Journal of Philosophical Logic (2008) 37:57–65
DOI: 10.1007/s10992-007-9058-y

CHRISTIAN LIST

WHICH WORLDS ARE POSSIBLE? A JUDGMENT AGGREGATION PROBLEM

Received 20 August 2006

ABSTRACT. Suppose the members of a group (e.g., committee, jury, expert panel) each form a judgment on which worlds in a given set are possible, subject to the constraint that at least one world is possible but not all are. The group seeks to aggregate these individual judgments into a collective judgment, subject to the same constraint. I show that no judgment aggregation rule can solve this problem in accordance with three conditions: "unanimity," "independence" and "non-dictatorship," Although the result is a variant of an existing theorem on "group identification" (Kasher and Rubinstein, Logique et Analyse 160:385–395, 1997), the aggregation of judgments on which worlds are possible (or permissible, desirable, etc.) appears not to have been studied yet. The result challenges us to take a stance on which of its conditions to relax.

KEY WORDS: judgment aggregation, possible worlds, propositions

1. Introduction

How can a group of individuals (e.g., committee, jury, expert panel) make collective judgments (true/false) on some propositions based on the group members' individual judgments on these propositions? This task – "judgment aggregation" – becomes non-trivial when the propositions are interconnected, as shown in a growing literature. Suppose a three-member group seeks to make collective judgments on p, q, and $p \wedge q$, where one member judges all three propositions to be true, a second judges p to be true but q and $p \wedge q$ to be false, and a third judges q to be true but p and $p \wedge q$ to be false. Then majorities judge p and q to be true and yet $p \wedge q$ to be false, an inconsistent set of majority judgments. This problem has been called the "discursive dilemma" (Pettit 2001 extending Kornhauser and Sager 1986) and shown to illustrate a more general impossibility result (List and Pettit 2002; 2004). Several extensions and generalizations have been given (e.g., Pauly and van Hees 2006; Dietrich 2006, 2007; Nehring and Puppe 2005; Bovens and Rabinowicz 2006; Pauly 2007; van Hees 2007). Crucially, in the

existing literature on judgment aggregation, propositions are modelled as sentences of a formal language, not as sets of possible worlds.

In this short paper, I want to draw attention to a different judgment aggregation problem not discussed in that literature, which arises when propositions are modelled as sets of possible worlds. Here each group member's judgment consists in the acceptance of a single non-contradictory and non-tautological proposition, represented by a set of worlds deemed possible by that individual. The group then seeks to make a collective judgment, which also consists in the acceptance of a single non-contradictory and non-tautological proposition, represented by a set of worlds deemed possible by the group. Thus the problem is to aggregate an n-tuple of non-contradictory and non-tautological propositions (across n individuals) into a single non-contradictory and non-tautological proposition (for the group as a whole).

The theorem I want to put on the table for discussion shows that it is impossible to solve this problem in accordance with some seemingly mild conditions: "unanimity," "independence," and "non-dictatorship." Mathematically, this result is a variant of an existing theorem on the so-called "group identification" problem by Kasher and Rubinstein (1997), but to the best of my knowledge the interpretation proposed here – i.e., the application to judgments on which worlds are possible – is new. Here I derive the theorem from a recent theorem on judgment aggregation in the standard sense (Dietrich and List 2007a; Dokow and Holzman 2005), which, in turn, has precursors in abstract aggregation theory (Wilson 1975; Rubinstein and Fishburn 1986; Nehring and Puppe 2002).

As the difficulties with judgment aggregation are usually thought to stem from the presence of multiple, interconnected propositions, it is surprising that an impossibility result can arise even when individual and collective judgments consist only in the acceptance of a single proposition. It is also worth noting that the result applies not only to judgments on which worlds are possible, but also to judgments on which worlds are permissible, or desirable, etc. In each of these cases, the theorem poses interesting questions.

2. Theorem

Let $\Omega = \{\omega_1, ..., \omega_k\}$ be a finite set of worlds, assuming $k > 2$; and let $N = \{1, ..., n\}$ be a finite group of individuals, assuming $n > 1$. Each individual $i \in N$ makes a judgment on which of the worlds $\omega \in \Omega$ are possible (e.g., relative to his/her evidence, but various interpretations of

the formalism can be given). Formally, a *judgment* is a subset $J \subseteq \Omega$. We say that J is *consistent* if $J \neq \emptyset$ and *informative* if $J \neq \Omega$. Making a consistent and informative judgment thus consists in accepting a non-contradictory and non-tautological proposition. Let $\mathcal{J}$ be the set of all consistent and informative judgments.

