Towards a unified theory of aggregation

Franz Dietrich
LSE & Maastricht

&

Christian List
LSE

SCW Conference, Moscow 2010
Towards a unified theory of aggregation

• Although the title of this lecture may sound very ambitious, the key word is “towards”.
• We will certainly not present a unified theory of aggregation here.
• We wish we had one, but we don’t!
• Our aim, much more modestly, is to sketch some of the conceptual ingredients of such a theory, and to illustrate some steps towards its development – in the hope that we can thereby inspire further work.
• Given the plenary format of this lecture, we want to focus on conceptual issues, and will try to be as non-technical as possible.
Plan for this lecture

- Part I (Christian): A basic perspective
- Part II (Franz): A more general perspective

The lecture is based on several works:

- D&L, “Arrow’s theorem in judgment aggregation”, *SCW* 2007
- Work in progress

(Further downloadable papers are on our two webpages.)
Plan for this lecture

• Part I (Christian): A basic perspective
• Part II (Franz): A more general perspective
Plan for this lecture

• Part I (Christian): A basic perspective
  – Background: the theory of aggregation and the challenge
  – A taxonomy of different objects that might be aggregated
  – The theory of judgment aggregation: a general theory of binary aggregation problems?

• Part II (Franz): A more general perspective
Plan for this lecture

• Part I (Christian): A basic perspective
  – Background: the theory of aggregation and the challenge
  – A taxonomy of different objects that might be aggregated
  – The theory of judgment aggregation: a general theory of binary aggregation problems?

• Part II (Franz): A more general perspective
Arrovian preference aggregation

• Ever since Arrow's book, *Social Choice and Individual Values* (1951/1963), the theory of aggregation has been thriving.
• Arrow focused on the aggregation of preferences, understood as the aggregation of multiple individual orderings over some mutually exclusive alternatives into a social ordering or choice.
• Since preference aggregation problems arise in many political and economic contexts (voting and welfare), Arrow's work struck a chord with scholars across the social sciences.
• But the interest in preference aggregation goes back at least to Condorcet in the 18th century, and in fact even to medieval times, e.g., to Ramon Llull (~1232-1315) and Nicolaus Cusanus (1401-1464).
Arrovian preference aggregation

- The enormous relevance of preference aggregation as well as the power and elegance of Arrow's axiomatic approach may explain why much (but not all) social-choice-theoretic research has shared Arrow's focus on preference aggregation.
- But preference orderings (or, in the limit, single votes or top-preferences) are not the only objects whose aggregation may be of interest, and a number of other potentially important aggregation problems have arguably received less attention than they deserve.
Continuous aggregation problems

- Some ‘continuous’ aggregation problems are notable exceptions. There are sizeable literatures (in decreasing order of size) on:
  - the aggregation of utility or welfare functions (e.g., Sen 1970, 1982), which are informationally richer than preference orderings in that they encode cardinal and/or interpersonally comparable information;
  - the aggregation of probabilities (e.g., McConway 1981, Genest and Zidek 1986, Dietrich and List forthcoming);
  - the combined aggregation of utilities and probabilities (e.g., Hylland and Zeckhauser 1979, Mongin 1995).
Discrete aggregation problems

• But unlike these ‘continuous’ aggregation problems, there are several ‘discrete’ aggregation problems that have received much less attention, for example:
  – the aggregation of binary relations other than orderings (e.g., Fishburn and Rubinstein 1986), such as partial orderings, equivalence relations, trees or networks;
  – the aggregation or merging of belief systems with a logical structure, such as sets of judgments in a court, expert panel, scientific community, or committee;
  – the aggregation of different set-membership criteria, capturing competing views on which objects belong to a particular set (e.g., Kasher and Rubinstein 1997), such as a cultural, professional or political group, a species in biology, or a category in a data base.
Why we need to go beyond the Arrovian framework

• The standard Arrovian framework – even when we focus on special domains such as “economic” domains – does not generally allow us to represent all these different aggregation problems, since the objects that are being aggregated often have a structure very different from (preference) orderings – in addition to a different interpretation.
The challenge

• Can we come up with a general theory of aggregation that subsumes these and other aggregation problems – and that still allows us to say something interesting and systematic about them?
Plan for this lecture

• Part I (Christian): A basic perspective
  – Background: the theory of aggregation and the challenge
  – A taxonomy of different objects that might be aggregated
  – The theory of judgment aggregation: a general theory of binary aggregation problems?

