Judgment Aggregation: Impossibilities, Solutions, Legal Examples

Philippe Mongin (CNRS & HEC Paris)

ESNIE Lecture, Cargèse
May 2012

Partly based on: "The Doctrinal Paradox, the Discursive Dilemma, and Logical Aggregation Theory", *Theory and Decision*, and "Judgment Aggregation", *Handbook of Formal Philosophy*, S.O. Hansson and V.F. Hendricks eds, Springer, both forthcoming.

PLAN OF THE TALK

1. The problem of judgment aggregation
2. The court example, viewed either as a doctrinal paradox or as a discursive dilemma
3. Basic formalism of judgment aggregation theory
4. An impossibility theorem
5. A possibility Solution: Experimenting with the agenda and the logic
6. A word on judiciary examples

The problem of judgment aggregation

Individual judgments that are well-behaved, in the sense of complying with relevant logical norms, can result in collective judgments that are not so well-behaved. In particular, logically consistent judgments at the individual level can result into collective inconsistencies.

This problem has been tackled by a new theory, *Judgment Aggregation Theory*, perhaps more accurately described as *Logical Aggregation Theory*, since its main concept, judgment, is formalized by logic.
See List & Pettit 2002, Nehring & Puppe 2002, 2010, Dietrich 2006, 2007, 2010, Dietrich & List 2007a,b, Mongin 2008, Dokow & Holzman 2009, 2010, Dietrich & Mongin 2010, etc.

Logical aggregation theory

Philosophically, to **judge that P** (=Corsica is an island, Corsica is beautiful, Corsica is smaller than Sardinia, Corsica is more beautiful than Sardinia) is to **assert the proposition that P**, and a proposition is a semantic entity that can receive truth values. See Frege and Russell.

This encompasses a preference judgment as a particular case, hence part of social choice theory (SCT) is also a particular case. Arrow's impossibility theorem and Sen's Paretian liberal theorem have been recovered as corollaries to more general impossibility theorems.

Most existing results take this **abstract and negative form of impossibility theorems**. Today's talk is a step towards **positive solutions** and (to a modest extent) **concrete applications**. We will discuss idealized judiciary examples, thus connecting with the law and economics literature.

The court example

LAT has a paradigmatic example which plays the motivating role of the Condorcet paradox for SCT. It is a semi-realistic judicial case due to legal theorists Kornhauser & Sager (1993). A tired example, but useful to introduce LAT for our purposes.

Three judges must decide a breach-of-contract case.
The unanimously agreed legal doctrine says that *damages are due to the plaintiff iff the contract with the defendant was valid and the defendant broke it*. So judges should resolve two **issues** (contract valid? contract broken?) in order to resolve the **case** (damages due?).

The court example (ctd)

Here are the judgments, both individual and collective. Simple majority voting is applied.

	v contract valid?	b contract broken?	d damages due?	$d \leftrightarrow v \wedge b$ legal doctrine
Judge 1	Y	Y	Y	Y
Judge 2	N	Y	N	Y
Judge 3	Y	N	N	Y

In the **issue-based method**, votes are tallied on each issue separately and the doctrine is applied to solve the case. In the **case-based method**, votes are tallied on the case directly. Each way leads to a different answer (**Y** for the first method, **N** for the second one).

The court example is a doctrinal paradox

K&S call the conflict of the two methods **the doctrinal paradox**.

There is a **paradox** because either method has something to say for itself, but neither seems sufficient. So one would like to have them both at a time.

- With the issue-based method, the court's decision respects the judges' reasons, but not their solutions, so it seems not to be sufficiently responsive to them.
- With the case-based method, the court's decision respects the judges' solutions, but not their reasons, and it may not have any rationale at all.

It is a **doctrinal** paradox because the judges' reasons are grounded in a commonly agreed legal doctrine, which is devised to apply to other cases as well.

The same structure can be found in many non-legal examples.

