Computational Social Choice 2021

Sirin Botan

28/04/21

Institute for Logic, Language and Computation
University of Amsterdam

Plan for Today

Logic meets social choice—let's do some Judgment Aggregation!

Judgment aggregation (JA) deals with aggregating yes/no opinions provided by agents into a collective opinion that should reflect the views of the group.

We'll see the (very) general framework of JA, and get a feel for how exactly it is more general than preference aggregation. Along the way we'll see an impossibility result and touch on strategic manipulation.

- Motivation: The doctrinal paradox
- Embedding preference aggregation in JA
- An impossibility result (how could we not?)
- Strategyproofness

Doctrinal Paradox/Discursive Dilemma

Aggregating judges' opinions in a legal case.

$p :=$ 'document is a valid contract'
$q :=$ 'the promise in the document was breached'
$r :=$ 'the defendant is liable'

All agents accept that $p \wedge q \leftrightarrow r$.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	✓	×	✓	×

✋ Should the court find 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.
C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Doctrinal Paradox/Discursive Dilemma

Aggregating judges' opinions in a legal case.

p := 'document is a valid contract'

q := 'the promise in the document was breached'

r := 'the defendant is liable'

All agents accept that $p \wedge q \leftrightarrow r$.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	✓	×	✓	×
	✓	✓	✓	✓

L.A. Kornhauser and L.G. Sager. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 1993.
C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Doctrinal Paradox/Discursive Dilemma

Aggregating judges' opinions in a legal case.

$p :=$ 'document is a valid contract'

$q :=$ 'the promise in the document was breached'

$r :=$ 'the defendant is liable'

All agents accept that $p \wedge q \leftrightarrow r$.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	✓	×	✓	×
				×

L.A. Kornhauser and L.G. Sager. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 1993.
C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Doctrinal Paradox/Discursive Dilemma

Aggregating judges' opinions in a legal case.

$p :=$ 'document is a valid contract'

$q :=$ 'the promise in the document was breached'

$r :=$ 'the defendant is liable'

All agents accept that $p \wedge q \leftrightarrow r$.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	✓	×	✓	×
majority	✓	✓	✓	×

L.A. Kornhauser and L.G. Sager. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 1993.
C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Doctrinal Paradox/Discursive Dilemma

Aggregating judges' opinions in a legal case.

$p :=$ 'document is a valid contract'

$q :=$ 'the promise in the document was breached'

$r :=$ 'the defendant is liable'

All agents accept that $p \wedge q \leftrightarrow r$.

	p	q	$(p \wedge q)$
agent 1	✓	✓	✓
agent 2	×	✓	×
agent 3	✓	×	×
majority	✓	✓	×

Kornhauser, L.A. and Sager, L.G. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 1993.
C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 2002.

Formal Framework

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

- An agenda Φ is a finite, nonempty set of propositional formulas (closed under complementation)
- A judgment set J is a subset of Φ. J is:
  - complete if $\varphi \in J$ or $\sim\varphi \in J$ for all $\varphi \in \Phi$
  - consistent if there is an assignment making all $\varphi \in J$ true

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

A set of agents $N = \{1, \ldots, n\}$ report their judgment sets, giving us a profile $\boldsymbol{J} = (J_1, \ldots J_n)$.

A (resolute) aggregation rule F is a function mapping a profile to a collective judgment set:

$$F : \mathcal{J}(\Phi)^n \to 2^{\Phi}$$

.

Example: Majority Rule

Notation: $N_\varphi^{\boldsymbol{J}}$ is the set of agents who accept φ in profile $\boldsymbol{J}$

The (strict) majority rule F_{maj} returns the set of formulas accepted by more than half the agents.

$$F_{maj} : \boldsymbol{J} \mapsto \{\varphi \mid |N_\varphi^{\boldsymbol{J}}| > \tfrac{n}{2}\}$$

	p	q	$p \to q$
Agent 1	✓	✓	✓
Agent 2	×	×	✓
Agent 3	×	✓	✓
Majority	×	✓	✓

$N = \{1, 2, 3\}$

$\Phi = \{p, q, (p \to q), \neg p, \neg q, \neg(p \to q)\}$

Profile $\boldsymbol{J} = (J_1, J_2, J_3)$

$J_1 = \{p, q, (p \to q)\}$

$J_2 = \{\neg p, \neg q, (p \to q)\}$

$J_3 = \{\neg p, q, (p \to q)\}$

✋ What is $F_{maj}(\boldsymbol{J})$?

