

Computational Social Choice 2024

Ulle Endriss

Institute for Logic, Language and Computation
University of Amsterdam

[<http://www.illc.uva.nl/~ulle/teaching/comsoc/2024/>]

Plan for Today

In *judgment aggregation* (JA) agents are asked to judge whether each of a given number of propositions is true or false, and we then need to aggregate this information into a single collective judgment.

Today's lecture will be an introduction to JA:

- motivating example: *doctrinal paradox*
- *formal model* for JA and relationship to *preference aggregation*
- some *specific aggregation rules* to use in practice
- two examples for results using the *axiomatic method*

Most of this material is covered in my book chapter cited below.

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

Example: The Doctrinal Paradox

A court with three judges is considering a case in contract law.

Legal doctrine stipulates that the defendant is *liable* (r) iff the contract was *valid* (p) and has been *breached* (q): $r \leftrightarrow p \wedge q$.

	p	q	r
Judge 1	Yes	Yes	Yes
Judge 2	No	Yes	No
Judge 3	Yes	No	No

Exercise: *Should this court pronounce the defendant guilty or not?*

L.A. Kornhauser and L.G. Sager. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 1993.

Why Paradox?

So why is this example usually referred to as a “paradox”?

	p	q	$p \wedge q$
Agent 1	Yes	Yes	Yes
Agent 2	No	Yes	No
Agent 3	Yes	No	No
Majority	Yes	Yes	No

Explanation 1: Two natural aggregation rules, the *premise-based rule* and the *conclusion-based rule*, produce *different* outcomes.

Explanation 2: Each individual judgment is *logically consistent*, but the collective judgment returned by the (natural) *majority rule* is *not*.

In philosophy, this is also known as the *discursive dilemma* of choosing between *responsiveness* to the views of decision makers (by respecting majority decisions) and the *consistency* of collective decisions.

The Model

Notation: Let $\sim\varphi := \varphi'$ if $\varphi = \neg\varphi'$ and let $\sim\varphi := \neg\varphi$ otherwise.

An *agenda* Φ is a finite nonempty set of propositional formulas (w/o double negation) closed under complementation: $\varphi \in \Phi \Rightarrow \sim\varphi \in \Phi$.

A *judgment set* J on an agenda Φ is a subset of Φ . We call J :

- *complete* if $\varphi \in J$ or $\sim\varphi \in J$ for all $\varphi \in \Phi$
- *complement-free* if $\varphi \notin J$ or $\sim\varphi \notin J$ for all $\varphi \in \Phi$
- *consistent* if there exists an assignment satisfying all $\varphi \in J$

Let $\mathcal{J}(\Phi)$ be the set of all complete and consistent subsets of Φ .

Now a finite set of *agents* $N = \{1, \dots, n\}$, with $n \geq 2$, express judgments on the formulas in Φ , producing a *profile* $\mathbf{J} = (J_1, \dots, J_n)$.

A (resolute) *aggregation rule* for an agenda Φ and a set of n agents is a function mapping a profile of complete and consistent individual judgment sets to a single collective judgment set: $F : \mathcal{J}(\Phi)^n \rightarrow 2^\Phi$.

Example: Majority Rule

Suppose three agents ($N = \{1, 2, 3\}$) express judgments on the propositions in the agenda $\Phi = \{p, \neg p, q, \neg q, p \vee q, \neg(p \vee q)\}$.

For simplicity, we only show the positive formulas in our tables:

	p	q	$p \vee q$	formal notation
Agent 1	Yes	No	Yes	$J_1 = \{p, \neg q, p \vee q\}$
Agent 2	Yes	Yes	Yes	$J_2 = \{p, q, p \vee q\}$
Agent 3	No	No	No	$J_3 = \{\neg p, \neg q, \neg(p \vee q)\}$

Under the (strict) *majority rule* we accept a formula if more than half of the agents do: $F_{\text{maj}}(\mathbf{J}) = \{p, \neg q, p \vee q\}$ [complete and consistent!]

Recall: F_{maj} does *not* guarantee *consistent* outcomes in general.

Exercise: Show that F_{maj} guarantees *complement-free* outcomes.

Exercise: Show that F_{maj} guarantees *complete* outcomes iff n is odd.

Embedding Preference Aggregation

In *preference aggregation*, agents express preferences (linear orders) over a set of alternatives A . We want a *SWF* $F : \mathcal{L}(A)^n \rightarrow \mathcal{L}(A)$.

