Lifting Rationality Assumptions in Binary Aggregation

Umberto Grandi and **Ulle Endriss**
Institute for Logic, Language and Computation (ILLC)
University of Amsterdam
The Netherlands
u.grandi@uva.nl, ulle.endriss@uva.nl

Abstract

We consider problems where several individuals each need to make a yes/no choice regarding a number of issues and these choices then need to be aggregated into a collective choice. Depending on the application at hand, different combinations of yes/no may be considered *rational*. We can describe such rationality assumptions in terms of a propositional formula. The question then arises whether or not a given aggregation procedure will *lift* the rationality assumptions from the individual to the collective level, i.e., whether the collective choice will be rational whenever all individual choices are. To address this question, for each of a number of simple fragments of the language of propositional logic, we provide an axiomatic characterisation of the class of aggregation procedures that will lift all rationality assumptions expressible in that fragment.

Introduction

Social choice theory, the study of methods for collective decision making, has recently received a lot of attention in AI. There are very good reasons for this trend. First, the methods of AI (and, more generally, of Computer Science) have turned out to be useful to deepen our understanding of social choice and, in some cases, can even suggest an entirely new perspective on classical problems. Examples include the complexity-theoretic analysis of optimisation problems arising in social choice and the adaptation of knowledge representation languages to support modelling of preferences. Second, the methods of social choice theory have important applications in AI, e.g., to achieve consensus amongst the autonomous software agents in a multiagent system or to aggregate the output of several search engines. Of particular interest to AI is the case of *social choice in combinatorial domains*, in which the space of alternatives from which the individuals have to choose has a multi-attribute structure (Chevaleyre et al. 2008).

In this paper, we analyse the problem of *binary aggregation*, which is an example for social choice in (boolean) combinatorial domains. In our model, a group of individuals each make a yes/no choice regarding a finite number of issues and then these choices need to be aggregated into a collective choice. This model goes back to work of (Wilson 1975) and has recently been studied in depth by (Dokow and Holzman 2008). The framework of *judgment aggregation* (List and Puppe 2009) is closely related to binary aggregation and we will occasionally refer to it.

(Dokow and Holzman 2008) characterise domains of aggregation over which every independent and unanimous procedure is dictatorial. This is a good example for the use of the axiomatic method in economic theory: the aim is to identify *the* appropriate set of axioms (e.g., to model real-word economies, specific moral ideals, etc.) and then to prove a characterisation (or impossibility) result for those axioms. AI suggests an alternative approach: with every new application the principles underlying a system may change; so we may be more interested in devising languages for expressing a range of different axioms rather than identifying the "right" set of axioms; and we may be more interested in developing methods that will help us to understand the dynamics of a range of different social choice scenarios rather than in technical results for a specific such scenario.

For this purpose we separate two parameters in the framework of binary aggregation. On the one hand, we introduce a propositional language to define the domain of aggregation by expressing a rationality assumption common to all individuals. On the other, we state a list of axioms to classify aggregation procedures over these domains. We call an aggregation procedure *collectively rational* with respect to a language if whenever all individuals submit ballots satisfying a formula in the language, so does the outcome of aggregation. We characterise, for several simple fragments of the language of propositional logic, the associated class of collectively rational procedures as the set of procedures satisfying a certain set of axioms. Towards the end of the paper, we relate our results to existing aggregation frameworks and we state future directions of research.

Basic definitions

Let $\mathcal{I} = \{1, \ldots, m\}$ be a finite set of *issues*, and let $\mathcal{D} = D_1 \times \cdots \times D_m$ be a boolean combinatorial *domain*, i.e., $|D_i| = 2$ for all $i \in \mathcal{I}$ (we assume $D_i = \{0, 1\}$). Let $PS = \{p_1, \ldots, p_m\}$ be a set of boolean variables/propositional symbols, one for each issue, and let $\mathcal{L}_{PS}$ be the corresponding propositional language. For any $\varphi \in \mathcal{L}_{PS}$, let $\mathrm{Mod}(\varphi)$ be the set of *models* that satisfy φ. For example, $\mathrm{Mod}(p_1 \wedge \neg p_2) = \{(1,0,0),(1,0,1)\}$ if $PS = \{p_1, p_2, p_3\}$.

A *language for integrity constraints* is a set of propositional formulas $\mathcal{L} \subseteq \mathcal{L}_{PS}$. Examples include the set of *literals*, the set of *cubes* (conjunctions of literals), and the set of *clauses of size at most* k. For a given language $\mathcal{L}$, any *integrity constraint* $\mathrm{IC} \in \mathcal{L}$ defines a *domain of aggregation* $\mathrm{Mod}(\mathrm{IC})$, which we shall often refer to as X.[1]

Integrity constraints can be used to define what tuples in $\mathcal{D}$ we consider *rational* choices. For example, as we shall explain in our discussion of related work at the end of this paper, $\mathcal{D}$ might be used to encode a binary relation, in which case we may want to declare only those elements of $\mathcal{D}$ rational that correspond to relations that are transitive. In the sequel, we shall therefore use the terms "integrity constraints" and "rationality assumptions" interchangeably.

