

Tutorial on Judgment Aggregation

Ulle Endriss

Institute for Logic, Language and Computation
University of Amsterdam

[<http://www.illc.uva.nl/~ulle/teaching/aamas-2013/>]

Example

Three agents hold different views on the truth of the propositions p , q , and $p \rightarrow q$ (e.g., p might stand for “the temperature is below 16°C” and q for “we should switch off the air conditioning”).

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

What should be the *collective decision* of the group?

Multiagent Systems

How is this relevant to Multiagent Systems?

- Agents need to make decisions together, which will often involve aggregating their distinct views on the state of the world.
JA offers a simple and elegant model for analysing this kind of multiagent decision making.
- Just as for other theoretical frameworks originating outside of the MAS research area (e.g., game theory, social choice theory), it turns out that the algorithmic perspective taken in MAS can make a contribution to the field of JA itself.

Tutorial Outline

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

Main topics to be covered:

- Basics: paradoxes, formal framework
- Axiomatic method: impossibilities and characterisations
- Concrete methods of aggregation
- Agenda characterisation results
- Strategic considerations
- ... and a few words on possible applications

The Doctrinal Paradox

Suppose 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 it 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

Paradox: Taking majority decisions on the *premises* (p and q) and then inferring the conclusion (r) yields a different result from taking a majority decision on the *conclusion* (r) directly.

L.A. Kornhauser and L.G. Sager. The One and the Many: Adjudication in Collegial Courts. *California Law Review*, 81(1):1–59, 1993.

The Discursive Dilemma

Our judges were expressing judgements on *atoms* (p, q, r) and consistency of a judgement set was evaluated wrt. an *integrity constraint* ($r \leftrightarrow p \wedge q$).

Alternatively, we could allow judgements on *compound formulas*. Examples:

	p	q	$p \wedge q$		p	q	$r \leftrightarrow p \wedge q$	r
Judge 1:	Yes	Yes	Yes	Judge 1:	Yes	Yes	Yes	Yes
Judge 2:	No	Yes	No	Judge 2:	No	Yes	Yes	No
Judge 3:	Yes	No	No	Judge 3:	Yes	No	Yes	No
Majority:	Yes	Yes	No	Majority:	Yes	Yes	Yes	No

From now on we will work with a framework without integrity constraints (“legal doctrines”), where all inter-relations between propositions stem from the logical structure of those propositions themselves.

In the philosophical literature, the term *doctrinal paradox* is reserved for the first version of our paradox, and the more general term *discursive dilemma* is used when there is no external “doctrine” that is responsible for the problem.

Why Paradox?

Again, what's paradoxical about our example?

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

Explanation 1: Two seemingly reasonable methods of aggregation, the *premise-based procedure* and the *conclusion-based procedure*, produce different outcomes.

Explanation 2: Each individual judgment set is logically consistent, but applying the seemingly reasonable *majority rule* to all propositions yields a collective judgment set that is inconsistent.

Formal Framework

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 *individuals* $\mathcal{N} = \{1, \dots, n\}$, with $n \geq 2$, express judgments on the formulas in Φ , producing a *profile* $\mathbf{J} = (J_1, \dots, J_n)$.

An *aggregation procedure* for an 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: $F : \mathcal{J}(\Phi)^n \rightarrow 2^\Phi$.

Example: Majority Rule

Suppose three agents ($\mathcal{N} = \{1, 2, 3\}$) 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	$p \vee q$	
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 propositions accepted by $> \frac{n}{2}$ agents.

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

In general, F_{maj} only ensures completeness and complement-freeness [and completeness only in case n is odd].

Example: Embedding Preference Aggregation

In *preference aggregation*, individuals express preferences (linear orders) over a set of alternatives \mathcal{X} and we need to find a collective preference.

We can embed this into JA (suppose $\mathcal{X} = \{A, B, C\}$):

- Take atomic propositions to be $[A \succ A]$, $[A \succ B]$, ...
- Suppose all individuals accept these propositions:
 - Irreflexivity: $\neg[A \succ A]$, $\neg[B \succ B]$, $\neg[C \succ C]$
 - Completeness: $[A \succ B] \vee [B \succ A]$ etc.
 - Transitivity: $[A \succ B] \wedge [B \succ C] \rightarrow [A \succ C]$, etc.

