This is a preprint of a paper published in Social Choice and Welfare 29(1): 19-33, 2007.

Arrow's theorem in judgment aggregation

Franz Dietrich and Christian List[1]

first version 4/2005, revised 7/2006
(Social Choice and Welfare, forthcoming)

In response to recent work on the aggregation of individual judgments on logically connected propositions into collective judgments, it is often asked whether judgment aggregation is a special case of Arrowian preference aggregation. We argue for the converse claim. After proving two impossibility theorems on judgment aggregation (using "systematicity" and "independence" conditions, respectively), we construct an embedding of preference aggregation into judgment aggregation and prove Arrow's theorem (stated for strict preferences) as a corollary of our second result. Although we thereby provide a new proof of Arrow's theorem, our main aim is to identify the analogue of Arrow's theorem in judgment aggregation, to clarify the relation between judgment and preference aggregation, and to illustrate the generality of the judgment aggregation model.
JEL Classification: D70, D71

1 Introduction

The problem of "judgment aggregation" has recently received much attention: How can the judgments of several individuals on logically connected propositions be aggregated into corresponding collective judgments? To illustrate, suppose a three-member committee has to make collective judgments (acceptance/rejection) on three connected propositions:

a: "Carbon dioxide emissions are above the threshold x."

$a \to b$: "If carbon dioxide emissions are above the threshold x, then there will be global warming."

b: "There will be global warming."

	a	$a \to b$	b
Individual 1	True	True	True
Individual 2	True	False	False
Individual 3	False	True	False
Majority	True	True	False

Table 1: The discursive paradox

[1]We thank Richard Bradley, Ruvin Gekker, Ron Holzman, Philippe Mongin, Klaus Nehring, Clemens Puppe and the referees for comments and suggestions. Addresses for correspondence: F. Dietrich, Department of Quantitative Economics, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands; C. List, Department of Government, London School of Economics, London WC2A 2AE, U.K.

As shown in Table 1, the first committee member accepts all three propositions; the second accepts a but rejects $a \rightarrow b$ and b; the third accepts $a \rightarrow b$ but rejects a and b. Then the judgments of each committee member are individually consistent, and yet the majority judgments on the propositions are inconsistent: a majority accepts a, a majority accepts $a \rightarrow b$, but a majority rejects b.

This so-called *discursive paradox* (Pettit 2001) has led to a growing literature on the possibility of consistent judgment aggregation under various conditions. List and Pettit (2002) have provided a first model of judgment aggregation based on propositional logic and proved that no aggregation rule generating consistent collective judgments can satisfy some conditions inspired by (but not equivalent to) Arrow's conditions on preference aggregation. This impossibility result has been extended and strengthened by Pauly and van Hees (forthcoming; see also van Hees forthcoming), Dietrich (2006), and Gärdenfors (forthcoming).[2] Abstracting from propositional logic, Dietrich (forthcoming) has provided a model of judgment aggregation in general logics, which we use in the present paper, that can represent aggregation problems involving propositions formulated in richer logical languages. Drawing on the related model of "property spaces", Nehring and Puppe (2002, 2005) have proved the first agenda characterization results, identifying necessary and sufficient conditions under which an agenda of propositions gives rise to an impossibility result under certain conditions.

Although judgment aggregation is different from the more familiar problem of preference aggregation, the recent results resemble earlier results in social choice theory. The discursive paradox resembles Condorcet's paradox of cyclical majority preferences, and the various recent impossibility theorems resemble Arrow's and other theorems on preference aggregation. This raises the question of how the work on judgment aggregation is related to earlier work in social choice theory. Provocatively expressed, is it just a reinvention of the wheel?

It can be replied that the logic-based model of judgment aggregation generalizes Arrow's classical model of preference aggregation. Specifically, preference aggregation problems can be modelled as special cases of judgment aggregation problems by representing preference orderings as sets of binary ranking judgments in predicate logic (List and Pettit 2001/2004; List 2003).[3] Less formally, this way of thinking about preferences goes back to Condorcet himself (see also Guilbaud 1966).

In this paper, we reinforce this argument. After introducing the judgment aggregation model in general logics and proving two impossibility results (using "systematicity" and "independence" conditions, respectively), we construct

[2]Possibility results, obtained by relaxing some of the conditions of these impossibility results, have been proved by List (2003, 2004); Dietrich (2006), Pigozzi (forthcoming), and Dietrich and List (2005). The relationship with the Condorcet jury theorem has been investigated by Bovens and Rabinowicz (2006) and List (2005).

[3]This embedding works only for the ordinal preference-relation-based part of Arrowian social choice theory, not for the cardinal welfare-function-based part. Wilson's (1975) aggregation model, as discussed in our concluding remarks, is another generalization of ordinal preference aggregation.

an explicit embedding of preference aggregation into judgment aggregation and prove Arrow's theorem (for strict preferences) as a corollary of our second impossibility result. We also point out that our first impossibility result has corollaries for the aggregation of other binary relations (such as partial orderings or equivalence relations).

Although we thereby provide a new proof of Arrow's theorem, our primary aim is to identify the analogue of Arrow's theorem in judgment aggregation, to clarify the logical relation between judgment and preference aggregation, and to illustrate the generality of the judgment aggregation model.

Related results were given by List and Pettit (2001/2004), who derived a simple impossibility theorem on preference aggregation from their (2002) impossibility result on judgment aggregation, and Nehring (2003), who derived an Arrow-like impossibility theorem from Nehring and Puppe's (2002) characterization result in the related model of "property spaces". But neither result exactly matches Arrow's theorem. Compared to Arrow's original theorem, List and Pettit's result requires additional neutrality and anonymity conditions, but no Pareto principle; Nehring's result requires an additional monotonicity condition. We highlight the connections of our present results with these and other results (including recent results by Dokow and Holzman 2005) throughout the paper.

