Judgment Aggregation

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

Guest Lecture for "Logic, Language and Computation"
Master of Logic, 11 September 2023

Computational Social Choice

My main area of research is computational social choice, which is concerned with questions of either a formal or an algorithmic nature related to collective decision making. Examples:

- voting and preference aggregation
- fair allocation of resources
- judgment aggregation [*today's topic*]

Example

Suppose three robots are in charge of climate control for this building. They need to make judgments on p (the temperature is above 22°C), q (we should switch on the AC), and the "policy" $p \rightarrow q$.

	p	$p \rightarrow q$	q
Robot 1	Yes	Yes	Yes
Robot 2	No	Yes	No
Robot 3	Yes	No	No

Exercise: *Should we switch on the AC?*

Outline

This will be an introduction to the theory of *judgment aggregation*.

- The *paradox* of judgment aggregation: a second example
- Main question: Is there a *reasonable method* of aggregation?
- Briefly: Examples for *recent research* on JA at the ILLC

Example

A defendant is accused of a breach of contract ...

Legal doctrine stipulates that you are *guilty* *if and only if* it is the case that the agreement was *binding* (p) *and* has *not* been *honoured* ($\neg q$).

	p	q	$p \wedge \neg q$
Judge 1	Yes	No	Yes
Judge 2	Yes	Yes	No
Judge 3	No	No	No

Exercise: *Should we pronounce the defendant guilty?*

The Paradox of Judgment Aggregation

Once again our two examples:

	p	$p \to q$	q
Robot 1	Yes	Yes	Yes
Robot 2	No	Yes	No
Robot 3	Yes	No	No

	p	q	$p \wedge \neg q$
Judge 1	Yes	No	Yes
Judge 2	Yes	Yes	No
Judge 3	No	No	No

Why do we call this a *paradox*? Two explanations:

- Premise-driven rule and conclusion-driven rule disagree
- Majority rule produces judgment set that is not consistent

Formal Framework

An *agenda* Φ is a set of propositional formulas (and their negations).

Example: $\Phi = \{p,\ \neg p,\ p \rightarrow q,\ \neg(p \rightarrow q),\ q,\ \neg q\}$

A *judgment set* J for the agenda Φ is a subset of Φ. We call J:

- *complete* if $\varphi \in J$ or $\neg\varphi \in J$ for all formulas $\varphi, \neg\varphi \in \Phi$
- *consistent* if J has a satisfying truth assignment

Now n individual *agents* each express judgments on the formulas in Φ, producing a *profile* $\boldsymbol{J} = (J_1, \ldots, J_n)$ of complete and consistent sets.

Example: $\boldsymbol{J} = (\{p, p \rightarrow q, q\},\ \{\neg p, p \rightarrow q, \neg q\},\ \{p, \neg(p \rightarrow q), \neg q\})$

An *aggregation rule* F for an agenda Φ and a group of n agents is a function mapping every given profile of complete and consistent sets to a single collective judgment set.

Example: Majority Rule

Suppose three agents express judgments on the formulas 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	$p \vee q$	
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	$J_3 = \{\neg p, \neg q, \neg(p \vee q)\}$

The (strict) *majority rule* F_{maj} takes a (complete and consistent) profile and returns the set of formulas accepted by $> \frac{n}{2}$ agents.

In our example: $F_{\text{maj}}(\boldsymbol{J}) = \{p, \neg q, p \vee q\}$ [*complete and consistent!*]

Other Rules

Instead of using the *majority rule*, we could also use:

- *Premise-driven rule:* use majority voting on literals and infer other formulas from the literals accepted
- *Quota-based rules:* e.g., accept a formula if $\geqslant \frac{2}{3}$ of the agents do
- *Distance-based rules:* select a judgment set that is as "close" as possible to what agents want—but without violating consistency

There are other options as well. *So how do you choose?*

The Axiomatic Method

What makes for a "good" aggregation rule F? The following so-called *axioms* all express intuitively appealing properties:

- *Anonymity:* Treat all individual agents the same!
- *Neutrality:* Treat all formulas the same!
- *Independence:* To decide whether to accept formula φ, you should only have to consider which individual agents accept φ!

Exercise: *How might you formalise these axioms?*

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

. . . but so do various other rules!

Exercise: *Can you think of some examples?*

Impossibility Theorem

We have seen that the majority rule does *not* preserve *consistency*.
Is there another "reasonable" rule that does not have this problem?

Surprise: *No!* (at least not for certain agendas)

Theorem 1 (List and Pettit, 2002) *No judgment aggregation rule for $\geqslant 2$ agents and an agenda Φ with $\{p, q, p \wedge q\} \subseteq \Phi$ that satisfies the axioms of anonymity, neutrality, and independence will always return a collective judgment set that is complete and consistent.*

Remark: Also true for other agendas (such as all those we saw today).

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

Proof

First, understand the impact of our three axioms:

- *Independence:* acceptance of φ only depends on *who* accepts φ.
- Add *anonymity:* it only depends on *how many* agents accept φ.
- Add *neutrality:* must use *same* acceptance criterion for all formulas.

