Computational Social Choice 2022

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

[http://www.illc.uva.nl/~ulle/teaching/comsoc/2022/]

Plan for Today

Today we focus on (existential) *agenda characterisation* results: given some axioms, for which agendas can we find a rule that is consistent?

We also discuss how *embedding preference aggregation* into JA allows us to recover classical results, and specifically *Arrow's Theorem*.

The results covered today have originally been proved for somewhat different formal models than we use here. For presentations close to what we do here, refer to the expository papers cited below.

C. List and C. Puppe. Judgment Aggregation: A Survey. In P. Anand *et al.* (eds.), *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

U. Endriss. Judgment Aggregation. In F. Brandt *et al.* (eds.), *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

Agenda Characterisation: Bigger Picture

Consider a *class of rules*, possibly determined by a set of *axioms:*

- *Existential Agenda Characterisation*

  Question: For which agendas can we find *some* rule that meets our axioms and that is consistent?

- *Universal Agenda Characterisation* ("*Safety of the Agenda*")

  Question: For which agendas is it the case that *every* rule meeting our axioms is consistent?

Note: If the axioms characterise a single rule, the two notions coincide.

Reminder: Consistent Aggregation under Majority

Recall that an agenda Φ satisfies the *median property* (MP) *iff* all its *minimally inconsistent subsets* have *size at most 2*.

We had proved:

Theorem 1 (Nehring and Puppe, 2007) *Let $n \geqslant 3$. The (strict) majority rule is consistent for a given agenda Φ iff Φ has the MP.*

If we look at a broader class of rules, can we make things work for a wider range of agendas?

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*, 2007.

Reminder: Axioms

- *Anonymity:* 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:* 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:* 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:* For any profile $\boldsymbol{J}$, agent i, judgment set J_i', and $\varphi \in J_i' \setminus J_i$, we should have $\varphi \in F(\boldsymbol{J}) \Rightarrow \varphi \in F(\boldsymbol{J}_{-i}, J_i')$.
- (Propositionwise) *Unanimity:* For any profile $\boldsymbol{J}$ and formula φ with $N_\varphi^{\boldsymbol{J}} = N$, we should have $\varphi \in F(\boldsymbol{J})$.

Dictatorships

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*.

Remark: Note that *anonymity* implies the absences of dictators. So being *nondictatorial* is like a very weak form of anonymity.

Reminder: Winning Coalitions

F is *independent* *iff* there exists a family of sets of *winning coalitions* $\mathcal{W}_\varphi \subseteq 2^N$, one for each $\varphi \in \Phi$, such that $\varphi \in F(\boldsymbol{J}) \Leftrightarrow N_\varphi^{\boldsymbol{J}} \in \mathcal{W}_\varphi$.

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

Now suppose F is independent and neutral, and defined by $\mathcal{W}$. Then:

- F is *anonymous* *iff* $\mathcal{W}$ is closed under equinumerosity: $C \in \mathcal{W}$ and $|C| = |C'|$ entail $C' \in \mathcal{W}$ for all $C, C' \subseteq N$
- 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 is *complete* *iff* $C \in \mathcal{W}$ or $\overline{C} \in \mathcal{W}$ for all $C \subseteq N$ $\quad [\overline{C} := N \setminus C]$
- F is *complement-free* *iff* $C \notin \mathcal{W}$ or $\overline{C} \notin \mathcal{W}$ for all $C \subseteq N$

Exercise: *What about unanimity? What about dictatorships?*

An Existential Agenda Characterisation Theorem

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

Theorem 2 (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*, 2007.

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 $\overline{C} \in \mathcal{W}$ for all $C \subseteq N$

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

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

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

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): $N \in \mathcal{W}$.
- From 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

Claim: $C \in \mathcal{W}$ or $\overline{C} \in \mathcal{W}$ for all $C \subseteq N$

We already saw that this holds, but let's go through the steps anyway.