Let $N = \{1, \ldots, n\}$ be a finite set of *individuals*. To simplify presentation, we shall assume that the number of individuals n is odd. A *ballot* B is an element of $\mathcal{D}$ (i.e., an assignment to the variables $p_1, \ldots, p_m$); and a *rational ballot* B is an element of $\mathcal{D}$ that satisfies the integrity constraints, i.e., an element of $\mathrm{Mod}(\mathrm{IC})$. A *profile* $\underline{B}$ is a vector of (rational) ballots, one for each individual in N. We write B_j for the jth element of a ballot B, and $\underline{B}_{i,j}$ for the jth element of ballot $\underline{B}_i$ within a profile $\underline{B} = (\underline{B}_1, \ldots, \underline{B}_n)$.

An *aggregation procedure* is a function $F : \mathcal{D}^N \to \mathcal{D}$, mapping each profile to an element of the domain $\mathcal{D}$. We are now ready to define one of the central concepts for this paper, *collective rationality* wrt. IC:

Definition 1. *An aggregation procedure $F : \mathcal{D}^N \to \mathcal{D}$ is called collectively rational (CR) for* IC*, if for all profiles $\underline{B} \in \mathrm{Mod}(\mathrm{IC})^N$ we have that $F(\underline{B}) \in \mathrm{Mod}(\mathrm{IC})$.*

Thus, F is CR if it can *lift* the rationality assumptions given by IC from the individual to the collective level.

Axioms

In social choice theory, aggregation procedures are studied using the axiomatic method. Axioms are used to express desirable properties of a procedure. In this section, we adapt the most important axioms familiar from standard social choice theory, and more specifically from judgment aggregation (List and Puppe 2009) and binary aggregation theory (Dokow and Holzman 2008), to our setting. We start with four common axioms:

Unanimity (U): For any profile $\underline{B} \in X^N$ and any $x \in \{0, 1\}$, if $\underline{B}_{i,j} = x$ for all $i \in N$, then $F(\underline{B})_j = x$.

Anonymity (A): For any profile $\underline{B} \in X^N$ and any permutation $\sigma : N \to N$, we have that $F(\underline{B}_1, \ldots, \underline{B}_n) = F(\underline{B}_{\sigma(1)}, \ldots, \underline{B}_{\sigma(n)})$

Issue-Neutrality ($\mathrm{N}^{\mathcal{I}}$): For any two issues $j, j' \in \mathcal{I}$ and any profile $\underline{B} \in X^N$, if for all $i \in N$ we have that $\underline{B}_{i,j} = \underline{B}_{i,j'}$, then $F(\underline{B})_j = F(\underline{B})_{j'}$.

Independence (I): For any issue $j \in \mathcal{I}$ and any two profiles $\underline{B}, \underline{B}' \in X^N$, if $\underline{B}_{i,j} = \underline{B}'_{i,j}$ for all $i \in N$, then $F(\underline{B})_j = F(\underline{B}')_j$.

[1] This definition is consistent with that of (Dokow and Holzman 2008) since every subset of $\mathcal{D}$ is of the form $\mathrm{Mod}(\varphi)$ for a certain propositional formula $\varphi \in \mathcal{L}_{PS}$.

Unanimity postulates that, if all individuals agree on issue j, then the aggregation procedure should implement that choice for j. Anonymity requires the procedure to be symmetric with respect to individuals. Issue-neutrality (a variant of the standard axiom of neutrality introduced in the literature on judgment aggregation) asks that the procedure be symmetric with respect to issues. Finally, independence requires the outcome of aggregation on a certain issue j to depend only on the individual choices regarding that issue. Combining independence with issue-neutrality, we get the axiom of systematicity (S) = (I) + ($\mathrm{N}^{\mathcal{I}}$).

It is important to remark that all axioms are domain-dependent. For instance, many aggregation procedures, such as the majority rule, are independent over the full combinatorial domain $\mathcal{D}$, while others, such as the one presented in the next example, are not. With two issues, let $\mathrm{IC} = (p_2 \to p_1)$ and let F be equal to the majority rule on the first issue, and accept the second issue only if the first one was accepted and the second one has the support of a majority of the individuals. This procedure is not independent on the full domain, but it is easy to see that it satisfies independence when restricted to $X^N = \mathrm{Mod}(\mathrm{IC})^N$.

As a generalisation of the axiom of neutrality introduced by (May 1952), we introduce the following:

Domain-Neutrality ($\mathrm{N}^{\mathcal{D}}$): For any two issues $j, j' \in \mathcal{I}$ and any profile $\underline{B} \in X^N$, if $\underline{B}_{i,j} = 1 - \underline{B}_{i,j'}$ for all $i \in N$, then $F(\underline{B})_j = 1 - F(\underline{B})_{j'}$.

The two notions of neutrality are uncorrelated but dual: issue-neutrality requires the outcome on two issues to be the same if all individuals agree on these issues; domain-neutrality requires it to be reversed if all the individuals make opposed choices on the two issues.

