The doctrinal paradox, the discursive dilemma, and logical aggregation theory

Mongin, Philippe

HEC Paris

2012

The Doctrinal Paradox, the Discursive Dilemma, and Logical Aggregation Theory[1]

Philippe Mongin[2]
Centre National de la Recherche Scientifique & HEC Paris

February 2012

Abstract

Judgment aggregation theory, or rather, as we conceive of it here, logical aggregation theory generalizes social choice theory by having the aggregation rule bear on judgments of all kinds instead of merely preference judgments. It derives from Kornhauser and Sager's doctrinal paradox and List and Pettit's discursive dilemma, two problems that we distinguish emphatically here. The current theory has developed from the discursive dilemma, rather than the doctrinal paradox, and the final objective of the paper is to give the latter its own theoretical development along the line of recent work by Dietrich and Mongin. However, the paper also aims at reviewing logical aggregation theory as such, and it covers impossibility theorems by Dietrich, Dietrich and List, Dokow and Holzman, List and Pettit, Mongin, Nehring and Puppe, Pauly and van Hees, providing a uniform logical framework in which they can be compared with each other. The review goes through three historical stages: the initial paradox and dilemma, the scattered early results on the independence axiom, and the so-called canonical theorem, a collective achievement that provided the theory with its specific method of analysis. The paper goes some way towards philosophical logic, first by briefly connecting the aggregative framework of judgment with the modern philosophy of judgment, and second by thoroughly discussing and axiomatizing the "general logic" built in this framework.

JEL Reference Numbers: D 70, D 71, D 79.

Keywords: Judgment Aggregation, Logical Aggregation, Doctrinal Paradox, Discursive Dilemma, General Logic, Premiss-Based vs Conclusion-Based Approach, Social Choice Theory, Impossibility Theorems

[1]The present English paper has evolved from an earlier French paper co-authored with Franz Dietrich ("Un bilan interprétatif de la théorie de l'agrégation logique", *Revue d'économie politique*, vol. 120, 2010, p. 929-972). Many thanks to him for allowing this author to present this new version. Thanks also for their comments to Brian Hill, Mikaël Cozic, Daniel Eckert, Itay Fainmesser, Jim Joyce, Lewis Kornhauser, Gabriella Pigozzi, Rohit Parikh, Roberto Serrano, Jan Sprenger, Jonathan Zvesper, and the participants to the many conferences or seminars where versions or variants of this paper were given.

[2]GREGHEC, 1 rue de la Libération, F-78350 Jouy-en-Josas. Email: mongin at greg-hec.com

1 Introduction

Contemporary aggregation theories have their roots in mathematical analyses of voting, developed in France from the end of the 18th century, as well as in the technical formulations of utilitarianism and its rarified variant, welfare economics, which were to follow in Great Britain from the 18th century to the middle of the 20th century. Classical and then neo-classical economists set great store by this second source, but were entirely unaware of the first, until Arrow masterfully orchestrated their reconciliation in *Social Choice and Individual Values* (1951). The title of his work fixed the accepted name of the theory it propounds, of *social choice*, a rather inaccurate name, because social choice theory starts with preference and not choice, as its fundamental concept, and it bears on all types of collectivity, the whole of society being just one particular instance. With the no less improperly named "social welfare" function, which is defined from individual to collective preferences, the Arrovian theory develops a formalism that can cover all of the aggregation problems that the two historical traditions, the French and the British, offered in isolation. Indeed, the notion of preference, individual or collective, can tend either towards the side of the utility function, which "represents" preferences according to economists' contemporary conception, or towards the side of choice, which "reveals" preferences according to the same conception. Voting is merely a particular kind of choice; it is in this way that Arrow and his successors were able to connect Bentham with Condorcet.

Considerable as that step of generalization might have been, it was still not enough, since the preferences of individuals or of the collectivity between two states of affairs do not exhaust the judgments they could make about those states, and it is just as legitimate to look at the aggregation of other kinds of judgments. "Bob prefers a long monthly meeting to short weekly meetings" can be paraphrased by saying that Bob judges one such meeting to be preferable to the others, and this makes us notice that preference is a special case in several ways. It is a *comparative* judgment made from the *evaluative* point of view that is specific to *preferability.* Concerning the same states of affairs, Bob can form other judgments, either absolute or again comparative: a long monthly meeting is tiring; more tiring than short weekly meetings; successful; more successful than short weekly meetings; and so on. It is even inevitable that Bob will form judgments other than of preference, since like any form of evaluative judgment, they require factual judgments to have already been made. When several Bobs express themselves concerning both of these levels of judgment, should the theory of aggregation only consider the final one without examining the intermediary one? Such a restriction would appear counter-productive, yet it is imposed in social choice theory, which does not admit judgments other than of preference.

A new theory, called *judgment aggregation theory*, overcomes this limitation. Its formalism allows it to represent individual and collective judgments of any kind, and thus to attack an entire class of neglected aggregation problems. The contributors to this theory add one step of generalization to that made by Arrow and his disciples in their time. From this group, they retain the axiomatic method of investigation. Indeed, they posit on individual and collective judgments certain normative properties which parallel the constraints imposed on preferences, and define a *collective judgment function*, going from admissible profiles of individual judgments to admissible collective judgments, which is the formal analogue of the social welfare function. They study properties that collective judgment functions could satisfy in terms of impossibility and possibility theorems, which is how Arrovian theory proceeds with the social welfare function.

There are already so many and diverse such axiomatic studies that it is impossible to account for them without fixing an angle of attack. The present account underlines the logical side of individual and collective judgments and sets out the new theory from this vantage point. Accordingly, we will refer to it as *logical aggregation theory.* As well as being more informative, this name also has the advantage of marking a clean separation with the preceding theory of probabilistic aggregation. In a broad view of judgments, which is in fact the common sense one, subjective certainty is not inherent to them, and the statements that express them may not have the full force of the values "true" and "false", the only ones considered by standard logic. If today's theory were also concerned with this view, it should include the theory of probabilistic aggregation; but that is far from the case, the former having been created without the support of the latter, and both still being unaware of each other apart from a few exceptions.[1]

As a matter of fact, it is logic to which the new theory has turned for technical help. Essentially, it assumes that the individual or collective subject who is making a judgment contemplates the proposition associated with it, and either affirms it in giving it the value "true" or denies it by giving it the value "false". The theory also allows the subject not to make any judgment, in which case he attributes no truth value to the proposition. That sketch is fleshed out by introducing a symbolic language with a formula for each proposition, and then, as in any logical work, a logical syntax or its semantic analogue to operate on these formulas. Propositional logic suffices

[1] The theory of probabilistic aggregation goes back to the 60s with the work of Stone, Raiffa and Winkler. The main results were obtained quickly, so that the survey by Genest and Zidekh (1986) remains up to date. In McConway (1981), the theory follows the axiomatic style of social choice theory, whereas in Lehrer and Wagner (1981), it takes the different form of a theory of consensus applicable to scientific activity.

even for some advanced results, but we will push this article in the direction of less elementary logics.

It is with List and Pettit (2002) that an aggregation theory incorporated – for the first time rigorously – the logical analysis of judgment that we just outlined. Before that, two American legal theorists, Kornhauser and Sager (1986, 1993) had outlined an aggregative conception of judgment, but only taken in the judiciary, and not the broader philosophical, sense. They showed that collegiate courts were subject to the so-called *doctrinal paradox*: individually coherent opinions of judges can lead the court as a group to inconsistency. In a seminal article that prepares the formalism later introduced with List, Pettit (2001) reexamines Kornhauser and Sager's paradox and, judging it too specific, proposes reformulating it as a universal problem that he calls the *discursive dilemma*. By and large, the logical aggregation literature has adhered to the discursive dilemma version of the problem, turning its back on the doctrinal paradox version, so it matters a great deal to understand how they differ. Unlike most contributors, we draw a sharp contrast between the two problems. This is the major guiding insight of the article, as it will influence both its review part and its more original conclusions. In the end, we will show that the much neglected doctrinal paradox can be used as a departure point for a branch of logical aggregation theory whose results are yet more general than those of the main trunk. We expand here on Dietrich and Mongin (2010) without reproducing their full technical exposition.

Without List and Pettit being aware of it, a French scholar of social mathematics, Guilbaud (1952), had already set about generalizing from preference to judgment. Inspired by Arrow, who had just published his book, but also by Condorcet, whose work he was to help rescue from oblivion, Guilbaud reformulated the former's theory of aggregation as the latter would have done, that is: not only for relations of preference, but for all sorts of "opinions".[1] Opposed to the Bourbakianism that then dominated French mathematics, Guilbaud rejected the axiomatic method and even eschewed general proofs, which makes it complicated to evaluate the extent of his contribution, but the trend today is to see him as a direct forerunner of logical aggregation theory.[2]

[1] Guilbaud prefers the term "opinion", from Condorcet, to "judgment", which he still uses sometimes.

[2] Monjardet (2003) singles out in Guilbaud a theorem still of the Arrovian style, but Eckert and Monjardet (2009) credit him with one of the judgment aggregation style, and this reading is more faithful; see also Mongin (2012). On the technical side, Guilbaud made the first use of filters and ultrafilters, notions which had just begun to enter the mathematics of his time. Since Kirman and Sondermann (1972), aggregative theories have often used this technique.

If we must find a first source for the current work, Condorcet is the only choice, with his *Essai sur l'application de l'analyse à la probabilité des décisions à rendre à la pluralité des voix* (1785) and his other treatises or articles on mathematical politics. His abiding method is to treat a preference as the accepting or rejecting of certain propositions. The voter who prefers A to B, B to C, and A to C, accepts "A is preferable to B", "B is preferable to C", "A is preferable to C", and rejects the contrary propositions. One would think, under Arrow's influence, that Condorcet only describes preference orderings in a roundabout way, but that is far from the case. He starts with propositions and their supposed logical connections, and he only attributes the ordering property to the preference relation in virtue of these logical connections. The greater generality of his reasoning is better seen in the jury theorem than in the voting paradox, but the latter is still representative. He analyzes it by writing that the propositions chosen by the majority constitute an "inconsistent system" (*Essai*, p. LV-LVI), and this suggests that logical coherence, and not preferability, is the crux of the matter. Reread in this manner, which is Guilbaud's, Condorcet becomes the distant precursor of the doctrinal paradox, of the discursive dilemma and of all the ongoing research.

Regrettably, we will touch on Condorcet only in connection with the paradox of voting, whereas the jury theorem would also be relevant to our theoretical purposes. First of all, the theorem illustrates even more clearly than the paradox that collective judgments raise problems of their own, beside what can be said of collective preferences. Second, unlike the paradox, it involves the two concepts of an objective truth and the probability of reaching it - remember Condorcet's statement of the judges' deliberation in *Essai*. New aggregative issues emerge when the framework of judgment aggregation is enriched to take these concepts into account. Third and not last, some writers have managed to connect the Condorcet jury theorem with the current theory, and this has even become a strand of literature by itself. To include it would have made the already substantial account untractable.[3]

This article essentially consists of a long circular development around the doctrinal paradox. Section 2 presents the paradox, returning to the initial judiciary source, then compares it to its reinterpretation as a discursive dilemma. The sharp contrast drawn between the two problems will colour the rest of the paper. Section 3 presents the formal framework of current logical aggregation theory. In a brief aside, we comment on the philosophy of judgment that underlies the chosen definitions, arguing that they

[3]See in particular Bovens and Rabinowicz (2004), Pigozzi (2006), Hartmann and Pigozzi (2007), Hartmann, Pigozzi and Sprenger (2010), Hartmann and Sprenger (forthcoming).

reflect a typically modern conception of judgment, as opposed to the old, Aristotelian one. From section 3 onwards, a sequence of impossibility results follows. That of List and Pettit (Theorem 1) imposes the questionable axiom of *systematicity* on the collective judgment function, whereas those stated in section 4, due to Pauly and van Hees, Dietrich, Mongin, and Nehring and Puppe (Theorems 2–5), only require *independence*, a normatively more defensible axiom that is close to the famous Arrovian one of independence of irrelevant alternatives. Section 5 is another interlude in the style of philosophical logic. It sets out, with improvements, Dietrich's general logic, which overcomes a limitation of Theorems 1-4 (they are formulated in propositional logics, and this is too specific). The general logic is axiomatized here in two equivalent ways that differ by their choices of primitive concepts - inference in one case and inconsistency in the other. (This section is technical and can be skipped by those willing to take for granted the logical apparatus of the theory: an executive summary awaits them at the beginning of the next.) In section 6, against the background of the general logic, we present those theorems which best structure the field today and can, because of this, be considered canonical. They have as their mathematical object the *agenda*, i.e., the set of logical formulas standing for the propositions about which the individuals and the group express opinions. The conditions placed on this object turn any collective judgment function that is subjected to certain axioms into a degenerate rule, such as dictatorship or oligarchy. Importantly, these agenda conditions are not only sufficient, but also necessary for the axiomatized function to degenerate, so that the theorems state possibilities no less than impossibilities. All researchers in the field have made some contribution here, but Nehring and Puppe on the one hand, and Dokow and Holzman on the other, stand out; in Theorems 6 and 7, we rely on the latter team's version. Section 7 returns to the doctrinal paradox to give that its own theoretical development, along the lines of Dietrich and Mongin's work. It shows how the framework of logical aggregation theory, which was initially intended for the discursive dilemma, can be adapted to the other problem. This leads to Theorems 8 and 9, which are then thoroughly applied to Kornhauser and Sager's court example. The last section summarizes conclusions and open problems.

2 From the doctrinal paradox to the discursive dilemma

Early forerunners aside, logical aggregation theory originates in the analysis of the legal institution laid out by Kornhauser and Sager (1986, 1993; see also Kornhauser, 1992). From this analysis, the *doctrinal paradox* emerged as a major finding. Here we restore it in its original judiciary terms, distinguishing it carefully from the transformations it underwent in Pettit (2001) and List and Pettit (2002). Presenting it as a *discursive dilemma*, these authors opened the way to logical aggregation theory, but - we will argue - they also swept aside some significant conceptual concerns.

The doctrinal paradox occurs for the first time in Kornhauser (1992) but becomes central only in Kornhauser and Sager (1993), where it is illustrated by the following - by now famous, and even a little tired - example. A plaintiff P, has brought a civil suit against a defendant D, alleging a breach of contract between them. The court is composed of three judges A, B and C, who must, based on contract law, determine whether or not D owes damages to P, a decision represented by the logical formulas d or $\neg d$. The case brings up two issues, i.e., whether the contract was valid or not in the first instance (v or $\neg v$), and whether D was or not in breach of it (b and $\neg b$), and the law decides for all possible responses, stipulating that D must pay damages to P if, and only if, both issues are answered in the affirmative. Suppose that the judges' deliberations lead them to the following responses and conclusions:

A	v	$\neg b$	$\neg d$
B	$\neg v$	b	$\neg d$
C	v	b	d

If the court rules directly on the case using simple majority voting, it will arrive at the conclusion $\neg d$, against the plaintiff. However it can, still using the same voting rule, first decide on the two issues, and then draw a conclusion about the case based on the law, and this will return the answers a and b, hence finally d, in favour of the plaintiff.