A non-legal example of the doctrinal paradox

Three experts must decide whether or not a celestial body X is a planet.

	m massive celestial body?	c clear neighbourhood?	p a planet?	$p \leftrightarrow m \wedge c$ astronomical doctrine
Expert 1	Y	Y	Y	Y
Expert 2	N	Y	N	Y
Expert 3	Y	N	N	Y

A semi-realistic example. In 2005-6, the AIA defined what a planet is by three criteria and decided that Pluto, like Eris, was a dwarf planet, not a planet.

We may define the DP in terms of *premisses* and *conclusions,* more generally than the issues and the case, and contrast the *premiss-based* with the *conclusion-based* method.

The court example is a discursive dilemma

Pettit (2001) and List & Pettit (2002) conceived of the court problem differently from K&S.

The court should accept or reject by whatever method each of the propositions/formulas represented by v, b, d, $d \leftrightarrow v \wedge b$, and the problem is simply that it ends up with an inconsistent set:

$$\{v, b, \neg d, d \leftrightarrow v \wedge b\}.$$

L&P call **discursive dilemma** this reformulation of the DP, and they argue that it is theoretically more apt.

- Since it does not differentiate the propositions, it is more general.
- Hence also more widely applicable. Many groups can be taken in a discursive dilemma: deliberative assemblies, experts committees, board of companies, etc, and even the Political Society as a whole (in Pettit's theory of deliberative democracy).
- While the DP is defined by a conflict of practical *methods*, the DD reflects a tension between theoretical *principles* (collective rationality vs individual responsiveness).

Two forms of the judgment aggregation problem

None of the previous arguments is really convincing.

- The premiss vs conclusion distinction can be an essential feature of the judgment aggregation problem, and the doctrine may not be treated like any other proposition.
- Many groups can also be taken in a DP, and indeed the very same as those mentioned by L&P for their DD.
- The DP also reflects a tension between theoretical principles (responsiveness to individuals vs collective rationality).

In sum, the judgment aggregation problem can take two forms, and neither seems theoretically superior to the other. The only requisite is to be clear about which form is selected for what set of applications. The judiciary and astronomical ones call for the DP form because of the specific roles played by the premisses, the conclusion and the doctrine, resp.

Two branches of LAT

Initially, LAT theorists disregarded the DP form of the judgment aggregation problem, and relying instead on its DD reformulation, turned it into a very general impossibility theorem, sometimes called the *canonical theorem*.

More recently, some revisited the DP in its own terms. One line (after Bovens and Rabinowicz, 2006) is to analyze the premiss-based and conclusion-based methods in terms of their *truth-tracking ability*. Not an axiomatic framework, rather a model with specific assumptions (in the style of Condorcet's jury theorem).

Another line (in Dietrich & Mongin, 2010, and Mongin, forthcoming) is to

restate the DP in the same axiomatic style as the DD and provide it with a canonical theorem of its own (our presentation below).

The formal framework

In either form, LAT fixes an **agenda of propositions** and analyzes individual and collective judgments as the acceptance or rejection of these propositions. Propositions represented by logical formulas, and acceptance/rejection either in terms of 0-1 values or of set membership (as done here).
The court agenda is

$$\overline{X} = \{v, b, d, d \leftrightarrow v \wedge b\} \cup \{\neg v, \neg b, \neg d, \neg(d \leftrightarrow v \wedge b)\}$$

In the DP form, unlike the DD form, X is partitioned into a set of **premisses** P and a set of **conclusions** C. For generality, $P \subseteq X$.
A **judgment set** B is a subset of X that collects the formulas accepted by an individual or the group.
Here we only consider the set D of j.s. that are **consistent** (according to the underlying logic) and **complete** (for each all $p \in X$, either p or $\neg p$). Eg, the judges' sets:

$$\{v, b, d, d \leftrightarrow v \wedge b\}, \{\neg v, b, \neg d, d \leftrightarrow v \wedge b\}, \{v, \neg b, \neg d, d \leftrightarrow v \wedge b\}$$

The formal framework (cntd)

For any **profile** of individual judgment sets $A_1, ..., A_n$ in D^n, there is an associated collective judgment set A. This is the **collective judgment function** F, which is reminiscent of Arrow's social welfare function. Typical examples:

Proposition-wise majority voting.

Voting with a quota (different from 1/2).