Introduce a propositional variable $p_{x \succ y}$ for every $x, y \in A$ with $x \neq y$.

Build $\Phi = \{p_{x \succ y}, \neg p_{x \succ y} \mid x \neq y\} \cup \{\Gamma, \neg\Gamma\}$, where Γ is conjunction of:

- Antisymmetry: $p_{x \succ y} \leftrightarrow \neg p_{y \succ x}$ for all distinct $x, y \in A$
- Transitivity: $p_{x \succ y} \wedge p_{y \succ z} \rightarrow p_{x \succ z}$ for all distinct $x, y, z \in A$

Now the *Condorcet Paradox* can be modelled in JA:

	Γ	$p_{a \succ b}$	$p_{b \succ c}$	$p_{a \succ c}$	<i>corresponding order</i>
Agent 1	Yes	Yes	Yes	Yes	$a \succ b \succ c$
Agent 2	Yes	No	Yes	No	$b \succ c \succ a$
Agent 3	Yes	Yes	No	No	$c \succ a \succ b$
Majority	Yes	Yes	Yes	No	<i>not a linear order</i>

Quota Rules

Let $N_\varphi^{\mathbf{J}}$ denote the *coalition* of *supporters* of φ in \mathbf{J} , i.e., the set of all those agents who accept formula φ in profile $\mathbf{J} = (J_1, \dots, J_n)$:

$$N_\varphi^{\mathbf{J}} := \{i \in N \mid \varphi \in J_i\}$$

The (uniform) *quota rule* F_q with quota $q \in \{0, 1, \dots, n+1\}$ accepts all propositions accepted by at least q of the individual agents:

$$F_q(\mathbf{J}) = \{\varphi \in \Phi \mid \#N_\varphi^{\mathbf{J}} \geq q\}$$

Example: The (*strict*) *majority rule* is the quota rule with $q = \lceil \frac{n+1}{2} \rceil$.

Intuition: high quotas good for consistency (but bad for completeness)

Exercise: Show that F_q with $q = n$ guarantees *consistent* outcomes!

Recall: The doctrinal paradox agenda is $\{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$.

Exercise: For the *doctrinal paradox agenda* and n agents, what is the *lowest uniform quota* q that will guarantee *consistent* outcomes?

Premise-Based Aggregation

Suppose we can divide the agenda into *premises* and *conclusions*:

$$\Phi = \Phi_p \uplus \Phi_c \quad (\text{each closed under complementation})$$

Then the *premise-based rule* F_{pre} for Φ_p and Φ_c is this function:

$$F_{\text{pre}}(\mathbf{J}) = \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\},$$

where $\Delta = \{\varphi \in \Phi_p \mid \#N_{\varphi}^{\mathbf{J}} > n/2\}$

A common assumption is that *premises* = *literals*.

Exercise: Show that this assumption guarantees *consistent* outcomes.

Exercise: Does it also guarantee *completeness*? What detail matters?

Remark: The *conclusion-based rule* is less attractive from a theoretical standpoint (as it is incomplete by design), but often used in practice.

Example: Premise-Based Aggregation

Suppose *premises = literals*. Consider this example:

	p	q	r	$p \vee q \vee r$
Agent 1	Yes	No	No	Yes
Agent 2	No	Yes	No	Yes
Agent 3	No	No	Yes	Yes
F_{pre}	No	No	No	No

So the *unanimously accepted* conclusion is *collectively rejected!*

Discussion: *Is this ok?*

The Kemeny Rule

Recall: The *Kemeny rule* in preference aggregation (as a *SWF*) returns linear orders that minimise the cumulative distance to the profile.

We can generalise this idea to JA:

$$F_{\text{Kem}}(\mathbf{J}) = \operatorname{argmin}_{J \in \mathcal{J}(\Phi)} \sum_{i \in N} H(J, J_i), \quad \text{where } H(J, J_i) = |J \setminus J_i|$$

Here the *Hamming distance* $H(J, J_i)$ counts the number of positive formulas in the agenda on which J and J_i disagree.

This is an attractive rule, but outcome determination is *intractable*.

Exercise: How would you generalise the *Slater rule* to JA?

