second type of situation would enrich the language by formulas of the form $[n]\varphi$ denoting that n people voted for φ. In the third scenario, we would even be able to express that individual i has voted for φ. Hence, the particular judgement aggregation scenario in question should determine which language, and hence also which kinds of models to use.

Standard axiomatisation results in social choice theory such as May's axiomatisation of majority voting [5] leave the axiomatic language implicit. And while subsequent axiomatisations of majority voting have opted for different axioms, e.g., avoiding May's positive responsiveness condition, there does not seem to be a discussion which problematises the axiomatic language, something that is of importance in much research in formal logic. May's axiom of anonymity, for instance, uses higher order quantification over permutation functions, and hence would usually require a rather expressive logical formalism. One advantage of making the axiomatic language explicit is that it allows for negative results like theorem 10 which show that something cannot be axiomatised within a particular language.

6 Conclusion

The present paper has provided axiomatisations of majority voting, consensus voting and dictatorship in the context of judgement aggregation. The axiomatisations are formulated in a common minimal language which can only express whether a proposition was accepted by the collective or not. This minimal language does not suffice to characterise these voting methods completely. Rather, the axiomatisations characterise the voting methods only w.r.t. the properties expressible within the language.

With respect to majority voting, the result obtained differs both from the classic result of May [5] and from more recent attempts to define a modal logic of majority [7]. The logic of majority considered in [7] aims at adding a majority operator to graded modal logic. The models considered are standard Kripke models rather than the models considered here, and infinite models are allowed. Furthermore, the language considered is much richer than the minimal language considered here, and no comparison between different voting procedures is made. As a consequence, the results obtained are at present more relevant to modal logic than to the axiomatisation of judgement aggregation considered here. Compared to May's result, on the other hand, we focus on judgement aggregation rather than preference aggregation. Furthermore, our aim is not to characterise majority voting completely, but only w.r.t. the properties expressible in a natural logical language. The language sensitivity offered by this logical style of axiomatisation allows one to obtain negative results like theorem 10, and more generally allows one to compare different axiomatisations on the level of linguistic expressiveness

There are a number of open issues arising out of the results presented in this paper. First, theorems 3 and 6 require a minimum number of individuals to be present. If there are too few individuals, additional properties hold, as we showed. In other words, the axiomatisation given is complete only for a big group of individuals. While these two result provide a bound, theorem 8 is really a result in the limit. A closer analysis of the proof should yield a bound here as well. In general, it remains to find out whether any *interesting* additional axioms become valid for small groups of individuals.

Second, the results obtained are weak axiomatisation results as explained in section 5. What are minimal languages for obtaining a strong axiomatisation of, e.g., the aggregation functions associated to majority voting? The present paper has taken the approach of syntactic minimalism paired with a formal-logic approach to axiomatisation, the advantage of the latter being that axiomatisations become possible even for languages which fail to completely characterise the voting procedure in question. Still, richer languages may yield stronger completeness result, more along the traditional lines of social choice theory. In any case, the language used for expressing axioms is crucial.

Third, one might want to consider 3-valued rather than 2-valued logic. This would allow individuals to abstain on propositions, where the third truth value represents abstention. While this seems natural enough at first sight, the problem arises how to define the truth value of complex propositions. Different definitions exist in 3-valued logic for how to define conjunction, implication, etc., and it is not clear a priori how an abstention on φ should influence the voting behaviour on $\varphi \rightarrow \psi$. This may depend on the voter's reasons for abstaining from φ, or on what exactly the voting procedure is supposed to elicit. If we are aiming to elicit voters' beliefs, voting on complex propositions will not be truth-functional at all, and even 3-valued logic will not help us. In that case, one would have to turn to a much more complex model capturing the voters' beliefs. To complicate the picture yet further, individual decision making may not be uniform across all individuals: one individual may apply a truth table of 3-valued logic, another may vote non-truth-functionally based on his belief, etc. So while the simple 2-valued approach does not allow for abstentions, it is a natural first step to consider before moving to philosophically and technically much more challenging territory. It will be interesting to see how stable the axiomatisations obtained for the simple 2-valued framework will turn out to be once we move to more complex and more realistic models.