Dictatorship, ie, there is $i \in \{1, ..., n\}$ s.t. $F(A_1, ..., A_n) = A_i$ for all $(A_1, ..., A_n)$.

Oligarchy, ie**,** there is $M \subseteq \{1, ..., n\}$ s.t. $F(A_1, ..., A_n) = A_i$ for all $(A_1, ..., A_n)$.
In the DP branch, F can be defined on either P alone, or C alone, or the whole of X.

The formal framework (cntd)

A set of axioms stating prima facie desirable or relevant properties of F.

Non Dictatorship, non-oligarchy.

Unanimity Preservation. For all p and all $(A_1, ..., A_n)$, if $p \in A_i$ for all i, then $p \in F(A_1, ..., A_n)$.

Independence. For all p and all $(A_1, ..., A_n)$, $(A'_1, ..., A'_n)$, if $p \in A_i \Leftrightarrow p \in A'_i$ for all i, then

$$p \in F(A_1, ..., A_n) \Leftrightarrow p \in F(A'_1, ..., A'_n),$$

Monotonicity. For all p and all $(A_1, ..., A_n)$, $(A'_1, ..., A'_n)$, if $p \in A_i \Rightarrow p \in A'_i$ for all i, and $p \notin A_i$ and $p \in A'_i$ for at least one i, then

$$p \in F(A_1, ..., A_n) \Rightarrow p \in F(A'_1, ..., A'_n).$$

In the DP branch, axioms can be defined on either P alone, or C alone, or the whole of X.

Back to the logic

LAT also states conditions on the agenda X.
To introduce them, we need to clarify the underlying logical concepts.
The first is $\mathcal{L}$, the set of all logical formulas (typically, a much larger set than X). Then, an **inference relation** $S \vdash p$ defines the permitted inferences from $S \subseteq \mathcal{L}$ to $p \in \mathcal{L}$, and each $S \subseteq \mathcal{L}$ is classified as either **consistent** or **inconsistent** (the latter iff $S \vdash p$ and $S \vdash \neg p$ for some p).

What $\vdash$ (or equivalently the inconsistent sets) must satisfy can be stated axiomatically (Dietrich 2007; Dietrich & Mongin 2010), but **you can assume that the standard propositional calculus applies**.

An agenda condition

First say that p **conditionally entails** q, denoted by $p \vdash^* q$, if

$$\{p\} \cup Y \vdash q$$

for some set of auxiliary assumptions $Y \subseteq X$ satisfying a proviso.

(I.e. $Y \cup \{p\}$ and $Y \cup \{\neg q\}$ are consistent, to ensure non-triviality and minimality.)

Second define the agenda condition
Path-connectedness. For all p, q, there are $p_1, ..., p_m$ s.t. $p = p_1 \vdash^* p_2 \vdash^* ... \vdash^* p_m = q$.
Again the precise form of the condition depends on the branch of LAT.
In the DP branch, $p, q, p_1, ..., p_m \in P$ (but formulas in $C = X/P$ may enter Y in the conditional entailments $\vdash^*$).

Back to the court example

In the court example, $\overline{X}$ satisfies (b_P) for the natural choice of premisses

$$P = \{b, v, b \wedge v \longleftrightarrow d\}^{\pm}.$$

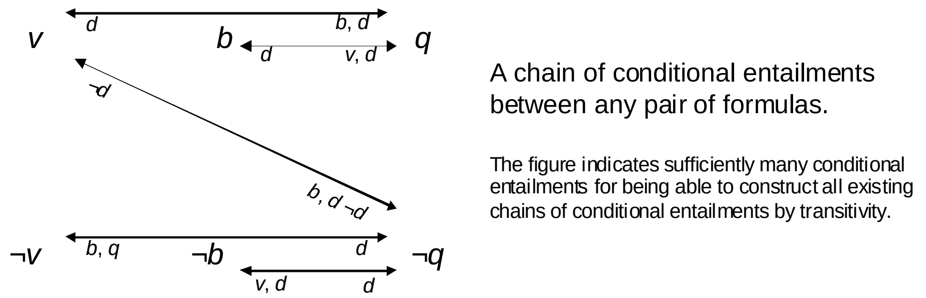

Figure 1: The agenda $\overline{X}$ with $P = \{b, v, q\}^{\pm}$ violates (b_P).