The group requires a method of aggregating each n-tuple of consistent and informative individual judgments $(J_1, ..., J_n) \in \mathcal{J}^n$ into a consistent and informative collective judgment $J \in \mathcal{J}$. Call such a method an *aggregation rule*, defined as a function $f : \mathcal{J}^n \to \mathcal{J}$.

Can we find an aggregation rule that meets some minimal conditions? Let me introduce three conditions. The first requires that if all individuals submit the same judgment, this judgment should also be the collective one. The second requires that the collective judgment on whether a given world ω is possible should depend only on individual judgments on whether ω is possible, not on individual judgments on whether other worlds $\omega' \neq \omega$ are possible. The third requires that the collective judgment should not always be determined by an antecedently fixed dictator.

Unanimity. For all $J \in \mathcal{J}, f(J, ..., J) = J$.
Independence. For all $(J_1, ..., J_n), (J_1', ..., J_n') \in \mathcal{J}^n$ and all $\omega \in \Omega$,

$$[(\forall i \in N)(\omega \in J_i \Leftrightarrow \omega \in J_i')] \Rightarrow [\omega \in f(J_1, ..., J_n) \Leftrightarrow \omega \in f(J_1', ..., J_n')].$$

Non-dictatorship. There is no $i \in N$ such that, for all $(J_1, ..., J_n) \in \mathcal{J}^n$, $f(J_1, ..., J_n) = J_i$.

Surprisingly, these three conditions cannot be satisfied simultaneously.

THEOREM 1. There is no aggregation rule $f : \mathcal{J}^n \to \mathcal{J}$ satisfying unanimity, independence and non-dictatorship.

A proof is given in the Appendix. In the case $n = k$, theorem 1 is equivalent to a theorem by Kasher and Rubinstein (1997) on the problem of "group identification." Here the n members of a group each make a judgment on which members of that group have a certain property (e.g., being a true scientist or having a particular religious identity), subject to the constraint that at least one individual has the property but not all individuals do. The group then seeks to aggregate the n individual judgments on who has the given property into a resulting collective judgment. Here, too, no aggregation rule can simultaneously satisfy unanimity, independence and non-dictatorship.

3. Illustration

While the full proof of theorem 1 is somewhat technical, the intuition behind it can be conveyed by considering a more restrictive class of aggregation rules and showing that this class is empty. Consider an aggregation rule that satisfies not only unanimity, independence and a strengthened version of non-dictatorship, namely "anonymity" (i.e., all individuals have equal weight in determining the collective judgment), but also "monotonicity" (i.e., if more individuals judge that a given world ω is possible, this cannot turn a collective judgment that ω is possible into one that ω is impossible) and "neutrality" (i.e., the criterion for determining the collective judgment on whether a given world ω is possible is the same for all worlds ω). (In standard judgment aggregation, various general characterizations of aggregation rules satisfying these conditions are given in Nehring and Puppe 2005 and Dietrich and List 2007b.) A necessary condition for an aggregation rule to meet all these conditions is that it is a *threshold rule* with some *acceptance threshold* $t \in \{1, 2, ..., n\}$, defined as follows. For all $(J_1, ..., J_n) \in \mathcal{J}^n$, $f(J_1, ..., J_n)$ is the set of all worlds ω judged possible by at least t individuals, formally

$$f(J_1, ..., J_n) = \{\omega \in \Omega : \text{number} \quad \text{of} \quad i \in N \quad \text{with} \quad \omega \in J_i \geq t\}.$$

Examples of threshold rules are the *union rule* $f(J_1, ..., J_n) = J_1 \cup ... \cup J_n$ (here $t = 1$), the *intersection rule* $f(J_1, ..., J_n) = J_1 \cap ... \cap J_n$ (here $t = n$), and *simple majority rule* (here t is the smallest integer greater than $n/2$).

Can we find an acceptance threshold such that the corresponding threshold rule assigns to every n-tuple of consistent and informative individual judgments a consistent and informative collective judgment? It is easy to see that, to ensure consistency, the threshold must not be too high, i.e., it must not happen that no world is deemed possible. A necessary and sufficient condition is $t < \frac{n}{k} + 1$; e.g., the union rule guarantees consistency. To ensure informativeness, the threshold must not be too low, i.e., it must not happen that all worlds are deemed possible. A necessary and sufficient condition is $t > n\frac{k-1}{k}$; e.g., the intersection rule guarantees informativeness. So the conjunction of consistency and informativeness requires the acceptance threshold to satisfy $n\frac{k-1}{k} < t < \frac{n}{k} + 1$. If there are $n > 1$ individuals and $k > 2$ possible worlds, as assumed, this inequality can easily be seen to have no solution. (In the degenerate case of $k = 2$ possible worlds and an odd number of individuals n, simple majority rule meets all conditions.) Therefore the class of aggregation rules satisfying the conditions of theorem 1 together with monotonicity and

neutrality (and non-dictatorship strengthened to anonymity) is empty when $n > 1$ and $k > 2$. The full proof shows that this impossibility continues to hold even without monotonicity, neutrality or the strengthening of non-dictatorship to anonymity.