Then the famous *Condorcet paradox* corresponds to this example in JA:

	$[A \succ B]$	$[A \succ C]$	$[B \succ C]$	<i>corresponding order</i>
Agent 1:	Yes	Yes	Yes	$A \succ B \succ C$
Agent 2:	No	No	Yes	$B \succ C \succ A$
Agent 3:	Yes	No	No	$C \succ A \succ B$
Majority:	Yes	No	Yes	<i>not a linear order</i>

Axioms

What makes for a “good” aggregation procedure F ? The following *axioms* all express intuitively appealing (yet, debatable) properties:

- *Anonymity*: Treat all individuals symmetrically!

Formally: for any profile \mathbf{J} and any permutation $\pi : \mathcal{N} \rightarrow \mathcal{N}$ we have $F(J_1, \dots, J_n) = F(J_{\pi(1)}, \dots, J_{\pi(n)})$.

- *Neutrality*: Treat all propositions symmetrically!

Formally: for any φ, ψ in the agenda Φ and any profile \mathbf{J} , if for all $i \in \mathcal{N}$ we have $\varphi \in J_i \Leftrightarrow \psi \in J_i$, then $\varphi \in F(\mathbf{J}) \Leftrightarrow \psi \in F(\mathbf{J})$.

- *Independence*: Only the “pattern of acceptance” should matter!

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

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

(But so do some other procedures! Can you think of some examples?)

Impossibility Theorem

We have seen that the majority rule is *not consistent*. Is there another “reasonable” aggregation procedure that does not have this problem? *Surprisingly, no!* (at least not for certain agendas)

Theorem 1 (List and Pettit, 2002) *No judgment aggregation procedure for 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 1: Note that the theorem requires $|\mathcal{N}| > 1$.

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

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

Proof: Part 1

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

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

We observe:

- Due to *independence*, whether $\varphi \in F(\mathbf{J})$ only depends on $N_\varphi^{\mathbf{J}}$.
- Then, by *anonymity*, whether $\varphi \in F(\mathbf{J})$ only depends on $|N_\varphi^{\mathbf{J}}|$.
- Finally, due to *neutrality*, the manner in which $\varphi \in F(\mathbf{J})$ depends on $|N_\varphi^{\mathbf{J}}|$ must itself *not* depend on φ .

Thus: if φ and ψ are accepted by the same number of individuals, then we must either accept both of them or reject both of them.

Proof: Part 2

Recall: For all $\varphi, \psi \in \Phi$, if $|N_{\varphi}^{\mathbf{J}}| = |N_{\psi}^{\mathbf{J}}|$, then $\varphi \in F(\mathbf{J}) \Leftrightarrow \psi \in F(\mathbf{J})$.

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

Consider a profile \mathbf{J} where $\frac{n-1}{2}$ individuals accept p and q ; one each accept exactly one of p and q ; and $\frac{n-3}{2}$ accept neither p nor q .

That is: $|N_p^{\mathbf{J}}| = |N_q^{\mathbf{J}}| = |N_{\neg(p \wedge q)}^{\mathbf{J}}|$. 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 any assumptions regarding the structure of the agenda:

Consider a profile \mathbf{J} with $|N_p^{\mathbf{J}}| = |N_{\neg p}^{\mathbf{J}}|$. Then:

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

Change of Perspective

Does the impossibility theorem mean that all hope is lost? *No.*

- We could analyse in more depth for *what agendas* this problem can actually occur. And if it can, we could analyse *how to detect* such a situation. We will follow this route a little later.
- We could argue that it is ok to *weaken those axioms*:
 - *Anonymity*: maybe some agents are smarter than others?
 - *Neutrality*: maybe it is actually ok to treat, say, atomic propositions differently from conjunctions?
 - *Independence*: there *are* logical dependencies between propositions; so why not allow them to affect aggregation?

Next we look at some practical aggregators that circumvent the noted impossibility (i.e., they all must violate at least one of the axioms).

Quota Rules

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

$$F_q(\mathbf{J}) = \{\varphi \in \Phi \mid |N_\varphi^{\mathbf{J}}| \geq q(\varphi)\}$$

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

- The *unanimous rule* F_n accepts φ iff everyone does.
- The *constant rule* F_0 (F_{n+1}) accepts all (no) formulas.
- 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 $q = \lceil \frac{n}{2} \rceil$.