To avoid misunderstandings, it is worth adding that the actual US contract law is more complex than is said here. Kornhauser and Sager also examine genuine cases, but they are too intricate to be strikingly paradoxical, and the authors therefore conceived of the present toy example, which came somewhat late in their joint work.[4] Their ultimate target is to investigate how the law changes when judiciary decisions have a collective form. Concretely, the only collegiate courts in the American legal system are the appellate courts of the States and of the Union. Everyone has heard of the Supreme Court, whose nine judges reach their decisions about federal cases through deliberation and - sometimes but not always - explicit voting. Less well known are the State appellate courts, often composed of three judges, and above them, the State supreme courts, which operate similarly.

Beside being exemplified, the doctrinal paradox can be defined in the abstract. This requires some relevant legal concepts to be introduced first,

[4]Logical aggregation theorists usually locate the doctrinal paradox already in Kornhauser and Sager (1986), but this is not the proper reference.

as in Kornhauser (1992), and we review them sketchily. A *case* brought before a court is subjected by this court to a *characterization*, which amounts to defining what part of the law, if any, is relevant to it. This inquiry leads the court to delineate the *legal doctrine*, which, once applied to the case, will provide its complete resolution. The doctrine relies on an admixture of statutes and common law, depending on the case at hand; given the Anglo-American tradition, the former will prevail in criminal matters, and the latter in civil matters. The more jurisprudence is involved, the closer the judges' activity to law-making, and the more entangled their deliberation. Logically, the doctrine does two things at once, i.e., it fixes the *issues* that the case presents, and it translates possible decisions on these issues into decisions on the case. This is captured by assuming that there are unambiguous questions - to be answered by yes or no - for both the case and each issue, and that the legal doctrine dictates an answer to the case once all answers to the issues have been collected.

The preceding concepts are sufficient only if there is a single judge. If the court is collegial, one has also to describe how individual answers are aggregated into a final judgment. Kornhauser and Sager contemplate two possibilities. The first has the court record directly *individual answers about the case* and apply some collective decision-making procedure, like simple majority voting, to them. In the second, *individual answers about each issue* are recorded, and the collective decision-making procedure is applied to each of these separately, after which the answer about the case follows from applying the legal doctrine. The doctrinal paradox arises any time that the first method, which is *case-by-case*, does not yield the same result as the second, which is *issue-by-issue*. That is the authoritative definition in Kornhauser (1992, p. 453, where it appears for the first time), Kornhauser and Sager (1993, p. 10-12) and subsequent legal theorists.[5]

There is an interesting contrast between Condorcet's voting paradox and Kornhauser and Sager's. The former does not already indicate where to search for solutions, but the latter does, since it is defined precisely in terms of their discord. It therefore has a structural quality which brings it already close to an impossibility theorem. Furthermore, being abstract, it allows for more than one interpretation. One may say that there are colleagues who decide simultaneously, but this is not the only possible view, nor is it the most interesting from the legal perspective. In the above example, the three judges might have sitted apart, each arriving at a decision for himself, whereas a fourth judge, involved after them, would ask how to make the best of the jurisprudence thus created. Can he only retain the answers on the case, or should he make use of the answers on the issues? As we

[5] Post and Salop's (1991–1992) work seems partly independent of Kornhauser and Sager's. For the subsequent law literature, see Nash's (2003) critical review.

read them, Kornhauser and Sager initially concerned themselves with the collective functioning of courts primarily from the angle of their *diachronic* consistency. If, once they had discovered the paradox, they focussed on *synchronic* consistency, we believe that this is simply for intellectual convenience. Of the two problems, the first is more important than the second, because it affects all courts, collective as well as individual, and all the more when common law tends to outweigh statutes in determining the law. However, the first problem being also more difficult, it was good policy to start with the second.[6]

If there is anything paradoxical in the clash between the case-by-case and issue-by-issue methods, it is because each can rely on a solid normative argument. By deciding case-by-case, the court fully respects the deliberations of individual judges, right up until the decisions they would make, were they alone in adjudicating the case. By deciding issue-by-issue, the court guarantees that its decision is based on the same type of reasons - those allowed by the legal doctrine - as the judges' individual decisions. According to Kornhauser and Sager, "where the doctrinal paradox arises, judgment and reason are immediately and inexorably pulled apart" (1993, p. 25). By "reason", they classically mean one's ability to justify conclusions using logic. As they assume that each judge exercises this capacity competently, the question is whether it holds at the group's level, and the issue-by-issue method arguably ensures that it does. By "judgment", they mean a conclusion obtained by the case-by-case method, and indirectly, the supporting argument just said that this method gives careful attention to individual judgments. In List and Pettit (2002, p. 94), the conflicting principles are called "collective rationality" and "individual responsiveness", a more explicit terminology that we will retain from now on.

Beyond the psychological shock of the paradox, the clash between the two methods poses a *dilemma*, in the usual sense of a forced choice between two unsatisfactory options, since the argument to want one is also an argument not to want the other, and even a *theoretical* dilemma, since two basic principles clash, as was just explained. Pettit (2001), then List and Pettit (2002), definitely move the doctrinal paradox in this abstract direction. However, the novelty of the *discursive dilemma*, which they promote as an alternative concept, cannot lie just in this reinterpretation. It must also go beyond the authors' claim that the doctrinal paradox occurs outside of the legal context, because this is so immediately obvious. Pettit points in particular to the deliberative entities of democratic institutions,

[6] When they call *case-by-case* one of the solutions to the synchronic problem, Kornhauser and Sager may still be echoing the diachronic version, for which this expression is more appropriate. List (2004) is the only logical aggregation theorist ever to have addressed that version.

review panels and authorities of economic regulation, clubs or other groups whose members coopt, and even, to some degree, political parties, unions and churches. Being a matter of empirical observation, the list can go on. The only problematic item is the whole of political society, which Pettit chooses to include, relying as he does on the theory of deliberative democracy that he has defended elsewhere (on this line, see also Brennan, 2001). This speculative extension of the doctrinal paradox also goes beyond the normal range of the discursive dilemma, and so cannot really mean a difference between the two. Nor is the distinction clarified by Pettit's labelling of the two methods as the "premiss-driven way" and the "conclusion-driven way" (2002, p. 274). This new contrast, which List and Pettit (2002) passed on to logical aggregation theorists, is simply a useful reminder that the legal distinction between the issues and the case can be put more abstractly (the currently received terms, to be used henceforth, are the *premiss-based* versus the *conclusion-based method*).

What substantially distinguishes the discursive dilemma from the doctrinal paradox is not to be found on the interpretive side, as in the previous suggestions, but rather on the formal side. Here is how List and Pettit (2002) reconstruct the judiciary example. They associate formulas of classical logic to all relevant considerations, including the legal doctrine, for which they take $d \longleftrightarrow v \wedge b$ ("d if and only if v and b") Supposing then that the court votes on each formula, they bring to light the logical contradiction that it would face:

A	v	$\neg b$	$\neg d$	$d \longleftrightarrow v \wedge b$
B	$\neg v$	b	$\neg d$	$d \longleftrightarrow v \wedge b$
C	v	b	d	$d \longleftrightarrow v \wedge b$
$Court$	v	b	$\neg d$	$d \longleftrightarrow v \wedge b$

Whereas the doctrinal paradox was defined in terms of two methods to relate conclusions to premisses, the discursive dilemma is defined by a contradiction within the overall collective judgment, without the need to distinguish between premisses and conclusions. Presented in this way, the problem falls within the scope of classical propositional logic, and it thus opens the way to an attractive formalism of judgment aggregation. In retrospect, the successful development of logical aggregation theory suggests that it was reformulated appropriately.

Still, if one is concerned to deepen the doctrinal paradox within its original context, there are some reasons to doubt that List and Pettit opened the right path. Indeed, in this context, their wide-ranging definition of collective judgment is questionable for two reasons. For one, the distinction between the issues and the case gave its fine structure to the judiciary decision problem, and by ignoring it, one simply destroys the connections with legal theory; our brief discussion of jurisprudential reasoning can flesh out this objection. For another, the legal doctrine calls for a separate logical analysis, but List and Pettit make it unspecific. To unpack the critique here, they take it for granted (i) that the doctrine can be represented by a formula of classical propositional logic, like the propositions describing the issues and the case, and (ii) that it falls under the scope of the same decision rule as these propositions. It transpires from Kornhauser and Sager's later writings that they have doubts about logical aggregation theory, but they have never expressed them fully.[7] As we suggest, the disagreement could relate to (i), (ii), or even more radically, the underlying claim (iii) that the doctrine is a proposition rather than a command or a rule. Section 7 touches on this last objection and takes the first two more thoroughly into account.

3 The formal framework of logical aggregation theory

The theory is developed from the specific notions of agenda, judgment sets, and the collective judgment function, as well as a small set of axioms to be put on this mapping, and various conditions to be put on the agenda and the judgment sets. This section and the following one present the theory with a minimum of logical details, only developing the formal language, and postponing until section 5 the full definitions of the logic properly (i.e., inference and associated notions).

By definition, a *language $\mathcal{L}$ of the theory* is a set of formulas $\varphi, \psi, \chi, \ldots$ containing logical symbols taken from a certain set $\mathcal{S}$. It is not necessary to specify the formulas beyond the minimal requirement that $\mathcal{S}$ contains the symbol for Boolean negation $\neg$ ("not") and $\mathcal{L}$ is closed for this symbol; i.e., if $\varphi \in \mathcal{L}$, then $\neg\varphi \in \mathcal{L}$. If $\mathcal{S}$ contains other elements, they will be symbols for the remaining Boolean connectives, $\vee$ ("or"), $\wedge$ ("and"), $\rightarrow$ ("if ..., then ..."), $\leftrightarrow$ ("if and only if ..., then ...") or for non-Boolean operators representing modalities (e.g., "it is obligatory that ...", "it is desirable that ...", "it is known that ...", or "if ..., then ..." taken in a non-Boolean sense, typically with a counterfactual interpretation). For each additional symbol of $\mathcal{S}$, the corresponding closure rule holds: if $\wedge \in \mathcal{S}$, the rule says that if $\varphi \in \mathcal{L}$ and $\psi \in \mathcal{L}$ then $\varphi \wedge \psi \in \mathcal{L}$, and so on. We distinguish *classical* and *non-classical* languages according to whether, respectively, $\mathcal{S}$ contains only Boolean symbols or others in addition.

In the very large class of permitted languages, the particular case of *propositional languages* $\mathcal{L}_{\mathcal{P}}$ stands out. They are defined in terms of a set

[7] A debate took place between Kornhauser and Sager (2004) and List and Pettit (2005), but it does not identify the disputes as clearly as one might like.

$\mathcal{P}$ of elementary formulas, or *propositional variables*, which do not contain any logical symbols, and a set $\mathcal{S}$ containing the five Boolean connective symbols. Since these connectives are inter-definable, we can equivalently have $\{\neg, .\} \subseteq \mathcal{S}$, replacing the dot with any one of $\vee$, $\wedge$, $\rightarrow$ or $\leftrightarrow$. Classical propositional languages are those for which $\{\neg, .\} = \mathcal{S}$.

By a *calculus* of the theory, we mean a language $\mathcal{L}$ together with a logic, i.e., a system of axioms and rules that determine the inferential and other logical links between the formulas of $\mathcal{L}$. Just as with the language, there is no need to specify this system in its entirety. Section 5 will show that it suffices to have an *inference relation* $B \vdash \psi$ defined for $B \subset \mathcal{L}$ and $\psi \in \mathcal{L}$, respecting some very general restrictions, but here and in the next section, we restrict ourselves to the special case of *classical propositional calculi*. Such calculi have classical propositional languages, and for their logic part, well-known systems of axioms and rules that fit the ordinary mathematical intuition; the interested reader may consult any logic text. These unstated, but obvious systems will for now fix the meaning of the inference rule $\vdash$ and of its associated notions, like logical truth, logical contradiction, logical equivalence, logical independence, consistency and inconsistency. Classical propositional calculi draw our attention only because they are so elementary. We used one of these like Jourdain used prose when formalizing the judiciary example in the discursive dilemma version. The language then was $\mathcal{L}_{\mathcal{P}}$, built from $\mathcal{P} = \{v, b, d\}$ and $S = \{\neg, \leftrightarrow\}$, and the set $\{v, b, \neg d, d \leftrightarrow v \wedge b\}$ was contradictory in the sense of any textbook system.

In $\mathcal{L}$, the theory fixes a subset X representing the propositions that are in question for the members of the group; this is the *agenda*. It can be large or small depending on the application, but in all generality it is only required to be non-empty and, as with $\mathcal{L}$, closed for negation. The judiciary agenda in discursive dilemma style is:

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

If we were to add to the agenda multiple negations $\neg\neg\varphi, \neg\neg\neg\varphi, \ldots$, the logic would in the end reduce them to either φ or $\neg\varphi$. It is better to anticipate that process and define agendas as sets:

$$X = \{\varphi, \chi, \ldots\}^{\pm},$$

where $\varphi, \chi, \ldots$ are *positive* (non-negated) formulas and the superscript $\pm$ indicates that they are accompanied with their negations $\neg\varphi, \neg\chi$, ... To simplify matters, we impose the restriction - going beyond what the theory needs - that agendas consist of *contingent* formulas, i.e., are neither logical truths, nor logical contradictions.

The theory represents individual and group judgments by subsets of X, *judgment sets*, which can be made to fulfil certain logical constraints, the primary one being consistency. They will be denoted by $B, B', \ldots$ generally, and by $A_i, A'_i, \ldots A, A', \ldots$ when they belong to, respectively, individuals i and the group they form. A formula φ from one of these sets represents a proposition, in the ordinary sense of a semantic object endowed with a truth value. If φ is used also to represent a judgment, in the sense of a cognitive operation, then it is in virtue of the natural interpretive rule:

(R) i (the group) judges that φ iff $\varphi \in A_i$ (resp. iff $\varphi \in A$).

We treat the formula φ in this statement as if it were itself the proposition; this terminological ease is commonplace in logic and will be taken for granted in what follows. Thanks to (R), judgments obey a distinction between two types of negation, internal and external, which has no analogue on the level of propositions or formulas, these being negated in only one way. Indeed, "judging that not" ($\neg\varphi \in B$) is different from "not judging that" ($\varphi \notin B$). Once the inference relation is defined, the logical consistency of judgment sets will relate one negation to the other as could be expected, i.e., "judging that not" will entail "not judging that", without the converse always holding.

