The aggregation of propositional attitudes: towards a general theory

Franz Dietrich   Christian List

October 2008, revised June 2009[1]
forthcoming in *Oxford Studies in Epistemology*

Abstract

How can the propositional attitudes of several individuals be aggregated into overall collective propositional attitudes? Although there are large bodies of work on the aggregation of various special kinds of propositional attitudes, such as preferences, judgments, probabilities and utilities, the aggregation of propositional attitudes is seldom studied in full generality. In this paper, we seek to contribute to filling this gap in the literature. We sketch the ingredients of a general theory of propositional attitude aggregation and prove two new theorems. Our first theorem simultaneously characterizes some prominent aggregation rules in the cases of probability, judgment and preference aggregation, including linear opinion pooling and Arrovian dictatorships. Our second theorem abstracts even further from the specific kinds of attitudes in question and describes the properties of a large class of aggregation rules applicable to a variety of belief-like attitudes. Our approach integrates some previously disconnected areas of investigation.

1   Introduction

On the Humean picture of rational agency, an agent is a system that has beliefs and desires about the world and acts in pursuit of its desires in accordance with its beliefs.[2] Beliefs and desires can be modelled as attitudes

---

[2] Our account of agency as well as our broad picture of the nature of propositional attitudes follows List and Pettit (forthcoming). See also Dennett (1987) and Pettit (1993).

that the agent holds towards certain propositions. In the simplest case, an agent believes a proposition if it represents the world as being such that the proposition is true; an agent desires a proposition if it would like the world to be such that the proposition is true. Many different theoretical models of propositional attitudes have been developed. In the standard theory of rational choice in economics, for example, beliefs are represented as probabilities, desires as utilities.[3] Alternatively, to give just one other example, beliefs can be represented as binary (acceptance/rejection) judgments, and desires as binary (acceptance/rejection) goals or preferences.

In many contexts, individual human beings are not the only agents holding propositional attitudes. Many collectives also need to form such attitudes, for instance when they are faced with certain joint decisions, or when they are required to speak with a single voice.[4] Legislatures, collegial courts, juries, multi-member governments, executive boards of companies, expert panels and groups of scientific collaborators are just a few examples of collectives that are commonly required to form propositional attitudes. They have to form beliefs on various facts about the world, as well as preferences on what actions to take. Even groups as large and diverse as the entire 'demos' within a democracy are sometimes said to hold beliefs and desires.

How can a group come to hold such attitudes? More specifically, how can the propositional attitudes of multiple individuals be combined into a single set of collective such attitudes? The aim of this paper is to investigate this problem of *attitude aggregation*. Notably, the need to aggregate propositional attitudes arises not only in the *interpersonal* contexts of group decision-making and collective agency just mentioned. It also arises in sev-

---

[3] Probability and utility functions are typically cardinal, but preferences, which are also commonly invoked in economics, are often represented ordinally.

[4] The most famous variant of this problem is Arrow's (1951/1963) problem of *social choice*, in which the preferences of multiple individuals are to be aggregated into an overall collective preference. The study of *preference aggregation* in Arrow's tradition has recently been complemented by the study of *judgment aggregation* (e.g., List and Pettit 2002, Dietrich 2007a and references below). Both problems can be traced back to Condorcet (1785). An account of *group agency* is developed in List and Pettit (forthcoming). There is also a relevant literature on *joint intentions*, e.g., Bratman (2007), Gilbert (1989) and Tuomela (2007).

eral *intrapersonal* contexts. When a single agent decides what opinion to form in light of the opinions of several others, he or she is effectively faced with a problem of attitude aggregation.[5] Furthermore, individual agents are sometimes conceptualized as composite systems whose overall attitudes result from the aggregation of the attitudes of its component 'selves'; this parallel between intrapersonal and interpersonal attitude aggregation is also implicit in Plato's famous analogy between the city and the soul.[6] Finally, an individual agent revising his or her beliefs in the light of new evidence can also be seen as merging his or her original attitudes with the attitudes warranted by the new evidence.[7] Thus attitude aggregation problems arise in contexts of social choice, individual deliberation based on the testimony of others, multiple selves, and belief revision.

While the aggregation of many particular kinds of propositional attitudes has already been investigated in great detail, the problem of attitude aggregation is seldom addressed in full generality. There are large literatures on the aggregation of preferences, the aggregation of judgments, the aggregation of probabilities, and the aggregation of utilities,[8] but these and other fields of aggregation theory are surprisingly disconnected from each other, and we still lack a unified theoretical framework that subsumes them all. Moreover, there are some kinds of propositional attitudes – for instance, those represented by ranking functions over propositions[9] – whose aggre-

---

[5] E.g., Lehrer and Wagner (1981), Pettit (2006), List (2008). We will now refer to an individual agent as *he* or *she* rather than *it*.

[6] E.g., Elster (1987), Minsky (1988), Pettit (2003), Plato (360 BC).

[7] E.g., van Benthem (2008).

[8] The literature on preference aggregation follows Arrow (1951/1963). On judgment aggregation, see, e.g., List and Pettit (2002), Pauly and van Hees (2006), Dietrich (2007a), Dietrich and List (2007a), Nehring and Puppe (2008) and surveys in List and Puppe (2009), List (2009). On probability aggregation, see, e.g., Lehrer and Wagner (1981), McConway (1981), Genest and Zidek (1986), Mongin (1995), Dietrich and List (2007b), Dietrich (2008). Ordinal probability aggregation has been studied by Weymark (1997). On utility aggregation, see, e.g., Sen (1982), d'Aspremont and Gevers (2002). List and Pettit (2004) and Dietrich and List (2007a) discuss the relationship between judgment aggregation and preference aggregation. In a conference talk, Mongin (2005) offered a comparison of judgment aggregation and probability aggregation.

[9] For an overview, see Spohn (forthcoming).

gation has never been studied. We are certainly unable to provide a fully general theory of propositional attitude aggregation in this paper, but we hope to sketch some ingredients of such a theory.

This paper has three main sections. In section 2, we informally introduce the problem of attitude aggregation and explain its generality. In section 3, we present a new theorem which simultaneously characterizes a class of aggregation rules in the cases of probability, judgment and preference aggregation. And in section 4, we formulate the ingredients of the problem of attitude aggregation more generally, describing how we can model propositional attitudes in a way that abstracts from the specific kinds of attitudes in question and presenting a more general theorem. The paper ends with some brief concluding remarks.[10]

2 The problem

As noted in the introduction, propositional attitudes are key concepts in any theory of rational agency. On the standard Humean picture of rational agency, an agent is characterized by having both beliefs and desires, which play certain roles in determining the agent's actions. On some alternative, less established pictures, an agent's attitudes fall into only one category whose functional role in determining action subsumes that of both beliefs and desires on the Humean picture.[11] There are a variety of theories of rational agency, which differ in the kinds of propositional attitudes they ascribe to agents and in their account of the nature of those attitudes.