Collective Rationality Requirements

Just as we can require judgment sets to be complete and consistent, we can also require that an aggregation rule "lifts" these requirements.

F is:

- complete if $F(\boldsymbol{J})$ is complete for all profiles $\boldsymbol{J}$
- consistent if $F(\boldsymbol{J})$ is consistent for all profiles $\boldsymbol{J}$

We already saw the majority rule is not consistent—the discursive dilemma. We will see now that this problem occurs more generally.

Other Rules (Better Rules?)

Premise-based rule: divide the agenda into premises and conclusions, aggregate opinion on premises, then accept a conclusion C if the accepted premises imply C.

Kemeny Rule:

$$F_{Kem}(\boldsymbol{J}) = \underset{J \in \mathcal{J}(\Phi)}{\text{argmin}} \sum_{i \in N} H(J, J_i)$$

Where $H(J, J') = |J \setminus J'|$ is the Hamming distance.

- Similar to Kemeny in voting (which minimises the sum of pairwise disagreements with agents' ballots).
- Guarantees consistency.

Embedding Voting in JA

We can use the JA framework to simulate the standard framework of preference aggregation.

Take the following preference profile 👀

voter 1: $a \succ b \succ c$
voter 2: $c \succ a \succ b$
voter 3: $b \succ c \succ a$

- for each pair of alternatives a and b: add p_{ab}—'a is preferable to b'
- We build the preference agenda Φ:
  - $p_{ab}, p_{ac}, p_{bc}, p_{ba}, p_{ca}, p_{ca} \in \Phi$—our propositional variables
  - $(p_{ab} \leftrightarrow \neg p_{ba}) \in \Phi$ for all pairs a, b—antisymmetry
  - $(p_{ab} \wedge p_{bc} \rightarrow p_{ac}) \in \Phi$ for all a, b, c—transitivity

These encode the properties of linear orders.

Embedding Voting in JA

We can use the JA framework to simulate the standard framework of preference aggregation.

Take the following preference profile

$$\begin{aligned} &\text{voter 1: } a \succ b \succ c \\ &\text{voter 2: } c \succ a \succ b \\ &\text{voter 3: } b \succ c \succ a \end{aligned}$$

- for each pair of alternatives a and b: add p_{ab}—'a is preferable to b'
- We build the preference agenda Φ:
  - $p_{ab}, p_{ac}, p_{bc}, p_{ba}, p_{ca}, p_{ca} \in \Phi$—our propositional variables
  - $(p_{ab} \leftrightarrow \neg p_{ba}) \in \Phi$ for all pairs a, b—antisymmetry
  - $(p_{ab} \wedge p_{bc} \rightarrow p_{ac}) \in \Phi$ for all a, b, c—transitivity

What would voter 1's judgment set be?

Embedding Voting in JA

We can use the JA framework to simulate the standard framework of preference aggregation.

Take the following preference profile 👀

voter 1: $a \succ b \succ c$
voter 2: $c \succ a \succ b$
voter 3: $b \succ c \succ a$

- for each pair of alternatives a and b: add p_{ab}—'a is preferable to b'
- We build the preference agenda Φ:
  - $p_{ab}, p_{ac}, p_{bc}, p_{ba}, p_{ca}, p_{ca} \in \Phi$—our propositional variables
  - $(p_{ab} \leftrightarrow \neg p_{ba}) \in \Phi$ for all pairs a, b—antisymmetry
  - $(p_{ab} \wedge p_{bc} \rightarrow p_{ac}) \in \Phi$ for all a, b, c—transitivity

$J_1 = \{p_{ab}, p_{ac}, p_{bc}, (p_{ab} \leftrightarrow \neg p_{ba}), (p_{ab} \wedge p_{bc} \rightarrow p_{ac}), \ldots\}$

Condorcet Paradox in JA

Preference Aggregation

voter 1: $a \succ b \succ c$
voter 2: $c \succ a \succ b$
voter 3: $b \succ c \succ a$

$\Rightarrow$ Condorcet cycle

Judgment Aggregation

	p_{ab}	p_{ac}	p_{bc}	$\dots$
agent 1:	✓	✓	✓	✓
agent 2:	✓	×	×	✓
agent 3:	×	×	✓	✓
majority:	✓	×	✓	✓

- Translating back to preferences: $a \succ b \succ c \succ a \cdots$
- In JA: the majority judgment is inconsistent: the majority accepts p_{ab}, p_{bc}, $p_{ab} \wedge p_{bc} \rightarrow p_{ac}$, but not p_{ac}.