From what we have said, it can be seen that logical aggregation theory is connected to a particular philosophical conception of judgments and propositions. The language $\mathcal{L}$ represents all expressible propositions, i.e., all propositions that can become the object of a judgment, but only those in X will actually become so. It is typical of the modern concepts of proposition and judgment – since Frege and Russell – that the former has a wider range than the latter; this definitely clashes with the ancient view – that of Aristotle, which is still to be found in Kant.[8] Logical aggregation theory uses the Fregean concept of *assertion*, here rendered as $\varphi \in B$, and as can be checked, it is faithful to the principle, also typically modern and Fregean, that assertion remains unaffected by logical operators. For it does not matter whether φ is positive or negative, conditional or unconditional, modal or non-modal; the indicator chosen for assertion – set membership – works always in the same way. Of course the modern conception does allow distinctions which are made by the logic to be lifted to judgments. Thus one may speak of a "positive judgment" or "negative judgment" of φ, to mean that φ or $\neg\varphi$, respectively, belongs to the judgment set in question; emphatically, it is only the negation sign, or lack thereof, in front of φ that differentiates the two cases.

[8] Kant's *Logic*, published in 1800, is a famous sample of the ancient view. The modern one is best exemplified by Frege's *Logical Investigations* (1918–1919). Vickers (1989) has a crisp summary of their differences.

Returning from this philosophical excursus to the formal framework, we index the individuals by $i = 1, \ldots, n$, assuming that $n \geq 2$, and define the *collective judgment function*, which associates a collective judgment set to each configuration, or *profile*, of judgment sets for the n individuals:

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

Like social choice theory, logical aggregation theory usually deals with finite sets of individuals.[9] As a generalization of the Arrovian social welfare function, F formalizes the decision rules that the group would apply to the formulas in the agenda. According to its standard definition, the only one considered here, F has a *universal domain*, i.e., is defined on the set of all possible profiles, given the logical constraints imposed on judgment sets. These constraints, to be explained in the next paragraph, may also affect the range of F.[10]

A judgment set B can be expected to be:

- *deductively closed*, i.e., for all $\varphi \in X$, if $B \vdash \varphi$ then $\varphi \in B$;
- *consistent*, i.e., for no $\varphi \in X$ do we have $B \vdash \varphi$ and $B \vdash \neg\varphi$;
- *complete*, i.e., for all $\varphi \in X$, either $\varphi \in B$ or $\neg\varphi \in B$.

Various families of judgment sets result from combining these properties. The main cases are:

- the set D of consistent and complete judgment sets, which satisfy the three properties or, equivalently, the last two (the first easily follows from them),
- the set $D^* \supsetneqq D$ of consistent and deductively closed judgment sets, as defined by the first two properties.

From there, one can define restrictions on the domain and range of F such as:
(i) $F : D^n \to D$,
(ii) $F : D^n \to D^*$,
(iii) $F : (D^*)^n \to D^*$,
(iv) $F : D^n \to 2^X$ or $F : (D^*)^n \to 2^X$.

[9] Dietrich and Mongin (2007), and then Herzberg and Eckert (2010) and Herzberg (2010), have looked at infinite sets. Their results translate those already obtained in social choice theory, in particular by Kirman and Sonderman (1972).

[10] Logical aggregation theory is only now beginning to look at restricted domains; see List (2003) and Dietrich and List (2010a).

In the beginning, only (i) was considered. It makes proofs easier but is called into question by cognitive psychology, which would favour weaker logical hypotheses. One can also – a more elaborate argument – question (i) by calling upon the modern notion of judgment that underlies the formalism. By ruling out abstention, D destroys the possibility it offers of dealing with a proposition *without having to assert it or its negation.* In other words, the theory loses the distinction between internal and external negation, since "not judging" becomes equivalent to "judging that not". There is therefore more than one reason to develop the options based on D^*, i.e., (ii) and (iii). Case (iv), in which individual and collective judgment sets are markedly different, is only given to help explain the others.

Two ways of formalizing collective judgment rules suggest themselves: one can either define F so that it coincides with a particular rule, or determine F by axiomatic conditions representing general principles to be satisfied by any rule. The same two possibilities occur concerning the social welfare function, and as its record shows, one gets the most by following both paths at the same time.[11] *Proposition-wise majority voting*, which is the decision rule associated with the judiciary example, will illustrate the process. This rule is defined here as the collective judgment function $F_{maj} : D^n \to 2^X$ such that, for every profile $(A_1, \ldots, A_n)$ of the domain,

$$F_{maj}(A_1, ..., A_n) = \{\varphi \in X : |\{i : \varphi \in A_i\}| \geq q\},$$
$$\text{with } q = \tfrac{n+1}{2} \text{ if } n \text{ is odd and } q = \tfrac{n}{2} + 1 \text{ if } n \text{ is even.}$$

Note that the range is not D because there can be unbroken ties between φ and $\neg\varphi$ when n is even, and it is not even D^* in view of the judiciary example, which exhibits an inconsistent collective judgment set. Having defined proposition-wise majority, we introduce its salient normative properties. This section mentions three such properties, defined abstractly for any F, that together allow for the easy proof of an impossibility theorem – the first to have occurred in the literature. The list will be extended in section 4 with more advanced results.

Systematicity. For every pair of formulas $\varphi, \psi \in X$, and for every pair of profiles $(A_1, \ldots, A_n)$, $(A'_1, \ldots, A'_n)$, if, for every $i = 1, \ldots, n$, the equivalence $\varphi \in A_i \Leftrightarrow \psi \in A'_i$ holds, then so does the equivalence

$$\varphi \in F(A_1, \ldots, A_n)) \Leftrightarrow \psi \in F(A'_1, \ldots, A'_n).$$

[11] It is because there are two such paths that the method of social choice theory bears some analogy with the axiomatic method of formal logic. For the definition of the rules, say majority voting, plays the role of *semantic* models with respect to the *syntax* constituted by the axioms, say IIA, and characterization theorems approximate completeness theorems proved in logic. More on this in Mongin (2003).

Systematicity means that the group, when confronted with a profile of individual judgment sets, gives the same answer concerning a formula as they would give concerning a *different* formula, when faced with a *different* profile, whenever the individual judgments concerning the first formula in the first profile are the same as the individual judgments concerning the second formula in the second profile. The rule F_{maj} clearly respects systematicity, whose analogue in social choice theory is neutrality (see, e.g., Sen, 1970, and d'Aspremont, 1985).

We will say that a collective judgment function F is a *dictatorship* if there is one individual j such that, for every profile $(A_1, \ldots, A_n)$,

$$F(A_1, \ldots, A_n) = A_j.$$

Given the universal domain assumption, there is only one such j per dictatorship, to be called the *dictator*. Obviously F_{maj} satisfies:

Non-dictatorship. F is not a dictatorship

and even more strongly:

Anonymity. For every profile $(A_1, \ldots, A_n)$, if $(A'_1, \ldots, A'_n)$ is obtained from $(A_1, \ldots, A_n)$ by permuting the individuals, then

$$F(A_1, \ldots, A_n) = F(A'_1, \ldots, A'_n).$$

The parallel with social choice theory is again clear. Note however that Arrow's (1963) dictator imposes only his strict preference, not his indifference, which means that dictatorship for him is not a projection property, as it is here.

The theory's first result made clear the conflict between anonymity and systematicity under a minor condition being imposed on X (List and Pettit, 2002, Theorem 1). In fact, the conflict can be expressed more strongly as that between non-dictatorship and systematicity (Pauly and van Hees, 2006, Theorem 4) and we therefore present that improved version of the result.

Theorem 1 (Pauly and van Hees, 2006, generalizing List and Pettit, 2002) Let $\mathcal{L}_\mathcal{P}$ be a classical propositional language with $\mathcal{S} = \{\neg, \wedge\}$; let $a, b \in \mathcal{P}$ be two distinct propositional variables such that $a, b, a \wedge b \in X$; then there is no $F : D^n \to D$ satisfying both non-dictatorship and systematicity.

Since F_{maj} satisfies non-dictatorship and systematicity on D^n, it must, by contraposition of the theorem, have a range other than D. When n is odd, the collective judgment sets are complete, so one of $F_{maj}(A_1, \ldots, A_n)$

must be inconsistent. This is exactly what the judiciary example in discursive dilemma form has taught, but the theoretical deduction supersedes the earlier finding, which was restricted to specific $\mathcal{L}_{\mathcal{P}}$, X and n. Beside generalizing its main example, Theorem 1 deepens the conceptual meaning of the discursive dilemma. "Collective rationality" is reflected in the assumption that the range of F is D, "individual responsiveness" in the non-dictatorship condition, but what about the systematicity axiom, which is related to neither? Although List and Pettit think of their formal analysis as mirroring their informal one, it rather seems that the value of Theorem 1 lies in its correcting primary intuitions. As it makes clear, the problem of collective judgment is in fact *a trilemma with systematicity as an additional element.* Unlike the other two, this property has no normative standing, and can only be defended in terms of its technical advantages. Nonetheless, it is involved just as much as the others in the impossibility conclusion.[12]

Among the many voting rules that Theorem 1 covers, we single out those which rely on a qualified majority. For any q such that $1 \leq q \leq n$ - the *quota* - define $F^q_{maj} : D^n \to 2^X$ thus: for every profile $(A_1, \ldots, A_n)$ of the domain,

$$F^q_{maj}(A_1, \ldots, A_n) = \{\varphi \in X : |\{i : \varphi \in A_i\}| \geq q\}.$$

In the limit case where $q = n$, a formula is collectively accepted if and only if all individuals accept it, a *unanimity rule* to be compared with the Pareto extension rule of social choice theory (see Sen, 1970, ch.5*, and 1986). Clearly, the F^q_{maj} functions cannot go to D, since some collective judgment sets are inconsistent for low q values, and others are incomplete for high q values (where proposition-wise majority voting defines the cut-off betwen "low" and "high"). Theorem 1 accounts for these failures at one go, thus illustrating the unifying power for which the axiomatic method is classically famous.[13]

In summary, against the straightforward background of a classical propositional calculus, new concepts take their shape: the agenda; individual and collective judgment sets; and the collective judgment function that connects them. The last concept permits dealing with both specific rules and general conditions. With this technology, the discursive dilemma was recast as an abstract impossibility of collective judgment. However, Theorem 1 only brushes the surface of the possible arguments.

[12]The basic weakness of the axiom is that it cancels out semantic differences between propositions (see Mongin, 2008). When premisses and conclusions are distinguished, another problem is that it makes them interchangeable, whereas the former serve as reasons for the latter and not vice-versa (see Chapman, 2002).

[13]Quota rules are defined here as in Dietrich and List (2007a), who, after Nehring and Puppe (2002, 2008), study them in detail; see also Dietrich (2010).

4 Theorems based on the independence axiom

Pursuing the example of voting rules, we now introduce three other salient properties that they typically satisfy, i.e., unanimity preservation, independence, monotonicity. The theorems below, which extend List and Pettit's in various ways, rely on these new axiomatic conditions. The first in the list requires the collectivity to reproduce the individuals' unanimous judgments. In the present framework, unanimity may be applied either to the judgment sets themselves, or – more strongly – to their formulas considered one by one. The parallel with systematicity, and indeed with other conditions that are to follow, is made clearer if we opt for the latter variant, which is also the closest analogue of the Pareto conditions in social choice theory.

Unanimity preservation. For every formula $\varphi \in X$ and every profile $(A_1, \ldots, A_n)$, if $\varphi \in A_i$ for every $i = 1, \ldots, n$, then $\varphi \in F(A_1, \ldots, A_n)$.

The second condition is a weakening of systematicity, hence the F^q_{maj} functions automatically satisfy it.

Independence. For every formula $\varphi \in X$ and every pair of profiles $(A_1, \ldots, A_n)$, $(A'_1, \ldots, A'_n)$, if for every $i = 1, \ldots, n$, the equivalence $\varphi \in A_i \Leftrightarrow \varphi \in A'_i$ holds, then so does the equivalence:

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

Independence is the same as restricting systematicity to the case where $\varphi = \psi$. It eliminates the conceptual element of *neutrality*, i.e., of indifference to the semantic content of propositions, while preserving another conceptual element which dovetails with it in the earlier condition, that is: the collective judgment of φ depends only on the individual judgments of φ. To put it differently, the set A is defined *formula-wise* from the sets $A_1, \ldots, A_n$. The theory can only express this idea by comparing a given profile with hypothetical profiles, in which the individual judgments on $\psi \neq \varphi$ may be different while those on φ stay the same. The axiom, including its multi-profile formulation, is closely related to Arrow's independence of irrelevant alternatives.[14]

Voting rules typically satisfy a strengthening property: when a collective result reflects the judgment of a group of voters, the result still holds if more voters join the group in their judgment. Like the related condition of positive responsiveness in one version of Arrow's theory, this requires a multi-profile formulation.[15]

[14] Despite the significant weakening of systematicity, some normative objections remain against independence (see Mongin, 2008).

[15] The 1951 version of Arrow's theorem relied on positive responsiveness, while the 1963 and still current version uses a Pareto condition. More on the former in Sen (1970).

Monotonicity. For every formula $\varphi \in X$, and for every pair of profiles $(A_1, \ldots, A_n)$, $(A'_1, \ldots, A'_n)$, if the implication $\varphi \in A_i \Rightarrow \varphi \in A'_i$ holds for every $i = 1, \ldots, n$, with at least one i such that $\varphi \notin A_i$ and $\varphi \in A'_i$, then the following implication holds:

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

Independence clearly does not imply monotonicity, and as the next example shows, neither does systematicity. A collective judgment function F is an *anti-dictatorship*, if there is j such that for every $(A_1, \ldots, A_n)$ and every $\varphi \in X$,

$$\varphi \notin A_j \Leftrightarrow \varphi \in F(A_1, \ldots, A_n).$$

Under the appropriate agenda restriction,[16] F has domain D^n and range D. It is systematic, but not monotonic, as illustrated by two profiles $(A_1, \ldots, A_n)$ and $(A'_1, \ldots, A'_n)$ such that $\neg\varphi \in A_j$, $\varphi \in A'_j$ and $A_i = A'_i$ for all $i \neq j$. Henceforth, *monotonic independence* and *monotonic systematicity* refer to the conjunction of the monotonicity axiom with the independence or systematicity axiom, respectively.