4. Avoiding the Impossibility

As with other impossibility results of social choice theory – notably, Arrow's (1951) theorem on preference aggregation – the significance of theorem 1 lies not primarily in establishing the impossibility of solving a particular aggregation problem, but rather in indicating which conditions must be relaxed in order to find a solution. The following escape routes are available:

Relaxing unanimity. If unanimity is dropped, a *constant aggregation rule* satisfies all other conditions. Such a rule assigns to every n-tuple $(J_1, ..., J_n) \in \mathcal{J}^n$ the same fixed collective judgment $J \in \mathcal{J}$. This is not an attractive solution, as it pays no attention to the judgments submitted by the individuals.

Relaxing independence. If independence is dropped, a *distance-based aggregation rule* can be constructed (inspired by the approach in Pigozzi 2006), which satisfies all other conditions. For each n-tuple $(J_1, ..., J_n) \in \mathcal{J}^n$, we here compare each "candidate" collective judgment $J \in \mathcal{J}$ with each submitted individual judgment J_i, using the following method. We say that J *agrees with* J_i *on a given world* ω if [$\omega \in J$ if and only if $\omega \in J_i$]. For each individual $i \in N$, we now count the total number of worlds $\omega \in \Omega$ on which J agrees with J_i and consider the sum-total of these counts across individuals $i \in N$. The collective judgment $J \in \mathcal{J}$ is then chosen so as to maximize this sum-total (with some additional provisions for breaking ties). Under such an aggregation rule, the collective judgment on whether a given world is possible depends not only on individual judgments regarding that world but also on individual judgments regarding other worlds. If this implication is accepted, a distance-based aggregation rule may be a satisfactory solution to the present aggregation problem.

Relaxing non-dictatorship. If non-dictatorship is dropped, a *dictatorial aggregation rule* satisfies all other conditions. Such a rule assigns to each n-tuple $(J_1, ..., J_n) \in \mathcal{J}^n$ the judgment J_i of the same antecedently fixed individual $i \in N$. For obvious reasons, this is not generally an attractive solution.

Restricting the domain of admissible inputs. If the aggregation rule is required to work not for every possible n-tuple of individual judgments but only for those that meet some additional constraints, then unanimity, independence and non-dictatorship can be satisfied simultaneously. For example, if only n-tuples $(J_1, ..., J_n)$ satisfying $J_1 \cap ... \cap J_n \neq \emptyset$ are admissible, then the intersection rule as defined above meets all conditions. Likewise, if only n-tuples $(J_1, ..., J_n)$ satisfying $J_1 \cup ... \cup J_n \neq \Omega$ are admissible, then the union rule as defined above meets them. A domain restriction of this kind is feasible in cases where the diversity in individual judgments is limited. If, for example, judgments represent the knowledge (as opposed to mere beliefs) of individuals, then it may seem reasonable to assume that only n-tuples $(J_1, ..., J_n)$ with a non-empty intersection occur, and the intersection rule may seem plausible.

Extending the range of admissible outputs. If the aggregation rule is permitted to generate collective judgments that violate consistency or informativeness, then it is possible to satisfy all other conditions simultaneously. For example, if judgments represent the beliefs of individuals (rather than their knowledge), then perhaps the requirements of consistency or informativeness are too strong at the collective level. As already noted, the union rule guarantees consistency (but not informativeness), and the intersection rule guarantees informativeness (but not consistency), while both rules satisfy unanimity, independence and non-dictatorship. More radically, one might permit collective judgments that take the form of continuous probability assignments across worlds as opposed to binary judgments of possibility or impossibility; but probability aggregation gives rise to some well-known problems of its own.

In conclusion, I have shown that a non-trivial impossibility result of aggregation arises not only for judgments on multiple, interconnected propositions, as in the much-discussed "discursive dilemma," but also for judgments that each consist in the acceptance of only a single proposition, modelled as a set of possible worlds. Whether or not we find this result compelling, it challenges us to take a stance on which of its conditions to relax.

