

Complexity of Judgment Aggregation: Safety of the Agenda

Ulle Endriss Umberto Grandi Daniele Porello

Institute for Logic, Language and Computation
University of Amsterdam

27 Gennaio 2010

Computational Complexity and Judgment Aggregation

Judgment aggregation is relevant for multi-agent systems

JA \iff **MAS**

Societies of agents have to **aggregate** their opinions/judgments

Is it feasible?

Computational complexity provides a first measure.

Feedback for JA:

- **precise formal** framework;
- new problems arise: **safety of the agenda**.
- inspired by work on complexity of manipulation in **voting theory**;

Computational Social Choice

This is a common approach in Computational Social Choice:

Apply tools from **theoretical computer science** to analyse methods for **collective decision making**.

A classical example from voting theory:

Theorem (Gibbard-Satterthwaite, 1973/5)

Every non-dictatorial voting rule for more than 3 candidates must be manipulable.

All reasonable voting rules gives incentives to misrepresent your preferences
BUT

For some of them manipulation is **very hard**:

Theorem (Bartholdi and Orlin 1991)

Manipulation for Single Transferable Vote is NP-hard.

Outline

1. **Basic framework** of judgment aggregation in propositional logic.
2. **Safety** of the agenda: is consistency guaranteed for all aggregators?
3. **Characterisation** results for agendas: when is safety guaranteed?
4. **Complexity** results: checking safety is Π_p^2 -complete.

Discursive Dilemma

JA was developed to generalise and study **paradoxical situations** that arises when a collective judgment has to be made on a set of **correlated** propositions.

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

Each individual is rational (i.e., has a consistent judgment)
but the majority is **contradictory!**

Kornhauser and Sager, Unpacking the court. Yale Law Journal, 1986.

List and Pettit, Aggregating sets of judgments. Economics and Philosophy, 2002

Basic Definitions

A set N of individuals expressing judgments on a set of correlated propositions:

Definition

An **agenda** is a finite subset of propositional formulas $\Phi \subseteq \mathcal{L}_{PS}$ closed under complementation and not containing double negations.

A **judgment set** on an agenda Φ is a subset $J \subseteq \Phi$. Call J :

Complete: if for all $\alpha \in \Phi$ either α or its complement is in J .

Complement-free: α and its complement are never both in J .

Consistent: there exists an assignment that makes all formulas in J true.

Call $J(\Phi)$ the set of all consistent and complete judgment sets.

Definition

An **aggregation procedure** for agenda Φ and a set N of n individuals is a function $F : J(\Phi)^n \rightarrow 2^\Phi$.

Axioms for Aggregation Procedures I

Aggregation procedures have been studied using the **axiomatic method**, listing axioms as desirable properties of the functions:

Weak Rationality (WR): $F(\mathbf{J})$ is complete and complement-free.

Other standard axioms:

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

Anonymity (A): F is symmetric with respect to individuals.

Neutrality (N): If φ and ψ share the same pattern of individuals' judgments then their outcome must be the same.

Independence (I): The outcome of F over φ depends only on the individuals' judgments over φ .

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

Two forms of **monotonicity** for independent (M^I) and neutral procedures (M^N).

Axioms for Aggregation Procedures II

We can play with this formalism to get (small) interesting results:

Lemma

*If an agenda Φ contains a **tautology**, then every aggregation procedure for Φ that satisfies (WR) and (S) is unanimous (U).*

Proof. φ^T is in all individual judgment sets and $\varphi^T \in F(\mathbf{J})$ for a profile \mathbf{J} . If a formula is unanimously accepted it behaves like φ^T and by (S) it must be accepted.

Lemma

*If the number of individuals is **even**, then there exists no aggregation procedure that satisfies (WR), (A) and (N).*

Proof. By (N) and (A) the acceptance of a formula φ depends only on the number of individuals supporting φ in profile \mathbf{J} . The profile where half of individuals accepts φ and half accepts $\neg\varphi$ is in contradiction with (WR).

Classes of Aggregation Procedures

Such axioms can be used to define different **classes of aggregation procedures**:

$$\begin{array}{ccc} \text{Set of axioms AX} & \Rightarrow & \text{Class of functions} \\ \text{Agenda } \Phi & & \mathcal{F}_\Phi[\text{AX}] \end{array}$$

Proposition (representation result)

$\mathcal{F}_\Phi[\text{WR}, A, S, M^1] = \mathcal{F}_\Phi[\text{WR}, A, N, M^N]$ is the majority rule.