If the impossibility conclusion of Theorem 1 could be derived from independence instead of systematicity, this would deepen the explanation of the discursive dilemma. The theory would then shift the problem of collective judgment to one of the two conceptual elements, namely formula-wise aggregation, from the other, neutrality. Systematicity would certainly remain in the conclusion – dictatorial functions, the only ones existing from Theorem 1, do satisfy this property – but it would be better not to have it in the assumptions. In that way, one would also re-establish the parallel with social choice theory, where the strongest results deal with independence of irrelevant alternatives as an assumption, neutrality serving only as an intermediary step.[17]

This programme was realized by Pauly and van Hees (2006, Theorem 4) and Dietrich (2006, Theorem 1, Corollary 2), who posit independence as their starting point, and also by Mongin (2008, Theorem 2), who, unlike them, also requires unanimity preservation to hold. Both Pauly and van Hees and Dietrich *derived* the latter condition in the course of their proofs. This leads to an impressive ratio of conclusions to assumptions, but one could want to make more explicit the two very different principles that are at work simultaneously. In order to have unanimity preservation as a separate

[16]For every consistent subset $B \subseteq X$, the negated subset $\{\neg\varphi : \varphi \in B\}$ is also consistent.

[17]Here, logical aggregation theory and social choice theory have followed opposite paths. Arrow's 1951 theorem started with independence of irrelevant alternatives, and it was only later and for special cases that some theorems proceeded from neutrality. Fleurbaey and Mongin (2005) reexamine the historical sequence.

assumption, Mongin *weakens* independence so that there is no entailment anymore. Each of the three axiom sets leads to an impossibility.

The three works have in common that they strengthen the very weak agenda conditions of Theorem 1; this is the price to pay for replacing systematicity by independence. Given a language $\mathcal{L}_{\mathcal{P}}$, let us say that X is *closed for propositional variables* if, for every formula $\varphi \in X$ and every propositional variable $a \in \mathcal{P}$ occurring in φ, $a \in X$. For example, $\overline{X}$ verifies this closure condition, since $\varphi = (d \leftrightarrow v \wedge b) \in \overline{X}$ and $v, b, d \in \overline{X}$. A *literal* is defined as some $a \in \mathcal{P}$ or its negation $\neg a$; it is denoted by $\pm a$. Given that X is closed by negation, the present closure condition is equivalently stated by putting $\pm a \in X$ instead of $a \in X$.

Theorem 2 (Pauly and van Hees, 2006). Let $\mathcal{L}_{\mathcal{P}}$ be a classical propositional language with $\mathcal{S} = \{\neg, \wedge\}$; let X be closed for propositional variables, with at least two distinct propositional variables, and such that, for all $\pm a, \pm b \in X$, $\pm a \wedge \pm b \in X$; then every $F : D^n \to D$ satisfying both non-dictatorship and independence is a constant function.[18]

A collective judgment function F is *constant* if there is a judgment set A such that, for every profile $(A_1, \ldots, A_n)$, $F(A_1, \ldots, A_n) = A$. The case arises when one moves from systematicity to independence; indeed, a constant collective judgment function into D or D^* satisfies the latter but not the former.[19]

Until now we have not made - and we in general will not make - any assumption concerning the number of propositional variables. However, the following theorem is best stated with $\mathcal{P}$ finite (and so, modulo logical equivalence, $\mathcal{L}_{\mathcal{P}}$ also finite). We can then define the *atoms* of $\mathcal{L}_{\mathcal{P}}$, which are the formulas $\pm a_1 \wedge \ldots \wedge \pm a_k$, in which each of the k distinct propositional variables of $\mathcal{P}$ occurs. The set of atoms of $\mathcal{L}_{\mathcal{P}}$, to be denoted by $\mathcal{AT}_{\mathcal{P}}$, is the finest logical partition – class of logically exclusive and logically exhaustive formulas – for this propositional language; in other words, each atom describes a conceivable state of affairs with maximal precision. Dietrich shows that, if the agenda contains the atoms, the same conclusion as in Theorem 2 follows, even though independence only applies to these formulas within the agenda.

Theorem 3 (Dietrich, 2006). Let $\mathcal{L}_{\mathcal{P}}$ be a classical propositional language with $\mathcal{S} = \{\neg, \wedge\}$ and $\mathcal{P}$ finite, containing at least two propositional

[18] While being classical in the sense of section 3, Pauly and van Hees's propositional calculus is unusual in allowing for any finite number of truth values. Van Hees (2007) and Duddy and Piggins (2009) also go beyond bivalent semantics (to which our syntactical formalism implicitly subscribes).

[19] The restriction to D or D^* is essential. Otherwise, the following F is both constant and systematic: $F(A_1, \ldots, A_n) = X$ for all $(A_1, \ldots, A_n)$.

variables; let X include $\mathcal{AT}_{\mathcal{P}}$; then every $F : D^n \rightarrow D$ satisfying non-dictatorship and independence restricted to $\mathcal{AT}_{\mathcal{P}}$ is a constant function.

For simplicity, we have stated only a special case of the original theorem, which we now explain in terms of an example. Take $\mathcal{P} = \{a, b\}$ and $X = \{a, \neg a \wedge b, \neg a \wedge \neg b\}^{\pm}$. This agenda does not contain all the atoms of $\mathcal{L}_{\mathcal{P}}$ but nonetheless satisfies a related property, i.e., for every judgment set $B \in D$, there is a formula in X that is logically equivalent to the conjunction of the elements of B. Indeed, D contains only three judgment sets:

$$\{a, \neg(\neg a \wedge b), \neg(\neg a \wedge \neg b)\}, \{\neg a, \neg a \wedge b, \neg(\neg a \wedge \neg b)\}, \{\neg a, \neg(\neg a \wedge b), \neg a \wedge \neg b\},$$

and these can respectively be identified with:

$$a, \neg a \wedge b, \neg a \wedge \neg b.$$

Given the restriction of $\mathcal{L}$ to X, judgment sets in D describe conceivable states of affairs with maximal precision; so their equivalent formulas may be defined as the atoms of $\mathcal{L}_{\mathcal{P}}$ *relative to* X. Dietrich's result in fact concerns this notion of atoms, which extends its scope beyond what has been stated formally.

In this theorem, independence holds only of a subset of the agenda. The next result restricts the axiom similarly, albeit to a different subset, i.e., the set PV_X of propositional variables occurring in X.

Theorem 4 (Mongin, 2008). Let $\mathcal{L}_{\mathcal{P}}$ be a classical propositional language; let X be closed for propositional variables, with at least two propositional variables, and moreover satisfying the agenda conditions stated in section 7. Then there is no $F : D^n \rightarrow D$ that satisfies non-dictatorship, unanimity preservation, and independence restricted to PV_X.

Theorems 2, 3 and 4 have a common ground, which is to clarify the negative role of the independence condition. The first two essentially say that a collective judgment function degenerates if it proceeds formula-wise on an agenda whose formulas are logically interconnected. The last theorem implicitly accepts this diagnosis, since it restricts independence to the only formulas that - in a classical propositional calculus - are not logically interconnected, i.e., to PV_X. The impossibility conclusion then follows from adding unanimity preservation, which makes this condition the relevant target of criticism. In the end, the discursive dilemma comes close to the problem of *spurious unanimity* that Mongin (1995, 1997) brought to light in the context of collective Bayesianism.[20]

[20] Individuals can make the same expected utility comparisons although they differ both

The judiciary example can serve to illustrate the two analyses just sketched. According to the first, the court is confronted with problems because it requires the judges to vote on each proposition considered in isolation, whereas they are logically connected by legal doctrine. According to the second, even if the court ensures that judges vote on logically independent propositions, it must still take care to apply unanimity preservation in the right way. As it happens, the judges are not in agreement about how to make use of the legal doctrine, and this undercuts the supposed normative force of their unanimity in this circumstance.

A formal example will make the two steps of this reasoning even more explicit. Let $\mathcal{L}_{\mathcal{P}}$ be a propositional language with $\mathcal{S} = \{\neg, \vee\}$ and $\mathcal{P} = \{a, b, c\}$; let the agenda be $X = \{a, b, c, a \vee b \vee c\}^{\pm}$, which fits the conditions of Theorem 4; finally, let $n = 3$ and the profile $(A_1, A_2, A_3) \in D^3$ be such that:

$$a, \neg b, \neg c \in A_1; \neg a, b, \neg c \in A_2; \neg a, \neg b, c \in A_3.$$

By deductive closure, $a \vee b \vee c \in A_i$ for all $i = 1, 2, 3$. If the collective judgment function is F_{maj}, the collective judgment set A contains $\neg a, \neg b, \neg c, a \vee b \vee c$ and is thus contradictory. This illustrates the difficulty of formula-wise aggregation, given the logical connection between a, b, c established by $a \vee b \vee c$, and it reflects the spirit of Theorems 2 and 3.[21] Now, the contradiction would still occur if F_{maj} were restricted to a, b, c and unanimity preservation were applied to $a \vee b \vee c$. This illustrates Theorem 4 and the critical role of spurious unanimity (since the three individuals have incompatible reasons to accept the same formula $a \vee b \vee c$).

As a matter of history, a theorem of Nehring and Puppe (2002), based on monotonic independence, came before the results just covered. It was not stated in the formalism of logical aggregation theory, but it is possible to translate it to there (see Nehring and Puppe, 2010). This theorem belongs to section 6, and here, we state another result by the same authors in order to illustrate their condition of monotonic independence at work. Let us say that a collective judgment function F has a *local veto power* if there is an individual j and a formula $\varphi \in X$ such that, for every profile $(A_1, \ldots, A_n)$,

$$\varphi \notin A_j \Longrightarrow \varphi \notin F(A_1, \ldots, A_n).$$

For a given F, there can be several veto holders j, each relative to a given φ. This is a weak technical variant of dictatorship, bearing some relation to

in their utility and their probability assignments. Mongin (1995) thereby explains the impossibility of collective Bayesianism. More generally, Mongin (1997) talks of *spurious unanimity* when a judgment is collectively agreed by individuals who have conflicting reasons for arriving at it. Nehring's (2005) abstract formalism of Paretian aggregation encapsulates related ideas.

[21] The spirit, not the letter, since the chosen X does not obey the agenda conditions of these theorems.

Gibbard's (1969) concept of an oligarchy in social choice theory (a tighter connection will be made in section 6).

Theorem 5. (Nehring and Puppe, 2008). Let $\mathcal{L}_{\mathcal{P}}$ be a classical propositional language; let X be closed for propositional variables, with at least one contingent formula that is not logically equivalent to a literal. Then there is no surjective F satisfying monotonic independence and having no local veto power.

Compared to Theorems 2, 3 and 4, the axioms on F are strengthened. Independence has been supplemented with monotonicity, which, in the presence of surjectivity, can be shown to entail unanimity preservation, and the absence of veto is clearly much more demanding than the absence of a dictator. At the same time, the constraints on X are definitely reduced. Thus, various trade-offs are possible between conditions placed on the agenda and on the axioms. Section 6 will develop this observation to the point of specifying meta-theoretical equivalences.

Theorem 5 is brought out here to limit technicalities and facilitate comparisons, but Nehring and Puppe have more powerful results (2008, Theorems 1 and 2). They define F to be an *oligarchy with default* if there are two non-empty subsets $J \subset X$ and $M \subset \{1, \ldots, n\}$ such that for every $(A_1, \ldots, A_n)$ and every $\varphi \in X$,

$$\begin{array}{lll} \varphi \in F(A_1, \ldots, A_n) & \text{iff} & \text{either } \varphi \in A_j \text{ for all } j \in M, \\ & & \text{or } \varphi \in J \text{ and } \varphi \in A_j \text{ for some } j \in M. \end{array}$$

The members $j \in M$ are called the *oligarchs*, and the set of formulas J the *default*. In essence, if the oligarchs agree about a formula φ, it goes through to the collective judgment set, and if they are divided, then the default makes the decision between φ or $\neg\varphi$. For certain agendas (we do not give the conditions here), Nehring and Puppe show that the only F satisfying monotonic independence and surjectivity are oligarchies with default.[22]

In summary, with Theorems 2–5, logical aggregation theory further deepens the discursive dilemma. Systematicity has given way to independence, sometimes posited by itself, sometimes modified by unanimity preservation or monotonicity. The more recent work favours the coupling of standard independence and unanimity preservation. Before we come to this, we will in the next section return to the formal framework of logical aggregation theory, which has not yet been defined in full generality. Readers unconcerned with these technicalities may go straight to section 6, at the beginning of which they are summarized informally.

[22] Nehring and Puppe (2010) drop the condition – called *truth-functionality* – which corresponds here to the closure of X for propositional variables. As a result, their agendas become compatible with non-oligarchic collective judgment functions.

5 A general logic for the theory

The theorems of sections 3 and 4 were formulated in terms of classical propositional calculi, which is restrictive. The question arises of extending them to *non-classical propositional calculi*, i.e., in which the language comprises of non-Boolean connectives and the logic has an inference relation different from the standard one. Equally, the question arises of extending them to *predicate calculi*, whether classical or not, which improve on the analysis of propositions by using symbols for predicates, variables and quantifiers. They strengthen the preceding logics in another direction, and on the application side, they are needed to obtain social choice theorems on preference relations as corollaries of the logical aggregation theorems.

Instead of working in two steps, first by proving a logical aggregation theorem for elementary calculi, and then checking that it holds for more advanced ones, it would be better to prove it once and for all *in a general logic that encompasses all the calculi one may be interested in.* This goal was set by Dietrich (2007a), who achieved it by axiomatizing the inference relation $\vdash$ without referring to any particular logic. We pursue the same approach using the improved axiomatization of Dietrich and Mongin (2010). Henceforth, once a theorem is proved for the general logic, it will suffice, in order for it to apply to a calculus whose language is of type $\mathcal{L}$ defined in section 3, that its inference relation obeys the axioms in question. The canonical theorem and the further results in section 6 and 7 are stated in this new formal framework.

Let us fix a binary relation $S \vdash \psi$, holding between certain sets $S \subseteq \mathcal{L}$ and certain formulas $\psi \in \mathcal{L}$. We define it to be an *inference relation*, with S being then called a *set of premisses* and ψ a *conclusion*, if it satisfies the following list of six axioms. In their statement, $S \nvdash \psi$ and $\varphi \vdash \psi$ mean, respectively, that $S \vdash \psi$ does not hold and that $\{\varphi\} \vdash \psi$.

(E1) There is no $\psi \in \mathcal{L}$ such that $\emptyset \vdash \psi$ and $\emptyset \vdash \neg\psi$ (non-triviality).

(E2) For every $\varphi \in \mathcal{L}$, $\varphi \vdash \varphi$ (reflexivity).

(E3) For every $S \subset \mathcal{L}$ and every $\varphi, \psi \in \mathcal{L}$, if $S \cup \{\varphi\} \vdash \psi$ or $S \cup \{\neg\varphi\} \vdash \psi$, then $S \vdash \psi$ (single-step completion).

(E4) For every $S \subset S' \subset \mathcal{L}$ and every $\psi \in \mathcal{L}$, if $S \vdash \psi$, then $S' \vdash \psi$ (monotonicity).

(E5) For every $S \subset \mathcal{L}$ and every $\psi \in \mathcal{L}$, if $S \vdash \psi$, then there is a finite subset $S_0 \subseteq S$ such that $S_0 \vdash \psi$ (compactness).