2 The judgment aggregation model

We consider a group of individuals $1, 2, \ldots, n$ ($n \geq 2$). The group has to make collective judgments on logically connected propositions.

Formal logic. Propositions are represented in an appropriate logic. A *logic (with negation symbol* $\neg$*)* is an ordered pair $(\mathbf{L}, \vDash)$, where (i) $\mathbf{L}$ is a non-empty set of formal expressions (*propositions*) closed under negation (i.e., if $p \in \mathbf{L}$ then $\neg p \in \mathbf{L}$), and (ii) $\vDash$ is an *entailment relation*, where, for each set $A \subseteq \mathbf{L}$ and each proposition $p \in \mathbf{L}$, $A \vDash p$ is read as "A entails p" (we write $p \vDash q$ to abbreviate $\{p\} \vDash q$).[4]

A set $A \subseteq \mathbf{L}$ is *inconsistent* if $A \vDash p$ and $A \vDash \neg p$ for some $p \in \mathbf{L}$, and *consistent* otherwise; $A \subseteq \mathbf{L}$ is *minimal inconsistent* if it is inconsistent and every proper subset $B \subsetneq A$ is consistent. A proposition $p \in \mathbf{L}$ is *contingent* if $\{p\}$ and $\{\neg p\}$ are consistent.

We require the logic to satisfy the following minimal conditions:

(L1) For all $p \in \mathbf{L}$, $p \vDash p$ (self-entailment).

(L2) For all $p \in \mathbf{L}$ and $A \subseteq B \subseteq \mathbf{L}$, if $A \vDash p$ then $B \vDash p$ (monotonicity).

(L3) $\emptyset$ is consistent, and each consistent set $A \subseteq \mathbf{L}$ has a consistent superset $B \subseteq \mathbf{L}$ containing a member of each pair $p, \neg p \in \mathbf{L}$ (completability).

[4]Formally, $\vDash \subseteq \mathcal{P}(\mathbf{L}) \times \mathbf{L}$, where $\mathcal{P}(\mathbf{L})$ is the power set of $\mathbf{L}$.

Many different logics satisfy conditions L1 to L3, including standard propositional logic, standard modal and conditional logics and, for the purpose of representing preferences, predicate logic, as defined below. For example, in standard propositional logic, $\mathbf{L}$ contains propositions such as a, b, $a \wedge b$, $a \vee b$, $a \rightarrow b$, $\neg(a \wedge b)$, and $\vDash$ satisfies $\{a, a \rightarrow b\} \vDash b$, $b \vDash a \vee b$, but not $b \vDash a \wedge b$.

The agenda. The *agenda* is a non-empty subset $X \subseteq \mathbf{L}$, interpreted as the set of propositions on which judgments are to be made, where X is a union of proposition-negation pairs $\{p, \neg p\}$ (with p not itself a negated proposition). For simplicity, we assume that double negations cancel each other out, i.e., $\neg\neg p$ stands for p.[5] In the discursive paradox example above, the agenda is $X = \{a, \neg a, b, \neg b, a \rightarrow b, \neg(a \rightarrow b)\}$, with $\rightarrow$ interpreted either as the material conditional in standard propositional logic or as a subjunctive conditional in a suitable conditional logic.

Agenda richness. Whether or not judgment aggregation gives rise to serious impossibility results depends on how the propositions in the agenda are interconnected. We consider agendas X with different types of interconnections. Our basic agenda assumption, which significantly generalizes the one in List and Pettit (2002), is *minimal connectedness.* An agenda X is *minimally connected* if (i) it has a minimal inconsistent subset $Y \subseteq X$ with $|Y| \geq 3$, and (ii) it has a minimal inconsistent subset $Y \subseteq X$ such that $(Y \backslash Z) \cup \{\neg z : z \in Z\}$ is consistent for some subset $Z \subseteq Y$ of even size.[6]

As Ron Holzman has indicated to us, part (ii) of minimal connectness is equivalent to Dokow and Holzman's (2005) assumption that the set of admissible yes/no views on the propositions in X is a non-affine subset of $\{0, 1\}^X$.[7]

To obtain a more demanding agenda assumption, we define *path-connectedness*, a variant of Nehring and Puppe's (2002) assumption of *total blockedness.*[8] For any $p, q \in X$, we write $p \vDash^* q$ if $\{p, \neg q\} \cup Y$ is inconsistent for some $Y \subseteq X$ consistent with p and with $\neg q$.[9] Now an agenda X is *path-connected* if, for every contingent $p, q \in X$, there exist $p_1, p_2, ..., p_k \in X$ (with $p = p_1$ and $q = p_k$) such that $p_1 \vDash^* p_2$, $p_2 \vDash^* p_3$, ..., $p_{k-1} \vDash^* p_k$.

The agenda of our example above is minimally connected, but not path-connected. As detailed below, preference aggregation problems can be represented by agendas that are both minimally connected and path-connected. The

[5] When we use the negation symbol $\neg$ hereafter, we mean a modified negation symbol $\sim$, where $\sim p := \neg p$ if p is unnegated and $\sim p := q$ if $p = \neg q$ for some q.

[6] Note that the set Y can be different in parts (i) and (ii).

[7] In the first version of this paper, we had used a more restrictive version of part (ii), requiring Z to be of size two rather than even size. The present version of part (ii) was introduced in Dietrich (forthcoming).

[8] For a compact logic, path-connectedness is equivalent to total blockedness; in the general case, path-connectedness is weaker.

[9] For non-paraconsistent logics (in the sense of L4 in Dietrich forthcoming), $\{p, \neg q\} \cup Y$ is inconsistent if and only if $\{p\} \cup Y \vDash q$.

aggregation of many other binary relations can be represented by minimally connected agendas.