Basic Axioms for Judgment Aggregation

What makes for a “good” aggregation rule F ? The following *axioms* all express intuitively appealing (but always debatable!) properties:

- *Anonymity*: Treat all agents symmetrically!

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

- *Neutrality*: Treat all propositions symmetrically!

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

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

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

Observe that the *majority rule* satisfies all of these axioms.

Exercise: *But so do some other rules! Can you think of examples?*

A Basic Impossibility Theorem

We saw that the majority rule cannot guarantee consistent outcomes. Is there some other “reasonable” aggregation rule that does not have this problem? *Surprisingly, no!* (at least not for certain agendas)

This is the main result in the original paper introducing the formal model of JA and proposing to apply the axiomatic method:

Theorem (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**.*

Note that the theorem requires $n \geq 2$. (*Why?*)

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

Proof: Part 1

Recall: $N_\varphi^{\mathbf{J}}$ is the set of agents who accept formula φ in profile \mathbf{J} .

Let F be any aggregator that is independent, anonymous, and neutral.

We observe:

- Due to *independence*, whether $\varphi \in F(\mathbf{J})$ only depends on $N_\varphi^{\mathbf{J}}$.
- Then, due to *anonymity*, whether $\varphi \in F(\mathbf{J})$ only depends on $|N_\varphi^{\mathbf{J}}|$.
- Finally, due to *neutrality*, the manner in which the status of $\varphi \in F(\mathbf{J})$ depends on $|N_\varphi^{\mathbf{J}}|$ must itself *not* depend on φ .

Thus: If φ and ψ are accepted by the same number of agents, then we must either accept both of them or reject both of them.

Proof: Part 2

Recall: For all $\varphi, \psi \in \Phi$, if $|N_{\varphi}^{\mathbf{J}}| = |N_{\psi}^{\mathbf{J}}|$, then $\varphi \in F(\mathbf{J}) \Leftrightarrow \psi \in F(\mathbf{J})$.

First, suppose the number n of agents is *odd* (and $n > 1$):

Consider a profile \mathbf{J} where $\frac{n-1}{2}$ agents accept p and q ; one accepts p but not q ; one accepts q but not p ; and $\frac{n-3}{2}$ accept neither p nor q .

That is: $|N_p^{\mathbf{J}}| = |N_q^{\mathbf{J}}| = |N_{\neg(p \wedge q)}^{\mathbf{J}}|$. Then:

- Accepting all three formulas contradicts consistency. ✓
- But if we accept none, completeness forces us to accept their complements, which also contradicts consistency. ✓

If n is *even*, we can get our impossibility even without having to make (almost) any assumptions regarding the structure of the agenda:

Consider a profile \mathbf{J} with $|N_p^{\mathbf{J}}| = |N_{\neg p}^{\mathbf{J}}|$. Then:

- Accepting both contradicts consistency. ✓
- Accepting neither contradicts completeness. ✓

Note: Neutrality only has “bite” here because we also have $q \in \Phi$.

Consistent Aggregation under the Majority Rule

An agenda Φ is said to have the *median property* (MP) iff every *MUS* (minimally unsatisfiable subset) of Φ has size ≤ 2 .

Intuition: MP means that all possible inconsistencies are “simple”.

Theorem (Nehring and Puppe, 2007): *The (strict) majority rule guarantees consistent outcomes for agenda Φ iff it has the MP (if $n \geq 3$).*

Remark: Note how $\{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$ violates the MP.

Exercise: *Is this a positive or a negative result?*

Checking whether Φ has the MP is *intractable* (Endriss et al., 2012).

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.

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

Proof

Claim: Φ is *safe* [$F_{\text{maj}}(\mathbf{J})$ is consistent] $\Leftrightarrow \Phi$ has the *MP* [MUSs ≤ 2]

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

- \leadsto By MP, there exists an inconsistent set $\{\varphi, \psi\} \subseteq F_{\text{maj}}(\mathbf{J})$.
- \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). \checkmark

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

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

Summary

This has been an introduction to the field of *judgment aggregation*, which (as we saw) is a *generalisation* of preference aggregation.

- examples for *rules*: quota rules, premise-based rule, Kemeny rule
- examples for *axioms*: anonymity, neutrality, independence
- examples for results: *impossibility* and *agenda characterisation*

JA is a powerful framework for modelling collective decision making that generalises several other models studied in COMSOC.

Topics not discussed: strategic behaviour, other logics, complexity, ...