(E6) For every $S \subset \mathcal{L}$, if there is $\psi \in \mathcal{L}$ such that $S \vdash \psi$ and $S \vdash \neg\psi$, then for every $\psi \in \mathcal{L}$, $S \vdash \psi$ (non-paraconsistency).

A further property follows from these:

(E7) For every $S, T, \subset \mathcal{L}$ and every $\psi \in \mathcal{L}$, if $T \vdash \psi$ and $S \vdash \varphi$ for every $\varphi \in T$, then $S \vdash \psi$ (transitivity).

From this list, (E4) is doubtless the most important condition. It expresses the monotonicity that is typical of *deductive* inferences, as opposed to the non-monotonicity typical of *inductive* inferences, which the following example illustrates. Suppose that S says that all ravens examined up to time t are black, and ψ that all ravens are black. Now, the inductive inference from S to ψ no longer holds if S is augmented with φ saying that a raven examined at time $t+1$ is not black. Neither the commonsensical, nor the philosophical concept of judgment appears to be analytically tied with the concept of deduction; rather, they both draw upon a broader idea of reasoning that can accommodate induction.[23] One should therefore see (E4) as a substantial restriction *on the judgments the theory is concerned with.* Incidentally, this is another reason to favour the label of *logical* aggregation promoted here.

Condition (E1) is essential for non-triviality, especially in the presence of (E4), and (E2) states a property that one would expect any inference, whether deductive or inductive, to have. (E3) permits suppressing unnecessary premisses, which is appropriate for deductive inferences. This condition corresponds to a more familiar one, which is stated below in terms of logical inconsistency. (E5) says that sets of premisses can be taken to be finite, a property that reflects a general concern among logicians for a tractable inference concept. This is a restrictive, if desirable, property, so logical aggregation theorists would be well-advised to mention it any time that they assume it (as is done, say, in the proofs of Dietrich and Mongin, 2010). (E6) imposes another restriction on the class of permitted inferences, but unlike (E4) and (E5), it appears to be unproblematic. It excludes a group of deductive calculi - the so-called *paraconsistent* ones - which have long vexed logicians and whose peculiar situation we will now explain.

Let $\mathcal{I}$ denote the set of *inconsistent* sets of $\mathcal{L}$-formulas; by definition, a set will be *consistent* if and only if it belongs to the complement of $\mathcal{I}$. One way of formalizing these notions is to define them in terms of the inference relation. According to the most standard definition in logic:

(Def*) $S \in \mathcal{I}$ if and only if for all $\psi \in \mathcal{L}$, $S \vdash \psi$.

However, paraconsistent logicians choose a weaker definition:

(Def**) $S \in \mathcal{I}$ if and only if there is $\psi \in \mathcal{L}$ such that $S \vdash \psi$ and $S \vdash \neg\psi$.

Either definition can object to the other on the ground that it gives rise to the wrong number of inconsistent sets - too many in the case of Def** and too few in the case of Def*. Mathematicians have implicitly pushed this debate aside by making the two definitions coincide; this is what (E6) achieves here. By adopting this axiom, logical aggregation theory complies with ordinary proof intuitions and only excludes a rather uncommon family

[23] For a similar view, see Makinson (2005, ch.1).

of logical calculi.[24]

Under either definition, the axiomatization (E1)–(E6) implies the following properties of $\mathcal{I}$:

(I1) $\emptyset \notin \mathcal{I}$ (non-triviality).

(I2) For every $\varphi \in \mathcal{L}$, $\{\varphi, \neg\varphi\} \in \mathcal{I}$ (reflexivity).

(I3) For every $S \subset \mathcal{L}$ and every $\varphi \in \mathcal{L}$, if $S \notin \mathcal{I}$, either $S \cup \{\varphi\} \notin \mathcal{I}$ or $S \cup \{\neg\varphi\} \notin \mathcal{I}$ (single-step completion).

(I4) For every $S \subset S' \subseteq \mathcal{L}$, if $S \in \mathcal{I}$ then $S' \in \mathcal{I}$ (monotonicity).

(I5) For every $S \subseteq \mathcal{L}$, if $S \in \mathcal{I}$, then there is a finite subset $S_0 \subseteq S$ such that $S_0 \in \mathcal{I}$ (compactness).

(I1) and (I4) restate their inferential counterparts. (I2) can be expected to hold when paraconsistency is put aside. (I3) permits completing a consistent set by a formula or its negation, or equivalently, by a finite set of formulas or their negations. In the presence of compactness - here (I5) - the step can be made to the corresponding infinite property:

(I3$^+$) For every $S \subset \mathcal{L}$, if $S \notin \mathcal{I}$, there is $T \subset \mathcal{L}$ such that (i) $S \subseteq T$, (ii) $T \notin \mathcal{I}$, and (iii), for every $\varphi, \neg\varphi \in \mathcal{L}$, either $\varphi \in T$ or $\neg\varphi \in T$ (full completion).

This is the so-called *Lindenbaum extension property*, which logicians usually prove from other premisses. The set D introduced in section 3 is now put on a firm logical basis: there exist consistent and complete judgment sets, no matter the cardinality of the language.

The general logic can be presented in the opposite order, i.e., starting from the set $\mathcal{I}$ axiomatized by (I1)–(I5), and treating the relation $\vdash$ as derived. It can then be checked that $\vdash$ satisfies (E1)–(E6), which become properties rather than axioms. A new connecting definition is needed if one follows this reverse order of doing things:

(Def***) $S \vdash \varphi$ if and only if $S \cup \{\neg\varphi\} \in \mathcal{I}$.

To reduce inference to inconsistency, as in (Def***), is no less common than to reduce inconsistency to inference, as in (Def*), and the fact that our general logic can rely on two axiomatizations instead of one makes it easier to use in proving aggregation theorems. Since axioms and properties reflect standard deductive practice, the proofs can be carried out at the intuitive level as far the logic goes (the only exception being compactness, which we argued must be mentioned).

Whichever of the two lists is taken as a criterion, classical propositional calculi fall under the general logic. *Non-classical propositional* logics, on the other hand, need to be examined one by one. Among them, there are many deductive logics, as opposed to inductive or non-monotonic ones, that fulfil the criterion, but some, especially with epistemic applications, turn

[24] See Priest (2002) for a survey of paraconsistent calculi and their motivations.

out not to be compact.[25] Any *classical predicate calculus* also obeys the general logic. In this case, the formulas of $\mathcal{L}$ are the closed formulas of the original language (i.e., those having no free variables in them) and the combinations of them obtained with the Boolean connectives. It is then routine to check that the inference relation of the calculus satisfies (E1)–(E6) when it is restricted to $\mathcal{L}$. There is another method to handle classical predicate calculi, which dispenses with such a direct check. It consists in extracting the propositional content from the predicate calculus by using a standard isomorphism construction (see Barwise, 1977) and then invoking the fact that classical propositional calculi agree with the general logic. Of course non-classical predicate calculi call for the same kind of reservations and checks as non-classical propositional calculi.

6 The canonical impossibility theorem

Section 5 has completed the unfinished work of section 3 in defining the formal framework of logical aggregation theory. As it turns out, it is unnecessary to fix a logical calculus in order to formalize judgments sets and collective judgment functions. Any logical calculus has two components, i.e., a language or set of formulas $\mathcal{L}$, and a logic as properly defined, which can be described equally well in terms of an inference relation $\vdash$ on $\mathcal{L}$ or a set $\mathcal{I}$ of inconsistent subsets of $\mathcal{L}$. Now, once $\mathcal{L}$ is fixed, logical aggregation theorists may not elaborate on these items, simply requiring that they have certain properties; the two salient ones are *monotonicity* and *compactness*. In terms of $\vdash$, monotonicity says that the premisses of an inference can be increased without any conclusion being lost, and in terms of $\mathcal{I}$, that the supersets of an inconsistent set are inconsistent. In terms of $\vdash$, compactness says that any inference can be drawn from a finite number of premisses, and in terms of $\mathcal{I}$, that any inconsistency occurs among a finite number of formulas. Because only such general conditions are needed for a proof in logical aggregation theory, the theory can free itself from classical propositional calculi. For given applications, it can envisage more expressive calculi, like the predicate calculus and non-classical propositional calculi, and when no specific application is intended, it can invoke *general logic* - the underlying laws of $\vdash$ or $\mathcal{I}$ - instead of any particular logical system.

Two derived logical notions enter the next theorem statements, and we now introduce them formally. First, a set of formulas $\mathcal{S} \subset \mathcal{L}$ is called *minimally inconsistent* if it is inconsistent and all its proper subsets are

[25] Probabilistic epistemic logics are not compact (see Heifetz and Mongin, 2001), nor are most logics of common knowledge (though some are, see Lismont and Mongin, 2002).

consistent. For instance, this is the case for

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

but not for

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

Observe that compactness is required to ensure minimality when $\mathcal{L}$ and X are infinite. Second, given $\varphi, \psi \in X$, we say that ψ *is inferred conditionally from* φ – denoted by $\varphi \vdash^* \psi$ – if there is a *set of auxiliary premisses* $Y \subset X$ such that (i) $Y \cup \{\varphi\} \vdash \psi$ and (ii) $Y \cup \{\varphi\}$ and $Y \cup \{\neg\psi\}$ are consistent. ($Y = \varnothing$ is permitted.) Conditional inference can be reformulated in terms of minimally inconsistent sets, and that is in effect how it first arose in the theory.[26] Two further properties need mentioning: first, conditional inference never relates a formula to its negation, and second, it satisfies contraposition, i.e.,

$$\varphi \vdash^* \psi \Leftrightarrow \neg\psi \vdash^* \neg\varphi.$$

Apart from not yet relying on general logic, Theorems 1-5 suffer from being somewhat imprecise. As they are formulated, they only state sufficient conditions on the agenda for there to exist no collective judgment functions – except for degenerate ones – that satisfy specified axiomatic properties. These hypotheses can be too strong for the conclusion, and if they are not, an additional proof of necessity should establish this. Influenced first by Nehring and Puppe (2002, 2010), and then by Dokow and Holzman (2009, 2010a and b), logical aggregation theory has taken on the task of *characterizing*, in the sense of necessary and sufficient conditions, the agendas which turn a list of axiomatic conditions into an impossibility. If the results of these authors deserve being called *canonical*, it is not so much because of their depth or generality, since they are far from unifying the whole theory, but rather because they have established a format of results that is now widely adopted. Here we follow Dokow and Holzman's analysis, not reproducing it as is, but rather rendering it into general logic; the difference with the original formalism will be explained at the end of the section.

Dokow and Holzman raise and solve the following aggregative problem: how to characterize the agendas X such that, if we define D with respect to X, there is no $F : D^n \to D$ that satisfies at once non-dictatorship, independence, and unanimity preservation? The answer to this problem - the mentioned canonical theorem - brings to the fore the following agenda conditions:

(a) *There exist a minimally inconsistent set of formulas* $Y \subseteq X$ *and distinct formulas* $\varphi, \psi \in Y$ *such that*

[26]Under compactness, $\varphi \vdash^* \psi$ is equivalent to requesting that $\varphi \neq \neg\psi$ and there be some minimally inconsistent $Y' \subset X$ with $\varphi, \neg\psi \in Y'$.

$$Y' = Y \setminus \{\varphi, \psi\} \cup \{\neg\varphi, \neg\psi\}$$

is consistent.

(b) *For every formulas* $\varphi, \psi \in X$*, there exist formulas* $\varphi_1, \ldots, \varphi_k \in X$ *such that*

$$\varphi = \varphi_1 \vdash^* \varphi_2 \vdash^* \ldots \vdash^* \varphi_k = \psi.$$

If an agenda X satisfies (a), it is said to be *even-number negatable*. As a notational shortcut, for any $Z \subseteq Y$, we write $Y_{\neg Z}$ for

$$Y' = (Y \backslash Z) \cup \{\neg\varphi : \varphi \in Z\}.^{27}$$

If X satisfies (b), it is said to be *path-connected* (another received expression is *totally blocked*).

Theorem 6 (Dokow and Holzman, 2010a; see also Nehring and Puppe, 2002 and 2010, and for sufficiency, Dietrich and List, 2007b). If X is even-number negatable and path-connected, there is no $F : D^n \to D$ that satisfies non-dictatorship, unanimity preservation and independence. When $n \geq 3$, the agenda conditions are also necessary for this conclusion.

To illustrate Theorem 1, we reexamine $F_{maj} : D^n \to 2^X$ when n is odd and it is thus equivalent to say that F_{maj} is not to D or to say that there exists $(A_1, \ldots, A_n)$ making $F_{maj}(A_1, \ldots, A_n)$ inconsistent; denote this property by (Inc). Given that F_{maj} satisfies the three axioms, Theorem 6 gives the implication (a),(b) $\Rightarrow$ (Inc). Let us illustrate this on the judiciary agenda in discursive dilemma form:

$$\overline{X} = \{v, b, d, d \leftrightarrow v \wedge b\}^{\pm}.$$

We see that (a) holds by taking:

$$\begin{array}{rcl} Y = \{\neg v, d, d \leftrightarrow v \wedge b\} & \text{and} & Z = \{\neg v, d\}, \quad \text{or} \\ Y = \{v, b, d, \neg(d \leftrightarrow v \wedge b)\} & \text{and} & Z = \{v, b\}, \end{array}$$

or yet more choices, which suggests that (a) is easy to fulfil despite being complex to specify. As for (b), Figure 1 shows that it is also satisfied. (In this figure and the next, q stands for $d \leftrightarrow v \wedge b$, and the arrows indicate conditional inferences, with the lower-case characters representing auxiliary premisses for these inferences.)

[27] To satisfy (a), one can generally take any $Z \subseteq Y$ of *even* size and request that $Y_{\neg Z}$ is consistent (for this equivalence, see Dokow and Holzman, 2010a, and Dietrich and Mongin, 2010).

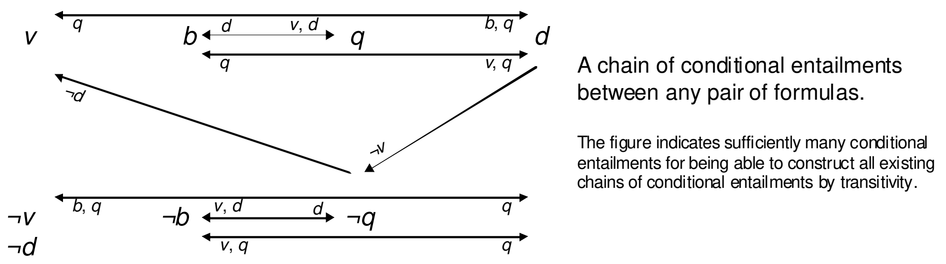

Figure 1: The agenda $\overline{X}$ satisfies (b).