The following axiom of monotonicity is often called *positive responsiveness*, and is formulated as an (inter-profile) axiom for independent aggregation procedures:[2]

I-Monotonicity (M^{I}): For any issue $j \in \mathcal{I}$ and any two profiles $\underline{B} = (\underline{B}_1, \ldots, \underline{B}_i, \ldots, \underline{B}_n)$ and $\underline{B}' = (\underline{B}_1, \ldots, \underline{B}'_i, \ldots, \underline{B}_n)$ in X^N, if $\underline{B}_{i,j} = 0$ and $\underline{B}'_{i,j} = 1$, then $F(\underline{B})_j = 1$ entails $F(\underline{B}')_j = 1$.

Every set of axioms identifies a class of aggregation procedures that satisfy these properties. A characterisation in mathematical terms can be obtained for some classes. One example is the class of *quota rules* $\mathcal{QR}$ introduced by (Dietrich and List 2007): an aggregation procedure F for n individuals is a quota rule if for every issue j there exists a quota $0 \leqslant q_j \leqslant n+1$ such that, if we denote by $N_j^{\underline{B}} = |\{i \mid \underline{B}_{i,j}=1\}|$, then $F(\underline{B})_j=1$ if and only if $N_j^{\underline{B}} \geqslant q_j$. The following representation result holds:

Proposition 1 (Dietrich and List, (Dietrich and List 2007)). *An aggregation procedure F satisfies* A*,* I*, and* M^{I} *if and only if it is a quota rule.*

A quota rule is called *uniform* if the quota is the same for all issues. By adding the axiom of issue-neutrality to Proposition 1 we get an axiomatisation of this class. The uniform

[2] A variant of this axiom for issue-neutral aggregators has been defined in previous work (Endriss, Grandi, and Porello 2010).

quota rule with $q_j = \lceil \frac{n}{2} \rceil$ for all issues j is the *majority rule*. It is interesting to link these results with May's Theorem ((May 1952)) on the axiomatic characterisation of the majority rule in voting. We can obtain a more general version of his result (which deals with the case of a single issue) by adding the axiom of domain-neutrality: this forces the quota to treat $N_j^{\underline{B}}$ and $n - N_j^{\underline{B}}$ symmetrically, and thus the only possibility is to fix the quota as the majority of the individuals.

Lifting individual rationality

We now want to establish connections between aggregation procedures characterised in terms of axioms and aggregation procedures characterised in terms of languages for integrity constraints for which they are collectively rational. To this end, we first define the class of procedures that can lift the integrity constraints belonging to a given language $\mathcal{L}$ (recall Definition 1).

Definition 2. *For any language $\mathcal{L} \subseteq \mathcal{L}_{PS}$, define the class $\mathcal{CR}[\mathcal{L}]$ of aggregation procedures that lift $\mathcal{L}$:*

$$\mathcal{CR}[\mathcal{L}] \quad = \quad \{F : \mathcal{D}^N \to \mathcal{D} \mid F \textit{ is CR for all } \mathrm{IC} \in \mathcal{L}\}$$

Next, we establish some basic properties of $\mathcal{CR}[\mathcal{L}]$. In our framework, we have made the assumption of IC being a single formula (rather than a set of formulas); we now provide a formal underpinning for this choice. For any $\mathcal{L} \subseteq \mathcal{L}_{PS}$, let $\mathcal{L}^\wedge$ be the language of conjunctions of formulas in $\mathcal{L}$.

Lemma 2. *$\mathcal{CR}[\mathcal{L}^\wedge] = \mathcal{CR}[\mathcal{L}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$.*

Proof. $\mathcal{CR}[\mathcal{L}^\wedge]$ is clearly included in $\mathcal{CR}[\mathcal{L}]$, since $\mathcal{L} \subseteq \mathcal{L}^\wedge$. It remains to be shown that, if an aggregation procedure F lifts every constraint in $\mathcal{L}$, then it lifts any conjunction of formulas in $\mathcal{L}$. Let $\bigwedge_k \mathrm{IC}_k$ be such a conjunction, and let $\underline{B} \in \mathrm{Mod}(\bigwedge_k \mathrm{IC}_k)^N$ be a profile satisfying this integrity constraint. Since $\mathrm{Mod}(\bigwedge_k \mathrm{IC}_k) = \bigcap_k \mathrm{Mod}(\mathrm{IC}_k)$, we have that $\underline{B} \in \mathrm{Mod}(\mathrm{IC}_k)$ for every k. Thus, if $F \in \mathcal{CR}[\mathcal{L}]$, then $F(\underline{B}) \in \mathrm{Mod}(\mathrm{IC}_k)$ for every k. Therefore, F will also be in $\mathrm{Mod}(\bigwedge_k \mathrm{IC}_k)$, and this concludes the proof. □

In particular, we have that $\mathcal{CR}[\textit{cubes}] = \mathcal{CR}[\textit{literals}]$ and $\mathcal{CR}[\textit{clauses}] = \mathcal{CR}[\mathcal{L}_{PS}]$. The latter holds, because for every propositional formula there is an equivalent formula in conjunctive normal form (CNF).

