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

Some otherwise reasonable aggregation rules (e.g., the majority rule) will sometimes return *inconsistent outcomes*. But for certain (simple) agendas this won't happen: they are *safe*.

Today we want to investigate the following question:

Which agendas are safe for a given aggregation rule?

More specifically, we want to:

- characterise agendas for which the majority rule is consistent
- characterise agendas for which all rules meeting certain axioms are

This is known as the problem of the *safety of the agenda*.

Reminder: Basic Impossibility Theorem

Recall from the first lecture on judgment aggregation:

Theorem 1 (List and Pettit, 2002) *No judgment aggregation rule for an agenda Φ with $\{p, q, p \wedge q\} \subseteq \Phi$ that is anonymous, neutral, and independent can guarantee outcomes that are complete and consistent.*

Raises an obvious question: *for which other agendas is this the case?*

Would like: *characterisation* of class of agendas for which "*reasonable*" (= A, N, I) and *consistent* aggregation is (im)possible.

C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Reminder: Characterisation of Majority

We also saw:

Proposition 2 (Dietrich and List, 2007) *For an odd number of agents, an aggregation rule is anonymous, neutral, independent, monotonic, complete, and complement-free iff it is the majority rule.*

This is a characterisation rule of the majority in terms of basic "reasonableness" requirements, but is agnostic about consistency. So this is different from what we want today.

Our characterisation results so far did not make any reference to logic *proper* (completeness and complement-freeness don't count).

F. Dietrich and C. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Journal of Theoretical Politics*, 2007.

Agenda Properties

Two useful properties of agendas Φ (i.e., of sets of formulas):

- *Median Property* (MP): Φ has the MP *iff* every minimally inconsistent subset of Φ has size $\leqslant 2$.
- *Simplified Median Property* (SMP): Φ has the SMP *iff* every non-singleton mi-subset of Φ is of the form $\{\varphi, \psi\}$ with $\models \varphi \leftrightarrow \neg\psi$.

Intuition: Observed inconsistencies must have simple explanations.

Exercise: *Show that, if Φ has the SMP, then it also has the MP.*

Example for an agenda with the MP but not the SMP:

$$\{p, \neg p, p \wedge q, \neg(p \wedge q)\}$$

Consistent Aggregation under the Majority Rule

Recall that n is the number of agents.

An agenda characterisation result for a specific aggregation rule:

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

Recall: MP = all mi-sets have size $\leqslant 2$

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*, 2007.

Proof

Claim: Φ is *safe* [$F_{\mathrm{maj}}(\boldsymbol{J})$ 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 $F_{\mathrm{maj}}(\boldsymbol{J})$ is *not* consistent.

- $\leadsto$ There exists an inconsistent set $\{\varphi, \psi\} \subseteq F_{\mathrm{maj}}(\boldsymbol{J})$.
- $\leadsto$ Each of φ and ψ must have been accepted by a strict majority.
- $\leadsto$ One individual 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 individual 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., $F_{\mathrm{maj}}(\boldsymbol{J})$ is inconsistent. ✓

Complexity Theory: The Polynomial Hierarchy

The *polynomial hierarchy* is an infinite sequence of complexity classes: $\Sigma_1^p :=$ NP and Σ_i^p (for $i > 1$) is the class of problems solvable in polynomial time by a nondeterministic machine that has access to an *oracle* that decides Σ_{i-1}^p-complete problems in constant time.

Also define: $\Pi_i^p := \mathrm{co}\Sigma_i^p$ (complements).

$\textsc{Sat}$ for *quantified boolean formulas* with $< i$ quantifier alterations is a *complete problem* for Σ_i^p (Π_i^p) if the first quantifier is $\exists$ ($\forall$).

We will work with Π_2^p (sometimes written coNP$^{\text{NP}}$). The satisfiability problem for formulas of the following type is complete for this class:

$$\forall x_1 \cdots x_r \exists y_1 \cdots y_s . \varphi(x_1, \ldots, x_r, y_1, \ldots, y_s)$$

S. Arora and B. Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.

Complexity of Safety of the Agenda

Deciding whether a given agenda is safe (guarantees consistency) for the majority rule is located at the 2nd level of the polynomial hierarchy:

Theorem 4 (Endriss et al., 2012) *Deciding whether a given agenda guarantees consistency of the majority rule is Π_2^p-complete.*

Discussion: Bad news. Not only are well-behaved agendas structurally simplistic, but recognising this simplicity is extremely hard.

By Theorem 3, the above result is equivalent to this:

Lemma 5 *Deciding the MP is Π_2^P-complete.*

Next, we give a proof of Π_2^p-membership and some basic intuitions regarding Π_2^p-hardness. The full proof is in the paper cited below.

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

Proof of Π_2^p-Membership

Claim: Deciding whether a set Φ has the MP [mi-sets $\leqslant 2$] is in Π_2^p.

That is: We need to show that a machine equipped with a Sat-*oracle* can, in *polynomial time*, verify the correctness of a *certificate* claiming to establish a *violation* of the median property ($\Pi_2^p = \text{coNP}^{\text{NP}}$).

Use as certificate a set $\Delta \subseteq \Phi$ with $|\Delta| > 2$ that is inconsistent but has no subset of size $\leqslant 2$ that is inconsistent.

We can verify the correctness of such a certificate using a polynomial number of queries to the Sat-oracle:

- one query to check that Δ is inconsistent
- $|\Delta|$ queries to check that each subset of size 1 is consistent
- $O(|\Delta|^2)$ queries to check that each subset of size 2 is consistent

Done. ✓

Intuition for Π_2^p-Hardness

We won't give a proof, only some intuition about what $\textsc{Sat}$ for QBF's of the form $\forall\vec{x}.\exists\vec{y}.\varphi$ has to do with the MP.