Acknowledgements

I would like to thank Bill Zwicker for his permission to incorporate theorem 2 together with its proof into the paper, and Christian List and Josep Freixas for their comments and suggestions.

Appendix

The following theorem generalises the result in [9] from simple games to simple games with ties. The proof closely follows [11].

Theorem 2 *For any simple game with ties $G = (N, W, T, L)$, the following are equivalent:*
(i) G is weighted.
(ii) G is trade robust.
(iii) G is 2^k-trade robust, with $k = 2^{|N|}$.

Proof. $(i) \Rightarrow (ii)$: Consider a weighted game $G = (N, W, T, L)$ with weight function w and quota q. Suppose that there are two sequences $\langle X_1, \ldots, X_k \rangle$ and $\langle Y_1, \ldots, Y_k \rangle$ which satisfy the antecedent of the trade-robustness condition. Since $w(X_i) \geq q$ and $w(Y_i) \leq q$ for every i, we have $q \leq \frac{1}{k} \sum_{i \leq k} w(X_i) = \frac{1}{k} \sum_{i \leq k} w(Y_i) \leq q$, where the equality is due to the fact $|\{i : p \in X_i\}| = |\{i : p \in Y_i\}|$ for every $p \in N$. Hence, the average weight of the Xs and the average weight of the Ys are both q. But then if some X_i were actually winning, the average weight of the Xs would have to be strictly bigger than q, analogously if some Y_i were actually losing.

The proof that $(ii) \Rightarrow (iii)$ is trivial, so it only remains to show that $(iii) \Rightarrow (i)$. We inductively construct the weighting in such a way that at each stage, the unweighted part $N - A$ acts like it is reasonably trade robust, and the weighted part A behaves as if it were part of a correct weighting. The following definition formalises this.

DEFINITION: Suppose that $G = (N, W, T, L)$ is a simple game with ties where $N = \{1, \ldots, n\}$. If $A \subseteq N$ and $f : A \to \mathbb{R}$, then we call f *trade robust for* A iff the following holds: Whenever $k \leq 2^s$, where $s = 2^{|N|-A} - 1$, and $\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle$ and $\langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle$ are two sequences of (not necessarily distinct) coalitions satisfying

(1) $\forall i \leq k : \; X_i \cap Y_i = X_i' \cap Y_i' = \emptyset$
(2) $\forall i \leq k : \; Y_i, Y_i' \subseteq A$
(3) $\sum_{i=1}^{k} \sum_{p \in Y_i} f(p) \; \leq \; \sum_{i=1}^{k} \sum_{p \in Y_i'} f(p)$
(4) $\forall p \in N : \; |\{i : p \in X_i\}| = |\{i : p \in X_i'\}|$
(5) $\forall i \leq k : \; X_i \cup Y_i \in W \cup T$ and $X_i' \cup Y_i' \in L \cup T$,
then
(A) $\forall i \leq k : \; X_i \cup Y_i \in T$
(B) $\forall i \leq k : \; X_i' \cup Y_i' \in T$
(C) $\sum_{i=1}^{k} \sum_{p \in Y_i} f(p) \; = \; \sum_{i=1}^{k} \sum_{p \in Y_i'} f(p)$.