The following lemma is an immediate consequence of our definitions:

Lemma 3. *$\mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2] = \mathcal{CR}[\mathcal{L}_1] \cap \mathcal{CR}[\mathcal{L}_2]$ for all $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathcal{L}_{PS}$.*

Next we introduce notation for defining classes of aggregation procedures in terms of axioms. As mentioned earlier, a particular axiom may be satisfied on a subdomain of interest, but not on the full domain. Here, we are interested in domains defined by means of integrity constraints (i.e., propositional formulas). We therefore need to be able to speak about the procedures that satisfy an axiom on the subdomain $\mathrm{Mod}(\mathrm{IC})^N$ induced by a given integrity constraint IC.

Let $F_{\restriction \mathrm{Mod}(\mathrm{IC})^N}$ denote the restriction of the aggregation procedure F to the subdomain $\mathrm{Mod}(\mathrm{IC})^N$.

Definition 3. *An aggregation procedure F satisfies a set of axioms* AX *wrt. a language $\mathcal{L} \subseteq \mathcal{L}_{PS}$, if for all constraints $\mathrm{IC} \in \mathcal{L}$ the restriction $F_{\restriction \mathrm{Mod}(\mathrm{IC})^N}$ satisfies the axioms in* AX. *This defines the following class:*

$$\mathcal{F}_{\mathcal{L}}[\mathrm{AX}] = \{F : \mathcal{D}^N \to \mathcal{D} \mid F_{\restriction \mathrm{Mod}(\mathrm{IC})^N} \textit{ sat. } \mathrm{AX} \textit{ for all } \mathrm{IC} \in \mathcal{L}\}$$

We write $\mathcal{F}[\mathrm{AX}]$ as a shorthand for $\mathcal{F}_{\{\top\}}[\mathrm{AX}]$, the class of procedures that satisfy AX over the *full* domain $\mathcal{D}$. It is easy to see that the following lemma holds:

Lemma 4. *$\mathcal{F}[\mathrm{AX}] \subseteq \mathcal{F}_{\mathcal{L}}[\mathrm{AX}]$ for all $\mathcal{L} \subseteq \mathcal{L}_{PS}$.*

We shall now seek to obtain results that link the two kinds of classes defined, i.e., results of the form

$$\mathcal{CR}[\mathcal{L}] \quad = \quad \mathcal{F}_{\mathcal{L}}[\mathrm{AX}],$$

for certain languages $\mathcal{L}$ and certain axioms AX.

Characterisation results

Our first characterisation result shows that the aggregation procedures that can lift all rationality constraints expressible in terms of a conjunction of literals (a cube) is precisely the class of unanimous procedures:

Proposition 5. *$\mathcal{CR}[\textit{cubes}] = \mathcal{F}_{\textit{cubes}}[\mathrm{U}]$.*

Proof. One direction is easy: If X is a domain defined by a cube, then every individual must agree on every literal in the conjunction, and, by unanimity, so will the collective. For the other direction, suppose that $F \in \mathcal{CR}[\textit{cubes}]$. Fix $j \in \mathcal{I}$. Pick a profile $\underline{B} \in \mathcal{D}^n$ such that $\underline{B}_{i,j} = 1$ (or 0) for all $i \in N$. That is, $\underline{B} \in \mathrm{Mod}(p_j)^N$ (or $\neg p_j$, respectively). Since F is collectively rational on every domain defined by a cube (and this includes literals), it must be the case that $F(\underline{B})_j = 1$ (or 0, respectively), proving unanimity of the aggregator. □

Observe that, as $\mathcal{F}_{\textit{cubes}}[\mathrm{U}] = \mathcal{F}[\mathrm{U}]$, the explicit mentioning of cubes on the righthand side of Proposition 5 is not needed. The statement takes therefore the following form: an aggregation procedure lifts cubes if and only if it is unanimous. That is, this result can be interpreted as a characterisation of the axiom of unanimity in terms of collective rationality with respect to the language of cubes. But since our starting point here is a logical language to express integrity constraints, we chose above form of presentation to focus more on determining minimal conditions for an aggregator to lift constraints of a certain form.[3]

By Lemma 2, we also get $\mathcal{CR}[\textit{literals}] = \mathcal{F}_{\textit{literals}}[\mathrm{U}]$ (it is easy to see that $\mathcal{F}_{\textit{literals}}[\mathrm{U}] = \mathcal{F}_{\textit{cubes}}[\mathrm{U}]$).