• Part II (Franz): A more general perspective
Preferences revisited

• Let’s begin by asking what preferences are.
• Preferences are special kinds of what philosophers call intentional attitudes, i.e., attitudes that are directed towards certain objects.
• Preference orderings, in particular, encode the relative desires that an agent has towards the alternatives that are being ranked by them.
  – The attitudes in this case are an agent’s desires (here taking a relational/comparative form);
  – and the objects towards which the attitudes are directed are the alternatives that are being ranked.
Intentional attitudes more generally

• But such relational or comparative desires towards certain options are not the only intentional attitudes of interest.

• On the standard picture of rational agency, which goes back at least to David Hume in the 18th century, there are two main types of intentional attitudes an agent may have:
  – **belief-attitudes** (those encode the way the agent represents the world as being); and
  – **desire-attitudes** (those encode the way the agent would like the world to be).

• While preferences are standard examples of desire-attitudes, judgments (e.g., that a defendant is guilty) are standard examples of belief-attitudes.
Intentional attitudes more generally

• Generally, we can think of an intentional attitude as a pair of things: (i) the attitude itself, and (ii) the object towards which it is held.

• The **attitude** can take a number of forms:
  – Interpretationally, it can be belief-like or desire-like.
  – Structurally, it can be ordinal or cardinal, binary or non-binary, and so on.

• The **object** can also take a number of forms:
  – It can be a possible world or state of the world, a good or bundle of goods, an election option or candidate, but more generally, it can be a proposition.
    (Unlike different worlds or states, propositions are not mutually exclusive but can be arbitrarily interconnected.)
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
Different objects of aggregation

• Each of these different kinds of intentional attitudes can be the objects of aggregation.
## A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th></th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
<td>Discrete</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
<td>Credence ratings</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
<td>Evaluation ratings</td>
</tr>
</tbody>
</table>
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
## A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preference orderings</td>
</tr>
</tbody>
</table>
## A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal</td>
</tr>
<tr>
<td></td>
<td>Ordinal</td>
<td>credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical</td>
<td>Preference</td>
</tr>
<tr>
<td></td>
<td>desires</td>
<td>orderings</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal</td>
</tr>
<tr>
<td></td>
<td>Ordinal</td>
<td>credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical</td>
<td>Preference</td>
</tr>
<tr>
<td></td>
<td>desires</td>
<td>orderings</td>
</tr>
</tbody>
</table>

- **Belief-attitudes**
  - Judgments
  - Ordinal credences
  - Credence ratings
  - Subjective probabilities

- **Desire-attitudes**
  - Categorical desires
  - Preference orderings
  - Evaluation ratings
  - Utilities
Plan for this lecture

• Part I (Christian): A basic perspective
  – Background: the theory of aggregation and the challenge
  – A taxonomy of different objects that might be aggregated
  – The theory of judgment aggregation: a general theory of binary aggregation problems?

• Part II (Franz): A more general perspective
Judgment aggregation

The theory of judgment aggregation can be traced back to at least three different intellectual origins:

– work on abstract algebraic aggregation (Wilson 1975, Fishburn and Rubinstein 1986);
– logic-based work on judgment aggregation (List and Pettit 2002/4, Pauly and van Hees 2006, Dietrich 2006/7);
– related work on strategy-proof social choice (Nehring and Puppe 2002/7).

(For a review of more recent work, see, e.g., List and Puppe 2009.)