Each premiss is connected with any other (note: through the conclusion d).

An impossibility theorem

Theorem (Dietrich and Mongin, 2010). *If the agenda X and the set of premisses P satisfy* (b_P), *no collective judgment function F can be $D^n \to D$ and satisfy Independence on P, Monotonicity on P, Non-dictatorship on P, and Unanimity preservation on the whole of X. Otherwise, for $n \geq 3$, there exist $F : D^n \to D$ with these properties.*

Corollary: *there is a logical inconsistency between simple majority voting on premisses and simple majority voting on conclusions.*
Proof of corollary: otherwise, there would be $F : D^n \to D$ satisfying both the properties of s.m.v. on premisses - hence, Independence on P, Monotonicity on P, Non-dictatorship on P, Unanimity preservation on P - and the properties of s.m.v. on non-premisses - hence, Unanimity preservation on C.
This is the canonical theorem for the DP form of the judgment aggregation problem.

Sketch of positive solutions

- A dramatic solution: to **give up the logical framework of judgment**, moving e.g. to a probabilistic framework, as in the previous work on probability aggregation (McConway 1981, Genest & Zidekh 1986, Mongin, 1995, and many others). Useful in clarifying one source of the difficulty (binary judgments), but not a fully convincing way out.
- A less dramatic solution: to **weaken the axiomatic conditions** on F. In a related framework, computer scientists weaken Independence (Konieczny & Pino-Pérez 1998, Pigozzi 2006, Miller & Osherson 2008). This is the main target condition, and without it, there is little structure left.
- An even less dramatic solution: to remain within the framework, including axioms, and **experiment with the logic**. This is the option here (after Dietrich, 2010, and Dietrich & Mongin, 2010). We show that possibilities result from importing into LAT a modicum of **non-classical** propositional logic.

Solutions based on the agenda and the logic

The strategy is to **replace the Boolean** (or material, or classical) **implication** $\rightarrow$ **by a non-Boolean implication** $\hookrightarrow$ **from conditional logic** (a form of non-classical propositional logic).
This will often turn an impossibility agenda into one with possibilities **by violating the (necessary!) condition ($\mathbf{b}_P$)**.
A strategy proposed by Dietrich (2010) and applied to the court agenda by Dietrich and Mongin (2010).
Its relevance is easy to defend, because Boolean (material) implication is an inadequate rendering of the implication of natural reasoning.

Against the material conditional

Some arguments against Boolean implication (see texts in logic and linguistics).

- $\to$ is a source of **classic paradoxes**, e.g., $(p \to q) \vee (q \to p)$ is a logical truth, and these are correct inferences:
  1. $(p \to q) \wedge (r \to s) \vdash (p \to s) \vee (r \to q)$
  2. $(p \wedge q) \to r \vdash (p \to r) \vee (q \to r)$.

(For the last) I can claim *if John comes and Ann comes, it will be a nice party* without implying that either of these holds: *if John comes, it will be a nice party*; *if Ann comes, it will be a nice party*.

Against material implication (contd)

- $\rightarrow$ **cannot handle at all counterfactual conditionals**. Compare:

If Obama lived in Paris, he would live in France.

If Obama lived in Paris, he would live in Spain.

One conditional is true and the other not, but both are true with $\rightarrow$.

- $\rightarrow$ **does not very well either with indicative (non-counterfactual) conditionals.** Compare:

If the sun does not come up tomorrow, it will not matter.

If the sun does not come up tomorrow, it will be the end of the world.

- **Negated conditionals are especially resistant**. Compare:

It is not the case that if the stock market crumbles again, bankers will get smaller bonuses.

The stock market will crumble again and bankers will *not* get smaller bonuses.

$\rightarrow$ makes them equivalent.

Moving to non-material implication

Classic explanation to the trouble: $\rightarrow$ is **truth-functional**, i.e., the truth value of $p \rightarrow q$ is a function of the truth values of p and q.

