Computational Social Choice: Spring 2019

Ulle Endriss
Institute for Logic, Language and Computation
University of Amsterdam

Plan for Today

Today we will see a couple of further applications of the axiomatic method in judgment aggregation, focusing on characterisation results:

- *axiomatic characterisation* of rules (namely the *quota rules*)
- *logical characterisation of agendas* for which a specific rule (namely the *majority rule*) is guaranteed to preserve consistency
- logical characterisation of agendas for which there *exist rules* that preserve consistency and that meet certain axioms
- logical characterisation of agendas for which *all rules* that meet certain axioms preserve consistency ("safety of the agenda")

You can read up on this material in my book chapter cited below.

U. Endriss. Judgment Aggregation. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A.D. Procaccia (eds.), *Handbook of Computational Social Choice*. CUP, 2016.

Reminder: Axioms

- *Anonymity:* Treat all agents symmetrically!

  For any profile $\boldsymbol{J}$ and any permutation $\pi : N \to N$, we should have $F(J_1, \ldots, J_n) = F(J_{\pi(1)}, \ldots, J_{\pi(n)})$.

- *Neutrality:* Treat all propositions symmetrically!

  For any φ, ψ in the agenda Φ and any profile $\boldsymbol{J}$ with $N_\varphi^{\boldsymbol{J}} = N_\psi^{\boldsymbol{J}}$ we should have $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.

- *Independence:* Only the "pattern of acceptance" should matter!

  For any φ in the agenda Φ and any profiles $\boldsymbol{J}$ and $\boldsymbol{J'}$ with $N_\varphi^{\boldsymbol{J}} = N_\varphi^{\boldsymbol{J}'}$ we should have $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \varphi \in F(\boldsymbol{J'})$.

- *Monotonicity:* Additional support should not be harmful!

  For any φ in the agenda, any profile $\boldsymbol{J}$, and any judgment set J_i', $\varphi \in J_i' \setminus J_i$ should entail $\varphi \in F(\boldsymbol{J}) \Rightarrow \varphi \in F(\boldsymbol{J}_{-i}, J_i')$.

Winning Coalitions

Observation: Rule F is *independent* if, for each proposition $\varphi \in \Phi$, there exists a set of $\mathcal{W}_\varphi \subseteq 2^N$ of *winning coalitions* of agents in N, such that for all profiles $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ we have $\varphi \in F(\boldsymbol{J}) \Leftrightarrow N_\varphi^{\boldsymbol{J}} \in \mathcal{W}_\varphi$.

Now suppose F is independent and defined by $\{\mathcal{W}_\varphi\}_{\varphi \in \Phi}$. Then:

- F is *anonymous* *iff* $\mathcal{W}_\varphi$ is closed under equinumerosity: $C \in \mathcal{W}_\varphi$ and $|C| = |C'|$ entail $C' \in \mathcal{W}_\varphi$ for all $C, C' \subseteq N$ and all $\varphi \in \Phi$.
- F is *monotonic* *iff* $\mathcal{W}_\varphi$ is upward closed: $C \in \mathcal{W}_\varphi$ and $C \subseteq C'$ entail $C' \in \mathcal{W}_\varphi$ for all $C, C' \subseteq N$ and all $\varphi \in \Phi$.
- F guarantees *complete* outcomes *iff* $C \in \mathcal{W}_\varphi$ or $(N \setminus C) \in \mathcal{W}_{\sim\varphi}$ for all $C \subseteq N$ and all $\varphi \in \Phi$.
- F guarantees *complement-free* outcomes *iff* $C \notin \mathcal{W}_\varphi$ or $(N \setminus C) \notin \mathcal{W}_{\sim\varphi}$ for all $C \subseteq N$ and all $\varphi \in \Phi$.

Exercise: *What about neutrality (in the presence of independence)?*

A Subtlety about Neutrality

Recall the formal definition of neutrality:

- For any φ, ψ in the agenda Φ and any profile $\boldsymbol{J}$ with $N_\varphi^{\boldsymbol{J}} = N_\psi^{\boldsymbol{J}}$ we should have $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.

Intuitively, this says that all formulas should be treated symmetrically. Thus, we (almost) get:

- For any independent rule F, it is the case that F is *neutral iff* $\mathcal{W}_\varphi = \mathcal{W}_\psi$ for all formulas $\varphi, \psi \in \Phi$.