Note first that if G is $2^{2^{|N|}}$ trade robust, then the empty function is trade robust for $A = \emptyset$. Second, if G is trade robust for $A = N$, then G must be weighted. To see this, observe that G is weighted if winning coalitions weigh more than both tied and losing coalitions, tied coalitions way more than losing coalitions, and all tied coalitions have the same weight. In this case, one may choose the quota q as the weight of a tied coalition, thereby showing that G is weighted. Now if G is trade robust for $A = N$, we can apply the definition for $k = 1$ to the sequences $\langle \emptyset \cup Y_1 \rangle$ and $\langle \emptyset \cup Y_1' \rangle$ in the cases just mentioned, e.g., with $Y_1 \in W$ and $Y_1' \in T$. Given these observations, in order to prove the theorem, it suffices to show the following:

MAIN CLAIM: Suppose that $G = (N, W, T, L)$ is a simple game with ties, $A \subseteq N$, f is trade robust for A and $a \in N - A$. Then there exists a $c \in \mathbb{R}$ such that $f \cup \{(a, c)\}$ is trade robust for $A \cup \{a\}$.

The idea behind the following argument is that when a real number fails to be an appropriate weight for the new individual a to be added, it can be classified as being either too light or too heavy. After showing that every failure is indeed either a low or a high failure (but never both), a number of claims are proved about these failures. In the end, we will have shown the existence of a weight which is less than all the high failures and greater than all the low failures. We now proceed to make this argument precise.

We call a number $c \in \mathbb{R}$ a *low failure of type A* if there exist sequences $\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle$ and $\langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle$ showing that $f \cup \{(a, c)\}$ is not trade robust for $A \cup \{a\}$ such that $|\{i : a \in Y_i\}| > |\{i : a \in Y_i'\}|$, clauses (1)-(5) of the definition are satisfied, and clause (A) of the definition is violated. Intuitively, c is chosen too low as a weight for a, and this is exploited in the witnessing sequences by using a excessively in the winning coalitions. Low failures of type B and C are defined similarly; a number can be a low failure of more than one type. Finally, we analogously define a *high failure* of type A, B and C, if violating sequences exist with the inequality reversed.

CLAIM 1: Every failure is either a high or a low failure.
Proof: Suppose c is a failure, but neither high nor low. Then if two sequences of coalitions witness the failure, there is an equal number of a among the Y_i and the Y_i'. But then the occurrences of a can be shifted from Y_i to X_i and from Y_i' to X_i' with conditions (3) and (4) of the definition preserved. This would show that f is not trade robust for A, a contradiction.

CLAIM 2: $c \in \mathbb{R}$ is a low failure iff there are witnessing sequences for which a occurs in none of the Y_i'.
Proof: Given any two sequences showing that c is a low failure, shift each occurrence of a among the Y_i' to X_i', and an equal number of occurrences of a among Y_i to X_i. Again, this modification preserves (3) and (4) of the definition, so the resulting sequences are as desired and still witness the low failure of c.

CLAIM 3: $c \in \mathbb{R}$ is a high failure iff there are witnessing sequences for which a occurs in none of the Y_i.
Proof: Analogous to the proof of claim 2.

CLAIM 4: If $c \in \mathbb{R}$ is a low failure and $c' < c$, then c' is also a low failure.
Proof: Sequences witnessing a low failure c will also witness a low failure c', given that condition (3) is preserved by decreasing the failure.

CLAIM 5: If $c \in \mathbb{R}$ is a high failure and $c' > c$, then c' is also a high failure.
Proof: Analogous to the proof of claim 4.

CLAIM 6: No failure can be both a high and a low failure.
Proof: Suppose that c is a low failure of type α as witnessed by

$$\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle \text{ and } \langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle,$$

chosen as in claim 2, and a high failure of type β as witnessed by

$$\langle U_1 \cup V_1, \ldots, U_l \cup V_l \rangle \text{ and } \langle U_1' \cup V_1', \ldots, U_l' \cup V_l' \rangle$$

chosen as in claim 3. Note also that $k, l \leq 2^z$ with $z = 2^{|N-(A\cup\{a\})|} - 1$. Let $|\{i : a \in Y_i\}| = s$ and $|\{i : a \in V_i'\}| = t$. We now construct two new sequences of coalitions, first the sequence of unprimed coalitions where we repeat each unprimed coalition $X_i \cup Y_i$ t times and $U_i \cup V_i$ s times, second the sequence of primed coalitions where we repeat each primed coalition $X_i' \cup Y_i'$ t times and $U_i' \cup V_i'$ s times. In these new sequences, a occurs $s \cdot t$ times among both Y_i and V_i', and not at all among Y_i' and V_i. Hence, we can shift a from the Y_is to the