An earlier precursor is Guilbaud 1966, and the so-called “doctrinal paradox” in a legal context (Roberto Vacca 1921, Kornhauser and Sager 1986/1993).
Judgment aggregation

• The theory of judgment aggregation focuses, in the first place, on the aggregation of yes/no or true/false judgments on some propositions, such as:
  – $\text{CO}_2$ emissions are above some critical threshold (proposition $p$).
  – If $\text{CO}_2$ emissions are above that threshold, then there will be a long-term temperature increase of more than 2 degrees Celsius (proposition $p \rightarrow q$).
  – There will be a long-term temperature increase of more than 2 degrees Celsius (proposition $q$).

• Generally, propositions can be represented by sentences from a suitable logic, which is endowed with a notion of consistency.
Judgment aggregation

- The challenges raised in that context are similar to those in Arrovian preference aggregation.
- In particular, majority voting and some other simple aggregation rules fail to guarantee logically consistent majority outcomes.
### A majority inconsistency

<table>
<thead>
<tr>
<th></th>
<th>( p )</th>
<th>( p \rightarrow q )</th>
<th>( q )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual 1</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Individual 2</strong></td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Individual 3</strong></td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Majority</strong></td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>
Preference aggregation as a special case

• But the model of judgment aggregation is more general than the Arrovian model of preference aggregation, as preference orderings can be formally represented as yes/no judgments on special kinds of propositions: propositions of binary preferability.

• For instance, the ordering $x > y > z$ corresponds to the acceptance of the propositions “$x$ is preferable to $y$”, “$y$ is preferable to $z$”, and “$x$ is preferable to $z$”.

• Logical consistency becomes consistency relative to the standard rationality constraints on preferences (such as transitivity).

• In this way, Condorcet’s paradox, for example, emerges as a special case of the problem of majority inconsistency in judgment aggregation.
# Condorcet’s paradox

<table>
<thead>
<tr>
<th></th>
<th>$xPy$</th>
<th>$yPz$</th>
<th>$xPz$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individual 1</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$(xPyPz)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Individual 2</strong></td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>$(yPzPx)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Individual 3</strong></td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>$(zPxPy)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Majority</strong></td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>
The earlier example

<table>
<thead>
<tr>
<th></th>
<th>$p$</th>
<th>$p \rightarrow q$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual 1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Individual 2</td>
<td>✓</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Individual 3</td>
<td>×</td>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>Majority</td>
<td>✓</td>
<td>✓</td>
<td>×</td>
</tr>
</tbody>
</table>
Subsuming binary aggregation problems in the JA model

• Any binary relations can be propositionally represented in this way, including preference orderings and credibility orderings, with the relevant constraints built into the logic.

• Similarly, any categorical beliefs and desires allow a propositional representation (e.g., the belief that \( p \) is true or the desire that \( p \) should be true).

• For this reason, all binary aggregation problems can in principle be formally represented as judgment aggregation problems.
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preferences</td>
</tr>
</tbody>
</table>
Subsuming binary aggregation problems in the JA model

• In this way, Arrow’s theorem and many other results can be generalized to a larger class of binary aggregation problems.

• As an illustration, we’ll informally state just one such result.
Arrow’s theorem in judgment aggregation

• Suppose we wish to aggregate yes/no or true/false judgments on a set of propositions (called the agenda) with the following properties, very informally stated:
  – The set has at least one minimal inconsistent subset of three or more propositions.
  – It is not isomorphic to a set of propositions whose only logical connectives are “not” and “if and only if”.
  – Any proposition in the set can be “reached” from any other via a sequence of conditional entailments (e.g., “x>z” can be reached from “x>y”, conditional upon “y>z”).
Arrow’s theorem in judgment aggregation

• Then the following holds:
  Any aggregation rule satisfying universal domain, collective rationality, propositionwise independence, and unanimity preservation is a dictatorship of one individual.

(Proved in D&L, SCW 2007, and in Dokow & Holzman, JET 2010, where it is also shown that the assumption on the agenda is minimal. An earlier related result with an additional monotonicity condition is in Nehring & Puppe, JET 2010.)

• Arrow’s theorem is a corollary.
Other general results

• This result is just illustrative of the large variety of results that can be obtained.