ACKNOWLEDGEMENTS

This work was presented at the Formal Epistemology Workshop 2006, UC Berkeley, 5/2006 and at the 2006 Conference of the Philosophy of Science Association, Vancouver, 11/2006. Much of the writing was done at the Australian National University, 7/2006. I am grateful for the hospitality and discussions I enjoyed in all three places. For comments

and criticism, I thank Richard Bradley, Franz Dietrich, Branden Fitelson, Robert Goodin, Alan Hayek, Marc Pauly, David Schmeidler and Timothy Williamson.

Appendix

I derive theorem 1 from a result in Dietrich and List (2007a), embedding the present aggregation problem into the standard judgment aggregation model (in Dietrich's 2007 "general logics" version), where propositions are modelled as sentences of a formal language. A notationally different, but mathematically equivalent derivation can be given from Dokow and Holzman's (2005) theorem or from Rubinstein and Fishburn's (1986) results, on which Kasher and Rubinstein's (1997) theorem on "group identification" is based. The present proof invokes the concept of "total blockedness" introduced by Nehring and Puppe (2002), whose main theorem is a precursor to the results in Dietrich and List (2007a) and Dokow and Holzman (2005). It is interesting to note that the general result from which theorem 1 follows also has Arrow's (1951) impossibility theorem on preference aggregation as a corollary.

Let $\mathbf{L}$ be a simple sentential language, with atomic sentences a_1, ..., a_k and connectives $\neg$, $\wedge$, $\vee$. To each world $\omega_j \in \Omega$, there corresponds a sentence a_j interpreted to mean that world ω_j is possible. For any $S \subseteq \mathbf{L}$ and any $p \in \mathbf{L}$, write $S \models p$ if and only if the set $S \cup Z \cup \{\neg p\}$ is inconsistent in the standard sense of sentential logic, where $Z = \{a_1 \vee ... \vee a_k, \neg(a_1 \wedge ... \wedge a_k)\}$. Informally, $S \models p$ means that S entails p relative to the constraint that the disjunction of a_1, ..., a_k is true (consistency) and their conjunction false (informativeness). Further, a set $S \subseteq \mathbf{L}$ is called *inconsistent* (in $\mathbf{L}$) if $S \models p$ and $S \models \neg p$ for some $p \in \mathbf{L}$, and *consistent* otherwise. The logic given by the pair $(\mathbf{L}, \models)$ captures the consistency and informativeness constraints on the present aggregation problem and satisfies axioms L1–L3 in Dietrich's (2007) general logics model.

Define the *agenda* of sentences on which judgments (in the standard sense) are to be made as $X = \{a_1, ..., a_k, \neg a_1, ... \neg a_k\}$. A judgment $J \in \mathcal{J}$ as defined above corresponds to a maximal consistent subset $A \subseteq X$, where, for each $\omega_j \in \Omega$, $a_j \in A$ if and only if $\omega_j \in J$ and $\neg a_j \in A$ if and only if $\omega_j \notin J$. An aggregation rule $f : \mathcal{J}^n \to \mathcal{J}$ as defined above corresponds to a judgment aggregation rule F for the agenda X in the standard sense (which, in addition, satisfies universal domain and collective rationality), i.e., a function F that maps each n-tuple of maximal consistent subsets of X (individual judgment sets in the standard sense) to a maximal consistent subset of X (collective judgment set in the standard sense). It is easy to see that F satisfies the standard unanimity and independence conditions if and only if f satisfies unanimity and independence as defined here.

One of the two main theorems in Dietrich and List (2007a) states that, when the agenda X satisfies the richness condition of "strong connectedness" (as defined by claims 1 to 3 below), a judgment aggregation rule F satisfies independence and unanimity (together with universal domain and collective rationality) if and only if it is dictatorial.

In order to see that the present theorem 1 follows from this result, it remains to be shown that the agenda X as defined here is "strongly connected." To show this, the following must be established, where (in)consistency means (in)consistency in $\mathbf{L}$:

Claim 1. There exists a minimal inconsistent subset $Y \subseteq X$ with $|Y| \geq 3$.

Claim 2. There exists 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.

Claim 3. For any $p, q \in X$, there exist $p_1, p_2, ..., p_k \in X$ (with $p = p_1$ and $q = p_k$) such that $p_1 \vDash^* p_2$, $p_2 \vDash^* p_3$, ..., $p_{k-1} \vDash^* p_k$, where $p_j \vDash^* p_h$ is defined to mean that $\{p_j, \neg p_h\} \cup Y$ is inconsistent for some $Y \subseteq X$ consistent with p_j and with $\neg p_h$ (claim 3 states that the agenda has the property of "total blockedness," as defined by Nehring and Puppe, 2002).