Let $\mathcal{L}_\leftrightarrow$ be the language of bi-implications of positive literals: $\mathcal{L}_\leftrightarrow = \{p_j \leftrightarrow p_k \mid p_j, p_k \in PS\}$. This language allows us to characterise issue-neutral aggregators:

Proposition 6. *$\mathcal{CR}[\mathcal{L}_\leftrightarrow] = \mathcal{F}_{\mathcal{L}_\leftrightarrow}[\mathrm{N}^{\mathcal{I}}]$.*

Proof. To prove the first inclusion ($\supseteq$), pick a positive bi-implication $p_j \leftrightarrow p_k$: issues j and k share the same pattern of acceptances/rejections and since the procedure is neutral over issues, we get $F(\underline{B})_j = F(\underline{B})_k$. The constraint is

[3] The same remark applies to Propositions 6 and 7 below.

therefore lifted. For the other direction ($\subseteq$), suppose that a profile $\underline{B}$ is such that $\underline{B}_{i,j} = \underline{B}_{i,k}$ for every $i \in N$. Then $\underline{B} \in \mathrm{Mod}(p_j \leftrightarrow p_k)^N$, and if F is in $\mathcal{CR}[\mathcal{L}_{\leftrightarrow}]$, then $F(\underline{B})_j$ must be equal to $F(\underline{B})_k$. Since this holds for every such $\underline{B}$, this proves that F is neutral over issues. □

Let $\mathcal{L}_{\not\leftrightarrow}$ be the language of bi-implications of one negative and one positive literal: $\mathcal{L}_{\not\leftrightarrow} = \{p_j \leftrightarrow \neg p_k \mid p_j, p_k \in PS\}$. That is, $\mathcal{L}_{\not\leftrightarrow}$ is the language of XOR-formulas over pairs of positive literals. With a proof analogous to the one above we can characterise domain-neutrality:

Proposition 7. $\mathcal{CR}[\mathcal{L}_{\not\leftrightarrow}] = \mathcal{F}_{\mathcal{L}_{\not\leftrightarrow}}[\mathrm{N}^{\mathcal{D}}]$.

Let $\mathcal{F} = \{F : \mathcal{D}^N \to \mathcal{D}\}$ be the class of *all* aggregation procedures (for fixed $\mathcal{D}$ and N). The next result is an immediate consequence of our definitions:

Proposition 8. $\mathcal{CR}[\{\bot\}] = \mathcal{CR}[\{\top\}] = \mathcal{F}$.

Hence, by Lemma 3, $\mathcal{CR}[\mathcal{L} \cup \{\bot\}] = \mathcal{CR}[\mathcal{L}]$, which shows that unsatisfiable formulas can be omitted from languages for integrity constraints.

We now move on to characterising two extreme cases of languages for integrity constraints: the case of formulas with a single model and the case of constraints in the full propositional language. A *dictatorship* is an aggregation procedure that copies in every profile the ballot of a certain fixed individual, the dictator. The class $\mathcal{F}_{\mathcal{L}}[\mathrm{DIC}]$ is composed by all functions that are dictatorships when restricted to $\mathrm{Mod}(\mathrm{IC})^N$ for all $\mathrm{IC} \in \mathcal{L}$. Note that on restricted domains this notion can differ significantly from the usual intuition of dictatorship. Now, let us call a language $\mathcal{L} \subseteq \mathcal{L}_{PS}$ *trivial*, if it is composed only of formulas having a single model each. Clearly:

Proposition 9. *If $\mathcal{L}$ is trivial, then $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\mathrm{DIC}]$.*

We propose the following definition of a class of aggregators that generalises the notion of dictatorship:

Definition 4. *An aggregation procedure $F : \mathcal{D}^N \to \mathcal{D}$ is a generalised dictatorship, if there exists a map $g : \mathcal{D}^N \to N$ such that $F(\underline{B}) = \underline{B}_{g(\underline{B})}$ for every $\underline{B} \in \mathcal{D}^N$.*

That is, a generalised dictatorship copies the ballot of a (possibly different) individual in every profile. Call this class $\mathcal{F}[\mathrm{GDIC}]$. This class fully characterises the aggregators that can lift *any* integrity constraint:

Proposition 10. $\mathcal{CR}[\mathcal{L}_{PS}] = \mathcal{F}[\mathrm{GDIC}]$.

Proof. Clearly, every generalised dictatorship lifts any arbitrary integrity constraint $\mathrm{IC} \in \mathcal{L}_{PS}$. To prove the other direction, suppose that $F \notin \mathcal{F}[\mathrm{GDIC}]$. Then there exists a profile $\underline{B} \in \mathcal{D}^N$ such that $F(\underline{B}) \neq \underline{B}_i$ for all $i \in N$. This means that for every i there exists an issue j_i such that $F(\underline{B})_{j_i} \neq \underline{B}_{i,j_i}$. Define now a literal ℓ_{j_i} to be equal to p_{j_i} if $\underline{B}_{i,j_i} = 1$, and to $\neg p_{j_i}$ otherwise. Consider as integrity constraint IC the following formula: $\bigvee_i \ell_{j_i}$. Clearly, $\underline{B}_i \models \mathrm{IC}$ for every $i \in N$, i.e., $\underline{B}$ is a rational profile for the integrity constraint IC. Since $F(\underline{B}) \not\models \mathrm{IC}$ by construction, F is not in $\mathcal{CR}[\{\mathrm{IC}\}]$ and therefore also not in $\mathcal{CR}[\mathcal{L}_{PS}]$. □

