Computational Social Choice: Spring 2019

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

Plan for Today

Our final topic for this course is *judgment aggregation* (JA), where agents judge several propositions to be either true or false and we need to aggregate this information into a single collective judgment.

Today will be an introduction to some of the basic concepts of JA:

- motivating example: *doctrinal paradox*
- general *formal model* for judgment aggregation
- relationship to *preference aggregation*
- a couple of *specific aggregation rules* to use in practice
- a basic *impossibility theorem* (using the *axiomatic method*)

Most of this material is covered 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.

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
Majority	Yes	Yes	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*, 81(1):1–59, 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* $N = \{1, \ldots, n\}$, with $n \geqslant 2$, express judgments on the formulas in Φ, producing a *profile* $\boldsymbol{J} = (J_1, \ldots, J_n)$.

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

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: $F_{\text{maj}}(\boldsymbol{J}) = \{p, \neg q, p \vee q\}$ [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.*

Variants of the Model

Our basic model of JA is due to List and Pettit (2002).

There are several variants where you use an *integrity constraint* Γ (a propositional formula) to further constrain admissible judgments:

$$J \in \mathcal{J}(\Phi, \Gamma) \quad \Leftrightarrow \quad J \cup \{\Gamma\} \text{ is consistent and } J \text{ is complete}$$

Most important instances:

- You get our basic model for $\Gamma = \top$.
- You get "binary aggregation with integrity constraints" when you restrict Φ to being a set of literals.

Refer to our KR-2016 paper for a comparison of these languages.

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

U. Endriss, U. Grandi, R. de Haan, and J. Lang. Succinctness of Languages for Judgment Aggregation. KR-2016.

Embedding Preference Aggregation

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

To translate into JA, make every *ordered pair of alternatives* a variable. Write $p_{x \succ y}$ for the variable corresponding to $(x, y) \in A \times A$.

Build an *integrity constraint* Γ as the conjunction of:

- Irreflexivity: $\neg p_{x \succ x}$ for all $x \in A$
- Completeness: $p_{x \succ y} \vee p_{y \succ x}$ for all $x, y \in A$ with $x \neq y$
- Transitivity: $p_{x \succ y} \wedge p_{y \succ z} \to p_{x \succ z}$ for all $x, y, z \in A$

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

	(x, y)	(x, z)	(y, z)	*corresponding order*
Agent 1	Yes	Yes	Yes	$x \succ y \succ z$
Agent 2	No	No	Yes	$y \succ z \succ x$
Agent 3	Yes	No	No	$z \succ x \succ y$
Majority	Yes	No	Yes	*not a linear order*

Useful Notation: Set of Supporters

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

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

Quota Rules

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

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

A quota rule F_q is called *uniform* if q maps any given formula to the same number λ. Examples:

- The *(strict) majority rule* F_{maj} is the quota rule with $q = \lceil \frac{n+1}{2} \rceil$.
- The *weak majority rule* is the quota rule with $q = \lceil \frac{n}{2} \rceil$.
- The *constant rule* F_0 (F_{n+1}) accepts all (no) formulas.
- The *unanimity rule* $F_n : \boldsymbol{J} \mapsto J_1 \cap \cdots \cap J_n$ accepts φ *iff* all do.
- The *nomination rule* $F_1 : \boldsymbol{J} \mapsto J_1 \cup \cdots \cup J_n$ accepts φ *iff* at least one of the agents does.

Observe that for *odd* n the majority rule and the weak majority rule coincide. For *even* n they differ (and only the weak one is complete).

Example: Supermajority Rules

Uniform quota rules with quota $\lambda > \frac{n}{2}$ are called *supermajority rules*.

Basic intuition:

- High quotas are good for collective consistency.
- Low quotas are good for collective completeness.

Exercise: *Show that the uniform quota rule F_n (with a uniform quota of $\lambda = n$) guarantees consistent outcomes for any agenda.*

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 λ that will guarantee consistent outcomes?*

Premise-Based Aggregation

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

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

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

$$F_{\text{pre}}(\boldsymbol{J}) \quad = \quad \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\},$$
$$\text{where } \Delta = \{\varphi \in \Phi_p \mid |N_\varphi^{\boldsymbol{J}}| > \frac{n}{2}\}$$

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:

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

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

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

Axiomatic Method

So how do you choose the right aggregation rule?

One way is to use the *axiomatic method*, as we saw earlier:

- identify normatively appealing properties of aggregators (*axioms*)
- cast those properties into mathematically rigorous definitions
- explore the consequences: *characterisations* and *impossibilities*

Basic Axioms

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

- *Anonymity:* Treat all agents symmetrically!

  For any profile $\boldsymbol{J}$ and any permutation $\pi : N \to N$, 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 $\boldsymbol{J}$ 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 $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \varphi \in F(\boldsymbol{J'})$.

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

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

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)

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.*

Remark 1: Note that the theorem requires $n \geqslant 2$. (*Why?*)

Remark 2: Similar impossibilities arise for other agendas with some minimal structural richness. (To be discussed later on in the course.)

Remark 3: This is the main result in the original paper introducing the formal model of JA and proposing to apply the axiomatic method.

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

Proof: Part 1

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

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

We observe:

- Due to *independence*, whether $\varphi \in F(\boldsymbol{J})$ only depends on $N_\varphi^{\boldsymbol{J}}$.
- Then, due to *anonymity*, whether $\varphi \in F(\boldsymbol{J})$ only depends on $|N_\varphi^{\boldsymbol{J}}|$.
- Finally, due to *neutrality*, the manner in which the status of $\varphi \in F(\boldsymbol{J})$ depends on $|N_\varphi^{\boldsymbol{J}}|$ 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 $|N_{\varphi}^{\boldsymbol{J}}| = |N_{\psi}^{\boldsymbol{J}}|$, then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.

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

Consider a profile $\boldsymbol{J}$ 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: $|N_p^{\boldsymbol{J}}| = |N_q^{\boldsymbol{J}}| = |N_{\neg(p \wedge q)}^{\boldsymbol{J}}|$. 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 $\boldsymbol{J}$ with $|N_p^{\boldsymbol{J}}| = |N_{\neg p}^{\boldsymbol{J}}|$. Then:

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

Remark: To be exact, you also need, say, $q \in \Phi$ for neutrality to "bite".

Summary

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

- variants of the *model:* $\pm$compound formulas, $\pm$integrity constraint
- examples for *rules:* quota rules, premise-based rule, Kemeny rule
- examples for *axioms:* anonymity, neutrality, independence
- example for a result: basic *impossibility* theorem

What next? Strategic behaviour in judgment aggregation.