We can characterize different kinds of propositional attitudes along at least two dimensions:[12] first, their functional role within the agent's rational

---

[10] A small disclaimer may be useful at the outset. Our terminology in this paper is based on the terminology in social choice theory and in other related branches of aggregation theory. From the perspective of epistemologists, some of our terminology will therefore seem somewhat non-standard: e.g., we use the term 'judgment' to denote a binary belief-like attitude and understand the term 'rational' in a thin, formal sense.

[11] For a discussion of some anti-Humean alternatives, see Lewis (1988, 1996), Bradley and List (forthcoming).

[12] In presenting this taxonomy, we follow closely the account of propositional attitudes developed in List and Pettit (forthcoming).

performance, e.g., whether they are belief-like or desire-like; and second, their formal structure, e.g., whether they are binary or admit of degrees. Propositions can also be represented in more than one way. On a syntactic approach, they are represented by sentences in a suitable language; on a semantic one, by sets of possible worlds. But since there exists a partial isomorphism between these two approaches, we here set this distinction aside, focusing instead on the variation in the role and structure of propositional attitudes.

An agent's attitude towards a proposition plays the role of a *belief* if it is part of the agent's cognitive representation of what the world is like; it plays the role of a *desire* if it is part of the agent's motivational state as to what the agent would want the world to be like. A belief-like attitude towards a proposition captures the degree to which the agent cognitively represents that proposition as being true or false; a desire-like attitude towards a proposition captures the degree to which the agent is emotively disposed in favour of or against its truth.[13]

An agent's attitude towards a proposition is structurally *binary* if the attitude is either on or off, i.e., the agent either takes an affirmative stance towards the proposition or a non-affirmative one. In the recent literature on judgment aggregation, binary belief-like attitudes are called *judgments*.[14] According to this terminology, an agent either does or does not judge a given proposition to be true. There is no such thing as judging to a greater or lesser extent that the proposition is true. Similarly, binary desire-like attitudes can be described as *goals* or, in a non-standard use of language, *categorical preferences* – not to be confused with *(ordinal) preferences*, as discussed below. On this picture, an agent either does or does not prefer a particular proposition to be true; the truth of that proposition either is or is not a goal for the agent.

On the other hand, an agent's attitude towards a proposition is *non-binary* if it admits of degrees, i.e., it can take a number of values, pos-

---

[13]As has become common, we distinguish between belief-like and desire-like attitudes in terms of the different 'directions of fit' between the attitudes and the world.

[14]This terminology was introduced in List and Pettit (2002).

sibly in between the extremes of full affirmation and full non-affirmation (and maybe there is not even a theoretical maximum or minimum to the strength of attitude). Non-binary propositional attitudes can take several forms. In the standard theory of rational choice, as mentioned above, non-binary beliefs are typically represented by subjective probability functions, and non-binary desires by utility functions. Generally, non-binary propositional attitudes can be represented by suitable real-valued functions whose properties capture the precise formal structure of the attitudes. If the functions representing the attitudes are unique only up to positive monotonic transformations – i.e., 'stretching', 'squeezing' and 'shifting' – then we say that the attitudes are *ordinal*; below we introduce an alternative relation-theoretic representation of ordinal attitudes. By contrast, if the numerical values of these functions have some significance over and above the ordering they induce over the propositions, then we say that the attitudes are *cardinal*.

Table 1 summarizes this simple taxonomy of different kinds of propositional attitudes.[15]

| structure | binary | non-binary | |
role		ordinal	cardinal
belief (cognitive)	judgments	ordinal probabilities	subjective probabilities
desire (emotive)	goals; categorical preferences	ordinal preferences	utilities

Table 1: Different kinds of propositional attitudes

For an agent to count as rational, the agent's propositional attitudes are usually required to satisfy certain conditions, which depend on the kinds of attitudes in question. For example, in the case of binary beliefs (i.e., judgments), full rationality requires truth-functional coherence. So judging that p, q and $p \wedge q$, which corresponds to a well-defined truth-value assignment, is rational, while judging that p, q and $\neg(p \wedge q)$, which is incompatible with any

---

[15] As noted, this draws on the taxonomy developed in List and Pettit (forthcoming).

such assignment, is not. In the case of cardinal non-binary beliefs (i.e., subjective probabilities), full rationality requires probabilistic coherence. For instance, assigning probability 0.5 to p and 0.7 to $p \vee q$ is probabilistically coherent and thus rational, while assigning probability 0.5 to p and 0.3 to $p \vee q$ is not. Full rationality in the case of ordinal non-binary desires (i.e., preferences), to give a final example, requires representability of those desires by an ordering over the relevant propositions. Thus a preference for p to q to r is rational (with p, q and r mutually exclusive), while a preference for p to q, for q to r and for r to p, which violates transitivity, is not. These rationality conditions are just illustrative; we give more formal definitions in subsequent sections. Any theory of rational agency will have to specify not only what kinds of propositional attitudes are ascribed to an agent, but also what the appropriate rationality conditions are.

We are now in a position to formulate the problem of attitude aggregation: How can a combination – called *profile* – of propositional attitudes across multiple individuals be aggregated into resulting 'collective' propositional attitudes? In particular, how can this aggregation be performed in such a way as to ensure the rationality of those collective attitudes? The answer to this question depends crucially on the kinds of attitudes and rationality conditions in question, and on how complex the set of propositions is over which the attitudes are held.

Notoriously, the aggregation of judgments or preferences runs into difficulties when rationality is understood as in our earlier examples and the attitudes are held over a set of propositions that is non-trivial in certain ways. Arrow's classic impossibility theorem, for instance, shows that when there are three or more distinct objects of preference, only dictatorial aggregation rules, where one fixed individual always determines the collective attitudes, can aggregate rational individual preferences into rational collective ones in accordance with some minimal conditions.[16] The recent impossibility theorems on judgment aggregation establish similar results for the aggregation of judgments.[17]

---

[16] See Arrow (1951/1963).

[17] Compare the earlier references.

The picture is more positive in the case of cardinal non-binary attitudes, whether belief-like or desire-like. There are some attractive aggregation rules based on the idea of *averaging*, which allow the aggregation of rational individual probability assignments into rational collective ones in accordance with some plausible conditions, and similarly the aggregation of rational individual utility assignments into rational collective ones. In the case of probability assignments – the belief case – those aggregation rules are the so-called *linear pooling rules*, as defined formally in the next section. In the case of utility assignments – the desire case – they are the *utilitarian rules* in the conventional sense.[18]

However, can we go beyond a case-by-case analysis of the aggregation of particular kinds of propositional attitudes and say something more general about the aggregation of propositional attitudes which applies to more than one box within Table 1? Our aim in the next section is to provide a first unified result about the aggregation of (i) subjective probabilities (i.e., cardinal non-binary beliefs), (ii) judgments (i.e., binary beliefs), and (iii) preferences (i.e., ordinal non-binary desires).

3   The first theorem: linear and dictatorial rules for probability, judgment and preference aggregation

Let X be a set of propositions, represented in some logic or language $\mathbf{L}$, on which attitudes are to be aggregated from the individual to the collective level.[19] We assume that X is finite, closed under negation and contains at

---

[18]The latter require not only the cardinality of attitudes but also their interpersonal comparability, but we set this issue aside in the present discussion.