All of the characterisation results presented thus far characterise a class of procedures determined by a *single* axiom (or apply to a very specific class of procedures) and by a *uniform* description of the language. So we might ask to what extent such results can be combined to allow us to make predictions regarding the collective rationality of procedures satisfying several such axioms, or in the case where the integrity constraints can be chosen from a more complex language. To illustrate the application of our results to such cases, suppose $\mathcal{CR}[\mathcal{L}_1] = \mathcal{F}_{\mathcal{L}_1}[\mathrm{AX}_1]$ and $\mathcal{CR}[\mathcal{L}_2] = \mathcal{F}_{\mathcal{L}_2}[\mathrm{AX}_2]$. Then Lemma 3 and the fact that $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2] \subseteq \mathcal{F}_{\mathcal{L}_1}[\mathrm{AX}_1] \cap \mathcal{F}_{\mathcal{L}_2}[\mathrm{AX}_2]$ entail $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2] \subseteq \mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2]$. (But note that the other inclusion is not always true.) Now, if we start from the language $\mathcal{L}_1 \cup \mathcal{L}_2$ or any of its sublanguages, then this shows that picking procedures from $\mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2]$ is a sufficient condition for collective rationality. If, instead, we start from the axioms in AX_1 and AX_2, then we can infer that the procedures we obtain will lift any language $\mathcal{L} \subseteq \mathcal{L}_1 \cup \mathcal{L}_2$, since by Lemma 4 $\mathcal{F}[\mathrm{AX}_1, \mathrm{AX}_2] \subseteq \mathcal{F}_{\mathcal{L}_1 \cup \mathcal{L}_2}[\mathrm{AX}_1, \mathrm{AX}_2]$ which in turn is included in $\subseteq \mathcal{CR}[\mathcal{L}_1 \cup \mathcal{L}_2] \subseteq \mathcal{CR}[\mathcal{L}]$.

Negative results

For two important classes of aggregators, it is not possible to obtain a characterisation result:

Proposition 11. *There is no language $\mathcal{L} \subseteq \mathcal{L}_{PS}$ such that $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\mathrm{I}]$.*

Proof. We prove this proposition by constructing, for any choice of a language $\mathcal{L}$, an independent function that is not collectively rational for a certain $\mathrm{IC} \in \mathcal{L}$. Fix a language $\mathcal{L}$. This language will contain a falsifiable formula φ (otherwise $\mathcal{CR}[\mathcal{L}]$=$\mathcal{F}$ by Proposition 8 and we are done, as $\mathcal{F} \neq \mathcal{F}_{\mathcal{L}}[\mathrm{I}]$). Choose a ballot/model $B^\star \in \mathcal{D}$ such that $B^\star \not\models \varphi$. Then the constant function $F \equiv B^\star$ is an independent function (on the full domain) that is not collectively rational. □

Proposition 12. *There is no language $\mathcal{L} \subseteq \mathcal{L}_{PS}$ such that $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\mathrm{A}]$.*

Proof. Employing a different technique than in the previous proof, we show that for every language $\mathcal{L}$ there exists a procedure that is collectively rational but not anonymous. First, in case $\mathcal{L}$ is trivial, by Proposition 9, $\mathcal{CR}[\mathcal{L}] = \mathcal{F}_{\mathcal{L}}[\mathrm{DIC}]$, which is strictly included in the class of all anonymous functions. Second, if $\mathcal{L}$ is not trivial, then a dictatorship is always collectively rational (cf. Proposition 10), and it is not anonymous since due to nontriviality there is an $\mathrm{IC} \in \mathcal{L}$ that allows for at least two different rational ballots. □

These results are coherent with the intuition that any assumption of collective rationality of an aggregator can only condition the outcome in view of a single profile at a time, without being able to express inter-profile requirements such as anonymity and independence. Similar remarks apply to the axiom of monotonicity (note that M^{I} is meaningful only in connection with I).

Results for clauses

In view of the negative results proved above, we now focus on procedures satisfying anonymity, independence and monotonicity, and analyse the ability of procedures to lift rationality assumptions *within* that class. This enables us to obtain interesting results concerning languages of clauses.

Recall from Proposition 1 that the independent, anonymous and monotone procedures are exactly the quota rules, i.e., procedures that assign a quota q_j to every issue j such that $F(\underline{B})_j = 1 \Leftrightarrow |\{i \mid \underline{B}_{i,j} = 1\}| \geqslant q_j$. That is, in our notation, $\mathcal{QR} = \mathcal{F}[\mathrm{A,I,M^I}]$.

By Proposition 10 and Lemma 2, we know that $\mathcal{CR}[clauses]$ is the collection of generalised dictatorships. Therefore, to obtain results for more attractive classes of procedures, we restrict attention to clauses of limited length. For $k \geqslant 1$, let *k-clauses* be the set of clauses of length $\leqslant k$, *k-pclauses* the set of positive k-clauses, i.e., disjunctions where all literals are positive, and *k-nclauses* the language of negative k-clauses, where all literals are negative.

Given a clause $\varphi = \ell_1 \vee \cdots \vee \ell_k$, we say that an issue j *occurs in* φ if one and only one of p_j and $\neg p_j$ is one of the disjuncts of φ.