But note that neutrality does not "bite" for trivial agendas such as $\Phi = \{p, \neg p\}$: it holds vacuously, as there exists no admissible profile in which the same agents accept p and $\neg p$. But for *nontrivial agendas*, the above characterisation indeed does hold.

Reminder: Quota Rules

A *quota rule* F_q is defined by a function $q : \Phi \to \{0, 1, \ldots, n{+}1\}$:

$$F_q(\boldsymbol{J}) \quad = \quad \{\varphi \in \Phi \mid |N_\varphi^{\boldsymbol{J}}| \geqslant q(\varphi)\}$$

F_q is *uniform* if q maps any given formula to the same number λ.

Axiomatic Characterisation of Quota Rules

Proposition 1 (Dietrich and List, 2007) *An aggregation rule is anonymous, independent, and monotonic iff it is a quota rule.*

Proof: Immediate from characterisation using winning coalitions. ✓

Thus, for nontrivial agendas (avoiding the subtlety with neutrality):

Corollary 2 *An aggregation rule is anonymous, neutral, independent, and monotonic (= ANIM) iff it is a uniform quota rule.*

High/low quotas good for complement-freeness/completeness. Thus:

Proposition 3 *For even n, no ANIM rule can guarantee complete and complement-free outcomes.*

Proposition 4 *For odd n, an ANIM rule guarantees complete and complement-free outcomes iff it is the (strict) majority rule.*

F. Dietrich and C. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007.

Agenda Characterisation Results

Remark: The previous results have nothing to do with logic, by which I mean that consistency does not feature in their statement.

Recall: The seminal List-Pettit impossibility theorem was a bit unsatisfactory in that it was just about agendas $\Phi \supseteq \{p, q, p \wedge q\}$.

So: Would like results that *really* engage with the logical structure of the agenda (as it determines the nature of the aggregation problem).

First, you might ask whether a given *agenda* is *safe* (guaranteed to return *consistent* outcomes) *for a specific rule*. More generally:

- *Existential Agenda Characterisation:* Which agendas are safe for at least one rule in a given class (satisfying certain axioms)?
- *Universal Agenda Characterisation:* Which agendas are safe for all rules in a given class (satisfying certain axioms)?

Discussion: *Which type of result is relevant for what type of situation?*

Consistent Aggregation under the Majority Rule

An agenda Φ is said to have the *median property* (MP) *iff* every *minimally inconsistent subset* (mi-subset) of Φ has size $\leqslant 2$.

Intuition: MP means that all possible inconsistencies are "simple".

An agenda characterisation result for a specific aggregation rule:

Theorem 5 (Nehring and Puppe, 2007) *Let $n \geqslant 3$. The (strict) majority rule guarantees consistent outcomes for a given agenda Φ iff this agenda Φ has the MP.*

Remark: Note how $\{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$ violates the MP. This was the agenda featuring in the List-Pettit impossibility theorem.

K. Nehring and C. Puppe. The Structure of Strategy-proof Social Choice. Part I: General Characterization and Possibility Results on Median Space. *Journal of Economic Theory*, 135(1):269–305, 2007.

Proof

Claim: Φ is *safe* [$F_{\mathsf{maj}}(\boldsymbol{J})$ is consistent] $\Leftrightarrow$ Φ has the *MP* [mi-sets $\leqslant 2$]

($\Leftarrow$) Let Φ be an agenda with the MP. Now assume that there exists an admissible profile $\boldsymbol{J} \in \mathcal{J}(\Phi)^n$ such that $F_{\mathsf{maj}}(\boldsymbol{J})$ is *not* consistent.

$\leadsto$ There exists an inconsistent set $\{\varphi, \psi\} \subseteq F_{\mathsf{maj}}(\boldsymbol{J})$.
$\leadsto$ Each of φ and ψ must have been accepted by a strict majority.
$\leadsto$ One agent must have accepted both φ and ψ.
$\leadsto$ Contradiction (individual judgment sets must be consistent). ✓

($\Rightarrow$) Let Φ be an agenda that violates the MP, i.e., there exists a minimally inconsistent set $\Delta = \{\varphi_1, \ldots, \varphi_k\} \subseteq \Phi$ with $k > 2$.

Consider the profile $\boldsymbol{J}$, in which agent i accepts all formulas in Δ except for $\varphi_{1+(i \bmod 3)}$. Note that $\boldsymbol{J}$ is consistent. But the majority rule will accept all formulas in Δ, i.e., $F_{\mathsf{maj}}(\boldsymbol{J})$ is inconsistent. ✓

An Existential Agenda Characterisation Theorem

F is a *dictatorship* if there exists an agent $i^\star \in N$ such that $F(\boldsymbol{J}) = J_{i^\star}$ for every profile $\boldsymbol{J}$. Otherwise F is *nondictatorial*.