• In a similar way, we can obtain generalizations of, for example:
  – Sen’s liberal paradox,
  – several results on domain restrictions,
  – Gibbard-Satterthwaite-style results on strategic voting,
  – some results on path-dependence and agenda manipulability, and so on.
Other general results

• In each case, there are some additional complexities because the logical interconnections between propositions can be more general than those generated by the standard rationality constraints on binary preference rankings (such as transitivity).

• For the purposes of this lecture, however, we want to move one step further and ask whether we can study the aggregation of intentional attitudes more generally, going beyond binary attitudes.
A taxonomy of intentional attitudes

<table>
<thead>
<tr>
<th>Structure Type</th>
<th>Binary</th>
<th>Non-binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute</td>
<td>Relational</td>
</tr>
<tr>
<td>Belief-attitudes</td>
<td>Judgments</td>
<td>Ordinal credences</td>
</tr>
<tr>
<td>Desire-attitudes</td>
<td>Categorical desires</td>
<td>Preferences</td>
</tr>
</tbody>
</table>
Plan for this lecture

• Part I (Christian): A basic perspective
  – Background: the theory of aggregation and the challenge
  – A taxonomy of different objects that might be aggregated
  – The theory of judgment aggregation: a general theory of binary aggregation problems?

• Part II (Franz): A more general perspective
Plan for this lecture

• Part I (Christian): A basic perspective
• Part II (Franz): A more general perspective
  – Model
  – An Arrow-type result
  – A Sen-type result
Plan for this lecture

- Part I (Christian): A basic perspective
- Part II (Franz): A more general perspective
  - Model
  - An Arrow-type result
  - A Sen-type result
• Individuals 1, 2, …, \( n \) \( (n \geq 2) \), faced with:
• A set \( X \) (**agenda**) of propositions under consideration (closed under negation, non-empty)
  → on these propositions collective attitudes must be formed

• E.g.: \( X \) contains the following prop.’s and their negations:
  \( a: \) Carbon dioxide emissions are above threshold \( x \)
  \( b: \) Global warming will continue
  \( a \rightarrow b: \) If carbon dioxide emissions are above threshold \( x \)
  \hspace{1cm} then global warming will continue

• Propositions (like \( a, b, a \rightarrow b \)) can be thought of as sentences (syntactic approach) or as sets of possible worlds (semantic approach).
Model

• $V$: set of values that an attitude on a proposition can take.
  
  – $V = \{0,1\} = \{\text{accept, reject}\}$ for binary attitudes
  
  – $V = [0,1]$ for probabilistic attitudes
  
  – ... 

• An **attitude function** is any function $X \rightarrow V$

• Not all attitude functions are rational!
  (More on rationality soon.)
Goal: aggregate individual attitude functions into (hopefully rational) collective attitude functions.

An aggregation rule/function $F$ maps any profile $(A_1,\ldots,A_n)$ of attitude functions (from some domain) to an aggregate attitude function $F(A_1,\ldots,A_n)$.

We will here focus on attitude functions satisfying universal domain and collective rationality:

- these are functions $F : \mathbb{R}^n \rightarrow \mathbb{R}$
- where $\mathbb{R} := \{\text{all rational attitude functions}\}$. 
E.g., an aggregation rule $F$ (with universal domain) is

- a **dictatorship** if $F$ is given by

  $$F(A_1, \ldots, A_n) = A_i$$

  for some fixed individual $i$.

- a **linear averaging rule** if $V \subseteq \mathbb{R}$ and $F$ is given by

  $$F(A_1, \ldots, A_n) = w_1A_1 + \ldots + w_nA_n$$

  for fixed weights $w_1, \ldots, w_n \geq 0$ of sum 1,
Model

• The notion of rationality is induced by the logical interconnections within the agenda.
  – E.g., \{a, b, \neg(a \rightarrow b)\} is inconsistent.