To prove claims 1 and 2, notice that, since $k \geq 3$, $Y = \{a_1, ..., a_k\}$ has the properties required by claims 1 and 2, with $Z = \{a_1, a_2\}$.

To prove claim 3, pick any pair of propositions $p, q \in X$. Consider the following exhaustive list of cases:

(a) $p = q$: trivially, $p \vDash^* q$ (with $Y = \emptyset$);
(b) $p = a_j$ and $q = \neg a_h$ with $j \neq h$: here $a_j \vDash^* \neg a_h$ (with $Y = \{a_l : l \neq j, h\}$);
(c) $p = \neg a_j$ and $q = a_h$ with $j \neq h$: here $\neg a_j \vDash^* a_h$ (with $Y = \{\neg a_l : l \neq j, h\}$);
(d) $p = a_j$ and $q = a_h$ with $j \neq h$: here $a_j \vDash^* \neg a_l$ for some $l \neq j, h$ (by b) and $\neg a_l \vDash^* a_h$ (by c);
(e) $p = \neg a_j$ and $q = \neg a_h$ with $j \neq h$: here $\neg a_j \vDash^* a_l$ for some $l \neq j, h$ (by c) and $a_l \vDash^* \neg a_h$ (by b);
(f) $p = a_j$ and $q = \neg a_j$: here $a_j \vDash^* a_h$ for some $h \neq j$ (by d) and $a_h \vDash^* \neg a_j$ (by b);
(g) $p = \neg a_j$ and $q = a_j$: here $\neg a_j \vDash^* a_h$ for some $h \neq j$ (by c) and $a_h \vDash^* a_j$ (by d).

This completes the proof. ■

REFERENCES

Arrow, K.: 1951, *Social Choice and Individual Values*, Wiley, New York.

Bovens, L. and Rabinowicz, W.: 2006, Democratic answers to complex questions: An epistemic perspective, *Synthese* **150**, 131–153.

Dietrich, F.: 2006, Judgment aggregation: (Im)possibility theorems, *Journal of Economics Theory* **126**, 286–298.

Dietrich, F.: 2007, A generalized model of judgment aggregation, *Social Choice and Welfare* **28**, 529–565.

Dietrich, F. and List, C.: 2007a, Arrow's theorem in judgment aggregation, *Social Choice and Welfare* **29**, 19–33.

Dietrich, F. and List, C.: 2007b, Majority voting generalized: Judgment aggregation by quota rules, *Journal of Theoretical Politics* **19**, in press.

Dokow, E. and Holzman, R.: 2005, *Aggregation of Binary Evaluations*, Working paper, Technion Israel Institute of Technology.

Kasher, A. and Rubinstein, A.: 1997, On the question "Who is a J?": A social choice approach, *Logique et Analyse* **160**, 385–395.

Kornhauser, L. A. and Sager, L. G.: 1986, Unpacking the court, *Yale Law Journal* **96**, 82–117.

List, C. and Pettit, P.: 2002, Aggregating sets of judgments: An impossibility result, *Economics and Philosophy* **18**, 89–110.

List, C. and Pettit, P.: 2004, Aggregating sets of judgments: Two impossibility results compared, *Synthese* **140**, 207–235.

Nehring, K. and Puppe, C.: 2002, *Strategy-proof Social Choice on Single-peaked Domains: Possibility, Impossibility and the Space Between*, Working paper, University of California at Davis.

Nehring, K. and Puppe, C.: 2005, *Consistent Judgement Aggregation: A Characterization*, Working paper, University of Karlsruhe.
Pauly, M.: 2007, Axiomatizing collective judgment sets in a minimal logical language, *Synthese* in press.
Pauly, M. and van Hees, M.: 2006, Logical constraints on judgment aggregation, *Journal of Philosophical Logic* **35**, 569–585.
Pettit, P.: 2001, Deliberative democracy and the discursive dilemma, *Philosophical Issues* **11**, 268–299.
Pigozzi, G.: 2006, Belief merging and the discursive dilemma: An argument-based account to paradoxes of judgment aggregation, *Synthese* **152**, 285–298.
Rubinstein, A. and Fishburn, P.: 1986, Algebraic aggregation theory, *Journal of Economic Theory* **38**, 63–77.
van Hees, M.: 2007, The limits of epistemic democracy, *Social Choice and Welfare* **28**, 649–666.
Wilson, R.: 1975, On the theory of aggregation, *Journal of Economic Theory* **10**, 89–99.

Department of Government,
London School of Economics,
London WC2A 2AE,UK
E-mail: C.List@lse.ac.uk