Having exemplified (a) and (b) in their role as *sufficient* conditions for the impossibility, we now illustrate why they are *necessary* by returning to the argument made about the doctrinal paradox. An easy way to pay attention to the legal theorists' insights is to keep a classical propositional formula for the legal doctrine and *make it part of the inference relation instead of the agenda.* Compare with the critical points listed at the end of section 2: this takes care not at all of (i), but fully of (ii), and it goes some way towards accommodating (iii). Put otherwise, the court determines the case from a common doctrine not putting it to vote, and using it rather like a rule of decision. Let, then, a new inference relation $\vdash_{d \leftrightarrow v \wedge b}$ be defined by:

$$S \vdash_{d \leftrightarrow v \wedge b} \text{ iff } S \cup \{d \leftrightarrow v \wedge b\} \vdash \varphi,$$

with a correspondingly reduced agenda:

$$\overline{\overline{X}} = \{v, b, d\}^{\pm}.$$

Given the changes in conditional entailments, $\overline{\overline{X}}$ satisfies (a) but not (b); this is shown by Figure 2. Thus, for this agenda, Theorem 6 entails a *possibility* result, i.e., there exists an $F : D^n \to D$ that is non-dictatorial, unanimity preserving, and independent. An example is the function $\widetilde{F}^n_{maj}$ defined on D^n as follows: for every positive formula $\varphi \in X$, it respects unanimity if either φ or $\neg\varphi$ belongs to all individual judgment sets, and in case of a split choice, it always chooses $\neg\varphi$. Compared with F^n_{maj} in section 3, $\widetilde{F}^n_{maj}$ makes collective judgment sets complete while keeping them consistent.[28]

The statement of Theorem 6 can be simplified when the focus of attention shifts from general F to specific rules. Consider again F_{maj}. We have already

[28] An inconsistent collective judgment set would have to include one of the following minimally inconsistent subsets of $\overline{\overline{X}}$: $\{v, r, \neg d\}$, $\{\neg v, d\}$ or $\{\neg r, d\}$. However, each case is ruled out by $\widetilde{F}^n_{maj}$.

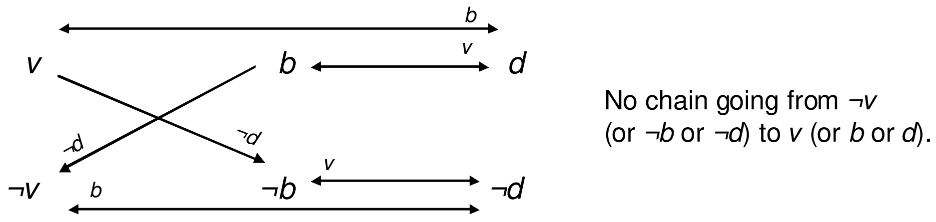

Figure 2: The agenda $\overline{\overline{X}}$ violates (b).

seen that (a), (b) $\Rightarrow$ (Inc). It turns out that (Inc) $\Longleftrightarrow$ (c), where this new condition is:

(c) *There exists a minimal inconsistent set of formulas $Y \subseteq X$ such that $|Y| \geq 3$.*

If X satisfies (c), it is said to be *non-simple*. To check that this is sufficient for (Inc), construct a profile $(A_1, \ldots, A_n)$ such that $Y \subset F_{maj}(A_1, \ldots, A_n)$. To show that it is also necessary, apply the necessity part of Theorem 6 and the fact that (b)$\Rightarrow$ (c), which is seen as follows. For any $\varphi \in X$, (b) entails that there is a chain of conditional inferences:

$$\varphi = \varphi_1 \vdash^* \varphi_2 \vdash^* \ldots \vdash^* \varphi_k = \neg\varphi.$$

In the absence of (c), this chain would reduce to:

$$\varphi = \varphi_1 \vdash \varphi_2 \vdash \ldots \vdash \varphi_k = \neg\varphi,$$

which is impossible because logical inference is a transitive relation by the general logic conditions.

Why has (a) disappeared and (b) been weakened so much when $F = F_{maj}$? Heuristically, this must relate to properties of that rule that Theorem 6 does not mention, and two of them stand out, which are monotonicity and systematicity. The following result, specifically part (iii), supports this analysis.

Theorem 6'. **(i)** If X is even-number negatable and non-simple, there is no $F : D^n \to D$ that satisfies non-dictatorship, unanimity preservation and systematicity. When $n \geq 3$, the conditions are also necessary for the conclusion.

(ii) If X is path-connected, there is no $F : D^n \to D$ that satisfies non-dictatorship, unanimity preservation and monotonic independence, and this is also necessary for the conclusion.

(iii) If X is non-simple, there is no $F : D^n \to D$ that satisfies non-dictatorship, unanimity preservation and monotonic systematicity. When $n \geq 3$, this is also necessary for the conclusion.

Each of these statements has been proved separately, and in particular, (ii) is the version of the canonical theorem established by Nehring and Puppe (in 2010 for the logical aggregation framework, but as early as in 2002 in a related framework of social choice). Today it is better to consider (i), (ii) and (iii) as being partial results leading to Theorem 6. Comparing it with Theorem 6' permits locating *what constraint on X is equivalent to a given axiom placed on F*, and in this way, the trade-off that is so typical of the new theory comes out most rigourously. More can be said to illustrate this trade-off.[29]

All of the preceding results allow for variants based on D^* rather than D. In a nutshell, these turn F into an oligarchic rather than a dictatorial collective judgment function, a somewhat less obvious form of degeneracy. By definition, F is an *oligarchy* if there is a non-empty subset $M \subseteq \{1, \ldots, n\}$ such that, for all $(A_1, \ldots, A_n)$,

$$F(A_1, \ldots, A_n) = \bigcap_{j \in M} A_j.$$

If F is an oligarchy, M is unique and will be called the *set of oligarchs*. Dictatorship is the particular case where M is a singleton. (In section 4, we encountered a stronger and less standard notion of oligarchy.) General logic secures the fact that the intersection of consistent and deductively closed sets retains these properties; as a result, if F is defined on D^n or $(D^*)^n$ and it is an oligarchy, then its range is D^*. As nothing is specified to settle disagreements between the oligarchs, F will often produce incomplete collective judgment sets. This can be seen, e.g., from the unanimity rule F^n_{maj}, which corresponds to the maximal set of oligarchs.

Formally, non-dictatorship is replaced by:

Non-oligarchy. F is not an oligarchy,

and the following impossibility theorems ensue.

Theorems 7 and 7'. The statements are those of Theorems 6 and 6', with $F : D^n \to D$ being replaced by $F : D^n \to D^*$ or $F : (D^*)^n \to D^*$, and non-dictatorship being replaced by non-oligarchy.

[29] Here are two more variants. Dietrich and List (2010b) weaken systematicity in statement (i) by requiring it only for pairs $\varphi, \neg\varphi \in X$ instead of, generally, pairs $\varphi, \psi \in X$. At the same time, they strengthen the agenda conditions of (i) by adding that X should be non-separable in some appropriate (and mild) sense. In their social-choice theoretic framework, Nehring and Puppe (2005) essentially prove a variant of (iii) relying on the same trade-off between weakening systematicity and adding non-separability.

These various extensions can be found in Dietrich and List (2008a) and Dokow and Holzman (2010b).[30] Like the initial results, they have counterparts in social choice theory, and we now turn to these comparisons. Put briefly, each logical aggregation theorem induces a social choice theorem *via a suitably selected logical calculus.* It will typically be a fragmentary classical predicate calculus, whose language has one or more binary relation symbols to represent preferences. Axioms formulated in this language will capture the properties of preferences that one is willing to assume, such as transitivity and the like. The inference relation of the chosen predicate calculus will have to be augmented with those preference axioms, in exactly the same way as the inference relation of the judiciary example was made to include the legal doctrine formula. That is the method followed by Dietrich and List (2007b) to derive from (the sufficiency part of) Theorem 6 a partial version of Arrow's theorem, in which there occur only *strict* preferences on both the individual and collective sides. They introduce a classical predicate calculus, whose language $\mathcal{L}_{\succ}$ is built from basic formulas $x \succ y$ ("x is strictly preferred to y") and whose inference relation $\vdash_{\succ}$ incorporates the three properties of asymmetry, transitivity and completeness. As an agenda $X_{\succ} \subset \mathcal{L}_{\succ}$, they simply take the set of basic formulas. The proof consists in showing, first, that conditions (a) and (b) hold of $X_{\succ}$, and second, that Arrow's "social welfare function", with its relevant set of axioms, can be associated with a collective judgment function F meeting the conditions of Theorem 6.[31]

With somewhat different techniques, Theorem 7 was also put to work on preference relations. Dokow and Holzman (2010b) show that it entails novel versions of Gibbard's (1969) theorem on oligarchies, and by a detour, that very theorem itself. Unlike the application just covered, this one involves *weak* preference relations, i.e., allows for the possibility of indifference. Recall that Gibbard proved that if the collective preference relation is required to satisfy *quasi-transitivity*, i.e., the transitivity of its strict preference part, regardless of the other forms of transitivity, then Arrow's conditions entail that there is an oligarchy rather a dictator. Here, an *oligarchy* is defined as any group of individuals which, for any preference profile, imposes strict preferences that are unanimously agreed in the group, and vetoes strict preferences that contradict the strict preferences of any member of the group.[32] As a somewhat unexpected by-product, Dokow and Holzman obtain Ar-

[30] The early oligarchic result of Gärdenfors (2006) imposes unnecessarily strong conditions on the agenda for impossibility.

[31] In their introductory article, List and Polak (2010) stepwise reformulate the standard proof of Arrow's theorem in order to get one for Theorem 6. This is another way of connecting the two results.

[32] In the above notation, $x \succ y$ holds if this is agreed, and $x \succeq y$ ("x is strictly preferred or indifferent with y") holds if at least one member agrees with $x \succ y$. Gibbard's unpublished theorem has gained fame owing to Sen (1970, ch.4, and 1986).

row's theorem in its integrity, that is for *weak* preferences, thus completing Dietrich and List's programme.[33]

If one adds another example from Dietrich and List (2008b), which concerns Sen's (1970) liberal paradox, and the already discussed work on voting rules, one has virtually exhausted the current stock of applications to social choice theory. Note that they all involve abstract domains of alternatives and preferences rather than specialized "economic" or "political" domains.[34] Clearly the more concrete the domain, the more problematic it is to describe by means of a logical language, and this sets a limit to the applications that can be hoped for. Still, within the scope of the theory, the technique that consists in specializing $\mathcal{L}$ to a preference language is both easy and promising, and much remains to be done along this line.

We have stated the results of this section in terms of the general logic, which gives them wide applicability, but this presentation does not accurately reflect the historical process of discovery, which went through various technical hypotheses, each of them more restrictive than ours. Dokow and Holzman, for their part, use a formalism called *abstract aggregation*, which goes back to Fishburn and Rubinstein (1986), Wilson (1975) and Guilbaud (1952). Starting from a finite number k of propositions that correspond to the positive formulas of our agendas, they render the individual and collective judgments concerning these propositions by the values 0 or 1 that the individuals or the collectivity attribute to them. Thus, after fixing an arbitrary order on propositions, they can reduce the aggregative problem to the study of subsets of $\{0,1\}^k$ and of functions defined from these subsets. If $\mathcal{E} \subset \{0,1\}^k$ represents the set of admissible judgment sets, then $G : \mathcal{E}^n \to \mathcal{E}$ represents a collective judgment function, the analogue of our $F : D^n \to D$. All conditions imposed on F can be redefined to bear on G.

Such a terse statement of the aggregative problem yields quick and elegant proofs, as Guilbaud had already foreshadowed, but it tends to erase the *logical* and *linguistic* properties of judgments, along with certain conceptual distinctions that flow from these properties. The stage of defining the agenda is absorbed into the - one would expect, later - stage of defining what judgment sets are allowed. Sometimes, the same $\mathcal{E}$ corresponds to different agendas. For example, with $k = 2$, take the set

$$\mathcal{E} = \{(1,1),(0,1),(0,0)\}.$$

[33] Dietrich (2007b) has an alternative derivation in a rich framework of logical aggregation, where he assumes that formulas have relevance relations in addition to their logical relations. In still a different framework, Nehring (2003) derives a version of Arrow's theorem for weak preferences that involves a monotonic addition to independence of irrelevant alternatives.

[34] Such as those described by Gaertner (2006) and Le Breton and Weymark (2011).

In classical propositional logic, there are at least two agendas that could give rise to this, i.e.,

$$X = \{a, a \vee b\}^{\pm} \text{ and } X' = \{a \wedge b, a \rightarrow b\}^{\pm},$$

and it would be a conceptual abuse to treat them as they were the same. (This example comes from List and Puppe, 2009.) Another relevant distinction, that between D and D^*, cannot be stated in the abstract aggregation framework as naturally as it is in the logical framework. As a secondary technical issue, the initial assumption of a finite number k of propositions is too sweeping. Thanks to the flexible use of compactness, the general logic here has an advantage, its cost being a certain unwieldiness.[35]

The set-theoretic framework just discussed should not be confused with those formalisms which differ from ours, much less drastically, by replacing the *syntactical* account of the logic (in terms of inference or inconsistent sets) by a *semantic* account (in terms of valuations or related model-theoretic ideas). For example, Pauly and van Hees (2006) describe individual and collective judgments in terms of Boolean valuations, rather than complete and consistent sets of Boolean formulas, but this is just an expository choice, and we have unproblematically translated their work into the present framework.[36]

This section brings the theoretical development of the discursive dilemma to a close. The initial insights were, in section 3, that it was rather a trilemma, and in section 4, that the omitted branch was independence, not systematicity. As it now appears from the canonical theorem and its variants, it is really a *tetralemma*, with the definition of the agenda as the fourth branch, because it can be also resolved by dropping either condition (a) or (b). It now remains to be seen if the doctrinal paradox can be submitted to a such a thorough analysis.

7 Back to the doctrinal paradox

Briefly put, the doctrinal paradox differs from the discursive dilemma by separating premisses and conclusions, and in its judicial application, by taking legal doctrine to be central and specific. We will reexamine both aspects in turn and show that logical aggregation theory can be revised so as to pay

[35] For more technical comparisons between the two frameworks, see Dokow and Holzman (2009).

[36] There are many other ways in which the theory lends itself to the logician's work. Here are two recent examples. Pauly (2007 and 2009) reformulates the acceptance of formulas in terms of a modal operator, rather than by set-theoretic membership, as is done here. Cariani, Pauly and Snyder (2008) define on collective judgment functions a condition of language invariance that leads to a new impossibility result.

attention to them. In this way, the doctrinal paradox will receive a theoretical treatment of its own, while the gap with the discursive dilemma will to some extent be filled.