• But where do these interconnections come from?
Model

- $X$ is a subset of $L$ (the logic), which also contains the propositions not under consideration (as, in the example, $a \land b, (a \land b) \rightarrow c, ...$)
  - $L$ has the structure of a Boolean algebra.
  - $X$ inherits its logical interconnections from $L$.
- Certain functions from $L$ to $V$ are admissible as valuation functions, e.g. the:
  - (binary logic) truth functions into $V = \{0,1\}$
  - (probability theory) probability measures into $V = [0,1]$
  - ...
Or, the valuation functions might be the

- (Dempster-Schafer theory) lower-probability functions into $V=[0,1]$
- (Spohnian ranking theory) ranking functions into $V = \{0,1,2,...\} \cup \{\infty\}$
- ($T$-valued logic) $T$-valued truth functions into $V = \{0,1,2,...,T-1\}$
- (why not?) functions $f$ into $V = \{0,1,\text{‘undecided out of conflicting info’}, \text{‘undecided out of conflicting intuition’}\}$ s.t. if $p$ entails $q$ then $f(q) \geq f(p)$, where $\geq$ is the partial order on $V$ given by $1 >$ each ‘undecidedness’ value $> 0$. 

Model
Model

• Now an attitude function $X \rightarrow V$ is rational if it is extendible to an admissible valuation function $L \rightarrow V$.
  
  – i.e., depending on the context, to a truth function, a probability measure, ...
Hope: general theorems on attitude aggregation, with corollaries for special aggregation problems.

I’ll present two theorems to you:

1) A result generalising Arrow (and more)
2) A result generalising Sen
Plan for this lecture

• Part I (Christian): A basic perspective
• Part II (Franz): A more general perspective
  – Model
  – An Arrow-type result
  – A Sen-type result
An Arrow-type result

**Independence.** For all propositions $p$ in $X$,

- (informally) the aggregate attitude on $p$ depends only on the individuals’ attitudes on $p$
- (formally) there is a fixed ‘decision method’ $D_p: V^n \rightarrow V$ such that
  
  $$F(A_1, \ldots, A_n)(p) = D_p(A_1(p), \ldots, A_n(p))$$

  for all admissible profiles $(A_1, \ldots, A_n)$.

**Motivation:** Local concept of democracy, preventing strategic manipulation.
An Arrow-type result

You might now expect this condition:

**Unanimity preservation.** If in an admissible profile \((A_1,\ldots,A_n)\) all \(A_i\) take the same attitude on some proposition \(p\in X\), then \(F(A_1,\ldots,A_n)\) takes this attitude on \(p\).

But UP doesn’t lead very far (with independence): it leads to

- *neutrality*, i.e., to a uniform decision method \(D = D_p\) (a ‘contagion’ or ‘field expansion’ lemma), for ‘most’ types of attitude and ‘many but far from all’ agendas \(X\)
- but not to more structure of \(D\), such as linearity or dictatorship.
An Arrow-type result

- We consider a different preservation condition than UP.
- It assumes that $V$ contains a *greatest* value (with respect to some relation of strength of belief/desire).
- The greatest value is 1 for binary or probabilistic attitudes.
Implication preservation.

• Informally: unanimous conditional attitudes are preserved.
  E.g., if all believe that if it rains then we get wet, then so does the collective.

• Formally: For all propositions $p, q$ in $X$ and admissible profiles $(A_1, \ldots, A_n)$, if every $A_i$ assigns the greatest value to $p \rightarrow q$ (in some extension to a valuation function in case $p \rightarrow q$ is not in $X$) then so does $F(A_1, \ldots, A_n)$.

IP doesn’t follow from UP because $p \rightarrow q$ needn’t belong to $X$. 

An Arrow-type result
An Arrow-type result

Theorem 1. Suppose attitudes are **logical or probabilistic**. If (and only if) the agenda $X$ is **non-simple**, the only independent and implication-preserving aggregation rules $F : \mathbb{R}^n \to \mathbb{R}$ are the linear averaging rules.