[19]Formally, $\mathbf{L}$ is set of *all* propositions in the logic or language. Propositions in $\mathbf{L}$ can be represented semantically or syntactically. In the semantic case, $\mathbf{L}$ consists of subsets of some set Ω of *possible worlds*, where $\mathbf{L}$ is an algebra, i.e., (i) if $p, q \in \mathbf{L}$, then $p \cap q \in \mathbf{L}$, (ii) if $p \in \mathbf{L}$, then $\Omega \backslash p \in \mathbf{L}$, and (iii) $\emptyset \in \mathbf{L}$. To capture logical interconnections, we define a subset Y of $\mathbf{L}$ to be *inconsistent* if $\cap_{p \in Y} p = \emptyset$. In the syntactic case, $\mathbf{L}$ is the set of sentences of some language which can express *at least* negation and conjunction, i.e., if $p, q \in \mathbf{L}$, then $\neg p, p \wedge q \in \mathbf{L}$. Logical interconnections in $\mathbf{L}$ are captured by a notion

least one contingent proposition.[20] The precise content of X depends on the aggregation problem in question. In one context, an expert panel may need to form a collective attitude on the proposition 'Annual CO_2 emissions are above 30000 million metric tons of carbon' but not on the proposition 'If CO_2 emissions are above this threshold, then the Greenland ice shield will melt', while in another context the panel may need to form collective attitudes on both of these propositions as well as their implication 'The Greenland ice shield will melt'. In the latter case, the set X could be expressed as containing propositions p, $p \rightarrow q$, q and their negations. It is perfectly imaginable that individuals hold attitudes also on some propositions outside the set X, but these attitudes are not taken into account in the aggregation.

For the purposes of our first theorem, we assume further that the propositions in X exhibit some *non-trivial logical interconnections*. By this, we mean that X has at least one minimal inconsistent subset of three or more propositions, i.e., an inconsistent subset of that size all of whose proper subsets are consistent. Practically any interesting set of propositions has these properties. For instance, if the set X contains propositions p, $p \rightarrow q$, q and their negations, as in the expert-panel example, then the assumptions are all met. In particular, the required minimal inconsistent subset is $\{p, p \rightarrow q, \neg q\}$.

We assume that there are $n \geq 2$ individuals, labelled $1, 2, ..., n$. Individual i's attitudes on the propositions in X are represented by an *attitude*

---

of *consistency* satisfying some regularity conditions (Dietrich 2007a): any pair $\{p, \neg p\}$ is inconsistent; supersets of inconsistent sets are inconsistent; the empty set $\emptyset$ is consistent and every consistent set of propositions S has a consistent superset T containing a member of each pair $p, \neg p$ in $\mathbf{L}$; and any conjunction $p \wedge q$ is logically equivalent with the pair $\{p, q\}$. A set of propositions S *entails* a proposition q if and only if $S \cup \{\neg q\}$ is inconsistent. We use the syntactic notation here, writing '$p \wedge q$', '$\neg p$', 'tautology', and 'contradiction' rather than '$p \cap q$', '$\Omega \backslash p$', 'Ω' and '$\emptyset$', respectively. Other connectives are defined in the usual way: e.g., $p \vee q$ stands for $\neg(\neg p \wedge \neg q)$, and $p \rightarrow q$ for $\neg(p \wedge \neg q)$.

[20]A proposition is *contingent* if it is neither a tautology nor a contradiction. *Closure under negation* means that if some proposition p is in X, then so is its negation $\neg p$ (more precisely, a proposition logically equivalent to $\neg p$, so as to prevent X from having to contain all of $\neg p, \neg\neg p, \neg\neg\neg p$ etc.). The set X is often called the *agenda*, and our non-triviality condition is an example of an *agenda condition* (often called *non-simplicity*).

function, denoted A_i, which assigns to each proposition p in X a value $A_i(p)$ in a set of possible values V. For the moment, we focus on two kinds of attitudes: subjective probabilities (cardinal non-binary beliefs) and judgments (binary beliefs). Below we also consider preferences (ordinal non-binary desires), which we will model differently. In the case of subjective probabilities, the value set V is the interval $[0, 1]$, and an attitude function is *rational* if it is probabilistically coherent, i.e., extendable to a well-defined probability function on the entire logic or language.[21] In the case of judgments, the value set V is the set $\{0, 1\}$, and an attitude function is *rational* if it is truth-functionally coherent, i.e., extendable to a well-defined truth-function on the entire logic or language. Note that truth-functions are special cases of probability functions restricted to the extremal values 0 and 1.

As already said, an n-tuple of attitude functions across individuals is called a *profile*. We are looking for an *aggregation rule*, denoted F, which assigns to each admissible profile $(A_1, A_2, ..., A_n)$ a collective attitude function $A = F(A_1, A_2, ..., A_n)$. Let us introduce four conditions that such an aggregation rule might be expected to satisfy:

Universal domain. F accepts as admissible inputs all possible profiles of rational individual attitude functions.

Collective rationality. F generates as its outputs rational collective attitude functions.

Independence. F generates the collective attitude on each proposition p in the set X as a function of individual attitudes on p.

To state the final condition, we say that an attitude function is *compatible* with the implication $p \rightarrow q$ if the attitude function can be extended to a probability or truth-function such that the value 1 is assigned to $p \rightarrow q$.

---

[21]Formally, A_i is *rational* if there exists a function $Pr : \mathbf{L} \rightarrow \mathbf{R}$ such that, for every proposition p in X, $Pr(p) = A_i(p)$, and Pr satisfies the axioms of probability theory, i.e., (i) $Pr(p) \geq 0$ for every $p \in \mathbf{L}$; (ii) $Pr(p) = 1$ for every tautology $p \in \mathbf{L}$; and (iii) $Pr(p \vee q) = Pr(p) + Pr(q)$ whenever p and q are mutually inconsistent.

Implication preservation. For all propositions p and q in the set X and all admissible profiles, if all individuals' attitude functions are compatible with the implication $p \rightarrow q$, then so is the collective attitude function generated by F.

While the requirements of universal domain and collective rationality are intuitively plausible, the main defence of independence is the democratic idea that the collective attitude on each proposition should be determined by the individuals' attitudes on it. This reflects a *local* notion of democratic aggregation, which underlies, for example, majority voting and other systems of direct democracy. Under a more *holistic* notion, by contrast, the collective attitude on a proposition p might also be influenced by individual attitudes on other propositions deemed relevant to p, e.g., on 'premises' or 'reasons' for p.[22] In many contexts, independence also turns out to be necessary for avoiding various forms of manipulability of an aggregation rule.[23]

The idea underlying implication-preservation is also intuitive: if all individuals' attitudes are compatible with some conditional proposition, for example that if there is an increase in CO_2 emissions, there will also be a temperature increase, then the collective attitudes should not be incompatible with it. Since the conditional $p \rightarrow q$ need not be among the propositions in X (even when p and q are in X), the statement of implication preservation is formulated not in terms of the assignment of an attitude of 1 to $p \rightarrow q$, but in terms of *compatibility* with this assignment.

