

# Complexity of Judgment Aggregation

Ulle Endriss

Institute for Logic, Language and Computation

University of Amsterdam

[ joint work with Umberto Grandi and Daniele Porello ]

## The Paradox of Judgment Aggregation

Story: three judges have to decide whether the defendant is guilty ...

	$p$	$p \rightarrow q$	$q$
Judge 1:	Yes	Yes	Yes
Judge 2:	No	Yes	No
Judge 3:	Yes	No	No
Majority:	Yes	Yes	No

Paradox: each *individual* judgment set is *consistent*, but the *collective* judgment arrived at using the *majority rule* is 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.

C. List and C. Puppe. Judgment Aggregation: A Survey. *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

## Talk Outline

- Introduction to Judgment Aggregation
- Decision Problems in JA and their Complexity:
  - Winner Determination
  - Strategic Manipulation
  - Safety of the Agenda
- More on Computational Social Choice

## Formal Framework

An *agenda*  $\Phi$  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  $\Phi$  is a subset of  $\Phi$ . 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  $\Phi$ .

Now a finite set of *individuals*  $\mathcal{N} = \{1, \dots, n\}$  express judgments on the formulas in  $\Phi$ , giving rise to a *profile*  $\mathbf{J} = (J_1, \dots, J_n)$ .

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

## Basic Aggregation Procedures

Under the *majority rule*, we accept  $\varphi$  if the number of individuals accepting  $\varphi$  is greater than the number of individuals accepting  $\sim\varphi$ .

	$p$	$p \rightarrow q$	$q$
Judge 1:	Yes	Yes	Yes
Judge 2:	No	Yes	No
Judge 3:	Yes	No	No

Under the *conclusion-based procedure*, we ignore the information on “premises” and decide on “conclusions” by majority.

Under the *premise-based procedure*, we decide on “premises” by majority and use logical inference for the “conclusions”.

Problems with these approaches:

- What’s a principled way of deciding what’s a “premise” and what’s a “conclusion”? (though this might be clear for concrete applications)
- Paradox will only go away for *some* examples.
- CBP: decisions cannot be explained in terms of premises anymore.

## Premise-Based Procedure

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

$$\Phi = \Phi_p \uplus \Phi_c$$

The *premise-based procedure PBP* for  $\Phi_p$  and  $\Phi_c$  is this function:

$$\begin{aligned} \text{PBP}(\mathbf{J}) &= \Delta \cup \{\varphi \in \Phi_c \mid \Delta \models \varphi\}, \\ &\text{where } \Delta = \{\varphi \in \Phi_p \mid \#\{i \mid \varphi \in J_i\} > \frac{n}{2}\} \end{aligned}$$

If we assume (as we shall) that

- the set of premises is the set of literals in the agenda,
- the agenda  $\Phi$  is closed under propositional letters, and
- the number  $n$  of individuals is odd,

then  $\text{PBP}(\mathbf{J})$  will always be *consistent* and *complete*.

## Winner Determination

The winner determination problem for a judgment aggregation procedure  $F$  is defined as follows:

$\text{WINDET}(F)$

**Instance:** Agenda  $\Phi$ , profile  $\mathbf{J} \in \mathcal{J}(\Phi)^n$ , formula  $\varphi \in \Phi$ .

**Question:** Is  $\varphi$  an element of  $F(\mathbf{J})$ ?

This is easy under the (simplified) premise-based procedure:

**Proposition 1**  $\text{WINDET}(\text{PBP})$  is in  $P$ .

Proof: some counting + model checking. ✓

## Example: Strategic Manipulation

Suppose we use the (simplified) premise-based procedure:

	$p$	$q$	$p \vee q$
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 she accepts  $p$ .



## Strategic Manipulation

Let us fix a notion of *preference* over outcomes:

- The *Hamming distance*  $H(J, J')$  between judgment sets  $J$  and  $J'$  is the number of positive agenda formulas on which they differ.
- We say that individual  $i$  *prefers*  $J$  to  $J'$  if  $H(J_i, J) < H(J_i, J')$ .

Now we can define the manipulability problem:

MANIPULABILITY( $F$ )

**Instance:** Agenda  $\Phi$ ,  $J_i \in \mathcal{J}(\Phi)$ , partial profile  $\mathbf{J}_{-i} \in \mathcal{J}(\Phi)^{n-1}$ .

**Question:** Is there a  $J'_i \in \mathcal{J}(\Phi)$  s.t.  $i$  prefers  $F(J'_i, \mathbf{J}_{-i})$  to  $F(J_i, \mathbf{J}_{-i})$ ?