We now prove the theorem for *odd* n (it's even easier for even n).

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

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: $|N_p^{\boldsymbol{J}}| = |N_q^{\boldsymbol{J}}| = |N_{\neg(p\wedge q)}^{\boldsymbol{J}}| = \frac{n+1}{2}$. Then:

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

So it indeed is *impossible* to satisfy all of our requirements. ✓

Recent Research on JA at ILLC

To conclude, let's review some recent research on JA at the ILLC, focusing mostly on research involving MoL and PhD students.

Strategic Behaviour

Suppose we use the premise-based rule (with premises = literals):

	p	q	$p \vee q$
Agent 1	No	No	No
Agent 2	Yes	No	Yes
Agent 3	No	Yes	Yes

If agent 3 only cares about the conclusion, then she has an incentive to *manipulate* and pretend that she actually accepts p.

With then-MoL students **Sirin Botan** (PhD Amsterdam 2021) and **Arianna Novaro** (PhD Toulouse 2019) we studied group manipulation.

S. Botan, A. Novaro, and U. Endriss. Group Manipulation in Judgment Aggregation. Proc. 15th Int'l Conf. on Auton. Agents and Multiagent Sys. (AAMAS-2016).

Computational Complexity

Generally applicable aggregation rules that can guarantee consistency tend to be algorithmically demanding. An example is the *Slater rule:*

> *Choose a consistent judgment set minimising the Hamming distance to the (possibly inconsistent) majority outcome.*

But how complex exactly? The cited paper with **Ronald de Haan** (faculty at ILLC) summarises a decade of research on this topic.

U. Endriss, R. de Haan, J. Lang, and M. Slavkovik. The Complexity Landscape of Outcome Determination in Judgment Aggregation. *Journal of AI Research*, 2020.

Truth-Tracking

Suppose each agent reports the correct truth value on each proposition with a given probability. *Can we recover the ground truth?*

We have studied this question in the specific context where we might not want to ask everyone for their judgment on everything.

Zoi Terzopoulou (PhD Amsterdam 2021) was a PhD student at the time, and previously a MoL student (also working on JA then).

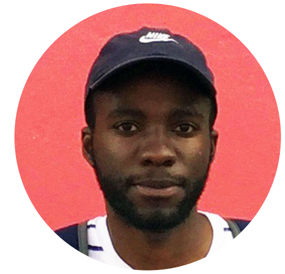
Z. Terzopoulou and U. Endriss. Optimal Truth-Tracking Rules for the Aggregation of Incomplete Judgments. Proc. 12th International Symposium on Algorithmic Game Theory (SAGT-2019).


Crowdsourcing

We designed JA-inspired methods for the *collective annotation* of linguistic corpora (e.g., to identify rhetorical questions) to support research in computational linguistics.

Ciyang Qing (PhD Stanford 2021) and **Justin Kruger** (PhD Paris 2019) were MoL students at the time.

C. Qing, U. Endriss, R. Fernández, and J. Kruger. Empirical Analysis of Aggregation Methods for Collective Annotation. Proc. 25th International Conference on Computational Linguistics (COLING-2014).

JA as a Modelling Tool

Due to its expressive power, JA can serve as a *lingua franca* to move between different frameworks of collective decision making. Example:

Working with agendas of propositions of the form $p_{A \succ B}$, we can simulate preference aggregation from within JA.

Simon Rey and **Julian Chingoma** are PhD students at the ILLC.

S. Rey, U. Endriss, and R. de Haan. Designing Participatory Budgeting Mechanisms Grounded in Judgment Aggregation. Proc. 17th International Conference on Principles of Knowledge Representation and Reasoning (KR-2020).

J. Chingoma, U. Endriss, and R. de Haan. Simulating Multiwinner Voting Rules in Judgment Aggregation. Proc. 21st International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2022).

Economics and Computation at ILLC

JA is highly interdisciplinary and, among others, studied by researchers working at the intersection of Economics and Computer Science.

Other *faculty members* at ILLC working in this domain:

- Davide Grossi
- Ronald de Haan
- Rebecca Reiffenhäuser
- Guido Schäfer

Relevant *courses* in the MoL:

- Game Theory
- Algorithmic Game Theory
- Seminar Economics and Computation
- Computational Social Choice

Relevant research *seminars:*

- Computational Social Choice Seminar
- FOAM Seminar of the TCS Research Unit [*one of multiple topics*]

Last Slide

This has been an introduction to *judgment aggregation*. We saw:

- Formal framework for aggregating views on complex matters
- Modelling *coherent* judgments as *consistent* sets of formulas
- *Paradox:* majority view of coherent judges may be incoherent
- Thus: need to carefully analyse the problem $\rightsquigarrow$ *axiomatic method*
- *Impossibility:* no "reasonable" rule can always be coherent
- *Recent research:* strategic behaviour, computational complexity, truth-tracking, crowdsourcing, modelling, ...