- A **non-simple** agenda has at least one minimal inconsistent subset $Y$ with $|Y| \geq 3$.
  - e.g. $Y = \{a, a \to b, \neg b\}$
- Simple agendas are trivial in the sense of not displaying any complex interconnections.
Application 1: probabilistic attitudes

- The probability aggregation (‘opinion pooling’) literature requires the agenda $X$ to form an algebra (or to take another special form)
  - very richly interconnected agenda
- This literature’s classic characterization of linear averaging rules (McConway 1981, Wagner 1982) is based on independence & UP.
- It can be obtained from ours by specialising to probabilistic attitudes and algebra agendas
  - since IP is then equivalent to UP.
Application 2: logical attitudes

**Corollary.** Suppose attitudes are logical. If (and only if) the agenda \( X \) is non-simple, the only independent and implication-preserving aggregation rules \( F : \mathbb{R}^n \rightarrow \mathbb{R} \) are the dictatorships.

**Proof.** If \( V = \{0,1\} \), ‘linear averaging’ is equivalent to ‘dictatorship’

– Because the only weight distributions which ensure aggregate attitudes in \( \{0,1\} \) are those assigning all weight to some individual.

**Remark.** To see why a simple agenda allows for non-dictatorial solutions, note that propositionwise majority rule among a fixed odd-sized non-singleton subgroup works.
Application 3: preference aggregation

• When applied to a preference agenda, Theorem 1 comes close to Arrow’s Theorem.

• Set of at least three alternatives $K$
  \[ X = \{xPy : x, y \text{ distinct alternatives in } K\}. \]

• The rationality conditions on strict linear preferences are built as axioms into the logic.
  – This creates the right logical interconnections;
  – E.g., the set \{xPy, yPz, zPx\} is inconsistent.

• $xPy$ is identified with the negation of $yPx$
  – So $X$ is closed under negation (as usual for agendas).
Application 3: preference aggregation

**Corollary.** For the preference agenda, the only independent and implication-preserving aggregation rules $F : \mathbb{R}^n \rightarrow \mathbb{R}$ are the dictatorships.

*Proof.* The preference agenda is non-simple, as the set $Y = \{xPy, yPz, zPx\}$ is minimal inconsistent with $|Y| \geq 3$, for any distinct $x,y,z$. 
Application 3: preference aggregation

This corollary comes close to Arrow’s Theorem:
• universal domain ⇐⇒ Arrow’s universal domain;
• aggregate rationality ⇐⇒ Arrow’s collective rationality;
• independence ⇐⇒ Arrow’s IIA;
• Implication-preservation ⇒ weak Pareto principle.
Plan for this lecture

• Part I (Christian): A basic perspective
• Part II (Franz): A more general perspective
  – Model
  – An Arrow-type result
  – A Sen-type result
Rights can clash with unanimous attitudes

• Two experts need collective judgments on our climate change agenda.
• Aggregation respects expert rights and unanimous attitudes.

Case 1: binary logical attitudes

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>a → b</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>expert on a</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>expert on a → b</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>group</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Rights can clash with unanimous attitudes

Case 2: probabilistic attitudes

<table>
<thead>
<tr>
<th></th>
<th>$a$</th>
<th>$a \rightarrow b$</th>
<th>$b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>expert on $a$</td>
<td>$\frac{3}{4}$</td>
<td>$\frac{1}{4}$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>expert on $a \rightarrow b$</td>
<td>$\frac{1}{4}$</td>
<td>$1$</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>group</td>
<td>$\frac{3}{4}$</td>
<td>$1$</td>
<td>$\frac{1}{4}$</td>
</tr>
</tbody>
</table>
A Sen-type result: impossibility of a Paretian liberal

**Definition.** Individual $i$ is *decisive* on proposition $p \in X$ if $F(A_1, \ldots, A_n)(p) = A_i(p)$ for all admissible profiles $(A_1, \ldots, A_n)$.

**Minimal rights.** There exist (at least) two individuals who are each decisive on (at least) one proposition.

**Theorem.** If (and only if) $X$ is connected, no aggregation rule $F : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies minimal rights and unanimity preservation.
**Connectedness**

*Definition.* The agenda $X$ is *connected* if any two propositions $p, p^*$ in $X$ are conditionally dependent.