Obviously, the precise meaning of the conditions – particularly universal domain and collective rationality – depends on whether the value set of the attitude functions is the interval $[0, 1]$ or just the set $\{0, 1\}$. In either case, however, the four conditions characterize a very particular class of aggregation rules. To state a general result, which applies to both subjective probabilities and judgments, call an aggregation rule *linear* if there exist non-negative weights $w_1, w_2, ..., w_n$ summing to 1 such that, for every profile $(A_1, A_2, ..., A_n)$ of rational individual attitude functions and every proposi-

---

[22] On holistic aggregation, see, e.g., List and Pettit (2006), Dietrich (2007b).

[23] On 'agenda' and 'strategic' manipulability, see Dietrich (2006) and Dietrich and List (2007c), respectively.

tion p in the set X, the collective attitude on p is the weighted average of the individual attitudes on it, i.e.,

$$A(p) = w_1 A_1(p) + w_2 A_2(p) + ... + w_n A_n(p).$$

Theorem 1. *An aggregation rule satisfies universal domain, collective rationality, independence and implication preservation if and only if it is linear.*

This theorem has radically different implications in the non-binary and binary cases. In the non-binary case of subjective probabilities, it is easy to see that *every* combination of weights $w_1, w_2, ..., w_n$ gives rise to a linear aggregation rule, which satisfies the theorem's conditions. The weighted average of several coherent probability assignments is still a coherent probability assignment.

Implication 1. *The probability aggregation rules satisfying the four conditions are precisely the linear rules defined for all possible combinations of weights. In particular, any combination of weights $w_1, w_2, ..., w_n$ defines such a rule.*[24]

In the binary case of judgments, by contrast, it is no longer true that *every* combination of weights $w_1, w_2, ..., w_n$ gives rise to a well-defined linear aggregation rule. In particular, unless all weight is assigned to a single individual, the average of the individual attitudes is not generally restricted to the extremal values 0 and 1, even in the presence of this restriction at the individual level. In other words, the well-defined linear rules in the binary case are all of a special form: they are *dictatorial*, i.e., one individual i has weight 1 while all others have weight 0.

Implication 2. *The judgment aggregation rules satisfying the four conditions are precisely the dictatorial rules, which are the only well-defined linear rules in this case.*

---

[24]This implication generalizes a classic result by McConway (1981); see also Lehrer and Wagner (1981). For a more technical account, see Dietrich and List (2007b).

Theorem 1 and implications 1 and 2 are of particular interest in relation to the recent literatures on probabilistic opinion pooling and judgment aggregation, respectively. Linear rules have gained particular prominence in recent debates on how to reconcile disagreements between epistemic peers.[25] The main rules proposed in this context, namely *splitting the difference*, *sticking to one's own opinion* (applicable only under the individual deliberation interpretation of attitude aggregation) and *deferring to another individual's opinion*, are all instances of linear rules. On the other hand, the recent literature on judgment aggregation is replete with characterizations of dictatorial rules.[26] It is therefore illuminating to see that both linear probability aggregation rules and dictatorial judgment aggregation rules can be characterized using the same conditions.

Can we extend this unified characterization further? In particular, does theorem 1 have any implication for the aggregation of preferences as well? Our route towards such an implication is less direct than our routes in the previous cases. While it is also possible to represent preferences by attitude functions, which are unique only up to positive monotonic transformations, a more elegant representation employs binary relations. Suppose that an agent holds preferences on some set of objects K. Since the focus of the present paper is on propositional attitudes, we assume that the objects of preferences are also propositions, but the set K could in principle be any set of objects: candidates, policy options, states of affairs etc. We assume that K has at least three, but finitely many, elements. We represent an agent's preferences on K – for simplicity, we assume they are strict – by a binary relation $\succ$ on K, where $p \succ q$ is interpreted to mean that the agent prefers p to q. The preference relation $\succ$ is rational just in case it satisfies the axioms of strict preferences.[27]

Let us now construct a special set of propositions X consisting of all binary ranking propositions of the form '$p \succ q$', where p and q are distinct elements of K. We call a set of binary ranking propositions *consistent* if

---

[25] E.g., Elga (2007). Earlier contributions include Lehrer and Wagner (1981).

[26] See the earlier references.

[27] For all p, q, r in K, (i) if $p \succ q$ then not $q \succ p$ (*asymmetry*), (ii) if $p \succ q$ and $q \succ r$, then $p \succ r$ (*transitivity*), and (iii) if $p \neq q$, then either $p \succ q$ or $q \succ p$ (*connectedness*).

it respects the axioms of strict preferences. For example, the set $\{$'$p \succ q$', '$q \succ r$'$\}$ (where p, q and r are distinct elements of K) is consistent, while the set $\{$'$p \succ q$', '$q \succ r$', '$r \succ p$'$\}$ is not. A preference relation $\succ$ on K can then be reexpressed as a binary attitude function A on X, where $A($'$p \succ q$'$) = 1$ if $p \succ q$, and $A($'$p \succ q$'$) = 0$ otherwise. The preference relation $\succ$ is rational if and only if the attitude function A is truth-functionally coherent relative to the axioms on strict preferences.

To apply theorem 1, we begin by observing that, since K has three or more elements, the set X we have constructed has a minimal inconsistent subset of three or more propositions and thereby meets the assumption required for the application of theorem 1. For instance, if p, q and r are three distinct elements of K, then a minimal inconsistent subset of X of size 3 is the set $\{$'$p \succ q$', '$q \succ r$', '$r \succ p$'$\}$. Moreover, the four conditions of theorem 1 can be restated as conditions on preference aggregation rules. Universal domain and collective rationality correspond to their namesakes in Arrow's classic theorem on preference aggregation. Independence corresponds to the *independence of irrelevant alternatives*, whereby the collective preference over any pair of propositions p and q in K depends only on individual preferences over this pair. Implication preservation, finally, strengthens the *Pareto principle*, according to which any unanimous individual preference for p over q should be preserved collectively.[28] We obtain the following result as an immediate implication of theorem 1, via implication 2:

Implication 3. *The preference aggregation rules satisfying the four conditions are precisely the dictatorial rules.*

Setting aside the strengthening of the Pareto principle, this result lies in the close vicinity of Arrow's theorem.[29] Theorem 1 thus provides a unified result by which we can characterize not only linear probability aggregation

---

[28] To see why the Pareto principle follows from implication preservation, note that a preference for p over q is equivalent to an attitude of 1 on '$p \succ q$', and hence to a value of 1 given to the implication '$\neg(p \succ q) \rightarrow (p \succ q)$' by any truth-function extending the attitude function.

[29] For an exact match of Arrow's theorem, see Dietrich and List (2007a).

and dictatorial judgment aggregation, but also Arrovian dictatorial preference aggregation using the same conditions.