Consider this QBF:

$$\forall x_1 \cdots x_r.\exists y_1 \cdots y_s.\varphi(x_1,\ldots,x_r,y_1,\ldots,y_s)$$

Now construct this agenda:

$$\Phi \quad := \quad \{x_1,\neg x_1,x_2,\neg x_2,\ldots,x_r,\neg x_r,\varphi,\neg\varphi\}$$

The QBF is not true *iff* there exists a subset of Φ (including φ) that is inconsistent but does not include complementary formulas. This latter property is similar to the MP (and even more so to the SMP).

Agenda Characterisation for Classes of Rules

Instead of a single rule, suppose we are interested in a *class of rules*, possibly determined by a set of *axioms*. Two types of results:

- *Existential Agenda Characterisation*

  Question: Is there *some* rule meeting certain axioms that is consistent for every agenda with a given property?

  Scenario: an economist looking for a rule meeting certain axioms

- *Universal Agenda Characterisation*

  Question: Is *every* rule meeting certain axioms consistent for every agenda with a given property?

  Scenario: multiagent system we have only partial knowledge about

Today we will look only into the latter, also called *safety results*.

Example for a Safety Theorem

Suppose we know that the group will use *some* aggregation rule meeting certain requirements, but we do not know which one exactly. Can we guarantee that the outcome will be consistent?

Assumption for the remainder of today: Φ contains no tautologies, and thus no contradictions (slightly simplifies statement of result).

A typical result (for the majority rule axioms, minus monotonicity):

Theorem 6 (Endriss et al., 2012) *An agenda Φ is safe for all anonymous, neutral, independent, complete and complement-free aggregation rules iff Φ has the SMP.*

Recall: SMP = all inconsistencies due to some $\{\varphi, \psi\}$ with $\models \varphi \leftrightarrow \neg\psi$

We now give a proof for the case where n is *odd*.

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

Proof

<u>Claim:</u> Φ is safe for every ANI/complete/comp-free rule $F \Leftrightarrow \Phi$ has SMP

($\Leftarrow$) Suppose Φ has the SMP. For the sake of contradiction, assume $F(\boldsymbol{J})$ is inconsistent. Then $\{\varphi, \psi\} \subseteq F(\boldsymbol{J})$ with $\models \varphi \leftrightarrow \neg\psi$. Now:

$\leadsto$ $\varphi \in J_i \Leftrightarrow {\sim}\psi \in J_i$ for each individual i (from $\models \varphi \leftrightarrow \neg\psi$ together with consistency and completeness of individual judgment sets)

$\leadsto$ $\varphi \in F(\boldsymbol{J}) \Leftrightarrow {\sim}\psi \in F(\boldsymbol{J})$ (from neutrality)

$\leadsto$ both ψ and ${\sim}\psi$ in $F(\boldsymbol{J}) \leadsto$ contradiction (with complement-freeness) ✓

($\Rightarrow$) Suppose Φ violates the SMP. Take any minimally inconsistent $X \subseteq \Phi$. If $|X| > 2$, then also the MP is violated and we already know that the majority rule is not consistent. ✓ So we can assume $X = \{\varphi, \psi\}$.

W.l.o.g., must have $\varphi \models \neg\psi$ but $\neg\psi \not\models \varphi$ (otherwise SMP holds).

But now we can find a rule that is not safe: the parity rule F_{par} accepts a formula *iff* an odd number of agents does. Consider a profile $\boldsymbol{J}$ with $J_1 \supseteq \{{\sim}\varphi, {\sim}\psi\}$, $J_2 \supseteq \{{\sim}\varphi, \psi\}$, $J_3 \supseteq \{\varphi, {\sim}\psi\}$. Then $F_{\text{par}}(\boldsymbol{J}) \supseteq \{\varphi, \psi\}$. ✓

Examples

- Let Φ_1 be an agenda consisting solely of *literals*.

  Φ_1 has the *SMP* and thus is *safe* for every rule that is anonymous, neutral, independent, complement-free, and complete.

- Let $\Phi_2 := \{p, \neg p, p \wedge q, \neg(p \wedge q), r, \neg r\}$.

  Φ_1 *violates* the *SMP*. Consider this profile and the *parity rule:*

	p	$p \wedge q$	r
Agent 1	Yes	Yes	Yes
Agent 2	No	No	Yes
Agent 3	Yes	No	Yes
F_{par}	No	Yes	Yes

  Note that the parity rule F_{par} is anonymous, neutral, independent, complement-free, and complete.

Complexity

Deciding whether an agenda has the SMP is also Π_2^p-complete.

So deciding safety of the agenda for the class of rules meeting the axioms of the List-Pettit impossibility theorem is highly intractable.

Summary

We have discussed the problem of the *safety of the agenda:*

- The *majority rule* is guaranteed to produce a *consistent* outcome *iff* every possible inconsistency in the agenda can be explained in terms of just two formulas (*median property*).
- To guarantee consistency for *all rules* that share the properties of the majority rule except for monotonicity, we need to simplify agendas even further and only permit inconsistencies arising from logical complements (*simplified median property*).
- Similar results hold for other combinations of axioms.
- Checking any of these agenda properties is highly intractable (complete for *coNP with NP-oracle*).

What next? Existential agenda characterisation. Deepest axiomatic results we'll cover. Prepare by recalling winning coalitions.