Individual judgment sets. Each individual i's *judgment set* is a subset $A_i \subseteq X$, where $p \in A_i$ means that individual i accepts proposition p. A judgment set A_i is *consistent* if it is a consistent set as defined above; A_i is *complete* if, for every proposition $p \in X$, $p \in A_i$ or $\neg p \in A_i$. A *profile (of individual judgment sets)* is an n-tuple $(A_1, \ldots, A_n)$.

Aggregation rules. A *(judgment) aggregation rule* is a function F that assigns to each admissible profile $(A_1, \ldots, A_n)$ a single collective judgment set $F(A_1, \ldots, A_n) = A \subseteq X$, where $p \in A$ means that the group accepts proposition p. The set of admissible profiles is called the *domain* of F, denoted $Domain(F)$. Examples of aggregation rules are the following.

- *Propositionwise majority voting.* For each $(A_1, ..., A_n)$, $F(A_1, ..., A_n) = \{p \in X :$ more individuals i have $p \in A_i$ than $p \notin A_i\}$.
- *Dictatorship of individual i.* For each $(A_1, ..., A_n)$, $F(A_1, ..., A_n) = A_i$.
- *Inverse dictatorship of individual i.* For each $(A_1, ..., A_n)$, $F(A_1, ..., A_n) = \{\neg p : p \in A_i\}$.

Regularity conditions on aggregation rules. We impose the following conditions on the inputs and outputs of aggregation rules.

Universal domain. The domain of F is the set of all possible profiles of consistent and complete individual judgment sets.

Collective rationality. F generates consistent and complete collective judgment sets.

Propositionwise majority voting, dictatorships and inverse dictatorships satisfy universal domain, but only dictatorships generally satisfy collective rationality. As the discursive paradox example of Table 1 shows, propositionwise majority voting sometimes generates inconsistent collective judgment sets. Inverse dictatorships satisfy collective rationality only in special cases (i.e., when the agenda is *symmetrical*: for every consistent $Z \subseteq X$, $\{\neg p : p \in Z\}$ is also consistent).

3 Two impossibility theorems on judgment aggregation

Are there any non-dictatorial judgment aggregation rules satisfying universal domain and collective rationality? The following conditions are frequently used in the literature.

Independence. For any proposition $p \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*) \in Domain(F)$, if [for all individuals i, $p \in A_i$ if and only if $p \in A_i^*$] then [$p \in F(A_1, \ldots, A_n)$ if and only if $p \in F(A_1^*, \ldots, A_n^*)$].

Systematicity. For any propositions $p, q \in X$ and profiles $(A_1, \ldots, A_n)$, $(A_1^*, \ldots, A_n^*) \in Domain(F)$, if [for all individuals i, $p \in A_i$ if and only if $q \in A_i^*$] then [$p \in F(A_1, \ldots, A_n)$ if and only if $q \in F(A_1^*, \ldots, A_n^*)$].

Unanimity principle. For any profile $(A_1, \ldots, A_n) \in Domain(F)$ and any proposition $p \in X$, if $p \in A_i$ for all individuals i, then $p \in F(A_1, \ldots, A_n)$.

Independence requires that the collective judgment on each proposition should depend only on individual judgments on that proposition. Systematicity strengthens independence by requiring in addition that the same pattern of dependence should hold for all propositions (a neutrality condition). The unanimity principle requires that if all individuals accept a proposition then this proposition should also be collectively accepted. The following result holds.

Proposition 1. For a minimally connected agenda X, an aggregation rule F satisfies universal domain, collective rationality, systematicity and the unanimity principle if and only if it is a dictatorship of some individual.

Proof. All proofs are given in the appendix. ■

Proposition 1 is related to an earlier result by Dietrich (forthcoming), which requires an additional assumption on the agenda X but no unanimity principle (the additional assumption is that X is also *asymmetrical*: for some inconsistent $Z \subseteq X$, $\{\neg p : p \in Z\}$ is consistent). This result, in turn, generalizes an earlier result on systematicity by Pauly and van Hees (forthcoming).

From Proposition 1, we can derive two new results of interest. The first is a generalization of List and Pettit's (2002) theorem on the non-existence of an aggregation rule satisfying universal domain, collective rationality, systematicity and anonymity (i.e., invariance of the collective judgment set under permutations of the given profile of individual judgment sets). Our result extends the earlier impossibility result to any minimally connected agenda and weakens anonymity to the requirement that there is no dictator or inverse dictator.

Theorem 1. For a minimally connected agenda X, every aggregation rule F satisfying universal domain, collective rationality and systematicity is a (possibly inverse) dictatorship of some individual.

The agenda assumption of Theorem 1 cannot be weakened further if the agenda is finite or the logic is compact (and $n \geq 3$ and X contains at least one contingent proposition), i.e., minimal connectedness is also necessary (and not just sufficient) for giving rise to (possibly inverse) dictatorships by the conditions of Theorem 1.[10]

The second result we can derive from Proposition 1 is the analogue of Arrow's theorem in judgment aggregation, from which we subsequently derive Arrow's theorem on (strict) preference aggregation as a corollary. We use the following lemma, which strengthens an earlier lemma by Nehring and Puppe (2002) by not requiring monotonicity.

Lemma 1. For a path-connected agenda X, an aggregation rule F satisfying universal domain, collective rationality, independence and the unanimity principle also satisfies systematicity.

Let us call an agenda *strongly connected* if it is both minimally connected and path-connected. Using Lemma 1, Proposition 1 now implies the following impossibility result.

Theorem 2. For a strongly connected agenda X, an aggregation rule F satisfies universal domain, collective rationality, independence and the unanimity principle if and only if it is a dictatorship of some individual.