This feature disappears with the non-Boolean $\hookrightarrow$. Many theories of counterfactuals and/or indicative conditionals can do the job here (e.g., Stalnaker 1968, Lewis 1973, and recent ones as in Edgington 2006).

Replacing $\rightarrow$ with $\hookrightarrow$ leads to **a smaller number of correct patterns of inference** (equivalently, of patterns of inconsistency). E.g., we lose the equivalences between a negated conditional and a conjunction (last point).
With less conditional entailments, the (necessary) agenda condition (b$_P$) is more difficult to meet.

Non-material implication (contd)

Take the agenda $\overline{X}$ and list all minimally inconsistent subsets $Y \subseteq \overline{X}$ of size greater than 2.

$$Y_1 = \{\neg v, d, q\}, Y_2 = \{\neg b, d, q\},$$
$$Y_3 = \{v, b, \neg d, q\}, Y_4 = \{v, b, d, \neg q\},$$
$$Y_5 = \{\neg v, \neg d, \neg q\}, Y_6 = \{\neg b, \neg d, \neg q\}.$$

$Y_3 = \{v, b, \neg d, q\}$ corresponds to the court example, and the other sets to related variants.

With $q' = d \hookleftarrow\!\hookrightarrow v \wedge b$ **instead of** q**, the list reduces to** Y_1, Y_2, Y_3.

Not every conditional logic theorist would exclude Y_4.

Non-material implication (contd)

The lost sets Y_4, Y_5, Y_6 relate to a theorem on Boolean equivalence that does not hold for non-Boolean ones:

$$\neg(d \longleftrightarrow v \wedge b) \vdash \neg d \longleftrightarrow v \wedge b.$$

Against this theorem, judges should be able to *reject* v (or b) and d without having to *accept* q.
They may take the legal doctrine to be $q' = d \hookleftarrow\hookrightarrow v \wedge b \wedge r$ (r for "registered") and reject $q = d \hookleftarrow\hookrightarrow v \wedge b$ on this theoretical basis.
And if they *reject* r, they have three *consistent* j.s. at their disposal (compare with Y_5,Y_6):

$$\left\{\neg v, b, \neg r, \neg d, \neg q, q'\right\}, \left\{v, \neg b, \neg r, \neg d, \neg q, q'\right\},$$
$$\left\{\neg v, \neg b, \neg r, \neg d, \neg q, q'\right\}.$$

Non-material implication (contd)

Similarly, judges should be able to *accept* v, b, and d without having to *accept* q.
Same argument: they may rely on $q' = d \hookleftarrow\!\!\hookrightarrow v \wedge b \wedge r$, reject $q = d \hookleftarrow\!\!\hookrightarrow v \wedge b$, and also reject r.
They have a *consistent* j.s. at their disposal: $\{v, b, \neg r, \neg d, \neg q, q'\}$ (compare with Y_4).

THE GREATER FLEXIBILITY OF $\hookleftarrow\!\!\hookrightarrow$ WILL BE PUT TO USE BY RETURNING TO THE THEOREM AND EXPLOITING ITS NECESSITY PART.

The court agenda violates the agenda conditions

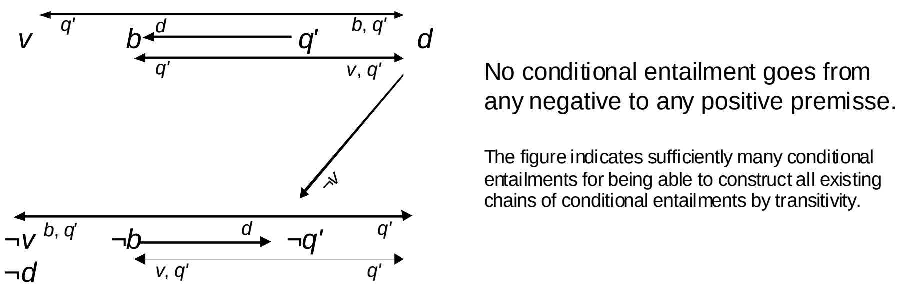

Figure 2: The agenda $\overline{X}'$ violates (b).