We will concentrate on various weakening of the axiomatisation of majority:

$$\mathcal{F}_\Phi[\text{WR}, A, S] \quad \mathcal{F}_\Phi[\text{WR}, A, N]$$

$$\mathcal{F}_\Phi[\text{WR}, A, I] \quad \mathcal{F}_\Phi[A, S, M^1]$$

Similar **representation results** can be proved for these classes.

E.g., $\mathcal{F}_\Phi[A, S, M^1]$ is the class of uniform quota rules, (Dietrich and List, 2007)

Safety of the Agenda I

Call a function **consistent** if $F(\mathbf{J})$ is consistent for all profiles \mathbf{J} in the domain.

A **discursive dilemma** for F is a profile $\mathbf{J} = (J_1, \dots, J_n)$ such that the outcome $F(\mathbf{J})$ is inconsistent.

Safety of the Agenda I

Call a function **consistent** if $F(\mathbf{J})$ is consistent for all profiles \mathbf{J} in the domain.

Possibility (characterisation) results:

There exists a consistent F in $\mathcal{F}_\Phi[AX]$ \Leftrightarrow Φ has a certain property.

These results characterise agendas for which consistent aggregation is **possible**.

E.g., Nehring and Puppe (2006), Dietrich and List (2007), Dokow and Holzman (2005).

Safety (characterisation) results:

All F in $\mathcal{F}_\Phi[AX]$ are consistent \Leftrightarrow Φ has a certain property.

These results characterise agendas for which consistent aggregation is **guaranteed**.

Safety of the Agenda II

Definition

An agenda Φ is *safe* with respect to a class of aggregation procedures \mathcal{F}_Φ if every function in \mathcal{F}_Φ is consistent.

A discursive dilemma is a counterexample for the safety of an agenda.

Proposition

Φ is safe for $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{S}, \text{M}^1]$ iff every inconsistent subset of Φ contains an inconsistent subset of size 2 (median property MP).

Proof.

Majority rule is the only function in $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{S}, \text{M}^1]$.

Therefore proving *safety* = proving *possibility*.

The proof follows from Nehring and Puppe (2006). □

Agenda Properties

An agenda Φ satisfies:

- the **k -median property** (k MP) for $k \geq 2$, if every inconsistent subset of Φ has itself an inconsistent subset of size at most k (2MP=MP).
- the **simplified median property** (SMP), if every nontrivially inconsistent subset of Φ (i.e., not containing any contradictory formula) has itself an inconsistent subset of the form $\{\varphi, \psi\}$ with $\models \varphi \leftrightarrow \neg\psi$.
- the **syntactic simplified median property** (SSMP), if every nontrivially inconsistent subset of Φ has an inconsistent subset of the form $\{\varphi, \neg\varphi\}$.

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

Agendas satisfying the SSMP are composed of **uncorrelated** formulas, i.e., they are essentially equivalent to agendas composed by atoms alone.

E.g. $\Phi = \{p \wedge q, \neg(p \wedge q), r \rightarrow (p \vee \neg p), \neg(r \rightarrow (p \vee \neg p))\}$

Characterisation Results I

Proposition

Φ is safe for $\mathcal{F}_\Phi[\text{WR}, A, S]$ if and only if Φ satisfies the SMP.

Proof.

\Leftarrow) If $F(\mathbf{J})$ is inconsistent then it contains a subset $\{\varphi, \psi\}$ with $\varphi \leftrightarrow \neg\psi$.

By individual rationality we know that $\varphi \in J_i \Leftrightarrow \neg\psi \in J_i$.

By systematicity this implies $\varphi \in F(\mathbf{J}) \Leftrightarrow \neg\psi \in F(\mathbf{J})$.

Since $\{\varphi, \psi\} \subseteq F(\mathbf{J})$ we should have also $\neg\psi \in F(\mathbf{J})$, contradicting (WR).

\Rightarrow) Suppose Φ violates the SMP: we can build a function in $\mathcal{F}_\Phi[\text{WR}, A, S]$ that will generate a discursive dilemma.



Safety Characterisation Results

We prove the following:

Proposition

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

Φ is safe for $\mathcal{F}_\Phi[\text{WR}, \text{A}, \text{I}]$ if and only if Φ satisfies the SSMP.

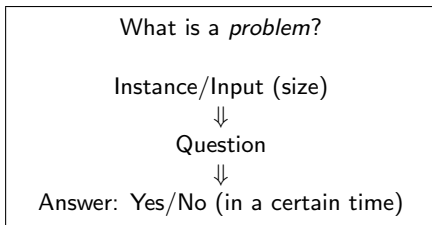
Reformulating work proved by Dietrich and List (2007) we obtain:

Proposition

Let $k \geq 2$. An agenda Φ is safe for the class of uniform quota rules F_m for n individuals satisfying $m > n - \frac{n}{k}$ if and only if Φ satisfies the k MP

Computational Complexity: what is a problem?