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).

Quota Rules with a High Quota

Clearly, a (uniform) quota rule with a sufficiently high quota will be consistent. Dietrich and List (2007) give lower bounds for the quota to ensure consistency as a function of n and the size of the largest *minimally inconsistent subset* of the agenda Φ . Example:

Let $\Phi = \{p, \neg p, q, \neg q, p \wedge q, \neg(p \wedge q)\}$. The largest mi-subset is $\{p, q, \neg(p \wedge q)\}$. Any quota $> \frac{2}{3}$ will ensure consistency.

But: We (may) lose completeness of the collective judgment set.

F. Dietrich and C. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Journal of Theoretical Politics*, 19(4):391–424, 2007.

Characterisation of Quota Rules

Quota rules are nice to demonstrate the axiomatic method ...

One more axiom:

- *Monotonicity*: If an accepted proposition gets additional support, then we should continue to accept it!

Formally: for any $\varphi \in \Phi$ and profiles \mathbf{J}, \mathbf{J}' , if $\varphi \in J'_{i^*} \setminus J_{i^*}$ for some i^* and $J_i = J'_i$ for all $i \neq i^*$, then $\varphi \in F(\mathbf{J}) \Rightarrow \varphi \in F(\mathbf{J}')$.

We can now *characterise* the class of quota rules:

Proposition 2 (Dietrich and List, 2007) *An aggregation procedure is anonymous, independent and monotonic iff it is a quota rule.*

F. Dietrich and C. List. Judgment Aggregation by Quota Rules: Majority Voting Generalized. *Journal of Theoretical Politics*, 19(4)391–424, 2007.

Proof

Claim: *anonymous* + *independent* + *monotonic* \Leftrightarrow *quota rule*

Clearly, any quota rule has these properties (right-to-left).

For the other direction (using the same technique as before):

- Independence means that acceptance of φ only depends on $N_\varphi^{\mathbf{J}}$.
- Anonymity means that, in fact, it only depends on $|N_\varphi^{\mathbf{J}}|$.
- Monotonicity means that acceptance of φ cannot turn to rejection as additional individuals accept φ .

Hence, it must be a quota rule. ✓

Reminder: $N_\varphi^{\mathbf{J}}$ is the set of individuals who accept φ in profile \mathbf{J} .

More Characterisations

Clearly, a quota rule F_q is uniform *iff* it is neutral. Thus:

Corollary 3 *An aggregation procedure is **anonymous, neutral, independent and monotonic** (= ANIM) **iff** it is a **uniform quota rule**.*

Now consider a uniform quota rule F_q with quota q . Two observations:

- For F_q to be **complete**, we need $q \leq \max_{0 \leq x \leq n} (x, n-x) \Rightarrow q \leq \lceil \frac{n}{2} \rceil$.
- For F_q to be **compl.-free**, we need $q > \min_{0 \leq x \leq n} (x, n-x) \Rightarrow q > \lfloor \frac{n}{2} \rfloor$.

For n **even**, no such q exists. Thus:

Proposition 4 *For n **even**, no aggregation procedure is **ANIM, complete and complement-free**.*

For n **odd**, such a q does exist, namely $q = \lceil \frac{n}{2} \rceil = \frac{n+1}{2}$. Thus:

Proposition 5 *For n **odd**, an aggregation procedure is **ANIM, complete and complement-free** **iff** it is the (strict) **majority rule**.*

The Premise-Based Procedure

Suppose we *can* divide the agenda into *premises* and *conclusions* (i.e., we are willing to give up *neutrality*):

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

The *premise-based procedure* PBP for Φ_p and Φ_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 that

- the set of *premises* is the set of *literals* in the agenda,
- the agenda Φ is *closed under propositional letters*, and
- the number n of individuals is *odd*,

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

Example: Violation of Unanimity

Consider the following basic axiom (satisfied by almost everything):

- *Unanimity*: Unanimous acceptance implies collective acceptance!

Formally: if $\varphi \in J_i$ for all $i \in \mathcal{N}$, then $\varphi \in F(\mathbf{J})$.

Curiously, the premise-based procedure violates unanimity:

	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
PBP:	No	No	No	No

Distance-Based Procedures

A general approach to designing aggregation procedures is to fix a notion of *distance* between judgments sets and then to use it to define what it means for a judgment set to be *closest* to the input profile amongst all consistent judgment sets and to then return that set.