Good news:

**Theorem 2** MANIPULABILITY(PBP) is NP-complete.

Proof by reduction from SAT.

## Distance-Based Procedure

A procedure that is more widely applicable than the premise-based procedure and that is intuitively appealing is *distance-based merging*:

$$\text{DBP}(\mathbf{J}) = \arg \min_{J \in \mathcal{J}(\Phi)} \sum_{i=1}^n H(J, J_i)$$

Remark: The DBP may return a set of tied winners.

Regarding complexity, we have:

**Theorem 3** *Winner determination for the DBP is NP-complete.*

Proof of hardness by reduction from a result in computational social choice (KEMENYScore). Membership via an integer program.

**Conjecture 4** *Manipulability for the DBP is  $\Sigma_2^p$ -complete.*

Membership is clear. Hardness is open.

## The Axiomatic Method

What makes for a “good” aggregation procedure? The standard approach in social choice theory is to formulate “*axioms*”, e.g.:

**Unanimity (U):** If  $\varphi \in J_i$  for all  $i$ , then  $\varphi \in F(\mathbf{J})$ .

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

**Neutrality (N):** For any  $\varphi, \psi$  in the agenda  $\Phi$  and profile  $\mathbf{J} \in \mathcal{J}(\Phi)$ , if for all  $i$  we have  $\varphi \in J_i \Leftrightarrow \psi \in J_i$ , then  $\varphi \in F(\mathbf{J}) \Leftrightarrow \psi \in F(\mathbf{J})$ .

**Independence (I):** For any  $\varphi$  in the agenda  $\Phi$  and profiles  $\mathbf{J}$  and  $\mathbf{J}'$  in  $\mathcal{J}(\Phi)$ , if  $\varphi \in J_i \Leftrightarrow \varphi \in J'_i$  for all  $i$ , then  $\varphi \in F(\mathbf{J}) \Leftrightarrow \varphi \in F(\mathbf{J}')$ .

## Impossibility Theorem

We have seen that the majority rule is not consistent.

Is there a reasonable procedure that is?

**Theorem 5 (List and Pettit, 2002)** *If the agenda contains at least  $p$ ,  $q$  and  $p \wedge q$ , then **no** aggregation procedure producing **consistent** and **complete** judgment sets satisfies all of (A), (N) and (I).*

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

## More Axioms

Two monotonicity axioms, one for independent rules (inter-profile) and one for neutral rules (intra-profile):

**I-Monotonicity** ( $M^I$ ): For any  $\varphi$  in the agenda  $\Phi$  and profiles  $\mathbf{J} = (J_1, \dots, J_i, \dots, J_n)$  and  $\mathbf{J}' = (J_1, \dots, J'_i, \dots, J_n)$  in  $\mathcal{J}(\Phi)$ , if  $\varphi \notin J_i$  and  $\varphi \in J'_i$ , then  $\varphi \in F(\mathbf{J}) \Rightarrow \varphi \in F(\mathbf{J}')$ .

**N-Monotonicity** ( $M^N$ ): For any  $\varphi, \psi$  in the agenda  $\Phi$  and profile  $\mathbf{J}$  in  $\mathcal{J}(\Phi)$ , if  $\varphi \in J_i \Rightarrow \psi \in J_i$  for all  $i$  and  $\varphi \notin J_k$  and  $\psi \in J_k$  for some  $k$ , then  $\varphi \in F(\mathbf{J}) \Rightarrow \psi \in F(\mathbf{J})$ .

Ideally, we'd like consistent outcomes, but instead we just demand:

**Weak Rationality** (WR):  $F(\mathbf{J})$  is complete and complement-free for *all* profiles  $\mathbf{J}$  [and  $F(\mathbf{J})$  includes no contradictions for *some*  $\mathbf{J}$ ]

Remark: the last condition (“non-nullity”) is a minor technicality (always satisfied if  $\Phi$  includes no tautologies) — please ignore

## Safety of the Agenda (SoA)

Given an agenda  $\Phi$  and a list of axioms AX, let  $\mathcal{F}_\Phi[\text{AX}]$  be the set of procedures  $F : \mathcal{J}(\Phi)^n \rightarrow 2^\Phi$  that satisfy all axioms in AX.