We saw that the majority rule works well only on "simple" agendas. *Do other rules do better?* Not if these are our requirements:

Theorem 6 (Nehring and Puppe, 2007) *Suppose $n \geqslant 3$ is odd. There exists a neutral, independent, monotonic, and nondictatorial aggregation rule that guarantees complete and consistent outcomes for a given agenda Φ iff this agenda Φ has the MP.*

Proof: The *possibility direction* ($\Leftarrow$) follows from our earlier results (majority rule does the job). Now for the *impossibility direction* ($\Rightarrow$).

K. Nehring and C. Puppe. The Structure of Strategy-proof Social Choice. Part I: General Characterization and Possibility Results on Median Space. *Journal of Economic Theory*, 135(1):269–305, 2007.

Reminder: Winning Coalitions

For nontrivial Φ, F is *independent* and *neutral* *iff* there exists a set of *winning coalitions* $\mathcal{W} \subseteq 2^N$ such that $\varphi \in F(\boldsymbol{J}) \Leftrightarrow N_\varphi^{\boldsymbol{J}} \in \mathcal{W}$.

So take any independent and neutral F with associated $\mathcal{W}$:

- F is *monotonic* *iff* $\mathcal{W}$ is upward closed: $C \in \mathcal{W}$ and $C \subseteq C'$ entail $C' \in \mathcal{W}$ for all $C, C' \subseteq N$.
- F guarantees *complete* outcomes *iff* $\mathcal{W}$ is maximal: $C \in \mathcal{W}$ or $(N \setminus C) \in \mathcal{W}$ for all $C \subseteq N$.
- F guarantees *complement-free* outcomes *iff* $C \notin \mathcal{W}$ or $(N \setminus C) \notin \mathcal{W}$ for all $C \subseteq N$.

The latter two claims follow from earlier observations and $\mathcal{W}_\varphi = \mathcal{W}_{\sim\varphi}$.

Exercise: *What does $\mathcal{W}$ look like for a dictatorship F?*

Proof Plan: Impossibility Direction

Note that the impossibility direction of our theorem is equivalent to:

Claim: If a *neutral*, *independent*, and *monotonic* rule F guarantees *complete* and *consistent* outcomes for agenda Φ *violating the MP*, then F must be a *dictatorship*.

So suppose Φ violates the MP (and thus is nontrivial) and that F has the properties mentioned above. Suppose $\mathcal{W}$ characterises F.

We will show that $\mathcal{W}$ is an *ultrafilter* on N, which means:

(i) The *empty coalition* is not winning: $\emptyset \notin \mathcal{W}$

(ii) Closure under *intersection:* $C, C' \in \mathcal{W} \Rightarrow C \cap C' \in \mathcal{W}$

(iii) *Maximality:* $C \in \mathcal{W}$ or $(N \setminus C) \in \mathcal{W}$

Appealing to the finiteness of N, this will allow us to show that $\mathcal{W} = \{C \subseteq N \mid i^\star \in C\}$ for some $i^\star \in N$, i.e., that F is *dictatorial*.

Proof: Noninclusion of the Empty Set

Exercise: *What would $\emptyset \in \mathcal{W}$ actually mean for F?*

Claim: $\emptyset \notin \mathcal{W}$.

We will use *monotonicity* as well as the requirement for outcomes to be *consistent* and thus also *complement-free:*

For the sake of contradiction, assume $\emptyset \in \mathcal{W}$.

From monotonicity (i.e., closure under supersets): $N \in \mathcal{W}$ as $\emptyset \subseteq N$.

From F guaranteeing complement-freeness: $C \notin \mathcal{W}$ or $(N \setminus C) \notin \mathcal{W}$ for all coalitions $C \subseteq N$. So we get a contradiction for $C = \emptyset$ ✓

Proof: Maximality

<u>Claim:</u> $C \in \mathcal{W}$ or $(N \setminus C) \in \mathcal{W}$ for all $C \subseteq N$.

We already saw that this corresponds to outcomes being *complete*. ✓

Proof: Closure under Taking Intersections

<u>Claim:</u> $C, C' \in \mathcal{W} \Rightarrow C \cap C' \in \mathcal{W}$ for all $C, C' \subseteq N$.