Computational Complexity provides a **measure** of how difficult a *problem* is, by measuring the amount of resources needed to solve it: space or **time**.



Important remarks:

- Complexity is **not** a property of a particular procedure to solve a problem.
- Complexity is **not** a property of an instance of a problem: it is a function that links the size of an input to the time required to compute the answer.

Safety of the Agenda: Identifying the problem

Our problem:
Is an agenda Φ **safe** for the class $\mathcal{F}_\Phi[AX]$?

We can define the following problem:

SAFETY[AX]	
Instance:	Agenda Φ of size n
Problem:	Is Φ safe for $\mathcal{F}_\Phi[AX]$?

Using our characterisation, e.g. for $\mathcal{F}_\Phi[WR, A, I]$, this boils down to checking the **agenda property**:

SSMP	
Instance:	Agenda Φ of size n
Problem:	Does Φ satisfy SSMP?

An example (an *instance*):

Is agenda $\{p, \neg p, p \wedge q, \neg(p \wedge q)\}$ safe for $\mathcal{F}_\Phi[WR, A, I]$ rules?

The same can be done for the remaining problems SMP, MP and k MP.

Complexity Classes: Reductions

How to **compare** two problems?

If P_1 can be “easily” translated into problem P_2
then P_2 is at least as difficult as P_1 .

Idea: there exists a **reduction** from $P_1 \leq_r P_2$
s.t. $x \in P_1 \Leftrightarrow r(x) \in P_2$

Every problem P defines the class C_P of problems that can be reduced to it.

Famous complexity classes:

- P is the class of problems solvable in polynomial time.
- NP is the class of problems solvable in non-deterministic polynomial time.

Complexity Classes: Complete Problems

Every (interesting) class C has complete problems.

P is **complete** for C if $C=C_P$

They are reference points that fully characterise their class.

SAT

Instance: Propositional formula φ

Problem: Is φ satisfiable?

SAT is **NP**-complete (Cook, 1971).

We are interested in Π_2^P , the class defined by the following satisfaction problem:

coQSAT₂

Instance: Quantified boolean formula (QBF)

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

Problem: Is the QBF satisfiable (i.e., true)?

Computational Complexity of SSMP

Checking safety turns out to be quite hard:

Theorem

SAFETY[WR, A, I] is Π_2^P -complete.

Proof.

- By characterisation this problem is equivalent to SSMP;
- SSMP is in Π_2^P (**membership**): producing an inconsistent subset of Φ not containing any $\{\varphi, \neg\varphi\}$ is a Π_2^P algorithm.
- SSMP is Π_2^P -complete (**hardness**): by reduction from co-QSAT₂. For every QBF Δ we can build an agenda Φ_Δ such that if Φ_Δ satisfies the SSMP iff the initial QBF Δ is satisfiable.



Complexity results

The following theorem is an overview of our complexity results:

Theorem

Checking the safety of an agenda is Π_2^P -complete for any of these classes of aggregation procedures:

- 1. the majority rule, corresponding to $\mathcal{F}_\Phi[\text{WR}, A, S, M^I]$ and $\mathcal{F}_\Phi[\text{WR}, A, N, M^N]$;*
- 2. systematic rules: $\mathcal{F}_\Phi[\text{WR}, A, S]$;*
- 3. neutral rules: $\mathcal{F}_\Phi[\text{WR}, A, N]$;*
- 4. independent rules: $\mathcal{F}_\Phi[\text{WR}, A, I]$;*
- 5. for any $k \geq 2$, the class of uniform quota rules F_m with $m > n - \frac{n}{k}$, where n is the number of individuals. $\mathcal{F}_\Phi^{k,m,n}[A, S, M^I]$.*

Proof.

SSMP, SMP, k MP for $k \geq 2$, are all Π_2^P -complete.



Conclusions and Future Work

In conclusion:

- To the best of our knowledge the **safety** of the agenda is a new problem relevant to an application-driven study of JA.
- We have proved several characterisation results to **recognise safe agendas** for different classes of functions.
- We have proved that this problem is **Π_2^P -complete** for all classes of functions we considered.

Guaranteeing safety for various classes of functions is a very **restrictive** requirement, and checking these simple properties is **computationally hard**.

Future work includes:

- Study the complexity of the safety problem for other classes of functions.
- Introduce and study the safety of a **profile**.
- Study the complexity of **strategic manipulation** in JA.