Before generalizing this picture further, it is worth noting that our initial assumption for the applicability of theorem 1 – namely that X has a minimal inconsistent subset of three or more propositions – is logically tight: it is not only sufficient for the theorem but also necessary.[30] Whenever X violates this assumption – i.e., all its minimal inconsistent subsets are of size at most two – there do exist aggregation rules that satisfy all of the theorem's conditions and yet are not linear or dictatorial. In the probabilistic case, those rules are somewhat complicated to construct, but in the case of judgment and preference aggregation, they include majority rule.[31] Thus our theorem offers a dual characterization: first, it characterizes the class of aggregation rules satisfying the theorem's conditions; and secondly, it characterizes the sets of propositions X for which this characterization of aggregation rules holds.

4   The second theorem: systematic rules more generally

In the previous section, we have offered a first unified result on three seemingly disparate aggregation problems: probability aggregation, judgment aggregation and preference aggregation. Can we obtain an even more general result, which abstracts further from the specific kinds of attitudes in question? In this section, we present a first result which applies to a very general class of belief-like attitudes, including not only judgments and subjective probabilities, but also non-binary truth-value assignments, fuzzy and vague truth-value assignments and ranking functions.

As before, X is the set of propositions on which attitudes are to be

---

[30]Assuming $n \geq 3$.

[31]For an explicit construction in the probabilistic case, see Dietrich and List (2007b). In the binary case, majority voting has the required properties when n is odd, i.e., there cannot be majority ties. For general group size, one can use majority voting among any non-singleton subgroup with odd size.

held, and an attitude function A_i assigns to each proposition p in X a value $A_i(p)$ in some set of possible values V. Again, the rationality conditions on attitude functions depend on the kinds of attitudes in question. While in the previous section we defined the rationality of an attitude function in terms of its extendability to a probability or truth-function, depending on whether the attitudes are subjective probabilities or judgments, we now define rationality more abstractly in terms of extendability to a so-called *valuation function*. This is a function, denoted f, which assigns to each proposition p in the logic or language **L** a value $f(p)$ in the value set V, subject to some relevant constraints. The precise constraints can take a number of different forms – the constraints on probability or truth-functions being special cases. But minimally, it must be possible to arrange the values within V in some complete or partial order of strength, denoted $\geq$ and bounded below by some minimal element v_{min}, such that the admissible valuation functions respect that order of strength:[32]

(i) For any admissible valuation function f and any propositions p and q in **L**, if p logically entails q, the value of q must be at least as great as that of p, i.e., $f(q) \geq f(p)$.

(ii) For any admissible valuation function f and any propositions p and q in **L**, if q has the minimal value, i.e., $f(q) = v_{min}$, the disjunction $p \vee q$ has the same value as p, i.e., $f(p \vee q) = f(p)$.

(iii) For any value v in V, any non-tautological proposition p in **L** and any proposition q in **L** that entails p but is not equivalent to p, there exists an admissible valuation function f such that $f(p) = v$ and $f(q) = v_{min}$.[33]

---

[32] A *partial ordering* $\geq$ on V is a reflexive, transitive and anti-symmetric binary relation. A *minimal element* with respect to $\geq$ is an element $v_{min} \in V$ such that, for all $v \in V$, $v \geq v_{min}$.

[33] The notion of 'admissible' valuation functions can be formally captured by the set of all such functions; call it $\mathcal{F}$. Conditions (i) to (iii) contain minimal requirements on the functions contained in $\mathcal{F}$. In particular, (i) and (ii) require each f in $\mathcal{F}$ to have certain properties, and (iii) requires $\mathcal{F}$ to be at least minimally rich, i.e., to contain at least certain functions f from **L** to V.

In the familiar cases of probability theory and classical logic, the value sets are the interval $[0, 1]$ and the set $\{0, 1\}$, respectively; the ordering on V is the usual order on the numbers, with minimal element $v_{min} = 0$, i.e., zero probability or falsehood; and the admissible valuation functions are the probability functions and truth-functions, respectively.[34] Below we discuss further examples.

In this general framework, as already indicated, an attitude function A_i is *rational* if it is extendable to an admissible valuation function.[35] It is easy to see that this definition reduces to the special cases of probabilistic coherence or truth-functional coherence when the admissible valuation functions are precisely the probability functions or truth-functions. As before, we are looking for an aggregation rule F which assigns to each admissible profile $(A_1, A_2, ..., A_n)$ a collective attitude function $A = F(A_1, A_2, ..., A_n)$. In particular, we impose the following conditions on such a function:

Universal domain. As before.

Collective rationality. As before.

Independence. As before.

Consensus preservation. If all individuals submit the same attitude function to the aggregation rule, then this attitude function is also the collective one, i.e., $F(A, A, ..., A) = A$ for every admissible unanimous profile $(A, A, ..., A)$.

---

[34] Probability and truth-functions satisfy property (i) because the probability or truth-value of propositions decreases (weakly) in their logical strength. They satisfy (ii) because the probability or truth-value of a proposition does not change by adding a disjunct that has zero probability or is false. They satisfy (iii) because there are enough degrees of freedom in how probabilities or truth-values can be assigned.

[35] Of course, this is to be understood as a thin, formal notion of rationality. Our model allows the analysis of more demanding notions of rationality as well, by imposing additional constraints on rational attitude functions or by requiring extendability to an admissible valuation function satisfying some additional requirements. We here focus on the simplest case.

The following result holds. Call an aggregation rule *systematic* if there exists some *propositionwise decision criterion* d which maps each n-tuple of values $v_1, v_2, ..., v_n$ in V to a single value $v = d(v_1, v_2, ..., v_n)$ such that, for every profile $(A_1, A_2, ..., A_n)$ of rational individual attitude functions and every proposition p in the set X, the collective attitude on p is determined by applying the decision criterion d to the individual attitudes on p, i.e.,

$$A(p) = d(A_1(p), A_2(p), ..., A_n(p)).$$

Examples of propositionwise decision criteria are various possible averaging criteria as in the case of linear rules.

Theorem 2. *Assuming the propositions in the set X exhibit sufficiently rich logical interconnections, every aggregation rule satisfying universal domain, collective rationality, independence and consensus preservation is systematic.*[36]

What do we mean by sufficiently rich logical interconnections? It is not enough for the theorem to require, as in the earlier results, that X has a minimal inconsistent subset of three or more propositions. Instead, we require, more strongly, that the propositions in X are *path-connected*: for any pair of propositions p and q in X, it is possible to 'reach' q from p via a sequence of pairwise conditional entailments.[37] Examples of path-connected sets are the set containing propositions p, q, $p \wedge q$, $p \vee q$ and their negations as well as the set of binary ranking propositions over three or more objects as defined in the previous section. Moreover, any subset of the logic or language **L** that is closed under conjunction or disjunction is path-connected.[38]

---

[36] This theorem generalizes earlier results in the binary cases of abstract aggregation (Nehring and Puppe 2002, Dokow and Holzman forthcoming) and judgment aggregation (Dietrich and List 2007a).