The most widely used distance is the *Hamming distance*:

$$H(J, J') = \frac{1}{2} \cdot |J \setminus J' \cup J' \setminus J|$$

There are several ways of turning this into an aggregator ...

G. Pigozzi. Belief Merging and the Discursive Dilemma: An Argument-based Account of Paradoxes of Judgment. *Synthese*, 152(2):285–298, 2006.

M.K. Miller and D. Osherson. Methods for Distance-based Judgment Aggregation. *Social Choice and Welfare*, 32(4):575–601, 2009.

The Standard Distance-Based Procedure

The standard *distance-based procedure* is defined as follows:

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

Recall that $\mathcal{J}(\Phi)$ is the set of all complete and consistent judgment sets. So the DBP is complete and consistent *by definition*.

Some remarks:

- The DBP may return a set of tied winners (“*irresolute*”).
- It is *not independent*.
- If the majority winner is consistent, then it is also the DBP-winner.

For those who know about voting and preference aggregation: the DBP is based on the same principle as the *Kemeny rule*.

Another Distance-Based Procedure

Rather than selecting a consistent outcome that is closest to the profile, we may select a consistent outcome that is closest to the (possibly inconsistent) majority outcome:

$$\text{DBP}'(\mathbf{J}) = \operatorname{argmin}_{J \in \mathcal{J}(\Phi)} H(F_{\text{maj}}(\mathbf{J}), J)$$

For those who know about voting and preference aggregation: this method is based on the same principle as the *Slater rule*.

Winner Determination

A *disadvantage* of the DBP is its high complexity. Consider the *winner determination* problem, asking whether a given partial judgement set can be extended to a winning judgment set for a given profile.

Theorem 6 *Winner determination for the DBP is Θ_2^p -complete.*

Proof: Omitted. [Θ_2^p is also known as “parallel access to NP”]

Compare this to the other aggregation procedures we have discussed:

Fact 7 *Winner determination for any quota rule F_q is in P .*

Proposition 8 *Winner determination for the PBP is in P .*

Proof: counting (for premises) + model checking (for conclusions) ✓

U. Endriss, U. Grandi, and D. Porello. Complexity of Judgment Aggregation. *Journal of Artificial Intelligence Research*, 45:481–514, 2012.

Representative-Voter Rules

The complexity of the DBP stems from the fact that we have to search through *all consistent judgment sets* to find the one that's closest to the profile. If we restrict this set, we can do better.

Idea: Only search through the *support*, i.e., judgment sets proposed by individuals. That is, identify “*the most representative voter*”.

One possible implementation of this idea is the *average-voter rule*:

$$\text{AVR}(\mathbf{J}) = \operatorname{argmin}_{J \in \text{SUPP}(\mathbf{J})} \sum_{i=1}^n H(J, J_i)$$

where $\text{SUPP}(\mathbf{J}) = \{J_1, J_2, \dots, J_n\}$

Fact 9 *Winner determination for the AVR is in P.*

U. Grandi. Binary Aggregation with Integrity Constraints. PhD thesis, ILLC, University of Amsterdam, 2012.

Support-Based Aggregation

Yet another idea for an aggregation procedure:

- Fix an order \gg on the agenda s.t. $\varphi \gg \psi$ whenever $|N_\varphi^J| > |N_\psi^J|$.
- Traverse the agenda according to \gg , and accept formulas one by one, except when that would render the set of accepted formulas inconsistent (in which case you skip to the next formula).

For those who know about voting and preference aggregation: this is based on the same principle as the *Tideman's Ranked Pairs method*.

J. Lang, G. Pigozzi, M. Slavkovik, and L. van der Torre. Judgment Aggregation Rules Based on Minimization. Proc. TARK-2011.

D. Porello and U. Endriss. Ontology Merging as Social Choice: Judgment Aggregation under the Open World Assumption. *J. Logic and Computation*. In press.

Agenda Characterisations

- The *impossibility result* we have seen showed that consistent aggregation is impossible under certain assumptions—but only for agendas including $\{p, q, p \wedge q\}$.

Instead we might ask: *for which agendas* does this impossibility arise? That is, we are after *agenda characterisations*.