Dokow and Holzman (2005) have independently shown that (for a finite agenda containing only contingent propositions) strong connectedness (in the form of the conjunction of non-affineness and total blockedness) is both necessary and sufficient for characterizing dictatorships by the conditions of Theorem 2 (assuming $n \geq 3$). A prior closely related result is Nehring and Puppe's (2002) characterization result, using total blockeness alone but imposing an additional monotonicity condition. In fact, within the general logics framework, the necessity holds if the agenda is finite or the logic is compact (and X contains at least one contingent proposition; again assuming $n \geq 3$).

Proposition 1 and Theorems 1 and 2 continue to hold under generalized definitions of minimally connected and strongly connected agendas.[11]

Of course, it is debatable whether and when independence or systematicity are plausible requirements on judgment aggregation. The literature contains

[10] It can then be shown that, if X is not minimally connected, there exists an aggregation rule that satisfies universal domain, collective rationality and systematicity and is not a (possibly inverse) dictatorship. Let M be a subset of $\{1, ..., n\}$ of odd size at least 3. If part (i) of minimal connectedness is violated, then majority voting among the individuals in M satisfies all requirements. If part (ii) is violated, the aggregation rule F with universal domain defined by $F(A_1, ..., A_n) := \{p \in X :$ the number of individuals $i \in M$ with $p \in A_i$ is odd$\}$ satisfies all requirements. The second example is inspired by Dokow and Holzman (2005).

[11] In the definition of minimal connectedness, (i) can be weakened to the following: (i*) there is an inconsistent set $Y \subseteq X$ with pairwise disjoint subsets Z_1, Z_2, Z_3 such that $(Y \backslash Z_j) \cup \{\neg p : p \in Z_j\}$ is consistent for any $j \in \{1, 2, 3\}$ (Dietrich forthcoming). In the definition of strong connectedness (by (i), (ii) and path-connectedness), (i) can be dropped altogether, as path-connectedness implies (i*). In the definitions of minimal connectedness and strong connectedness, (ii) can be weakened to (ii*) in Dietrich (forthcoming).

extensive discussions of these conditions and their possible relaxations. In our view, the importance of Theorems 1 and 2 lies not so much in establishing the impossibility of consistent judgment aggregation, but rather in indicating what conditions must be relaxed in order to make consistent judgment aggregation possible. The theorems describe boundaries of the logical space of possibilities.

4 Arrow's theorem

We now show that Arrow's theorem (stated here for strict preferences) can be restated in the judgment aggregation model, where it is a direct corollary of Theorem 2. We consider a standard Arrowian preference aggregation model, where each individual has a strict preference ordering (asymmtrical, transitive and connected, as defined below) over a set of options $K = \{x, y, z, ...\}$ with $|K| \geq 3$. We embed this model into our judgment aggregation model by representing preference orderings as sets of binary ranking judgments in a simple predicate logic, following List and Pettit (2001/2004). Although we consider strict preferences for simplicity, we note that a similar embedding is possible for weak preferences.[12]

A simple predicate logic for representing preferences. We consider a predicate logic with constants $x, y, z, ... \in K$ (representing the options), variables v, w, v_1, v_2 , ..., identity symbol =, a two-place predicate P (representing strict preference), logical connectives $\neg$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (if-then), and universal quantifier $\forall$. Formally, **L** is the smallest set such that

- **L** contains all propositions of the forms $\alpha P \beta$ and $\alpha = \beta$, where α and β are constants or variables, and
- whenever **L** contains two propositions p and q, then **L** also contains $\neg p$, $(p \wedge q)$, $(p \vee q)$, $(p \rightarrow q)$ and $(\forall v)p$, where v is any variable.

Notationally, we drop brackets when there is no ambiguity. The entailment relation $\vDash$ is defined as follows. For any set $A \subseteq \mathbf{L}$ and any proposition $p \in \mathbf{L}$,

$$A \vDash p \text{ if and only if } \begin{array}{l} A \cup Z \text{ entails } p \text{ in the standard} \\ \text{sense of predicate logic,} \end{array}$$

where Z is the set of rationality conditions on strict preferences:

$(\forall v_1)(\forall v_2)(v_1 P v_2 \rightarrow \neg v_2 P v_1)$ (asymmetry);
$(\forall v_1)(\forall v_2)(\forall v_3)((v_1 P v_2 \wedge v_2 P v_3) \rightarrow v_1 P v_3)$ (transitivity);
$(\forall v_1)(\forall v_2)(\neg\, v_1{=}v_2 \rightarrow (v_1 P v_2 \vee v_2 P v_1))$ (connectedness).[13]

[12] If we represent *weak* preference aggregation in the judgment aggregation model using the embedding indicated below, the independence condition and the unanimity principle become stronger than Arrow's independence of irrelevant alternatives and the weak Pareto principle. So, in the case of weak preferences unlike that of strict ones, Theorem 2 only implies a slightly weaker form of Arrow's theorem.

[13] For technical reasons, Z also contains, for each pair of distinct constants x, y, the condition $\neg\, x{=}y$, reflecting the mutual exclusiveness of the options.

To represent weak preferences rather than strict ones, Z simply needs to be redefined as the set of rationality conditions on weak preferences (i.e., reflexivity, transitivity, and connectedness); see also Dietrich (forthcoming).[14] Binary relations with other properties can be represented analogously, by defining Z as the set of appropriate rationality conditions, e.g., the set containing reflexivity (respectively, asymmetry) and transitivity for weak (respectively, strict) partial orderings, and the set containing reflexivity, transitivity and symmetry for equivalence relations.

The agenda. The *preference agenda* is the set X of all propositions of the forms $xPy, \neg xPy \in \mathbf{L}$, where x and y are distinct constants.[15] Note the following lemma (which holds for strict as well as weak preferences). The path-connectedness part of the result is equivalent to a lemma by Nehring (2003).

Lemma 2. The preference agenda X is strongly connected.