X_is and from the V_i's to the U_i's while preserving conditions (3) and (4) of the definition. Also, a violation of condition (A), (B), or (C) by the original sequences will carry over to the new combined sequences. Finally, the length of the new sequences is at most $2 \cdot (2^z)^2$, where $z = 2^{|N-(A\cup\{a\})|} - 1$ which is equivalent to 2^w with $w = 2^{|N-A|} - 1$. Consequently, f cannot be trade robust for A, a contradiction.

CLAIM 7: Let $c = 2 \cdot 2^{2^{|N|}} \cdot \sum_{p\in A} |f(p)|$. If $c' > c$, then c' is not a low failure. If $c' < -c$, then c' is not a high failure
Proof: Assume that c' is a low failure with $c' > c$ and witnessing sequences chosen as in claim 2. Since a appears among the Y_i but not among the Y_i',

$$\sum_{i=1}^{k} \sum_{p\in Y_i} f(p) \ > \ c - 2^{2^{|N|}} \sum_{p\in A} |f(p)| \ = \ 2^{2^{|N|}} \sum_{p\in A} |f(p)| \ \geq \ \sum_{i=1}^{k} \sum_{p\in Y_i'} f(p),$$

thus contradicting condition (3). The argument for the case where $c' < -c$ is analogous.

CLAIM 8: The low failures are bounded above, and hence there is a supremum c_L of the low failures. Analogously, the high failures are bounded below, and hence there is an infimum c_H of the high failures.
Proof: A direct consequence of claim 7.

CLAIM 9: If c_L is the supremum of the low failures of type A, then c_L is itself a low failure of type A. The same holds for low failures of type B.
Proof: Since we have a bound on the length of coalition sequences, only finitely many can witness failures. Hence, there must be a sequence $c_1 < c_2 < \ldots$ converging on c_L for which there is a single pair of coalition sequences witnessing that each c_i is a low failure of type A. But this means that we have $\sum_{i=1}^{k} \sum_{p\in Y_i} f(p) \leq \sum_{i=1}^{k} \sum_{p\in Y_i'} f(p)$ for each $c_j = f(a)$, and so the inequality must still hold for $f(a) = c_L$. Hence, c_L must be a low failure of type A. The case for failures of type B is analogous.

CLAIM 10: If c_L is the supremum of the low failures of type C, then c_L is itself not a low failure of type C. Instead, there is a *lower margin witness* pair of coalition sequences for $A \cup \{a\}$ and $f \cup \{(a, c_L)\}$ satisfying conditions (1)-(5) of the definition such that $|\{i : a \in Y_i\}| > |\{i : a \in Y_i'\}|$.
Proof: If c_L were a low failure of type C, then by continuity, there is also some sufficiently small $\varepsilon > 0$ such that $c_L + \varepsilon$ is also a low failure of type C, contradicting our assumption that c_L is the supremum. Like in claim 9, there must be a sequence $c_1 < c_2 < \ldots$ converging on c_L for which there is a single pair of coalition sequences witnessing that each c_i is a low failure of type C. This now means that we have $\sum_{i=1}^{k} \sum_{p\in Y_i} f(p) < \sum_{i=1}^{k} \sum_{p\in Y_i'} f(p)$ for each $c_j = f(a)$, and so the weak inequality $\sum_{i=1}^{k} \sum_{p\in Y_i} f(p) \leq \sum_{i=1}^{k} \sum_{p\in Y_i'} f(p)$ must still hold for $f(a) = c_L$. Since c_L, is not a low failure of type C, this inequality cannot be strict for $f(a) = c_L$. Hence, the pair of coalition sequences is a lower margin witness.