- Recall that we have seen several *characterisation results* already (for quota rules). They only use *choice-theoretic axioms* (independence, etc.) and *syntactic conditions* on the outcome (completeness and complement-freeness). No logic so far (that is, no use of consistency for characterisation results so far).

Safety of the Agenda under Majority Voting

Previously we saw that the majority rule can produce an inconsistent outcome for *some* (not all) profiles based on agendas $\Phi \supseteq \{p, q, p \wedge q\}$. How can we *characterise* the class of agendas with this problem?

An *agenda* Φ is said to be *safe* for an aggregation procedure F if the *outcome* $F(\mathbf{J})$ is *consistent* for *any* admissible profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$.

Theorem 10 (Nehring and Puppe, 2007) *Let $n \geq 3$. An agenda Φ is safe for the (strict) majority rule iff Φ has the median property.*

A set of formulas Φ satisfies the *median property* if every inconsistent subset of Φ does itself have an inconsistent subset of size ≤ 2 .

K. Nehring and C. Puppe. The Structure of Strategy-proof Social Choice. Part I: General Characterization and Possibility Results on Median Space. *Journal of Economic Theory*, 135(1):269–305, 2007.

Proof

Claim: Φ is safe [$F_{\text{maj}}(\mathbf{J})$ is consistent] $\Leftrightarrow \Phi$ has the median property

(\Leftarrow) Let Φ be an agenda with the median property. Now assume that there exists an admissible profile \mathbf{J} such that $F_{\text{maj}}(\mathbf{J})$ is *not* consistent.

\rightsquigarrow There exists an inconsistent set $\{\varphi, \psi\} \subseteq F_{\text{maj}}(\mathbf{J})$.

\rightsquigarrow Each of φ and ψ must have been accepted by a strict majority.

\rightsquigarrow One individual must have accepted both φ and ψ .

\rightsquigarrow Contradiction (individual judgment sets must be consistent). \checkmark

(\Rightarrow) Let Φ be an agenda that violates the median property, i.e., there exists a minimally inconsistent set $\Delta = \{\varphi_1, \dots, \varphi_k\} \subseteq \Phi$ with $k > 2$.

Consider the profile \mathbf{J} , in which individual i accepts all formulas in Δ except for $\varphi_{1+(i \bmod 3)}$. Note that \mathcal{J} is consistent. But the majority rule will accept all formulas in Δ , i.e., $F_{\text{maj}}(\mathbf{J})$ is inconsistent. \checkmark

Agenda Characterisation for Classes of Rules

Now instead of a single aggregator, suppose we are interested in a *class of aggregators*, possibly determined by a set of *axioms*.

We might ask:

- *Possibility*: Does there exist an aggregator meeting certain axioms that will be consistent for any agenda with a given property?
- *Safety*: Will every aggregator meeting certain axioms be consistent for any agenda with a given property?

Possibility Theorem for Monotonic Rules

Theorem 11 (Nehring and Puppe, 2010) *There exists a **unanimous, anonymous, independent and monotonic** aggregator for the agenda Φ that is **complete and consistent** iff Φ is **not blocked**.*

Here an agenda Φ is called **blocked** if there exists a $\varphi \in \Phi$ with a **conditional path** from φ to $\sim\varphi$ and *vice versa*, where a conditional path is a sequence $\varphi_1, \varphi_2, \dots, \varphi_k$ such that $\varphi_i \neq \sim\varphi_{i+1}$ and $\{\varphi_i, \sim\varphi_{i+1}\}$ is part of some mi-subset of Φ for every $i < k$.

Proof: Omitted.

List and Puppe (2009) give an overview of known possibility theorems.

K. Nehring and C. Puppe. Abstract Arrovian Aggregation. *Journal of Economic Theory*, 145(2):467–494, 2010.

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

Safety Theorem for Systematic Rules

Suppose we know that the group will use *some* aggregation procedure meeting certain requirements, but we do not know which procedure exactly. Can we guarantee that the outcome will be consistent?

A typical result (for the majority rule axioms, minus monotonicity):

Theorem 12 (Endriss et al., 2010) *An agenda Φ is **safe** for any **anonymous, neutral, independent, complete and complement-free** aggregation procedure iff Φ has the **simplified median property**.*

An agenda Φ has the **simplified median property** if every inconsistent subset of Φ has itself an inconsistent subset $\{\varphi, \psi\}$ with $\models \varphi \leftrightarrow \neg\psi$.