[37] Formally, there exists a sequence of propositions $p_1, p_2, ..., p_k \in X$ with $p_1 = p$ and $p_k = q$ such that, for each j, $\{p_j\} \cup Y$ logically entails p_{j+1}, where Y is some subset of X consistent with each of p_j and $\neg p_{j+1}$. This condition was first introduced in an abstract aggregation setting by Nehring and Puppe (2002) under the name *total blockedness*.

[38] Assuming that the subset contains more than two propositions and that tautologies and contradictions are removed from it.

Theorem 2 basically tells us that, in the general case, any aggregation function satisfying the theorem's four conditions is definable in terms of a single propositionwise decision criterion.[39] Linear or dictatorial rules, as discussed in the previous section, are of course special cases of such rules.

How general is the present theorem? As we have already noted, probability functions and truth-functions are instances of valuation functions, and thus the theorem covers the cases of probability and judgment aggregation. Other important examples of valuation functions are non-binary truth-functions, fuzzy or vague truth-functions, and ranking functions. Non-binary truth-functions take values in the set $V = \{0, 1, ..., T - 1\}$, where T is the number of possible truth-values, and are defined as in T-valued logic. Fuzzy or vague truth-functions take values in the interval $V = [0, 1]$ and are defined as in fuzzy logic or a suitable semantics for vagueness.[40] Ranking functions, as defined by Spohn (forthcoming), take values in the set $V = \{0, 1, ...\} \cup \{\infty\}$ of non-negative integers *together with infinity*. The rank of a proposition p can be interpreted as a degree of rejectability or *dis*belief, and the order $\geq$ is thus defined as the reverse of the natural order, leading to the minimal value $v_{min} = \infty$.[41] To give an example in which the order $\geq$ is incomplete, let the set of values V have four elements: v_{min} ('rejectable'), v_{max} ('acceptable'), u ('undecidable out of conflicting information') and u' ('undecidable out of conflicting intuition'). Let $\geq$ be the partial order on V that ranks v_{min} below all other values and v_{max} above all

---

[39]In the special case of probabilities or judgments, the theorem's path-connectedness assumption on X is not only sufficient for the result but also necessary, i.e., in its absence, there exist non-systematic rules satisfying the theorem's conditions. In the binary case, see also Nehring and Puppe (2002).

[40]If **L** is a propositional logic or language with $\neg$ and $\wedge$ as its only connectives, a T-valued truth-function is usually defined as a function $f : \mathbf{L} \to \{0, 1, ..., T - 1\}$ satisfying $f(p \wedge p) = \min\{f(p), f(q)\}$ and $f(\neg p) = T - f(p)$, and a fuzzy truth-function as a function $f : \mathbf{L} \to [0, 1]$ satisfying $f(p \wedge p) = \min\{f(p), f(q)\}$ and $f(\neg p) = 1 - f(p)$. In both cases, the order $\geq$ on V is the usual one.

[41]A ranking function gives tautologies maximal rank 0, contradictions rank ∞, and disjunctions $p \vee q$ the minimum of the ranks of p and q (with 'minimum' understood here in terms of the natural order). Ranking functions are particularly useful in belief revision theory, since it is possible to define conditional ranks as well as unconditional ones.

other values while leaving u and u' unranked relative to each other. Let a valuation function be simply *any* function from $\mathbf{L}$ to $V = \{v_{min}, v_{max}, u, u'\}$ satisfying conditions (i) and (ii) above. For instance, the proposition 'it will rain' may be deemed undecidable out of conflicting information (u), the proposition 'a party in the rain is fun' may be deemed undecidable out of conflicting intuition (u'), the conjunction 'it will rain *and* a party in the rain is fun' may be deemed rejectable (v_{min}) and the corresponding disjunction acceptable (v_{max}).[42]

In each of these cases, our theorem implies that, when the rationality of propositional attitudes is defined in terms of extendability to a valuation function with the specified properties, the only aggregation rules satisfying universal domain, collective rationality, independence and consensus preservation (for a set of propositions X with sufficiently rich interconnections) are the systematic rules.

5 Concluding remarks

After presenting a taxonomy of different kinds of propositional attitudes – in terms of their role and structure – we have formulated the problem of propositional attitude aggregation as follows. There is a *set of propositions* on which attitudes are to be held, represented in some suitable *logic* or *language*. An agent's attitudes are usually represented by an *attitude function*, which assigns to each proposition in that set a particular value from some *set of possible values*. The value assigned to a proposition could be its truth-

---

[42]Some notable exceptions of functions violating our conditions for valuation functions above are *capacity* functions and Dempster-Shafer belief and plausibility functions (Dempster 1967; Shafer 1976). Such functions are vaguely described as being non-additive variants of probability functions. More precisely, they resemble probability functions in that they take values in the interval $[0, 1]$, assign a value of 0 to contradictions and a value of 1 to tautologies, and are monotonic (i.e., they satisfy (i)), but crucially the value of disjunctions $p \vee q$ of mutually exclusive disjuncts may differ from the sum of values of p and q. For instance, someone may assign value 0 to the proposition that it will rain on Monday and also to the proposition that it will rain on Tuesday, but a positive value to the proposition that it will rain on either Monday or Tuesday. This may cause a violation of our condition (ii) on valuation functions.

value, probability value or some other measure of its acceptability or rejectability. An agent's attitudes are deemed *rational* if the attitude function representing them is *extendable to a well-defined valuation function* on the entire logic or language. Important special cases of such valuation functions are truth-functions and probability functions. An *aggregation rule*, finally, maps each admissible combination of individual attitude functions to a resulting collective attitude function. Aggregation rules can be axiomatically characterized in terms of certain *conditions*; we have here considered only on a few examples of such conditions, namely universal domain, collective rationality, independence, and implication or consensus preservation.

To illustrate our approach, we have proved two new theorems. The first theorem simultaneously characterizes linear and dictatorial aggregation rules in the important cases of probability, judgment and preference aggregation and thus offers a first step towards integrating the literatures on probabilistic opinion pooling, judgment aggregation and Arrovian preference aggregation. It is illuminating to see that such seemingly disparate things as splitting-the-difference in probabilistic opinion pooling and dictatorships in preference aggregation can be axiomatically characterized using the same conditions. The second theorem abstracts even further from the specific kinds of attitudes in question and describes a large class of aggregation rules – the systematic rules – applicable to a wide variety of belief-like attitudes.

Of course, one may wish to generalize these results further, by studying an even broader variety of propositional attitudes and considering aggregation rules other than systematic, linear or dictatorial ones. The present results, however, are only meant to be illustrative of our more general approach. We hope that this paper will not only establish the ingredients of a general theory of propositional attitude aggregation, but also open up new avenues for further research.

21

6 References

Arrow, K. (1951/1963) *Social Choice and Individual Values.* New York (Wiley).

d'Aspremont, C. and L. Gevers (2002) 'Social welfare functionals and interpersonal comparability.' In K. J. Arrow, A. K. Sen and K. Suzumura, eds., *Handbook of Social Choice and Welfare*, volume 1, Amsterdam (Elsevier): ch. 10.

van Benthem, J. (2008) 'Updating Beliefs as Social Choice.' Working paper, ILLC, University of Amsterdam.