CLAIM 11: If c_H is the infimum of the high failures of type A, then c_H is itself a high failure of type A. The same holds for high failures of type B.
Proof: Analogous to the proof of claim 9.

CLAIM 12: If c_H is the infimum of the high failures of type C, then c_H is itself not a high failure of type C. Instead, there is an *upper margin witness* pair of coalition sequences

for $A \cup \{a\}$ and $f \cup \{(a, c_H)\}$ satisfying conditions (1)-(5) of the definition such that $|\{i : a \in Y_i\}| < |\{i : a \in Y_i'\}|$.
Proof: Analogous to the proof of claim 10.

To complete the proof of our main claim, we consider four different cases depending on whether c_L and c_H are themselves failures.

Case (i) Assume that c_L is a low failure and c_H is a high failure. Then the high failures constitute an interval closed on the left and the low failures constitute a disjoint interval closed on the right. Hence, there is an open interval between the two which contains real numbers that are not failures and that can be chosen as a suitable weight for a.

Case (ii) Assume that c_L is not a low failure and c_H is not a high failure. Then the low failures constitute an interval open on the right, and the high failures a disjoint interval open on the left. Hence, there is a nonempty closed interval between the two, and so there must be at least one point which is not a failure.

Case (iii) Assume that c_L is not a low failure but c_H is a high failure. Then c_H must be a high failure of type A or B, and c_L must be the supremum of low failures of type C. It suffices to show that $c_L < c_H$ for obtaining a nonempty half open interval of values which are not failures. In order to obtain a contradiction, suppose that $c_L = c_H$. Suppose that

$$\langle X_1 \cup Y_1, \ldots, X_k \cup Y_k \rangle \text{ and } \langle X_1' \cup Y_1', \ldots, X_k' \cup Y_k' \rangle,$$

chosen as in claim 2, serve as a lower margin witness for c_L, and

$$\langle U_1 \cup V_1, \ldots, U_l \cup V_l \rangle \text{ and } \langle U_1' \cup V_1', \ldots, U_l' \cup V_l' \rangle,$$

chosen as in claim 3, witness that c_H is a failure of type A or B. Now apply the construction of claim 6, multiplying, combining and shifting a from the Y_i to the X_i and from V_i' to the U_i'. The combined system satisfies conditions (1)-(5) of the definition, and fails to satisfy (A) or (B), depending on the type of failure of c_H. Consequently, f cannot be trade robust for A, a contradiction.

Case (iv) Assume that c_L is a low failure but c_H is not a high failure. Then the proof is analogous to case (iii). ⊣

References

[1] M. Anthony. Boolean functions and artificial neural networks. CDAM Research Report LSE-CDAM-2003-01, London School of Economics, 2003.

[2] B. Chellas. *Modal Logic - An Introduction.* Cambridge University Press, 1980.

[3] C. List. http://personal.lse.ac.uk/LIST/doctrinalparadox.htm.

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

[5] K.O. May. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20:680–684, 1952.

[6] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* John Wiley and Sons, 1944.

[7] Eric Pacuit and Samer Salame. Majority logic. In *Proceedings of the 9th Intl. Conference on Principles of Knowledge Representation and Reasoning (KR)*, pages 598–605, 2004.

[8] M. Pauly and M. van Hees. Logical constraints on judgement aggregation. *Journal of Philosophical Logic*, ??(?):??, ???

[9] A. Taylor and W. Zwicker. A characterization of weighted voting. *Proceedings of the American Mathematical Society*, 115(4):1089–1094, 1992.

[10] A. Taylor and W. Zwicker. *Simple Games: Desirability Relations, Trading, Pseudoweightings.* Princeton University Press, 1999.

[11] W. Zwicker. Personal communication, 2005.