The correspondence between preference orderings and judgment sets. It is easy to see that each (asymmetrical, transitive and connected) preference ordering over K can be represented by a unique consistent and complete judgment set in X and vice-versa, where individual i strictly prefers x to y if and only if $xPy \in A_i$. For example, if individual i strictly prefers x to y to z, this is uniquely represented by the judgment set $A_i = \{xPy, yPz, xPz, \neg yPx, \neg zPy, \neg zPx\}$.

The correspondence between Arrow's conditions and conditions on judgment aggregation. For the preference agenda, the conditions of universal domain, collective rationality, independence ("independence of irrelevant alternatives") and the unanimity principle ("the weak Pareto principle"), as stated above, exactly match the standard conditions of Arrow's theorem, where an Arrowian preference aggregation rule is represented by a judgment aggregation rule.

As the preference agenda is strongly connected, Arrow's theorem now follows from Theorem 2.

Corollary 1. (Arrow's theorem) For the preference agenda X, an aggregation rule F satisfies universal domain, collective rationality, independence and the unanimity principle if and only if it is a dictatorship of some individual.

[14] Transitivity and connectedness are as defined above. Reflexivity can be stated by the proposition $(\forall v)(vPv)$. For aesthetic reasons, one might also replace the predicate symbol P by R in the logic.

[15] xPy is interpreted as "x is better than/preferable to y". Note that this represents preference aggregation as the aggregation of *beliefs of betterness/preferability*. One might argue that preferences are desire-like rather than belief-like and thus object to re-interpreting them as beliefs of preferability. To respond to this objection, we might, for example, interpret xPy as "x is socially preferred to y", and interpret an individual judgment set $A_i \subseteq X$ as the set of propositions that individual i *desires* (rather than *believes*), while interpreting a collective judgment set $A \subseteq X$ as a set of propositions about social preference.

Corollary 1 strengthens Nehring's (2003) corollary by dropping monotonicity; it also strengthens List and Pettit's (2001/2004) corollary by weakening systematicity to independence and (effectively) anonymity to non-dictatorship, at the expense of imposing, in addition, the unanimity principle.

The correspondence between preference and judgment aggregation concepts under the constructed embedding is summarized in Table 2.

Preference aggregation	Judgment aggregation
Preference ordering over a set of options	Judgment set in the preference agenda
Three or more options	Strong connectedness of the preference agenda
Asymmetry, transitivity and connectedness of the preference ordering	Consistency and completeness of the judgment set
Preference aggregation rule	Judgment aggregation rule
Universal domain	Universal domain
Collective rationality	Collective rationality
Independence of irrelevant alternatives	Independence
Weak Pareto principle	Unanimity principle
Arrowian dictator	(Judgment) dictator
Arrow's theorem	Corollary of Theorem 2

Table 2: The embedding of concepts

5 Concluding remarks

After proving two impossibility theorems on judgment aggregation – Theorem 1 with systematicity and a weak agenda assumption, Theorem 2 with independence and a stronger agenda assumption – we have shown that Arrow's theorem (for strict preferences) is a corollary of Theorem 2, applied to the aggregation of binary ranking judgments in a simple predicate logic. In the case of binary relations other than preference orderings, Theorem 2 does not necessarily apply, as the resulting agenda is not necessarily path-connected. For example, if the binary relations in question are partial orderings or equivalence relations (as briefly mentioned above), the agenda is merely minimally connected; but Theorem 2 still yields an immediate corollary for the aggregation of profiles of such binary relations into corresponding collective binary relations: here every aggregation rule satisfying universal domain, collective rationality and systematicity is a (possibly inverse) dictatorship of some individual.

These findings illustrate the generality of judgment aggregation. Impossibility and possibility results such as Theorems 1 and 2 can apply to a large class of aggregation problems formulated in a suitable logic – any logic satisfying conditions L1 to L3 – of which a predicate logic for representing preferences is a special case. Other logics to which the results apply are propositional, modal or conditional logics, some fuzzy logics as well as different predicate logics.

An alternative, very general model of aggregation is the one introduced by Wilson (1975) and used by Dokow and Holzman (2005), where a group has to determine its yes/no views on several issues based on the group members' views on these issues (subject to feasibility constraints). Wilson's model can also be represented in our model; Dokow and Holzman's results for Wilson's model apply to a logic satisfying L1 to L3 and a finite agenda.[16]

Although we have constructed an explicit embedding of preference aggregation into judgment aggregation, we have not proved the impossibility of a converse embedding. We suspect that such an embedding is hard to achieve, as Arrow's standard model cannot easily capture the different informational basis of judgment aggregation. It is unclear what an embedding of judgment aggregation into preference aggregation would look like. In particular, it is unclear how to specify the *options* over which individuals have preferences. The *propositions* in an agenda are not candidates for options, as propositions are usually not mutually exclusive. Natural candidates for options are perhaps entire *judgment sets* (consistent and complete), as these are mutually exclusive and exhaustive. But in a preference aggregation model with options thus defined, individuals would feed into the aggregation rule not a single judgment set (option), but an entire preference ordering over all possible judgment sets (options). This would be a different informational basis from the one in judgment aggregation. In addition, the explicit logical structure within each judgment set would be lost under this approach, as judgment sets in their entirety, not propositions, would be taken as primitives. However, the construction of a useful converse embedding or the proof of its non-existence remains a challenge.

6 References

Bovens L, Rabinowicz W (2006) Democratic Answers to Complex Questions: An Epistemic Perspective. Synthese 150(1): 131-153

Dietrich F (2006) Judgment Aggregation: (Im)Possibility Theorems. Journal of Economic Theory 126(1): 286-298

Dietrich F (forthcoming) A generalized model of judgment aggregation. Social Choice and Welfare

Dietrich F, List C (2005) Judgment aggregation by quota rules. Working paper, LSE

[16] In Wilson's model, the notion of consistency (feasibility) rather than that of entailment is a primitive. While the notion of entailment in our model fully specifies a notion of consistency, the converse does not hold for all logics satisfying L1 to L3.