Bradley, R. and C. List (forthcoming) 'Desire-as-belief revisited.' *Analysis.*

Bratman, M. (2007) *Structures of Agency: Essays.* Oxford (Oxford University Press).

Condorcet, M. d. (1785). *Essay sur l'Application de l'Analyse à la Probabilité des Décisions Rendue à la Pluralité des Voix.* Paris.

Dempster, A. P. (1967) 'Upper and lower probabilities induced by a multivalued mapping.' *Annals of Mathematical Statistics* 38: 325-399.

Dennett, D. (1987) *The Intentional Stance.* Cambridge, Mass. (MIT Press).

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

Dietrich, F. (2007a) 'A generalised model of judgment aggregation.' *Social Choice and Welfare* 28: 529-565.

Dietrich, F. (2007b) 'Aggregation and the relevance of some issues for others.' Working paper, University of Maastricht.

Dietrich, F. (2008) 'Bayesian group belief.' Working paper, University of Maastricht.

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

Dietrich, F., and C. List (2007b) 'Opinion pooling on general agendas.' Working paper, London School of Economics.

Dietrich, F., and C. List (2007c) 'Strategy-proof judgment aggregation.' *Economics and Philosophy* 23: 269-300.

Dietrich, F., and C. List (2008) 'Propositionwise judgment aggregation.' Unpublished manuscript.

Dokow, E., and R. Holzman (forthcoming) 'Aggregation of binary evaluations.' *Journal of Economic Theory*.

Elga, A. (2007) 'Reflection and Disagreement.' Noûs, 41: 478-502.

Elster, J., ed. (1987) *The Multiple Self*. Cambridge (Cambridge University Press).

Genest, C. and J. V. Zidek (1986) 'Combining Probability Distributions: A Critique and Annotated Bibliography.' *Statistical Science* 1: 113-135.

Gilbert, M. (1989) *On Social Facts*. London / New York (Routledge).

Lehrer, K. and C. Wagner (1981) *Rational Consensus in Science and Society*. Dordrecht/Boston: Reidel.

Lewis, D. (1988) 'Desire as Belief.' *Mind* 97: 323-332.

Lewis, D. (1996) 'Desire as Belief II.' *Mind* 105: 303-313.

List, C. (2007) 'Group Deliberation and the Transformation of Judgments: An Impossibility Result.' Working paper, London School of Economics.

List, C. (2009) 'Judgment aggregation: a short introduction.' In U. Mäki, ed., *Handbook of the Philosophy of Economics*, Amsterdam (Elsevier).

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

List, C. and P. Pettit (2004) 'Aggregating Sets of Judgments: Two Impossibility Results Compared.' *Synthese* 140: 207-235.

List, C. and P. Pettit (2006) 'Group Agency and Supervenience.' *Southern Journal of Philosophy* XLIV (Spindel Supplement): 85-105; reprinted in J. Hohwy and J. Kallestrup, eds., *Being Reduced*. Oxford (Oxford University Press), 2008.

List, C. and P. Pettit (forthcoming) *Group Agency: The Possibility, Design and Status of Corporate Agents.*

List, C. and C. Puppe (2009) 'Judgment aggregation: a survey.' In P. Anand, C. Puppe and P. Pattanaik, eds., *Oxford Handbook of Rational and Social Choice*, Oxford (Oxford University Press).

McConway, K. J. (1981) 'Marginalization and Linear Opinion Pools.' *Journal of the American Statistical Association* 76: 410-414.

Minsky, M. (1988) *The Society of Mind.* New York (Simon and Schuster).

Mongin, P. (1995) 'Consistent Bayesian aggregation.' *Journal of Economic Theory* 66: 313-351.

Mongin, P. (2005) 'Logical aggregation, probabilistic aggregation and social choice.' Lecture given at the 4th International Conference on Logic, Game Theory and Social Choice, University of Caen, June 2005.

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

Nehring, K. and C. Puppe (2008) 'Consistent judgement aggregation: the truth-functional case.' *Social Choice and Welfare* 31: 41-57.

Pauly, M., and M. van Hees (2006) 'Logical Constraints on Judgment Aggregation.' *Journal of Philosophical Logic* 35: 569-585.

Pettit, P. (1993) *The Common Mind: An Essay on Psychology, Society and Politics.* New York (Oxford University Press).

Pettit, P. (2003) 'Akrasia: Collective and Individual.' In S. Stroud and C. Tappolet, eds., *Weakness of Will and Practical Irrationality.* Oxford (Oxford University Press): 68-96.

Pettit, P. (2006) 'When to defer to majority testimony – and when not.' *Analysis* 66: 179-187.

Plato (360 BC) *The Republic.*

Shafer, G. (1976) *A mathematical theory of evidence.* Princeton (Princeton University Press).

Sen, A. K. (1982) *Choice, Welfare and Measurement.* Oxford (Blackwell).

Spohn, W. (forthcoming) 'A Survey of Ranking Theory.' In F. Huber and C. Schmidt-Petri, eds., *Degrees of Belief. An Anthology.* Oxford (Oxford University Press).

Tuomela, R. (2007) *The Philosophy of Sociality: The Shared Point of View.* Oxford (Oxford University Press).

Weymark, J. (1997) 'Aggregating Ordinal Probabilities on Finite Sets.' *Journal of Economic Theory* 75: 407-432.

A   Appendix: Proofs

The notation is as introduced above. The proofs presented here are formulated for the case that propositions are represented syntactically, but the proofs in the semantic case are very similar (and, if anything, simpler; for example, Lemma 1 becomes unnecessary).

Theorem 1 is derived from two earlier results of ours.

Proof of theorem 1. (1) First, consider a linear rule. Obviously, the rule satisfies independence and universal domain. In the binary case, we have already explained that linear rules are dictatorial, which implies that implication preservation and collective rationality are satisfied as well. In the probabilistic case, collective rationality and implication preservation follow from the fact that a weighted average of probability functions on **L** is again a probability function.

(2) Conversely, suppose a rule F satisfies all four conditions. In the case of binary attitudes, these conditions become equivalent to those in an earlier impossibility theorem;[43] and by this theorem F is dictatorial; so F is linear (with all weight assigned to the dictator). In the case of probabilistic attitudes, our four conditions become equivalent to conditions of another

---

[43]Dietrich and List (2008).

earlier theorem after some translation work,[44] and by this theorem F is linear. ∎

Our proof of theorem 2 is self-contained and builds on the following lemmas.