Axioms

The following three axioms have obvious counterparts in voting.

- Anonymity: for any profile $\boldsymbol{J}$ and any permutation $\pi : N \to N$, we have that $F(J_1, \dots J_n) = F(J_{\pi(1)}, \dots J_{\pi(n)})$.
- Neutrality: for any $\varphi, \psi \in \Phi$ and any profile $\boldsymbol{J}$, if $\varphi \in J_i \Leftrightarrow \psi \in J_i$ for all $i \in N$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.
- Independence: for any $\varphi \in \Phi$ and any two profiles $\boldsymbol{J}$ and $\boldsymbol{J'}$, if $\varphi \in J_i \Leftrightarrow \varphi \in J'_i$ for all $i \in N$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \varphi \in F(\boldsymbol{J'})$.

Axioms

The following three axioms have obvious counterparts in voting.

- Anonymity: Treating all agents symmetrically.
- Neutrality: Treating all formulas the same.
- Independence: Outcome on φ depends only on judgments on φ.

Note that the majority rule satisfies all three axioms.

Is the premise-based rule anonymous? neutral? independent?

An Impossibility Result

Theorem 1 (List and Pettit, 2002) No judgment aggregation rule for an agenda s.t. $\{p, q, p \wedge q\} \subseteq \Phi$ satisfies all of anonymity, neutrality, independence, completeness and consistency.

C. List and P. Pettit. Aggregating Sets of Judgments: An Impossibility Result. *Economics and Philosophy*, 18(1), 89-110, 2002.

Proof. . .

Notation: $N_\varphi^{\boldsymbol{J}}$ is the set of agents who accept φ in profile $\boldsymbol{J}$

Let F be some aggregation rule that is anonymous, neutral, and independent.

- F is independent: whether $\varphi \in F(\boldsymbol{J})$ depends only on $N_\varphi^{\boldsymbol{J}}$.
- F is anonymous: we actually only need to look at $|N_\varphi^{\boldsymbol{J}}|$.
- F is neutral: the way in which the status of $\varphi \in F(\boldsymbol{J})$ depends on $|N_\varphi^{\boldsymbol{J}}|$, cannot depend on φ.

So, if φ and ψ are accepted by the same *number* of individuals, F must either accept both or reject both.

. . . Proof.

Let $\{p, q, p \wedge q\} \subseteq \Phi$. 👀

<u>If n is even:</u>

take any profile $\boldsymbol{J}$ where $|N_p^{\boldsymbol{J}}| = |N_{\neg p}^{\boldsymbol{J}}|$.

- Accept both: not consistent ✓
- Accept neither: not complete ✓
- Accept just one: contradicts previous slide ✓

. . . Proof.

If n is odd:

Consider the following profile

	p	q	$p \wedge q$
$\frac{n-1}{2}$ agents	✓	✓	✓
$\frac{n-3}{2}$ agents	×	×	×
1 agent	✓	×	×
1 agent	×	✓	×

Then $|N_p^{\boldsymbol{J}}| = |N_q^{\boldsymbol{J}}| = |N_{\neg(p \wedge q)}^{\boldsymbol{J}}| = \frac{n-1}{2} + 1 = \frac{n-3}{2} + 2$, and by the previous (previous) slide, we have to accept either all or none of them.

- Accept all: not consistent ✓
- Accept none: If the rule is complete, then we must accept all complements ($\neg p, \neg q$, and $(p \wedge q)$), so then not consistent ✓

Domain Restrictions to the Rescue (Again)

A profile is unidimensionally aligned if we can order agents such that for any φ, the agents accepting φ are either all to the left or all to the right of those rejecting φ.

	p	q	$(p \to q)$
agent 1	✓	×	×
agent 2	✓	×	×
agent 3	×	×	✓
agent 4	×	×	✓
agent 5	×	✓	✓
majority	×	×	✓

Idea similar to single-peakedness in voting. We can also defined restrictions based on orderings of the agenda.

C. List. A Possibility Theorem on Aggregation over Multiple Interconnected Propositions. Mathematical Social Sciences, 2003.

Domain Restrictions to the Rescue (Again)

A profile is unidimensionally aligned if we can order agents such that for any φ, the agents accepting φ are either all to the left or all to the right of those rejecting φ.

Proposition 1 (List, 2003) The majority rule is consistent on unidimensionally aligned profiles.