... where propositions $p, p^* \in X$ are *conditionally dependent* (in $X$) if

(informally) the attitudes taken on them constrain each other *conditional* on taking certain attitudes on certain other propositions;

(formally) there exist values $v, v^* \in V$ and a function $g$ from some set $Y \subseteq X$ to $V$ such that, among all rational valuation functions $A$ matching $g$ on $Y$, none satisfies both $A(p) = v$ and $A(p^*) = v^*$ (but some satisfies just $A(p) = v$ and some satisfies just $A(p^*) = v^*$).
Many agendas are connected:

- The preference agenda $X = \{xPy, yPz, \ldots\}$ is connected (whether attitudes are binary or probabilistic).
  
  E.g., $xPy$ and $yPz$ are dependent conditional on the attitude on $xPz$.

- The agenda $X = \{a, \neg a, a \rightarrow b, \neg(a \rightarrow b), b, \neg b\}$ is connected (whether attitudes are binary or probabilistic).
  
  E.g., $a$ and $b$ are dependent conditional on a non-zero attitude on $a \rightarrow b$. 
**Theorem (repeated).** If (and only if) $X$ is connected, no aggregation rule $F : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies minimal rights and unanimity preservation.

**Proof.** Part 1. **Sufficiency.** Assume for a contradiction that $X$ is connected and $F$ has all properties. By minimal rights, some individuals $i, j$ are decisive on some propositions $p, q \in X$, respectively. Since $X$ is connected, there are values $v, w \in V$ and a function $g$ from some set $Y \subseteq X$ such that

(i) no rational valuation function $C$ satisfies $C(p)=v$, $C(q)=w$ and $C|_Y = g$;

(ii) some rational valuation function $A$ satisfies $A(p) = v$ and $A|_Y = g$;

(iii) some rational valuation function $B$ satisfies $B(q) = w$ and $B|_Y = g$. 

Proof
Proof

Construct a profile \((A_1, \ldots, A_n)\) such that \(A_i\) is \(A\), \(A_j\) is \(B\), and each other judgment set is either \(A\) or \(B\).

By universal domain, this profile is admissible.

- \(F(A_1, \ldots, A_n)(p) = v\) because \(i\) is decisive on \(p\);
- \(F(A_1, \ldots, A_n)(q) = w\) because \(j\) is decisive on \(q\);
- \(F(A_1, \ldots, A_n)|_{\gamma} = g\) because \(A_1|_{\gamma} = \ldots = A_n|_{\gamma} = g\).

This contradicts (i).
Proof

Part 2. *Necessity*. Assume $X$ is not connected. So some $p, q \in X$ aren’t conditionally dependent. Let $F$ be an aggregation rule with universal domain given as follows. Consider any profile $(A_1, \ldots, A_n)$ of rational attitude functions. Since $p$ and $q$ are not connected, there exists a rational attitude function $A$ such that

1. $A(p) = A_1(p)$,
2. $A(q) = A_2(q)$, and
3. $A$ agrees with all $A_i$ on $Y := A_1 \cap \ldots \cap A_n$.

(This uses the fact that (3) is consistent with (1) alone (take $A=A_1$) and with (2) alone (take $A = A_2$).)

Put $F(A_1, \ldots, A_n) := A$. 
Proof

The so-defined aggregation rule satisfies

- minimal rights, since individuals 1 and 2 are decisive on $p$ and $q$ by (1) and (2), respectively,
- unanimity preservation by (3),
- universal domain and aggregate rationality by definition.
How our result generalises Sen’s

<table>
<thead>
<tr>
<th>Our result</th>
<th>Sen’s result</th>
</tr>
</thead>
<tbody>
<tr>
<td>any connected agenda</td>
<td>the preference agenda</td>
</tr>
<tr>
<td>any kind of attitude</td>
<td>binary attitude</td>
</tr>
<tr>
<td>Minimal rights</td>
<td>Minimal Liberalism</td>
</tr>
<tr>
<td>Unanimity Principle</td>
<td>Pareto Principle</td>
</tr>
<tr>
<td>Universal Domain</td>
<td>Universal Domain</td>
</tr>
<tr>
<td>Aggregate rationality</td>
<td>Aggregate rationality</td>
</tr>
</tbody>
</table>
Thank you!