The Problem of the Safety of the Agenda in Judgment Aggregation

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

[joint work with Umberto Grandi and Daniele Porello]

Social Choice Theory

SCT studies collective decision making: how should we aggregate the preferences of the members of a group to obtain a "social preference"?

$$\triangle \succ_1 \bigcirc \succ_1 \square$$

$$\square \succ_2 \triangle \succ_2 \bigcirc$$

$$\bigcirc \succ_3 \square \succ_3 \triangle$$

$$\overline{}$$

$$?$$

SCT is traditionally studied in Economics and Political Science, but now also by "us": *Computational Social Choice*.

Talk Outline

- Introduction to Judgment Aggregation
- A new problem: Safety of the Agenda
- Some Results: Characterisation and Complexity

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

The Doctrinal Paradox

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.

Formal Framework

An *agenda* Φ is a finite nonempty set of propositional formulas not containing any double negations such that $\alpha \in \Phi \;\Rightarrow\; {\sim}\alpha \in \Phi$.

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

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

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

Now a finite set of *individuals* $N = \{1, \ldots, n\}$ with $n \geqslant 3$ express judgments on Φ, giving rise to a *profile* $\mathbf{J} = (J_1, \ldots, J_n)$.

An *aggregation procedure* for agenda Φ 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: $F : J(\Phi)^n \to 2^\Phi$.

Axioms

Use *axioms* to express desiderata for F. Examples:

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

Neutrality (N): For any φ, ψ in the agenda Φ and profile $\mathbf{J} \in J(\Phi)$, 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 φ in the agenda Φ and profiles $\mathbf{J}$ and $\mathbf{J}'$ in $J(\Phi)$, if $\varphi \in J_i \Leftrightarrow \varphi \in J'_i$ for all i, then $\varphi \in F(\mathbf{J}) \Leftrightarrow \varphi \in F(\mathbf{J}')$.

Systematicity (S) = (N) + (I)

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

Impossibility Theorem

We have seen that the majority rule is not consistent.

Is there a reasonable procedure that is?

Theorem 1 (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 both* (A) *and* (S)*.*

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

Weak Rationality

Instead of always requiring consistent outcomes, use this axiom:

> **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 1: the second condition ("non-nullity") is a minor technicality (always satisfied if Φ includes no tautologies) — please ignore

Remark 2: the majority rule does satisfy all of (WR), (A), (S)

Monotonicity Axioms

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

I-Monotonicity (M^{I}): For any φ in the agenda Φ and profiles $\mathbf{J} = (J_1, \ldots, J_i, \ldots, J_n)$ and $\mathbf{J}' = (J_1, \ldots, J_i', \ldots, J_n)$ in $J(\Phi)$, 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 φ, ψ in the agenda Φ and profile $\mathbf{J}$ in $J(\Phi)$, 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 $\varphi \in F(\mathbf{J}) \Rightarrow \psi \in F(\mathbf{J})$.

<u>Remark:</u> only (M^{I}) seems to show up in the literature

Classes of Aggregation Procedures

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

Proposition 2 $\mathcal{F}_\Phi[\text{WR,A,S,M}^\text{I}] = \mathcal{F}_\Phi[\text{WR,A,N,M}^\text{N}]$ *is empty if n is even and it is a set including only the majority rule if n is odd.*

Further interesting combinations of axioms:

- dropping monotonicity: $\mathcal{F}_\Phi[\text{WR,A,S}]$, $\mathcal{F}_\Phi[\text{WR,A,N}]$, $\mathcal{F}_\Phi[\text{WR,A,I}]$
- $\mathcal{F}_\Phi[\text{A,S,M}^\text{I}]$, the *uniform quota rules* (Dietrich and List, 2007)

F. Dietrich and Ch. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Theoretical Politics*, 19(4):529–565, 2007.

Safety of the Agenda

A new concept in JA, practice-inspired:

Definition 1 *An agenda Φ is safe wrt. a class of procedures $\mathcal{F}$, if $F(\mathbf{J})$ is consistent for every $F \in \mathcal{F}$ and every $\mathbf{J} \in J(\Phi)$.*

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 $\Leftrightarrow$ Φ 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 $\Leftrightarrow$ Φ has such-and-such property

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

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

Agenda Properties

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

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

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

Characterisation Theorems I

The first is a known result (Nehring and Puppe, 2007):

Theorem 3 Φ *is safe for* $\mathcal{F}_\Phi$[WR,A,S,M$^{\mathsf{I}}$] *iff it satisfies the MP.*

Remark: $\mathcal{F}_\Phi$[WR,A,S,M$^{\mathsf{I}}$] includes just one rule (the majority rule), thus possibility theorem and characterisation theorem coincide.

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

Characterisation Theorems II

Three new characterisation results:

Theorem 4 Φ *is safe for* $\mathcal{F}_\Phi[\text{WR,A,S}]$ *iff it satisfies the SMP.*

Theorem 5 Φ *is safe for* $\mathcal{F}_\Phi[\text{WR,A,N}]$ *iff it satisfies the SMP and does not contain a contradictory formula.*

Theorem 6 Φ *is safe for* $\mathcal{F}_\Phi[\text{WR,A,I}]$ *iff it satisfies the SSMP.*

Reformulation of a result by Dietrich and List (2007):

Theorem 7 *Let* $k \geqslant 2$*.* Φ *is safe for the class of uniform quota rules with a quota* m *satisfying* $m > n - \frac{n}{k}$ *iff* Φ *satisfies the* $k\mathrm{MP}$*.*

F. Dietrich and Ch. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Theoretical Politics*, 19(4):529–565, 2007.

Complexity Results

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

Theorem 8 *Checking the safety of the agenda is Π_2^p-complete for any of the classes of aggregation procedures considered.*

Approach:

- the typical Π_2^p-complete problem is SAT for QBFs of the form

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

- reduce that problem to the problem of checking the SSMP, to establish Π_2^p-*hardness* of the latter (similarly for SMP, MP, kMP)
- prove that checking the SSMP, SMP, MP, kMP are all *in* Π_2^p
- apply the characterisation theorems

Last Slide

- New problem in JA: *Safety of the Agenda*
- *Characterisation results* for safe agendas for classes of aggregation procedures induced by natural axioms
- *Complexity results* showing how hard it is to check safety: second level of the polynomial hierarchy (probably worse than NP)
- Conclusion: ensuring safety requires simplistic agendas; checking that those simplistic properties hold is hard (but not impossible)
- The technical results are from a paper due to be presented at AAMAS-2010 and available from my website:

`http://www.illc.uva.nl/~ulle/pubs/`

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