$\overline{X}' = \{v, b, d, q\}^{\pm}$ with $q = d \hookleftarrow\hookrightarrow v \wedge b$, $P = \{v, b, q\}^{\pm}$.

No path from negative to positive formulas of P for the reason said above.

$\Rightarrow$ **by Thm possibilities exist**.

A solution to the DP?

In sum, the proposed solution to DP is to exploit the possibility side of the impossibility theorem devised for DP, and more precisely to change the underlying logic.

Two possible reservations.

- Not a waterproof strategy. It works if there is a doctrine, i.e., a conditional proposition that is not correctly analyzed by Boolean conditionals. All right for legal and scientific applications, but some agendas have no doctrine.
- We predict the existence of possibilities, no that they are attractive. Simple majority is excluded anyway (falls prey to a weak sufficient condition for impossibility).

With the court agenda, *the modified unanimity rule* works (for all non-negated $p \in X$, respect unanimity of p or $\neg p$, and in case of disagreement, choose $\neg p$). Unattractive? But it generalizes to *quota rules* with more individuals.

Legal examples (crude sketch)

The court example of DP is a toy example, but K& S (1993) and followers (e.g., Nash, 2003) found actual occurrences **in the US appellate courts** (collegial courts unlike the lower ones).
Best examples with **the Supreme Court** and its 9 judges (because so well documented).

- *Pennsylvania v. Union Gas Co* (1989). Union Gas suited the state of Pennsylvania in a federal court. The *case*: Is Pennsylvania vulnerable to the suit in question? The *issues*: Can Congress make states vulnerable to private suits in federal courts despite 11th Amendment? And if Congress can do so, is a certain 1986 Act an exercise of this power? There was a DP and one of the 9 changed his vote to suppress it in the sense of the premiss-based method. The court decided that the suit could go forward.

Legal sketch (cntd)

PENNSYLVANIA V. UNION GAS CO (from K&S, 1993)

a = Power to abrogate, b= Power exercised in statute, c= Pennsylvania can be sued, $a \wedge b \rightarrow c$

Justices:

Blackmun	Y Y Y
Brennan	Y Y Y
Marshall	Y Y Y
Stevens	Y Y Y
White	Y N **Y(?)**
Kennnedy	N N N
O'Connor	N N N
Rehnquist	N N N
Scalia	N Y N
Court	**Y** (5-4) **Y** (5-4) **Y** (5-4) *instead of N(5-4)*

Legal sketch (cntd)

- *Arizona v. Fulminante* (1991). Fulminante had confessed and been convicted by an Arizonan court. The *case*: Is conviction to be reaffirmed? The *issues*: Was Fulminante's confession voluntary? If not, was it harmless to accept it as evidence in the trial? Again, there was a DP and one of the 9 changed his vote to suppress it in the same direction of the PBM. The court invalidated the conviction.

Rare, but empirically significant examples. Show that the USSC discovers the DP when stumbling on it and has no deliberate strategy to respond to it. Show that it leans towards PBM when understanding the choice.

Legal sketch (cntd)

Hence the two normative questions: (i) Should the USSC fix a preventive

strategy or adopt ad hoc strategies as here? (ii) If the former, should it be the PBM or CBM?

Standard legal answers: (i) **yes** for rationality reasons, (ii) **the PBM**, because it makes the legal doctrine effective at the court's level, not just the individual level.

Objection 1: more complicated cases in which the legal doctrine is not agreed on. A whole new area of problems! Note that the non-classical logic modelling becomes indispensable when there are several doctrines.

Objection 2: PBM and CBM only considered and discussed in terms of simple majority voting, but other schemes are possible (sequential voting) and other voting rules as well (eg, the quota rules mentioned above). Another area of discussion for legal theory.

What have we achieved?

Given the DP its own theoretical development and thus proposed a bridge between judgment aggregation theory and legal theory.

Generalized the DP in the abstract way of an impossibility theorem, which makes it both more threatening (not an accidental phenomenon) and more open to solutions (the other face of the coin with impossibility theorems).

Crudely sketched an application that should be pursued, using perhaps some of the formal technology presented here.