Dokow E, Holzman R (2005) Aggregation of binary evaluations, Working paper, Technion Israel Institute of Technology

Gärdenfors P (forthcoming) An Arrow-like theorem for voting with logical consequences. Economics and Philosophy

Guilbaud GT (1966) Theories of the General Interest, and the Logical Problem of Aggregation. In Lazarsfeld PF, Henry NW (eds.) Readings in Mathematical Social Science, Cambridge/MA (MIT Press): 262-307

List C (2003) A Possibility Theorem on Aggregation over Multiple Interconnected Propositions. Mathematical Social Sciences 45(1): 1-13

List C (2004) A Model of Path Dependence in Decisions over Multiple Propositions. American Political Science Review 98(3): 495-513

List C (2005) The Probability of Inconsistencies in Complex Collective Decisions. Social Choice and Welfare 24: 3-32

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

List C, Pettit P (2001/2004) Aggregating Sets of Judgments: Two Impossibility Results Compared. Social and Political Theory Paper W20 (technical report ID 931), Australian National University, published in Synthese 140(1-2): 207-235

Nehring K (2003) Arrow's theorem as a corollary. Economics Letters 80: 379-382

Nehring K, Puppe C (2002) Strategyproof Social Choice on Single-Peaked Domains: Possibility, Impossibility and the Space Between. Working paper, University of California at Davis

Nehring K, Puppe C (2005) Consistent Judgment Aggregation: A Characterization. Working paper, University of Karlsruhe

Pauly M, van Hees M (forthcoming) Logical Constraints on Judgment Aggregation. Journal of Philosophical Logic

Pettit P (2001) Deliberative Democracy and the Discursive Dilemma. Philosophical Issues 11: 268-299

Pigozzi G (forthcoming) Belief merging and the discursive dilemma: an argument-based account to paradoxes of judgment aggregation. Synthese

van Hees M (forthcoming) The limits of epistemic democracy. Social Choice and Welfare

Wilson R (1975) On the Theory of Aggregation. Journal of Economic Theory 10: 89-99

A Appendix

Proof of Proposition 1. Let X be minimally connected and let F be any aggregation rule. Put $N := \{1, ..., n\}$. If F is dictatorial, F obviously satisfies universal domain, collective rationality, systematicity and the unanimity principle. Now assume F satisfies the latter conditions. Then there is a set

$\mathcal{C}$ of ("winning") coalitions $C \subseteq N$ such that, for every $p \in X$ and every $(A_1, ..., A_n) \in Domain(F)$, $F(A_1, ..., A_n) = \{p \in X : \{i : p \in A_i\} \in \mathcal{C}\}$. For every consistent set $Z \subseteq X$, let A_Z be some consistent and complete judgment set such that $Z \subseteq A_Z$.

Claim 1. $N \in \mathcal{C}$, and, for every coalition $C \subseteq N$, $C \in \mathcal{C}$ if and only if $N \backslash C \notin \mathcal{C}$.

The first part of the claim follows from the unanimity principle, and the second part follows from collective rationality together with universal domain.

Claim 2. For any coalitions $C, C^* \subseteq N$, if $C \in \mathcal{C}$ and $C \subseteq C^*$ then $C^* \in \mathcal{C}$.

Let $C, C^* \subseteq N$ with $C \in \mathcal{C}$ and $C \subseteq C^*$. Assume for contradiction that $C^* \notin \mathcal{C}$. Then $N \backslash C^* \in \mathcal{C}$. Let Y be as in part (ii) of the definition of minimally connected agendas, and let Z be a *smallest* subset of Y such that $(Y \backslash Z) \cup \{\neg z : z \in Z\}$ is consistent and Z has even size. We have $Z \neq \emptyset$, since otherwise the (inconsistent) set Y would equal the (consistent) set $(Y \backslash Z) \cup \{\neg z : z \in Z\}$. So, as Z has even size, there are two distinct propositions $p, q \in Z$. Since Y is minimal inconsistent, $(Y \backslash \{p\}) \cup \{\neg p\}$ and $(Y \backslash \{q\}) \cup \{\neg q\}$ are each consistent. This and the consistency of $(Y \backslash Z) \cup \{\neg z : z \in Z\}$ allow us to define a profile $(A_1, ..., A_n) \in Domain(F)$ as follows. Putting $C_1 := C^* \backslash C$ and $C_2 := N \backslash C^*$ (note that $\{C, C_1, C_2\}$ is a partition of N), let

$$A_i := \begin{cases} A_{(Y \backslash \{p\}) \cup \{\neg p\}} & \text{if } i \in C \\ A_{(Y \backslash Z) \cup \{\neg z : z \in Z\}} & \text{if } i \in C_1 \\ A_{(Y \backslash \{q\}) \cup \{\neg q\}} & \text{if } i \in C_2. \end{cases} \quad (1)$$

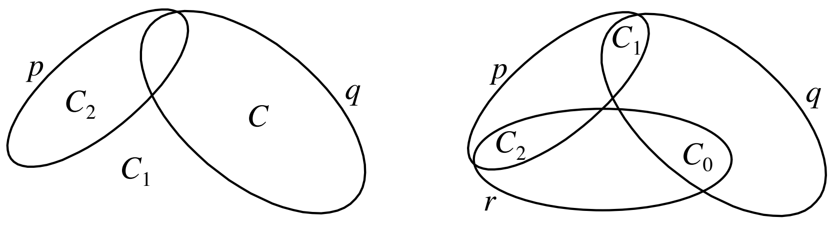

Figure 1: The profiles constructed in the proofs of claims 2 (left) and 3 (right).