We call an agenda  $\Phi$  is *safe* wrt. a class of procedures  $\mathcal{F}_\Phi[\text{AX}]$ , if  $F(\mathbf{J})$  is consistent for every  $F \in \mathcal{F}_\Phi[\text{AX}]$  and every  $\mathbf{J} \in \mathcal{J}(\Phi)$ .

Goal: We want to be able to check the safety of a given agenda for a given class of procedures (characterised in terms of a set of axioms).

We approach this by proving *characterisation results*:

*all*  $F \in \mathcal{F}_\Phi[\text{AX}]$  are consistent  $\Leftrightarrow \Phi$  has such-and-such property

This is similar to *possibility results* proven in the JA literature:

*some*  $F \in \mathcal{F}_\Phi[\text{AX}]$  is consistent  $\Leftrightarrow \Phi$  has such-and-such property

## Majority Rule

It is known (Nehring and Puppe, 2007) that the *majority rule* is consistent on agendas that satisfy the *median property*.

$\Phi$  satisfies the median property (MP), if every inconsistent subset of  $\Phi$  has itself an inconsistent subset of size  $\leq 2$ .

It is also known (folk theorem?) that

$$\mathcal{F}_\Phi[\text{WR}, A, N, I, M^I] = \mathcal{F}_\Phi[\text{WR}, A, N, M^N] = \{\text{majority rule}\}$$

We thus get our first characterisation theorem for free:

**Theorem 6**  $\Phi$  is safe for  $\mathcal{F}_\Phi[\text{WR}, A, N, I, M^I]$  (and thus also for  $\mathcal{F}_\Phi[\text{WR}, A, N, M^N]$ ) iff it satisfies the MP.

K. Nehring and C. Puppe. The Structure of Strategy-proof Social Choice. *Journal of Economic Theory*, 135(1):269–305, 2007.

## Agenda Properties

Call a set of formulas *nontrivially inconsistent* if it is inconsistent but does not contain an inconsistent formula. An agenda  $\Phi$  satisfies

- the *median property* (MP), if every nontrivially inconsistent subset of  $\Phi$  has itself an inconsistent subset of size 2;
- the *simplified MP* (SMP), if every nontrivially inconsistent subset of  $\Phi$  has itself an inconsistent subset  $\{\varphi, \psi\}$  with  $\models \varphi \leftrightarrow \neg\psi$ ;
- the *syntactic SMP* (SSMP), if every nontrivially inconsistent subset of  $\Phi$  has itself an inconsistent subset  $\{\varphi, \neg\varphi\}$ .

$$\text{SSMP} \Rightarrow \text{SMP} \Rightarrow \text{MP}$$



## Characterisation Theorems

We have looked for characterisation theorems for sets of axioms that are a little weaker than those defining the majority rule.

**Theorem 7**  $\Phi$  is safe for  $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{N}, \text{I}]$  iff it satisfies the SMP.

**Theorem 8**  $\Phi$  is safe for  $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{N}]$  iff it satisfies the SMP and does not contain a contradictory formula.

**Theorem 9**  $\Phi$  is safe for  $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{I}]$  iff it satisfies the SSMP.

## Complexity Results

For a given agenda, how hard is it to check safety?

**Theorem 10** *Checking the safety of the agenda is  $\Pi_2^p$ -complete for any of the classes of aggregation procedures considered.*

Approach:

- the typical  $\Pi_2^p$ -complete problem is SAT for QBFs of the form

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

- reduce that problem to the problem of checking the SSMP, to establish  $\Pi_2^p$ -hardness of the latter (similarly for SMP and MP)
- prove that checking the SSMP, SMP, MP are all *in*  $\Pi_2^p$
- apply the characterisation theorems

## Last Slide

- We have have seen several results in *judgment aggregation*:
  - *Manipulation* tends to be harder than *winner determination* (good)
  - *SoA* requires simplistic agendas and is hard to check (bad)
- This is an example for work in *Computational Social Choice*, combining ideas from economics (particularly social choice theory) and CS.
  - SCT: preference aggregation, voting, fair division, matching, ...
  - CS: algorithms, complexity, logic, knowledge representation, ...

For more information, see the COMSOC website:

<http://www.illc.uva.nl/COMSOC/>

U. Endriss, U. Grandi, and D. Porello. Complexity of Winner Determination and Strategic Manipulation in Judgment Aggregation. Proc. COMSOC-2010.

U. Endriss, U. Grandi, and D. Porello. Complexity of Judgment Aggregation: Safety of the Agenda. Proc. AAMAS-2010.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. A Short Introduction to Computational Social Choice. Proc. SOFSEM-2007.