In the framework adopted here, the distinction between premisses and conclusions can only be made clear if at least one axiom applies differently to the corresponding sets of formulas. For this axiom, the canonical theorem leaves the choice between independence and unanimity preservation, since non-dictatorship is used only to state impossibility. Let us represent the premisses and conclusions by two subsets P and C of the agenda X. *We will discriminate between P and C in terms of independence, reserving it to P, while keeping unanimity preservation applied to the whole of X.* In the judiciary example, the decisions about the issues are taken following a majority vote on each of them, and by imposing both independence and unanimity preservation on P, we merely generalize this fact abstractly. By contrast, according to the legal theorists, a decision on the case can be taken differently from by a formal vote, be it simple majority or otherwise. Respecting consensus seems to be the only norm that the procedure must then guarantee, and this is what our single condition placed on C conveys.

Also, Kornhauser and Sager, as opposed to List and Pettit, do not always want to subject the legal doctrine to a formal vote by the judges. In their view, agreement may result on it as informally as it does on the case. Then, unanimity preservation strikes one again as being *the* suitable condition - no less than this, because it is a bare normative minimum, and no more either, because the informal procedure is left unspecificied. Thus, if it makes sense to locate the legal doctrine in C rather than P, our assumptions will capture another significant legal insight.

Theorem 8 below shows that, under agenda conditions close to those of the canonical theorem, the set of axioms just discussed forces F to be a dictatorship, and Theorem 9 delivers the corresponding oligarchic result. Further results that we do not state here reproduce the two impossibility conclusions when independence gives way to systematicity, monotonic independence, or monotonic systematicity. This analysis, due to Dietrich and Mongin (2010), actually *includes* the canonical theorem and its variants, because the chosen axiomatization is weaker - independence or its reinforced versions being only applied to $P \subseteq X$ - and because the agenda conditions turn out to be the canonical ones in the special case $P = X$.

Formally, we define $P = \{p, q, \ldots\}^{\pm}$ as any non-empty set of X that is closed by negation, and put $C = X \setminus P$. The new axioms on F restate existing ones in terms of this partition.

Systematicity (resp. Independence) on premisses P: now only for each pair of formulas $p, q \in P$ (resp. every formula $p \in P$).

Unanimity preservation on premisses P (resp. conclusions C): now only for every formula $p \in P$ (resp. every formula $p \in C$).

Non-dictatorship on premisses P: there is no $i = 1, \ldots, n$ such that $F(A_1, \ldots, A_n) \cap P = A_i \cap P$ for every $(A_1, \ldots, A_n) \in D^n$.

Non-oligarchy on premisses P: there is no nonempty subset $M \subseteq \{1, \ldots, n\}$ such that $F(A_1, \ldots, A_n) \cap P = (\bigcap_{j \in M} A_j) \cap P$ for every $(A_1, \ldots, A_n) \in D^n$.

The new agenda conditions also take the partition of P and C into account.

(a_P) *There exist a minimally inconsistent set of formulas $Y \subset X$ and distinct premiss formulas $p, q \in Y \cap P$ such that $Y_{\neg\{p,q\}}$ is consistent.*

(b_P) *For every premiss formulas $p, q \in P$, there exist premiss formulas $p_1, \ldots, p_k \in P$ such that*

$$p = p_1 \vdash^* p_2 \vdash^* \ldots \vdash^* p_k = q.$$

(Remember that $p \vdash^* q$ means $Y \cup \{p\} \vdash q$ for some $Y \subset X$; thus, the formulas of P can be logically related by means of those of C.)

(c_P) *There exists a minimally inconsistent subset of formulas $Y \subset X$ such that $|Y \cap P| \geq 3$.*

Let us say that X is *even-number negatable* (resp. *path-connected*, *non-simple*) *in the premisses* P if (a_P) (resp. (b_P), (c_P)) holds.

Theorem 8. If X is both even-negatable and path-connected in the premisses P, there is no $F : D^n \to D$ that satisfies:

- on premisses P, non-dictatorship, independence and unanimity preservation,
- on conclusions C, unanimity preservation.

When $n \geq 3$, the agenda conditions are necessary for this conclusion.

Theorem 9. The statement is the same with $F : D^n \to D^*$ instead of $F : D^n \to D$, and non-oligarchy on premisses P instead of non-dictatorship on premisses P.

As pointed out, Theorems 6 and 7 follow from setting $P = X$ in Theorems 8 and 9, and Theorems 6' and 7' are similarily recovered (see Dietrich and Mongin, 2010, for details). The agenda characterization for F_{maj} is also included; for it (b_P)$\Rightarrow$(c_P) holds, and (c_P) is necessary and sufficient for $F_{maj}(A_1, \ldots, A_n) \cap P$ to be inconsistent for some profile $(A_1, \ldots, A_n)$.

On the conceptual level, the analysis puts the doctrinal paradox to the test, and it escapes consolidated. There is no need for majority voting in order for the premiss-based and the conclusion-based methods to conflict: it is enough to state some of the broad conditions that they satisfy. This parallels the statement already made on the discursive dilemma in sections 3 and 4, but here is a more specific comment. Many writers - among them Pettit (2001) and a majority of legal theorists[37] - have expressed a considered preference for the premiss-based over the conclusion-based method. In their view, the objection that the former is not fully responsive to the individuals carries less weight than the objection that the latter is open to collective irrationality. These were the opposite considerations adduced in section 2. What Theorems 8 and 9 add to this discussion is that the premiss-based method cannot be made more responsive to individuals, since it collapses into impossibilities with a modicum of extra-responsiveness (unanimity preservation on conclusions); so its democratic weakness appears to be irreparable.

Thus far, in order to keep the analysis more general, we have made no restriction on what distinguishes a premiss from a conclusion. Now, to recover ordinary notions, it is enough to assume that formulas in C bear relevant logical connections with those in P. The following agenda condition does the job: it requires that complete and consistent judgments sets be axiomatized by the premisses they contain.

(d_P) *For every $B \in D$,*

$$B = \{\varphi \in X \mid B \cap P \vdash \varphi\}.$$

Let us say that X is *logicaly split by premisses* P if (d_P) holds. Adding this agenda conditions to the others strengthens the impossibilities, for instance turning dictatorship on P into dictatorship (i.e., on the whole of X).

Corollary to Theorem 8. If X is logically split by premisses P, and X is both even-negatable and path-connected in P, there is no $F : D^n \to D$ that satisfies non-dictatorship, and

- for P, independence and unanimity preservation,
- for C, unanimity preservation.

When $n \geq 3$, the agenda conditions are necessary for this conclusion.

From this Corollary, Theorem 4 follows if one specializes the general logic into a classical propositional calculus. In this case, the choice of $P = PV_X$ automatically satisfies (d_P). Section 4 had not fully stated the conditions that Theorem 4 placed on X; the missing ones are actually (a_P) and (b_P).

[37] See the review of legal theorists in Nash (2003). For their part, Kornhauser and Sager choose differently according to the circumstances. They also recommend that the court take a "meta-vote" on the procedure first.

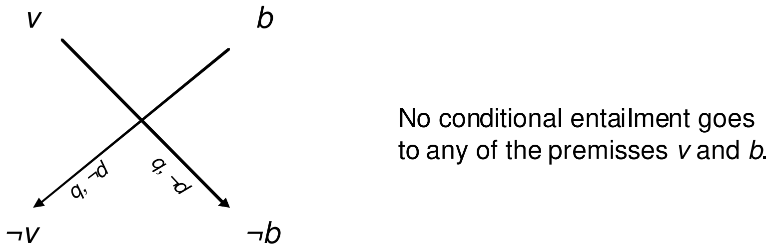

Figure 3: The agenda $\overline{X}$ with $P = \{b, v\}^{\pm}$ violates (b_P)

This ex post derivation establishes the generality of the analysis in another way.

The agenda conditions (a_P), (b_P), (c_P), (d_P) can be illustrated by new variants of the judiciary example, and this discussion is also useful in clarifying models of the legal doctrine. To begin with, take the agenda $\overline{X}$, which contains the Boolean formula $q = d \leftrightarrow v \wedge b$. On the one hand, if $P = \{v, b\}^{\pm}$, then (b_P) is violated, and the court's decision escapes from the impossibility, contrary to what the canonical theorem would predict here. This is illustrated in Figure 3. The choice of P amounts to making the legal doctrine part of the conclusions, which is permissible since (d_P) does not hold and the partition between P and C can be interpreted procedurally (the doctrine is not submitted to a vote or any formal procedure). Note that (c_P) is also violated, hence F_{maj} no longer has any drawback (at least if n is odd). On the other hand, if $P = \{v, b, q\}^{\pm}$, then (b_P) is satisfied. This is illustrated by Figure 4. Also, (a_P) is satisfied; so, the court's decision falls back into the impossibility predicted by the canonical theorem. Unlike the previous one, this choice P satisfies (d_P) and thus turns the partition with C into an intuitive division of premisses and conclusions.

Now, take $\overline{\overline{X}} = \{v, b, d\}^{\pm}$, i.e., the agenda that is associated with the modified inference relation $\vdash_q$ of last section. With $P = \{v, b\}^{\pm}$, a natural choice, (b_P) is violated, as Figure 5 shows. Even (c_P) is violated, which makes F_{maj} unobjectionable. Also, (d_P) is satisfied, which makes the partition interpretable in terms of premisses and conclusions as normally understood. This case captures the popular solution to the doctrinal paradox that consists in applying the premiss-based method alone.

The agendas just discussed illustrate Theorem 8 while having some legal relevance, but it is doubtful that any of them represents legal doctrine

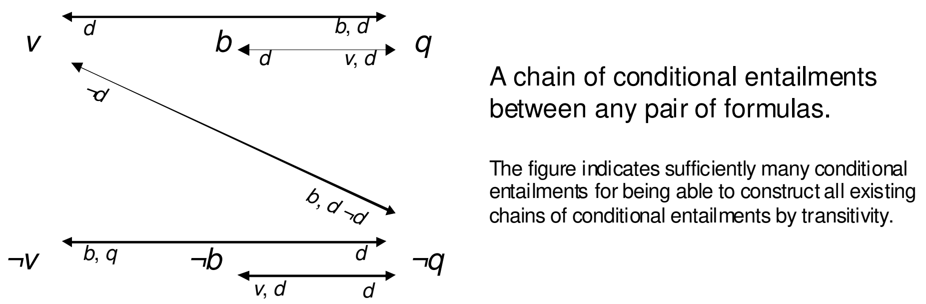

Figure 4: The agenda $\overline{X}$ with $P = \{b, v, q\}^{\pm}$ satisfies (b_P).

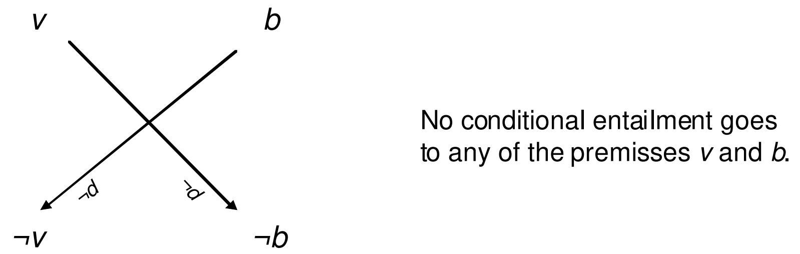

Figure 5: The agenda $\overline{\overline{X}}$ with $P = \{b, v\}^{\pm}$ violates (b_P).

appropriately. At the end of section 2, we asked - see question (i) - whether it could be rendered by classical logic. We now argue that the Boolean biconditional $\leftrightarrow$ in the formula $q = d \leftrightarrow v \wedge b$ has undesirable effects that can be avoided by the non-Boolean operator $\hookleftarrow\!\hookrightarrow$ of conditional logic. The argument parallels that which Dietrich (2010) uses more generally in favour of such logics.[38]

To see what goes wrong with agenda $\overline{X}$, let us pick up those minimal inconsistent subsets which contain q:

$$\begin{array}{ll} Y_1 = \{\neg b, d, q\}, & Y_2 = \{\neg v, d, q\}, \\ Y_3 = \{v, b, \neg d, q\}, & Y_4 = \{q, \neg q\}, \\ Y_5 = \{v, b, d, \neg q\}, & Y_6 = \{\neg v, \neg d, \neg q\}, \\ Y_7 = \{\neg b, \neg d, \neg q\}. & \end{array}$$

That Y_5, Y_6, Y_7 are inconsistent is somewhat counterintuitive given the legal context. In each of these sets, judges deny that d is equivalent to v and b,

[38] As is well-known, the calculi of conditional logic overcome the paradoxes of "material" (Boolean) implication, another example of which is given in the next paragraph. We may only refer to the classic work by Stalnaker (1968) and Lewis (1973). For a review, see Edgington (2008) or Nuete and Cross (2001).

and this allegedly clashes with certain positions they take on v, b or d. It seems that they may *consistently* deny the equivalence and accept these positions. Specifically, suppose that they have in mind another issue s that is not mentioned here and hold that d is equivalent to $v \wedge b \wedge s$. In this case, they may deny that d is equivalent to $v \wedge b$ and nonetheless:

- accept v, b, s and d (contrast with the alleged inconstency of Y_5);
- reject v and d regardless of how they judge b and s (contrast with the alleged inconsistency of Y_6);
- reject b and d regardless of how they judge v and s (contrast with the alleged inconsistency of Y_7).

To put it otherwise, the following theorem of classical propositional logic:

$$(*)\ \neg(d \leftrightarrow v \wedge b) \dashv\vdash \neg d \leftrightarrow v \wedge b$$

contradicts normal intuitions of legal deliberation. Now, returning to the table, we see there is nothing intuitively wrong with the sets Y_1 to Y_4, or with the corresponding theorems:

$$\begin{aligned}(**)\ & d \leftrightarrow v \wedge b, v, b \vdash d \\ & d \leftrightarrow v \wedge b, d \vdash v \\ & d \leftrightarrow v \wedge b, d \vdash b.\end{aligned}$$

One would indeed expect that *accepting* the doctrine, as opposed to *refusing* it, entail the consequences claimed formally by classical propositional logic.

The calculi of conditional logic axiomatize the conditional, $\hookrightarrow$, and so the biconditional, $\hookleftarrow\!\hookrightarrow$, in a way that exactly fits these divided intuitions. They give rise to a list of minimally inconsistent subsets that is reduced to Y_1–Y_4, or equivalently, only retain $(**)$ excluding $(*)$. It is not necessary to decide between the various different systems, since all satisfy the general logic and any of them can do for the purpose. We would have then to replace $\overline{X}$ with $\overline{X}' = \{v, b, d, q'\}$, where $q' = d \hookleftarrow\!\hookrightarrow v \wedge b$. Remarkably, when Theorem 6 is applied to this agenda, it is seen to be even-number negatable, but not path-connected, so that *the negative conclusion obtained for* $\overline{X}$ *no longer holds*; see Figure 6 and compare it with Figure 1. Similarly with Theorem 8, $\overline{X}'$ is *not path-connected in premisses* for any of the choices for P that we have envisaged. Thus, the negative conclusion is again beaten back.