C. List. A Possibility Theorem on Aggregation over Multiple Interconnected Propositions. Mathematical Social Sciences, 2003.

Strategic Manipulation

As in voting, agents can manipulate (submit untruthful judgment sets) to get a more favourable outcome.

Suppose we are using the premise-based rule.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	✓	×	✓	×
	✓	✓	✓	✓

If judge 3 only cares about the outcome wrt. the conclusion she has incentive to manipulate.

Strategic Manipulation

As in voting, agents can manipulate (submit untruthful judgment sets) to get a more favourable outcome.

Suppose we are using the premise-based rule.

	p	q	$(p \wedge q) \leftrightarrow r$	r
judge 1	✓	✓	✓	✓
judge 2	×	✓	✓	×
judge 3	×	×	✓	×
	×	✓	✓	×

If judge 3 only cares about the outcome wrt. the conclusion she has incentive to manipulate.

Strategyproofness

Agent i's preferences are modelled as a weak order $\succeq_i$ over judgment sets. We will only look at closeness-respecting preferences.

- $\succeq_i$ is closeness-respecting iff $(J' \cap J_i) \subset (J \cap J_i) \Rightarrow J \succeq_i J'$.

Example: If $J_i = \{p, q, r\}$, $J' = \{p, \neg q, \neg r\}$, $J = \{p, \neg q, r\}$, then $J \succeq_i J'$ because $J' \cap J_i = \{p\} \subset \{p, q\} = J \cap J_i$.

✋ Were judge 3's preferences closeness-respecting?

- Agent i manipulates if she reports a judgment set $J \neq J_i$.
- She has an incentive to do so if $F(\boldsymbol{J}_{-i}, J_i') \succ_i F(\boldsymbol{J})$ for some $J_i' \in \mathcal{J}(\Phi)$.

An aggregation rule F is strategyproof for a given class of preferences if no agent (with such preferences) has incentive to manipulate.

Strategyproof Rules

Independence: outcome on φ depends only on judgments on φ.

Monotonicity: for any $\varphi \in \Phi$ and profiles $\boldsymbol{J}$ and $\boldsymbol{J'}$, $\boldsymbol{J} =_{-i} \boldsymbol{J'}$, and $\varphi \in J'_i \setminus J_i$ for some agent $i \in N$, then $\varphi \in F(\boldsymbol{J}) \Rightarrow \varphi \in F(\boldsymbol{J'})$.

Theorem 7 (Dietrich and List, 2007) F is independent and monotonic iff F is strategyproof for all closeness-respecting preferences.

Proof.

- Independence means we (and the manipulator) can consider one formula at a time. Monotonicity means it can never help to add support to unwanted formulas, or remove support from wanted formulas. Thus, it is always in her best interest to report her truthful judgment set. ✓

Dietrich, F. and List, C. Strategy-Proof Judgment Aggregation. *Economics and Philosophy*, 23(3), 269-300, 2007.

Proof...

- ⇐ Suppose F is strategyproof for all closeness-respecting preferences.

For monotonicity: Let $\boldsymbol{J} =_{-i} \boldsymbol{J}$. Since $J_i \neq J_i'$ there must be some $p \in J_i'$ $p \notin J_i$. Suppose agent i only cares about p. If $p \in F(\boldsymbol{J})$ then $p \in F(\boldsymbol{J'})$—otherwise F would not be strategyproof. ✓

For independence: Take two profiles $\boldsymbol{J}$ and $\boldsymbol{J'}$ s.t. J_i and J_i' agree on p for all agents i. All agents only care about p. Step by step move from $\boldsymbol{J}$ to $\boldsymbol{J'}$.

$$(J_1, J_2, \ldots, J_n) \rightsquigarrow (J_1', J_2, \ldots, J_n) \rightsquigarrow (J_1', J_2', \ldots, J_n) \rightsquigarrow \cdots \rightsquigarrow (J_1', J_2', \ldots, J_n')$$

Each step must preserve collective judgment on p. Otherwise either J_i or J_i' must disagree with outcome. If J_i disagree with outcome, then manipulation possible from J_i ("truthful") to J_i'. If J_i' disagree with outcome, then manipulation possible from J_i' ("truthful") to J_i. ✓

Summary

Things we talked about:

- The Doctrinal Paradox & failure of collective rationality
- Embedding PA in JA
- Domain restrictions (similar to single-peakedness)
- Strategic manipulation

Things we didn't talk about:

- Axiomatic characterisation of aggregation rules
- Agenda characterisations
- Complexity results: there are many