Note: This is more restrictive than the median property: $\{\neg p, p \wedge q\}$.

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

Proof

Claim: Φ is safe for any ANI/complete/comp-free rule $F \Leftrightarrow \Phi$ has SMP

(\Leftarrow) Suppose Φ has the SMP. For the sake of contradiction, assume $F(\mathbf{J})$ is inconsistent. Then $\{\varphi, \psi\} \subseteq F(\mathbf{J})$ with $\models \varphi \leftrightarrow \neg\psi$. Now:

$\rightsquigarrow \varphi \in J_i \Leftrightarrow \sim\psi \in J_i$ for each individual i (from $\models \varphi \leftrightarrow \neg\psi$ together with consistency and completeness of individual judgment sets)

$\rightsquigarrow \varphi \in F(\mathbf{J}) \Leftrightarrow \sim\psi \in F(\mathbf{J})$ (from neutrality)

\rightsquigarrow both ψ and $\sim\psi$ in $F(\mathbf{J}) \rightsquigarrow$ contradiction (with complement-freeness) \checkmark

(\Rightarrow) Suppose Φ violates the SMP. Take any minimally inconsistent $\Delta \subseteq \Phi$. If $|\Delta| > 2$, then also the MP is violated and we already know that the majority rule is not consistent. \checkmark So we can assume $\Delta = \{\varphi, \psi\}$.

W.l.o.g., must have $\varphi \models \neg\psi$ but $\neg\psi \not\models \varphi$ (otherwise SMP holds).

But now we can find a rule F that is not safe: accept a formula if at most one individual does and take a profile with $J_1 = \{\sim\varphi, \sim\psi, \dots\}$, $J_2 = \{\sim\varphi, \psi, \dots\}$, and $J_3 = \{\varphi, \sim\psi, \dots\}$. Then $F(\mathbf{J}) = \{\varphi, \psi, \dots\}$. \checkmark

Comparing Possibility and Safety Results

Possibility theorems and safety theorems are closely related:

- Possibility: *some* aggregator in the class determined by the given axioms will produce consistent outcomes *iff* the agenda has a given property
- Safety: *all* aggregators in the class determined by the given axioms will produce consistent outcomes *iff* the agenda has a given property

In what situations do we need these results?

- Possibility: a mechanism designer wants to know whether she can design an aggregation rule meeting a given list of requirements
- Safety: a system might know certain properties of the aggregator users will employ (but not all properties) and we want to be sure there won't be any problem (we might want to check this again and again)

For safety problems in particular we might want to develop *algorithms*, i.e., *complexity* plays a role.

Complexity of Safety of the Agenda

Deciding whether a given agenda is safe for the majority rule (as well as several classes of rules we get by relaxing the axioms defining the majority rule) is located at the second level of the polynomial hierarchy.

Proving those results involves the following lemma (and variations):

Lemma 13 (Endriss et al., 2010) *Deciding whether a given agenda has the **median property** is Π_2^p -complete.*

Proof: Omitted.

Recall that $\Pi_2^p = \text{coNP}^{\text{NP}}$ is the class of problems for which we can verify a certificate for a negative answer in polynomial time if we have access to an NP oracle. A typical problem in the class is deciding truth of formulas of the form $\forall x \exists y \varphi(x, y)$. So: very hard.

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

Example: Strategic Manipulation

Suppose we use the premise-based procedure:

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

If the this agent only cares about the conclusion $p \vee q$, she could *manipulate* the aggregation by claiming to believe that p is true.

Strategic Manipulation

Note that in pure JA, we cannot talk about strategic behaviour, as there is no notion of preference. We need to add one! *How?*

This is still underexplored territory. Main definition in use so far:

- Your true judgment set is your most preferred outcome.
- The closer an outcome to your true judgment set, in terms of the *Hamming distance*, the more you prefer that outcome.

Thus: Agent i with true judgment set J_i *prefers* J to J' ($J \succ_i J'$) *iff*

$$H(J_i, J) < H(J_i, J').$$

Now i has an *incentive to manipulate* if $F(\mathbf{J}_{-i}, J) \succ_i F(\mathbf{J})$ for some $J \neq J_i$.

We say that F is *immune to manipulation* if it never provides an incentive to manipulate to any of the agents.