By (1), we have $Y \backslash Z \subseteq F(A_1, ..., A_n)$ as $N \in \mathcal{C}$. Also by (1), we have $q \in F(A_1, ..., A_n)$ as $C \in \mathcal{C}$, and $p \in F(A_1, ..., A_n)$ as $C_2 = N \backslash C^* \in \mathcal{C}$. In summary, writing $Z^* := Z \backslash \{p, q\}$, we have (*) $Y \backslash Z^* \subseteq F(A, ..., A_n)$. We distinguish two cases.

Case $C_1 \notin \mathcal{C}$. Then $C \cup C_2 = N \backslash C_1 \in \mathcal{C}$. So $Z^* \subseteq F(A_1, ..., A_n)$ by (1), which together with (*) implies $Y \subseteq F(A_1, ..., A_n)$. But then $F(A_1, ..., A_n)$ is inconsistent, a contradiction.

Case $C_1 \in \mathcal{C}$. So $\{\neg z : z \in Z^*\} \subseteq F(A_1, ..., A_n)$ by (1). This together with (*) implies that $(Y \backslash Z^*) \cup \{\neg z : z \in Z^*\} \subseteq F(A_1, ..., A_n)$. So $(Y \backslash Z^*) \cup \{\neg z :$

$z \in Z^*\}$ is consistent. As Z^* also has even size, the minimality condition in the definition of Z is violated.

Claim 3. For any coalitions $C, C^* \subseteq N$, if $C, C^* \in \mathcal{C}$ then $C \cap C^* \in \mathcal{C}$.

Consider any $C, C^* \in \mathcal{C}$. Let $Y \subseteq X$ be as in part (i) of the definition of minimally connected agendas. As $|Y| \geq 3$, there are pairwise distinct propositions $p, q, r \in Y$. As Y is minimally inconsistent, each of the sets $(Y\backslash\{p\}) \cup \{\neg p\}$, $(Y\backslash\{q\}) \cup \{\neg q\}$ and $(Y\backslash\{r\}) \cup \{\neg r\}$ is consistent. This allows us to defined a profile $(A_1, ..., A_n) \in Domain(F)$ as follows. Putting $C_0 := C \cap C^*$, $C_1 := C^*\backslash C$ and $C_2 := N\backslash C^*$ (note that $\{C_0, C_1, C_2\}$ is a partition of N), let

$$A_i := \begin{cases} A_{(Y\backslash\{p\})\cup\{\neg p\}} & \text{if } i \in C_0 \\ A_{(Y\backslash\{r\})\cup\{\neg r\}} & \text{if } i \in C_1 \\ A_{(Y\backslash\{q\})\cup\{\neg q\}} & \text{if } i \in C_2. \end{cases} \quad (2)$$

By (2), $Y\backslash\{p, q, r\} \subseteq F(A_1, ..., A_n)$ as $N \in \mathcal{C}$. Again by (2), we have $q \in F(A_1, ..., A_n)$ as $C_0 \cup C_1 = C^* \in \mathcal{C}$. As $C \in \mathcal{C}$ and $C \subseteq C_0 \cup C_2$, we have $C_0 \cup C_2 \in \mathcal{C}$ by claim 2. So, by (2), $r \in F(A_1, ..., A_n)$. In summary, $Y\backslash\{p\} \subseteq F(A_1, ..., A_n)$. As Y is inconsistent, $p \notin F(A_1, ..., A_n)$, and hence $\neg p \in F(A_1, ..., A_n)$. So, by (2), $C_0 \in \mathcal{C}$.

Claim 4. There is a dictator.

Consider the intersection of all winning coalitions, $\widetilde{C} := \cap_{C \in \mathcal{C}} C$. By claim 3, $\widetilde{C} \in \mathcal{C}$. So $\widetilde{C} \neq \emptyset$, as by claim 1 $\emptyset \notin \mathcal{C}$. Hence there is a $j \in \widetilde{C}$. As j belongs to every winning coalition $C \in \mathcal{C}$, j is a dictator: indeed, for each profile $(A_1, ..., A_n) \in Domain(F)$ and each $p \in X$, if $p \in A_j$ then $\{i : p \in A_i\} \in \mathcal{C}$, so that $p \in F(A_1, ..., A_n)$; and if $p \notin A_i$ then $\neg p \in A_i$, so that $\{i : \neg p \in A_i\} \in \mathcal{C}$, implying $\neg p \in F(A_1, ..., A_n)$, and hence $p \notin F(A_1, ..., A_n)$. ■

Proof of Theorem 1. Let X be minimally connected, and let F satisfy universal domain, collective rationality and systematicity. If F satisfies the unanimity principle, then, by Proposition 1, F is dictatorial. Now suppose F violates the unanimity principle.

Claim 1. X is symmetrical, i.e., if $A \subseteq X$ is consistent, so is $\{\neg p : p \in A\}$.

Let $A \subseteq X$ be consistent. Then there exists a consistent and complete judgment set B such that $A \subseteq B$. As F violates the unanimity principle (but satisfies systematicity), the set $F(B, ..., B)$ contains no element of B, hence contains no element of A, hence contains all elements of $\{\neg p : p \in A\}$ by collective rationality. So, again by collective rationality, $\{\neg p : p \in A\}$ is consistent.

Claim 2. The aggregation rule $\widehat{F}$ with universal domain defined by $\widehat{F}(A_1, ..., A_n) := \{\neg p : p \in F(A_1, ..., A_n)\}$ is dictatorial.

As F satisfies collective rationality and systematicity, so does $\widehat{F}$, where the consistency of collective judgment sets follows from claim 1. $\widehat{F}$ also satisfies the unanimity principle: for any $p \in X$ and any $(A_1, ..., A_n)$ in the universal domain, where $p \in A_i$ for all i, $p \notin F(A_1, ..., A_n)$, hence $\neg p \in F(A_1, ..., A_n)$, and so $p = \neg\neg p \in \widehat{F}(A_1, ..., A_n)$. Now Proposition 1 applies to $\widehat{F}$, and hence $\widehat{F}$ is dictatorial.