We'll use *MP-violation*, *monotonicity*, *consistency*, and *completeness*.

MP-violation means: there's a *mi-subset* $X = \{\varphi_1, \ldots, \varphi_k\} \subseteq \Phi$ with $k \geqslant 3$.

We can construct a complete and consistent profile $\boldsymbol{J}$ with these properties:

- $N^{\boldsymbol{J}}_{\varphi_1} = C$
- $N^{\boldsymbol{J}}_{\varphi_2} = C' \cup (N \setminus C)$
- $N^{\boldsymbol{J}}_{\varphi_3} = N \setminus (C \cap C')$
- $N^{\boldsymbol{J}}_{\psi} = N$ for all $\psi \in X \setminus \{\varphi_1, \varphi_2, \varphi_3\}$

Thus: everyone accepts $k-1$ of the propositions in X. And $N^{\boldsymbol{J}}_{\sim\varphi_3} = C \cap C'$.

- $C \in \mathcal{W} \;\Rightarrow\; \varphi_1 \in F(\boldsymbol{J})$
- From monotonicity: $C' \in \mathcal{W} \;\Rightarrow\; C' \cup (N \setminus C) \in \mathcal{W} \;\Rightarrow\; \varphi_2 \in F(\boldsymbol{J})$
- From maximality: $\emptyset \notin \mathcal{W} \;\Rightarrow\; N \in \mathcal{W} \;\Rightarrow\; X \setminus \{\varphi_1, \varphi_2, \varphi_3\} \subseteq F(\boldsymbol{J})$

Thus: for consistency we need $\varphi_3 \notin F(\boldsymbol{J})$, i.e., for completeness $\sim\!\varphi_3 \in F(\boldsymbol{J})$.

In other words: $N^{\boldsymbol{J}}_{\sim\varphi_3} = (C \cap C') \in \mathcal{W}$ ✓

Proof: Dictatorship

We have shown that the set of winning coalitions $\mathcal{W}$ is an *ultrafilter* on the (*finite!*) set of agents N:

(i) The *empty coalition* is not winning: $\emptyset \notin \mathcal{W}$
(ii) Closure under *intersection:* $C, C' \in \mathcal{W} \Rightarrow C \cap C' \in \mathcal{W}$
(iii) *Maximality:* $C \in \mathcal{W}$ or $(N \setminus C) \in \mathcal{W}$

From (i) and the completeness of outcomes: $N \in \mathcal{W}$.

Contraction Lemma: If $C \in \mathcal{W}$ and $|C| \geqslant 2$, then $C' \in \mathcal{W}$ for *some* $C' \subset C$.

> Proof: Split C into two proper subsets: $C_1 \uplus C_2 = C$.
> By maximality, $C_1 \notin \mathcal{W}$ implies $(N \setminus C_1) \in \mathcal{W}$, which by closure under taking intersections implies $C_2 = (C \cap (N \setminus C_1)) \in \mathcal{W}$. ✓

By induction: $\{i^\star\} \in \mathcal{W}$ for *one* $i^\star \in N$, i.e., $\mathcal{W} = \{C \subseteq N \mid i^\star \in C\}$.

That is, $i^\star$ is a *dictator*. ✓

Remark: The above just spells out the well-known fact that every ultrafilter on a finite set must be *principal*, i.e., of the form $\mathcal{W} = \{C \subseteq N \mid i^\star \in C\}$.

Safety of the Agenda

There also are *universal* agenda characterisation results that establish when an agenda guarantees consistent outcomes for *all rules*.

The (algorithmic) problem of deciding whether a given agenda can offer this guarantee is called the problem of the *safety of the agenda*.

U. Endriss, U. Grandi, and D. Porello. Complexity of Judgment Aggregation. *Journal of Artificial Intelligence Research (JAIR)*, 45:481–514, 2012.

Summary

We saw a number of characterisation results in judgment aggregation that make heavy use of the axiomatic method:

- *quota rules* = anonymous, independent, monotonic
  - neutrality forces quotas to be uniform
  - complement-freeness and completeness bound the quota
- *median property* (mi-sets $\leqslant 2$) of agenda necessary and sufficient for *majority rule* to be safe (returning consistent outcomes)
- *ultrafilter method* (similar to what we used for Arrow's Theorem) to prove *existential agenda characterisation theorem*
- we briefly mentioned *universal agenda characterisation results* as a means to ensure the *safety of the agenda*

What next? Computational complexity of judgment aggregation.