Recall: $H(J, J') = \frac{1}{2} \cdot |J \setminus J' \cup J' \setminus J|$

Immunity to Manipulation

Some aggregators are immune to manipulation:

Proposition 14 (Dietrich and List, 2007) *An aggregator is immune to manipulation iff it is both independent and monotonic.*

Proof: We only sketch the right-to-left direction. *Independence* means that we can consider each formula in isolation. But then *monotonicity* just means that it is never against an agent's interest to accept a formula she holds true. ✓

But note that the class of independent and monotonic aggregators is *not that attractive* (dictatorships, inconsistent majority rule, ...).

F. Dietrich and C. List. Strategy-proof Judgment Aggregation. *Economics and Philosophy*, 23(3):269–300, 2007.

Hardness of Manipulation

When there is no immunity from manipulation, we at least would like manipulation to be computationally hard.

MANIPULABLE(F)

Instance: Agenda Φ , profile $\mathbf{J} \in \mathcal{J}(\Phi)^n$, agent $i \in \mathcal{N}$.

Question: Is there a $J \in \mathcal{J}(\Phi)$ such that $F(\mathbf{J}_{-i}, J) \succ_i F(\mathbf{J})$?

Good news:

Proposition 15 MANIPULABLE(PBP) is NP-complete.

Proof omitted (works by reduction from SAT). Only case we know where winner determination is easy and manipulation is hard.

U. Endriss, U. Grandi, and D. Porello. Complexity of Judgment Aggregation. *Journal of Artificial Intelligence Research*, 45:481–514, 2012.

Bribery and Control

The computational complexity of other problems related to undesirable behaviour (besides strategic manipulation) has also been studied:

- *Bribery*: Can an outsider affect the outcome in their interest by bribing some individuals to change their judgment set?
- *Control*: Can an outsider affect the outcome in their interest by strategically adding/deleting/replacing individuals?

D. Baumeister, G. Erdélyi, O.J. Erdélyi, and J. Rothe. Bribery and Control in Judgment Aggregation. Proc. COMSOC-2012.

Applications

This tutorial has focussed on theory, both because that is where one should start when learning about a new topic and because there has not yet been much work on applications.

Still, some recent work has at least suggested possible directions for using judgment aggregation techniques in applications. Examples:

- Collective decision making in multiagent systems
- Ontology merging on the Semantic Web
- Aggregating crowdsourced data (e.g., for computational linguistics)

M. Slavkovik. Judgment Aggregation for Multiagent Systems. PhD thesis, University of Luxembourg, 2012.

D. Porello and U. Endriss. Ontology Merging as Social Choice: Judgment Aggregation under the Open World Assumption. *J. Logic and Computation*. In press.

U. Endriss and R. Fernández. Collective Annotation of Linguistic Resources: Basic Principles and a Formal Model. Proc. ACL-2013.

Further Reading

Possible starting points:

- Easy-going tutorial paper on basic JA:
C. List. The Theory of Judgment Aggregation: An Introductory Review. *Synthese*, 187(1):179–207, 2012.
- More advanced tutorial paper focusing on impossibilities:
D. Grossi and G. Pigozzi. Introduction to Judgment Aggregation. In: *Lectures on Logic and Computation*, Springer, 2012.
- Paper on complexity of JA (same notation as in this tutorial):
U. Endriss, U. Grandi, and D. Porello. Complexity of Judgment Aggregation. *Journal of Artificial Intell. Research*, 45:481–514, 2012.

Or visit the website for my Amsterdam course on computational social choice for more material:

<http://www.illc.uva.nl/~ulle/teaching/comsoc/>

Last Slide

We have covered the basics of the theory of judgment aggregation:

- Questions first raised by real-world problems in legal scholarship with philosophical ramifications (*doctrinal paradox*); later formalised in the style of *social choice theory*
- *Formal basics*: axioms, List-Pettit impossibility, characterisations
- *Procedures*: based on quota, premises, distances, repr. voters, support
- *Agenda characterisations*: possibility and safety; complexity
- *Strategic considerations*: manipulation, control, bribery

The full potential of judgment aggregation for multiagent systems and other applications in computer science is yet to be explored.

These slides, together with supporting materials, will remain available here:

<http://www.illc.uva.nl/~ulle/teaching/aamas-2013/>