Claim 3. F is inverse dictatorial.
The dictator for $\widehat{F}$ is an inverse dictator for F. ■

Proof of Lemma 1. Let X and F be as specified. To show that F is systematic, consider any $p, q \in X$ and any $(A_1, ..., A_n), (A_1^*, ..., A_n^*) \in Domain(F)$ such that $C := \{i : p \in A_i\} = \{i : q \in A_i^*\}$, and let us prove that $p \in F(A_1, ..., A_n)$ if and only if $q \in F(A_1^*, ..., A_n^*)$. If p and q are both tautologies ($\{\neg p\}$ and $\{\neg q\}$ are inconsistent), the latter holds since (by collective rationality) $p \in F(A_1, ..., A_n)$ and $q \in F(A_1^*, ..., A_n^*)$. If p and q are both contradictions ($\{p\}$ and $\{q\}$ are inconsistent), it holds since (by collective rationality) $p \notin F(A_1, ..., A_n)$ and $q \notin F(A_1^*, ..., A_n^*)$. It is impossible that one of p and q is a tautology and the other a contradiction, because then one of $\{i : p \in A_i\}$ and $\{i : q \in A_i^*\}$ would be N and the other $\emptyset$.

Now consider the remaining case where both p and q are contingent. We say that C is *winning for* r ($\in X$) if $r \in F(B_1, ..., B_n)$ for some (hence by independence any) profile $(B_1, .., B_n) \in Domain(F)$. with $\{i : r \in B_i\} = C$. We have to show that C is winning for p if and only if C is winning for q. Suppose C is winning for p, and let us show that C is winning for q (the converse implication can be shown analogously). As X is path-connected and p and q are contingent, there are $p = p_1, p_2, ..., p_k = q \in X$ such that $p_1 \vDash^* p_2$, $p_2 \vDash^* p_3$, ..., $p_{k-1} \vDash^* p_k$. We show by induction that C is winning for each of $p_1, p_2, ..., p_k$. If $j = 1$ then C is winning for p_1 by $p_1 = p$. Now let $1 \leq j < k$ and assume C is winning for p_j. We show that C is winning for p_{j+1}. By $p_j \vDash^* p_{j+1}$, there is a set $Y \subseteq X$ such that (i) $\{p_j\} \cup Y$ and $\{\neg p_{j+1}\} \cup Y$ are each consistent, and (ii) $\{p_j, \neg p_{j+1}\} \cup Y$ is inconsistent. Using (i) and (ii), the sets $\{p_j, p_{j+1}\} \cup Y$ and $\{\neg p_j, \neg p_{j+1}\} \cup Y$ are each consistent. So there exists a profile $(B_1, ..., B_n) \in Domain(F)$ such that $\{p_j, p_{j+1}\} \cup Y \subseteq B_i$ for all $i \in C$ and $\{\neg p_j, \neg p_{j+1}\} \cup Y \subseteq B_i$ for all $i \notin C$. Since $Y \subseteq A_i$ for all i, $Y \subseteq F(A_1, ..., A_n)$ by the unanimity principle. Since $\{i : p_j \in A_i\} = C$ is winning for p_j, we have $p_j \in F(A_1, ..., A_n)$. So $\{p_j\} \cup Y \subseteq F(A_1, ..., A_n)$. Hence, using collective rationality and (ii), we have $\neg p_{j+1} \notin F(A_1, ..., A_n)$, and so $p_{j+1} \in F(A_1, ..., A_n)$. Hence, as $\{i : p_{j+1} \in A_i\} = C$, C is winning for p_{j+1}. ■

Proof of Lemma 2. Let X be the preference agenda. X is minimally connected, as, for any pairwise distinct constants x, y, z, the set $Y = \{xPy, yPz, zPx\} \subseteq X$ is minimal inconsistent, where $\{\neg xPy, \neg yPz, zPx\}$ is consistent.

To prove path-connectedness, note that, by the axioms of our predicate logic for representing preferences, (*) $\neg xPy$ and yPx are equivalent (i.e., entail each other) for any distinct $x, y \in K$. Now consider any (contingent) $p, q \in X$, and let us construct a sequence $p = p_1, p_2, ..., p_k = q \in X$ with $p \models^* p_2, ..., p_{k-1} \models^* q$. By (*), if p is a negated proposition $\neg xPy$, then p is equivalent to the non-negated proposition yPx ; and similarly for q. So we may assume without loss of generality that p and q are non-negated propositions, say p is xPy and q is $x'Py'$. We distinguish three cases, each with subcases.

Case $x = x'$. If $y = y'$, then $xPy \vDash^* xPy = x'Py'$ (take $Y = \emptyset$). If $y \neq y'$, then $xPy \vDash^* xPy' = x'Py'$ (take $Y = \{yPy'\}$).

Case $x = y'$. If $y = x'$, then, taking any $z \in K \backslash \{x, y\}$, we have $xPy \vDash^* xPz$ (take $Y = \{yPz\}$), $xPz \vDash^* yPz$ (take yPx), and $yPz \vDash^* yPx = x'Py'$ (take $Y = \{zPx\}$). If $y \neq x'$, then $xPy \vDash^* x'Py$ (take $Y = \{x'Px\}$) and $x'Py \vDash^* x'Py'$ (take $Y = \{yPy'\}$).

Case $x \neq x', y'$. If $y = x'$, then $xPy \vDash^* xPy'$ (take $Y = \{yPy'\}$) and $xPy' \vDash^* x'Py'$ (take $Y = \{x'Px\}$). If $y = y'$, then $xPy \vDash^* x'Py = x'Py'$ (take $Y = \{x'Px\}$). If $y \neq x', y'$, then $xPy \vDash^* x'Py'$ (take $Y = \{x'Px, yPy'\}$). ■