We will use the fact that F is supposed to be *complete:*

- take any coalition $C \subseteq N$ and any formula $\varphi \in \Phi$
- construct a profile $\boldsymbol{J}$ with $N_\varphi^{\boldsymbol{J}} = C$
- from completeness: $\varphi \in F(\boldsymbol{J})$ or $\sim\varphi \in F(\boldsymbol{J})$
- from $\mathcal{W}$-determination of F: $N_\varphi^{\boldsymbol{J}} \in \mathcal{W}$ or $N_{\sim\varphi}^{\boldsymbol{J}} \in \mathcal{W}$
- from $\boldsymbol{J}$ being complete and complement-free: $N_{\sim\varphi}^{\boldsymbol{J}} = \overline{N_\varphi^{\boldsymbol{J}}}$
- putting everything together: $C \in \mathcal{W}$ or $\overline{C} \in \mathcal{W}$ ✓

Interlude: Unanimity

At this point we have established:

- $\emptyset \notin \mathcal{W}$
- $C \in \mathcal{W}$ or $(N \setminus C) \in \mathcal{W}$ for all $C \subseteq \mathcal{W}$ (including $C = \emptyset$)

So we can infer:

- $N \in \mathcal{W}$

So the grand coalition N is a winning coalition. Observe that this means that F must satisfy (propositionwise) *unanimity*.

Proof: Closure under Taking Intersections

Claim: $C, C' \in \mathcal{W} \Rightarrow (C \cap C') \in \mathcal{W}$ 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

So the set of winning coalitions $\mathcal{W}$ is an *ultrafilter* on the (*finite*) set 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 $\overline{C} \in \mathcal{W}$ for all $C \subseteq N$

We also saw that these findings imply that $N \in \mathcal{W}$ (*unanimity*).

Contraction Lemma: $C \in \mathcal{W}$ & $|C| \geqslant 2 \Rightarrow C' \in \mathcal{W}$ for *some* $C' \subset C$

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

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*. ✓

Second Example for a Characterisation Result

Call an agenda Φ *well-behaved* [on this slide only] if it is not both *totally blocked* and *even-number-negatable*.

Theorem 3 (Dokow and Holzman, 2010) *There exists a unanimous, independent, and nondictatorial rule that is complete and consistent for a agenda Φ iff Φ is well-behaved.*

Proof and exact definition of agenda properties omitted.

Remark: An added challenge here is to prove neutrality, so we can work with (one family of) winning coalitions. Rest of the proof is similar.

E. Dokow and R. Holzman. Aggregation of Binary Evaluations. *Journal of Economic Theory*, 2010.

Relevance to Preference Aggregation

What makes the impossibility direction of Dokow and Holzman's result particularly interesting is that it may be considered a generalisation of the most famous theorem in social choice theory:

Arrow's Theorem for *preference aggregation* (*next slide*)

We have already seen how to embed preference aggregation into JA. For a proof using this embedding, see Dietrich and List (2007).

For a direct proof (directly in preference aggregation) using the same ultrafilter technique we have seen, refer to my paper cited below.

F. Dietrich and C. List. Arrow's Theorem in Judgment Aggregation. *Social Choice and Welfare*, 2007.

U. Endriss. Logic and Social Choice Theory. In A. Gupta and J. van Benthem (eds.), *Logic and Philosophy Today*. College Publications, 2011.

Arrow's Theorem

Let X be a finite set of *alternatives*, $\mathcal{L}(X)$ the set of *linear orders* on X, and $N = \{1, \ldots, n\}$ a set of *agents*.

A *preference aggregator* $F : \mathcal{L}(X)^n \to \mathcal{L}(X)$ might satisfy these axioms:

- *Pareto:* if every agent prefers x over y, then so should the collective preference order returned by F
- *IIA* (independence of irrelevant alternatives): the relative ranking of x and y by F should only depend on how agents rank x and y

Theorem 4 (Arrow, 1951) *For $|X| \geqslant 3$, every preference aggregator that satisfies the Pareto and IIA conditions must be a dictatorship.*

Here as well: *dictatorship* $=$ output always copied from dictator in N

K.J. Arrow. *Social Choice and Individual Values*. John Wiley and Sons, 2nd edition, 1963. First edition published in 1951.

Summary

We saw two agenda characterisation theorems of this form:

> There exists a nondictatorial complete and consistent *rule* meeting certain *axioms* $\Leftrightarrow$ the *agenda* has a certain *property*.

Both directions are of interest:

($\Leftarrow$) *Possibility direction:* If the agenda property holds for your problem, then "reasonable" and consistent aggregation is possible.

($\Rightarrow$) *Impossibility direction:* For structurally rich domains, all seemingly "reasonable" rules are in fact dictatorial.

Possibility is proved by providing a concrete rule doing the job.

Impossibility is (sometimes) proved using the *ultrafilter technique*.

What next? Overview of what we covered in JA. Then other topics.