Lemma 13. *If* IC $\in$ *k-pclauses and* n *is the number of individuals, then every quota rule with* $q_j \leqslant \lceil \frac{n}{k} \rceil$ *for every issue* j *that occurs in* IC *is collectively rational.*

Proof. Since the clause IC is accepted by every individual, there exists a literal that is accepted by at least $\lceil \frac{n}{k} \rceil$ of them. IC is made of positive literals, therefore by restricting the quota q_j to be at most $\lceil \frac{n}{k} \rceil$, we guarantee that that literal (and the disjunction) will be lifted. □

The analogous version for negative clauses holds too:

Lemma 14. *If* IC $\in$ *k-nclauses and* n *is the number of individuals, then every quota rule with* $q_j > n - \lceil \frac{n}{k} \rceil$ *for every issue* j *that occurs in* IC *is collectively rational.*

If we denote with $\mathcal{QR}_{c(q_j)}$ the set of quota rules such that the quotas q_j satisfy the constraint in the subscript for all issues j, then by the previous lemmas we obtain the following characterisations:

Proposition 15. $\mathcal{CR}[k\text{-}pclauses] \cap \mathcal{QR} \supseteq \mathcal{QR}_{q_j \leqslant \lceil \frac{n}{k} \rceil}$.

Proposition 16. $\mathcal{CR}[k\text{-}nclauses] \cap \mathcal{QR} \supseteq \mathcal{QR}_{q_j > n - \lceil \frac{n}{k} \rceil}$.

As may easily be checked, in Propositions 15 and 16 above the inclusion is strict (only) for $k > 1$.

Let us now turn to the general case of arbitrary k-clauses. We say that an issue j occurs positively in a clause, if it does occur in that clause and the corresponding literal is positive; otherwise we say that it occurs negatively. With a similar proof as above, we can show:

Lemma 17. *Suppose* IC $\in$ *k-clauses and* n *is the number of individuals. A quota rule is collectively rational for* IC*, if* $q_j \leqslant \lceil \frac{n}{k} \rceil$ *for every issue* j *that occurs positively in* IC *and* $q_j > n - \lceil \frac{n}{k} \rceil$ *for every issue* j *that occurs negatively in* IC.

In the special case of $k = 2$, we get the following:[4]

[4]Recall the assumption that the set of individuals is odd.

Proposition 18. *The majority rule is in* $\mathcal{CR}[2\text{-}clauses]$.

Proof. The quota relative to every issue has to satisfy both types of constraints from Lemma 17. But these are incompatible unless $k = 2$, in which case $q_j = \lceil \frac{n}{2} \rceil$. □

If there are only two issues, then the majority rule can lift any kind of rationality assumption:

Corollary 19. *If there are at most two issues (*$|\mathcal{I}| \leqslant 2$*), then the majority rule is in* $\mathcal{CR}[\mathcal{L}_{PS}]$.

Proof. This follows immediately from Proposition 18 and Lemma 2, together with the observation that the CNF of any formula involving at most two distinct propositional symbols is a conjunction of 2-clauses. □

From Lemma 17, we can also extract a general method for constructing a collectively rational quota rule (if one exists), given an arbitrary constraint IC $\in \mathcal{L}_{PS}$:

- Rewrite IC in CNF;[5] call the result $\mathrm{IC}_{\mathrm{CNF}}$.
- For each issue j, write down these two constraints:
  - $q_j \leqslant \lceil \frac{n}{k} \rceil$, where k is the size of the longest clause in $\mathrm{IC}_{\mathrm{CNF}}$ in which j occurs positively.
  - $q_j > n - \lceil \frac{n}{k'} \rceil$, where k' is the size of the longest clause in $\mathrm{IC}_{\mathrm{CNF}}$ in which j occurs negatively.
- Every solution to this system defines a quota rule that is collectively rational for IC.

See (Dietrich and List 2007) for further discussion of quota rules.

Related work

The framework of binary aggregation was introduced by (Wilson 1975) and further developed by (Rubinstein and Fishburn 1986) and (Dokow and Holzman 2008). As already mentioned, this work relates closely to the results we proved here, and we already explained the novelty of our results, pertaining to a language for the syntactic specification of rationality assumptions, with respect to that line of research. In this section we review some of the most important frameworks for aggregation, and we show a natural translation between the paradoxical behaviour described in these contexts and the notion of collective rationality of an aggregation procedure.

In *preference aggregation* individuals express a linear order over a set of alternatives A. We can go back to the work of Condorcet in the 18th century to find the first occurrence of the following paradoxical situation (called Condorcet cycle). For three individuals, let their preferences be $a{>}b{>}c$, $b{>}c{>}a$ and $c{>}a{>}b$. Pairwise majority aggregation leads to accepting $a{>}b$ and $b{>}c$ but also $c{>}a$, i.e, an intransitive (hence irrational) outcome. A linear order can be encoded as a rational ballot in binary aggregation in the following way: given a set of alternatives A, introduce a boolean variable p_{ab} for every ordered pair of alternatives $a \neq b$. The condition of antisymmetry can be enforced with

[5]Note that this step may give rise to an exponential growth in the size of the formula.