Lemma 1 *Every valuation function $f : \mathbf{L} \to V$ satisfies $f(p) = f(q)$ for all logically equivalent p, q in $\mathbf{L}$.*

Proof. Let f be a valuation function and let p, q in $\mathbf{L}$ be logically equivalent. As q entails p and p entails q, and by property (i) of valuation functions, $f(p) \geq f(q)$ and $f(q) \geq f(p)$; so, by anti-symmetry of $\geq$, $f(p) = f(q)$. ∎

Lemma 2 *For every valuation function f and all $p_1, ..., p_k$ in $\mathbf{L}$ $(k \geq 1)$ if $f(p_1) = ... = f(p_k) = v_{min}$ then $f(p_1 \vee ... \vee p_k) = v_{min}$.*

Proof. Let f be a valuation function. The proof is by induction on k. For $k = 1$ the claim is obvious. Now suppose the claim holds for k, and consider $p_1, ..., p_{k+1}$ in $\mathbf{L}$ such that $f(p_1) = ... = f(p_{k+1}) = v_{min}$. By induction hypothesis, $f(p_1 \vee ... \vee p_k) = v_{min}$, and hence by one of our assumptions on valuation functions

$$f(p_1 \vee ... \vee p_{k+1}) = f((p_1 \vee ... \vee p_k) \vee p_{k+1}) = f(p_{k+1}) = v_{min}. \quad ∎$$

Lemma 3 *Let an aggregation rule F satisfy universal domain, independence and unanimity preservation. For all p in X, all v in V, and all profiles $(A_1, ..., A_n)$ of rational attitude functions, if $A_1(p) = ... = A_n(p) = v$ then $F(A_1, ..., A_n)(p) = v$.*

Proof. Let F be as specified, consider any p in X and any $(A_1, ..., A_n)$ in the universal domain, and assume that $A_1(p) = ... = A_n(p)$. The (universal)

domain of F also contains the unanimous profile $(A_1, ..., A_1)$. Applying independence and then unanimity preservation, we have $F(A_1, ..., A_n)(p) = F(A_1, ..., A_1)(p) = A_1(p)$. $\blacksquare$

Lemma 4 *Let an aggregation rule F satisfy universal domain, collective rationality, independence and unanimity preservation. For all p, q in X, if p conditionally entails q, then $D_p \leq D_q$, where $D_p, D_q : V^n \to V$ are (by independence, existent) local decision rules for p and q, respectively.*

Proof. Let F be as specified, and let p in X conditionally entail q in X, say in virtue of $Y \subseteq X$. Let $(v_1, ..., v_n)$ in V^n. We show that $D_p(v_1, ..., v_n) \leq D_q(v_1, ..., v_n)$.

Check that $q \vee (\neg \wedge_{y \in Y} y)$ is not a tautology and that $[\neg p \wedge q \wedge (\wedge_{y \in Y} y)] \vee [\neg \wedge_{y \in Y} y]$ entails it without being equivalent to it. So, by property (iii) of valuation functions, there are valuation functions $f_1, ..., f_n$ such that, for all i,

$$f_i(q \vee (\neg \wedge_{y \in Y} y)) = v_i \text{ and } f_i([\neg p \wedge q \wedge (\wedge_{y \in Y} y)] \vee [\neg \wedge_{y \in Y} y]) = v_{min}.$$

Consider any individual i. By property (i) of valuation functions, it follows that

$$
\begin{aligned}
f_i(\neg p \wedge q \wedge (\wedge_{y \in Y} y)) &\leq v_{min}, \\
f_i(r \wedge [\neg \wedge_{y \in Y} y]) &\leq f_i(\neg \wedge_{y \in Y} y) \leq v_{min} \text{ for all } r \text{ in } \mathbf{L}, \\
f_i(y) &\leq v_{min} \text{ for all } y \text{ in } Y,
\end{aligned}
$$

and hence, as v_{min} is minimal,

$$
\begin{aligned}
f_i(\neg p \wedge q \wedge (\wedge_{y \in Y} y)) &= v_{min}, \\
f_i(r \wedge [\neg \wedge_{y \in Y} y]) &= f_i(\neg \wedge_{y \in Y} y) = v_{min} \text{ for all } r \text{ in } \mathbf{L}, \\
f_i(\neg y) &= v_{min} \text{ for all } y \text{ in } Y.
\end{aligned}
$$

The first two of these three equation lines imply (using Lemma 1 and one

of our assumptions on valuation functions) that

$$
\begin{aligned}
f_i(p) &= f_i([p \wedge (\wedge_{y \in Y} y)] \vee [p \wedge (\neg \wedge_{y \in Y} y)]) \\
&= f_i(p \wedge (\wedge_{y \in Y} y)) = v_i, \\
f_i(q) &= f_i([p \wedge (\wedge_{y \in Y} y))] \vee [\neg p \wedge q \wedge (\wedge_{y \in Y} y))] \vee [q \wedge \neg \wedge_{y \in Y} y]) \\
&= f_i([p \wedge (\wedge_{y \in Y} y))] \vee [\neg p \wedge q \wedge (\wedge_{y \in Y} y))]) \\
&= f_i(p \wedge (\wedge_{y \in Y} y))) = v_i.
\end{aligned}
$$

Define $A_i : X \to V$ as $f_i|_X$, the restriction of f_i to X. By universal domain, the profile $(A_1, ..., A_n)$ is in the domain of F. As we have just shown,

$$
A_i(p) = A_i(q) = v_i \text{ for all individuals } i = 1, ..., n.
$$

So, letting A denote the output attitude function $F(A_1, ..., A_n)$ and $f : \mathbf{L} \to V$ a (by collective rationality existent) extension of A to a valuation function, we have

$$
f(p) = D_p(v_1, ..., v_n) \text{ and } f(q) = D_q(v_1, ..., v_n).
$$

Further, for each y in Y we have shown that $A_i(\neg y) = v_{min}$ for all i, so that by Lemma 3 $A(\neg y) = v_{min}$, i.e., $f(\neg y) = v_{min}$. So $f(\vee_{y \in Y} \neg y) = v_{min}$ by Lemma 2, i.e., $f(\neg \wedge_{y \in Y} y) = v_{min}$ by Lemma 1. Hence, by property (i) of valuation functions, $f(r \wedge (\neg \wedge_{y \in Y} y)) = v_{min}$ for all r in $\mathbf{L}$. So, using another one of our assumptions on valuation functions and then Lemma 1,

$$
f(r \wedge (\wedge_{y \in Y} y)) = f([r \wedge (\wedge_{y \in Y} y)] \vee [r \wedge (\neg \wedge_{y \in Y} y)]) = f(r) \text{ for all } r \text{ in } \mathbf{L}.
$$

In particular,

$$
f(p \wedge (\wedge_{y \in Y} y)) = f(p) \text{ and } f(q \wedge (\wedge_{y \in Y} y)) = f(q).
$$

So, by property (i) of valuation functions and as $p \wedge (\wedge_{y \in Y} y)$ entails $q \wedge (\wedge_{y \in Y} y)$, we have $f(p) \leq f(q)$, i.e., $D_p(v_1, ..., v_n) \leq D_q(v_1, ..., v_n)$. ∎

Proof of Theorem 2. Let the aggregation rule satisfy all four conditions. The last lemma tells us that, for all p, q in X, we have $D_p \geq D_q$ and $D_q \geq D_p$, and hence $D_p = D_q$ as $\geq$ is anti-symmetric. ∎