the formulas $p_{ab} \leftrightarrow \neg p_{ba}$ for all $a \neq b$ and transitivity with $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$ for all a, b, c. The conjunction of these formulas form the integrity constraint IC. The Condorcet cycle presented above forms a profile $\underline{B}$ that yields an outcome where all three variables p_{ab}, p_{bc} and p_{ca} are accepted. This outcome does not satisfy IC, therefore the pairwise majority rule is not collectively rational for this IC.[6]

Not only can we express the framework of preference aggregation, but we can also write classical impossibility theorems (and potentially devise new proofs) in terms of collective rationality of aggregation. Arrow's Theorem ((Arrow 1963)), for instance, takes the following form:

$$\mathcal{CR}[\mathcal{L}_{pref}] \cap \mathcal{F}_{\mathcal{L}_{pref}}[\text{U,I,NDIC}] = \emptyset,$$

where $\mathcal{L}_{pref}$ denotes the language representing the set of linear orders and NDIC the axiom of non-dictatorship. Another example is Wilson's Theorem ((Wilson 1975)), which states that $\mathcal{F}_{\mathcal{L}_{pref}}[\text{I}]$ consists only of dictatorships, antidictatorships, and constant functions. Moreover, integrity constraints can be seen as domain restrictions: while it is likely that a possibility result can be proved by restricting the domain using propositional formulas, classical restrictions like single-peakedness are difficult to express in our framework in view of the fact that they usually are inter-profile conditions.

The model we presented is clearly very expressive in situations of *voting in boolean combinatorial domains* e.g., (Brams, Zwicker, and Kilgour 1998), (Lang 2007). To cite an example, consider the case of three agents having to decide which of 3 tasks p_1, p_2 and p_3 to fulfil, with the resource constraint that a maximum of two tasks can be supported. This is translated into the constraint $\neg(p_1 \wedge p_2 \wedge p_3)$. Suppose that agent 1 votes for the first two tasks, agent 2 for the last two, and the third one votes for the first and the last task. All agents are submitting rational ballots, but the majority rule will accept all three tasks, against the integrity constraint.

Judgment aggregation can also be expressed as a binary aggregation problem (Dokow and Holzman 2008). This suggests that many possibility results (List and Puppe 2009), as well as safety results (Endriss, Grandi, and Porello 2010), can be easily related to the characterisations established in the previous section.

Conclusions

We introduced a simple propositional language to express individual rationality constraints in the framework of boolean aggregation, and we defined an aggregation procedure to be collectively rational if the collective outcome satisfies a certain constraint whenever all individuals do. We proved several results to characterise, for various subsets of the language, a set of axioms that guarantees the collective rationality of a procedure for all constraints in this subset, and we have outlined an approach for how to apply these results in more complex situations. Finally, we showed how several existing frameworks for aggregation have a natural translation using our definitions, and we pointed out the potential of such a unified treatment.

The use of logic in this work is limited to its expressivity as a language interpreted on boolean combinatorial domains. A promising direction is to develop this use further, enabling us to exploit all the power of the logical formalism to devise new proofs of (im)possibility results in the area. This work can also be seen as a first step in the construction of a model for the more complex problem of combinatorial aggregation (Lang 2004) where the aggregation is performed over a product of arbitrary domains. Sequential voting (Lang 2007) represents a clear trend in this area: given an integrity constraint, we might be able to devise an order of aggregation over the set of issues that will guarantee the rationality of the outcome. Finally, by using more powerful languages to express rationality assumptions we can move towards more complex logical models of artificial agents.

Acknowledgements. We would like to thank three anonymous reviewers and a senior PC member for their thorough and useful comments.

References

Arrow, K. J. 1963. *Social Choice and Individual Values*. John Wiley & Sons, 2nd edition.

Brams, S. J.; Zwicker, W. S.; and Kilgour, D. M. 1998. The paradox of multiple elections. *Social Choice and Welfare* 15(2):211–236.

Chevaleyre, Y.; Endriss, U.; Lang, J.; and Maudet, N. 2008. Preference handling in combinatorial domains: From AI to social choice. *AI Magazine* 29(4):37–46.

Dietrich, F., and List, C. 2007. Judgment aggregation by quota rules: Majority voting generalized. *Journal of Theoretical Politics* 19(4):391–424.

Dokow, E., and Holzman, R. 2008. Aggregation of binary evaluations. *J. Economic Theory*. In press.

Endriss, U.; Grandi, U.; and Porello, D. 2010. Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of AAMAS-2010*.

Lang, J. 2004. Logical preference representation and combinatorial vote. *Annals of Mathematics and Artificial Intelligence* 42(1–3):37–71.

Lang, J. 2007. Vote and aggregation in combinatorial domains with structured preferences. In *Proceedings of IJCAI-2007*.

List, C., and Puppe, C. 2009. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press.

May, K. O. 1952. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica* 20(4):680–684.

Rubinstein, A., and Fishburn, P. C. 1986. Algebraic aggregation theory. *J. Economic Theory* 38(1):63–77.

Wilson, R. B. 1975. On the theory of aggregation. *J. Economic Theory* 10(1):89–99.

[6] A similar correspondence between paradoxes in different frameworks is given by (Brams, Zwicker, and Kilgour 1998).