Now, what about the agenda $\overline{\overline{X}} = \{v, b, d\}^{\pm}$ when the equivalence formula placed in the inference rule is non-classical? The rule defined by

$$S \vdash_{q'} \varphi \text{ iff } S \cup \{q'\} \vdash \varphi,$$

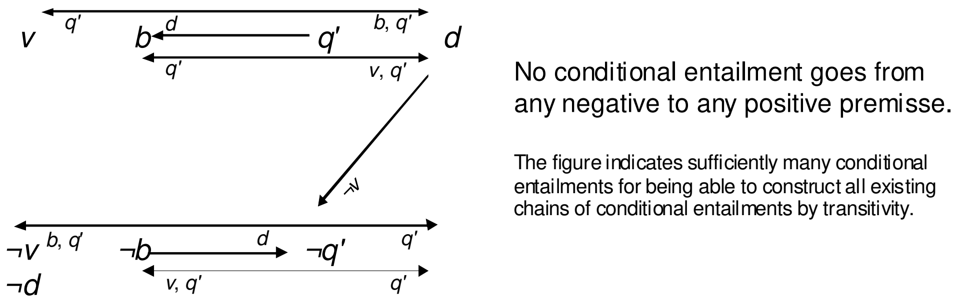

Figure 6: The agenda $\overline{X}'$ violates (b).

leads to the same violations of path-connectedness as $\vdash_q$ did above. Indeed, only *accepting* the doctrine is a possibility, and as transpires from what has been said, classical and non-classical equivalences collapse onto each other in this case.

To summarize the main results of this section, when legal doctrine is *internal* to the agenda (case $\overline{X}'$ versus $\overline{X}$), departing from classical logic gives way to additional possibilities, but nothing is gained when legal doctrine is *external* (case $\overline{\overline{X}}$ with either $\vdash_q$ or $\vdash_{q'}$). Is there some reason to decide between internal and external representations, as there is one to decide between classical and non-classical calculi? We do not think so. The discursive dilemma automatically imposes the internal representation, and section 2 questioned this choice, but this was not to say that the opposite one was compelling. Actually, each may be justified according to the circumstances: legal theory only suggests that judges do not *normally* vote on the doctrine, not that they never do so. In sum, the usually best model is given by $\overline{\overline{X}} = \{v, b, d\}^{\pm}$ with the rule $\vdash_{q'}$ (here equivalent to $\vdash_q$), but $\overline{X}'$ will nonetheless be sometimes appropriate. The only agenda we exclude is $\overline{X}$, that of the discursive dilemma, which signals where our analysis departs from the more familiar one.

8 Conclusion and some open questions

Throughout this account of logical aggregation theory, our guiding heuristic was that the doctrinal had been underrated, compared with its discursive dilemma variant, and that it called for its own analytical treatment. We have shown how the current work, as epitomized by the canonical theorem, could be revised so as to take notice of the paradox and deepen its explanation. This move illustrates the flexibility and expressive power of the framework collectively put in place in the 2000s. Notice however that using logic is essential to the changes we suggest, and not all current contributors approve of the logical turn taken by judgment aggregation theory.

By and large, the doctrinal paradox appears to be less of a problem for collective judgment than does the discursive dilemma. The simplest reason is that the premiss-based approach, which is not even definable in the context of the latter problem, offers a satisfactory way out in many occurrences of the former. As the last section showed, the impossibility part of our theorem applies to the judiciary example only for a dubious logical rendering of the logical doctrine. However, this unfavourable case is a warning that the premiss-based method is not immune to impossibilities, contrary to what is generally believed. Legal theorists and political philosophers still have to take this finding into account.

Returning to the more standard work, we should emphasize that it has not yet reached its final stage. To begin with, the canonical theorem has fixed a format of results that is not yet applied everywhere. In particular, the early theorems should be revisited. They provide only sufficient agenda conditions for the impossibilities they state, and because they derive them without the help of unanimity preservation, they are covered neither by the canonical theorem nor by our generalization. In bringing them to the format, one may hope to clarify two theoretical issues, i.e., what agenda conditions are both necessary and sufficient for independence to entail unanimity preservation, and what impossibilities, if any, surround independence when this entailment does not hold.

In social choice theory, independence of irrelevant alternatives and the Pareto conditions are logically independent conditions, and an impossibility has famously been derived by Wilson (1972) from the former condition alone. Thus, by answering the previous group of questions, one would further tighten the connection with the antecedent theory. As the existing derivations of Arrow's, Gibbard's and Sen's theorems indicate, it is natural to work from logical aggregation theory to social choice theory, and much remains to be done in this way despite the limitations imposed by logical languages. However, one may wonder whether the other direction is feasible. Could a suitably doctored variant of Arrow's impossibility theorem entail a corresponding result in logical aggregation theory? Many believe that this reverse programme is a non-starter, but few have actually tried their hands at it.

A no doubt more pressing task would be to complement today's negative conclusions by a richer array of positive solutions. It would be nonsense to complain that the theory is *exclusively* negative, since every theorem stated in the canonical format can be read in the positive way, as we illustrated at length with judiciary examples. However, after so much emphasis laid on agenda conditions, more work should be done on the axioms put on the collective judgment function. Computer scientists have opened an interesting avenue when defining *merging rules on belief sets*, and several writers - starting with Pigozzi (2002) - have recommended that logical aggregation theory borrow from this technology. A belief set is essentially the same as a judgment set in the syntactical formulation developed here, and an important class of merging rules, i.e., the *distance-based* ones, are especially easy to accommodate by this formulation. They amount to minimizing the distance from the collective set to the given profile, where the notion of distance between two sets can be defined variously from the logical language; see Konieczny and Pino-Perez (2002) for basic principles and Miller and Osherson (2008), as well as Lang, Pigozzi, Slavkovik, and van der Torre (2011), for relevant elaborations.

In these schemes, the impossibility of logical aggregation is circumvented by giving up independence, and there are normative arguments to support this line. One of them is that formula-wise aggregation leaves no room for taking into account *the reasons* that individuals have to accept or reject a proposition (Mongin, 2008). This should be balanced against the technical advantage that independence prevents strategic manipulations of agendas (Dietrich, 2006). Social choice theorists have also debated independence of irrelevant alternatives, but the two discussions are not parallel, and this can be explained by their different objects.[39] To relax unanimity preservation would lead to different possibilities, but regrettably, these have been hardly discussed thus far. There is also a normative argument to support this line, as was mentioned earlier.

The most drastic resolution of all is to move from the logical to the *probabilistic* framework, as many economists brought up in the Bayesian tradition would no doubt recommend. This move is somehow comparable with the change undergone by social choice theory when the "social welfare functional", defined on profiles of individual utility functions, replaced the "social welfare function" defined on profiles of individual preference relations.[40] Inspection of the existing results for this richer framework shows that the same selection of axioms, mutatis mutandis, leads to non-trivial combinations of probability measures instead of dictatorships or oligarchies. However, these positive solutions degenerate for suitable strengthenings of the axioms, and the richer framework needs justifying anyhow. In social

[39] The received argument against independence of irrelevant alternatives is that it prevents taking into account preference intensities, and the received argument for it that it blocks the individuals' strategic misreporting of their preferences. For a critical reexamination, see Lehtinen (2011).

[40] See d'Aspremont (1985), Sen (1986), and specifically, d'Aspremont and Gevers (2002).

choice theory, the observer or planner may be unable to define the "informational basis" that would allow him to make trade-offs between the individuals' conflicting interests, and by the same token, the group's representative may be unable to quote the numerical degrees of certainty that would allow him to balance the individuals' conflicting opinions against each other.

9 REFERENCES

Arrow, K.J. (1951), *Social Choice and Individual Values*, New York, Cowles Foundation and Wiley, 2nd ed., 1963.

d'Aspremont, C. (1985), Axioms for Social Welfare Orderings, in *Social Goals and Social Organization*, ed. by Hurwicz, L., Schmeidler, D., Sonnenschein, H., Cambridge, Cambridge University Press.

d'Aspremont, C., Gevers, L. (2002), Social Welfare Functionals and Interpersonal Comparability, in *Handbook of Social Choice and Welfare*, ed. by Arrow, K.J., Sen, A.K., Suzumura, K., Elsevier, I, 449-541.

Barwise, J. (1977), An Introduction to First-Order Logic, in *Handbook of Mathematical Logic*, ed. by Barwise, J., Amsterdam, North Holland, 5-46.

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

Brennan, G. (2001), Collective Coherence?, *International Review of Law and Economics*, 21, 197-211.

Cariani, F., Pauly, M., Snyder, J. (2008), Decision Framing in Judgment Aggregation, *Synthese*, 163, 1-24.

Chang, C.C., Keisler, H.J. (1973), *Model Theory*, Amsterdam, Elsevier.

Chapman, B. (2002), Rational Aggregation, *Politics, Philosophy and Economics*, 1, 337-354.

Condorcet (Nicolas de Caritat, Marquis de) (1785), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Paris, Imprimerie Royale. Also in *Oeuvres*, ed. by Condorcet-O'Connor, A., Arago, M.F., Paris, Firmin Didot, 1847-1849. Partial English transl. by McLean, I., Urkenas, A.B., as *An Essay on the Application of Analysis to the Probability of Decisions Rendered by a Plurality of Votes*, Classics of Social Choice, Ann Arbor, University of Michigan Press, 1995, 91-112.

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

Dietrich, F. (2007a), A Generalized Model of Judgment Aggregation, *Social Choice and Welfare,* 28, 529-565.

Dietrich, F. (2007b), Aggregation and the Relevance of Some Issues for Others, Research Memorandum 002, Maastricht Research School of Economics of Technology and Organization.

Dietrich, F (2010), The Possibility of Judgment Aggregation on Agendas with Subjunctive Implications, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 603-638.

Dietrich, F., List, C. (2007a), Judgment Aggregation by Quota Rules: Majority Voting Generalized, *Journal of Theoretical Politics*, 19, 391-424.

Dietrich, F., List, C. (2007b), Arrow's Theorem in Judgment Aggregation, *Social Choice and Welfare,* 29, 19-33.

Dietrich, F., List, C. (2008a), Judgment Aggregation Without Full Rationality, *Social Choice and Welfare,* 31, 15-39.

Dietrich, F., List, C. (2008b), A Liberal Paradox for Judgment Aggregation, *Social Choice and Welfare,* 31, 59–78

Dietrich, F., List, C. (2010a), Majority Voting on Restricted Domains, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 512-543.

Dietrich, F. and List, C. (2010b), The Impossibility of Unbiased Judgment Aggregation, *Theory and Decision*, 68, 281-299.

Dietrich, F., Mongin, P. (2007), The Premiss-Based Approach to Logical Aggregation, Cahier de recherche HEC 886/2008, Paris.

Dietrich, F., Mongin, P. (2010), The Premiss-Based Approach to Judgment Aggregation, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 562-582.

Dokow, E., Holzman, R. (2009), Aggregation of Binary Evaluations for Truth-functional Agendas, *Social Choice and Welfare,* 32, 221–241.

Dokow, E., Holzman, R. (2010a), Aggregation of Binary Evaluations, *Journal of Economic Theory*, Symposium on Judgment Aggregation, 145, 495-511.

Dokow, E., Holzman, R. (2010b), Aggregation of Binary Evaluations with Abstentions, *Journal of Economic Theory*, Symposium on Judgment Aggregation, 544-561

Duddy, C., Piggins, A. (2009), Many-valued Judgment Aggregation: Characterizing the Possibility/impossibility Boundary for an Important Class of Agendas, Department of Economics WP 0154, National University of Ireland Galway.

Eckert, D., Monjardet, B. (2009), Guilbaud's Theorem: An Early Contribution to Judgment Aggregation, Document de travail 2009.47, Centre d'Economie de la Sorbonne, Paris.

Edgington, D., Conditionals, *The Stanford Encyclopedia of Philosophy* (Winter 2008 Edition), ed. by Zalta, E.N., online (http://plato.stanford.edu/archives/win2008/entries

Fishburn, P. C., Rubinstein, A. (1986), Algebraic Aggregation Theory, *Journal of Economic Theory,* 38, 63-77.

Fleurbaey, M., Mongin, P. (2005), The News of the Death of Welfare Economics is Greatly Exaggerated, *Social Choice and Welfare*, 25, 381-418.

Frege, G. (1918-1919-1923), Logische Untersuchungen, *Beiträge zur Philosophie der deutschen Idealismus*, Hefte I/2, II/3-4, III/1. English transl. in

Gottlob Frege: Collected Papers on Mathematics, Logic, and Philosophy, ed. by McGuinness, B., Oxford, Blackwell, 1984.

Gaertner,W. (2006), *Domain Conditions in Social Choice Theory*, Cambridge, Cambridge University Press

Gärdenfors, P. (2006), An Arrow-like Theorem for Voting with Logical Consequences, *Economics and Philosophy*, 22, 181-190.

Genest, C., Zidek, J.V. (1986), Combining Probability Distributions: a Critique and an Annotated Bibliography, *Statistical Science,* 1, 114-148.

Gibbard, A. (1969), Social Choice and the Arrow Conditions, Unpublished, Department of Philosophy, University of Michigan (http://www-personal.umich.edu/∼gibbard/).

Guilbaud, G.T. (1952), Les théories de l'intérêt général et le problème logique de l'agrégation, *Economie appliquée*, 5, 501-551. To be republished in *Revue économique*, 2012.

Hartmann, S., Pigozzi, G., Sprenger, J. (2010), Reliable Methods of Judgment Aggregation, *Journal of Logic and Computation*, 20, 603-617.

Hartmann, S., Sprenger, J. (forthcoming), Judgment Aggregation and the Problem of Tracking the Truth, *Synthese*.

Herzberg, F.S. (2010), Judgment Aggregators and Boolean Algebra Homomorphisms, *Journal of Mathematical Economics*, 46, 132-140.

Herzberg, F.S., Eckert, D. (2010), Impossibility Results for Infinite-electorate Abstract Aggregation Rules, IMW WP 427, Universität Bielefeld.

Heifetz, A., Mongin, P. (2001), Probability Logic for Type Spaces, *Games and Economic Behavior*, 35, 31-53.

Kant, E. (1800), *Logik, ein Handbuch für Vorlesungen*, herausgegegen by Jäsche, G.B., Königsberg. English transl. in *Kant Lectures on Logic*, *The Cambridge Edition of the Works of Immanuel Kant*, Cambridge, Cambridge University Press, 1992, Part IV, *The Jäsche Logic*, 521-642.

Kirman, A.P., Sondermann, D. (1972), Arrow's Theorem, Many Agents, and Invisible Dictators, *Journal of Economic Theory*, 267–277.

Konieczny, S., Pino-Perez, R. (2002), Merging Information Under Constraints: a Logical Framework, *Journal of Logic and Computation*, 12, 773-808.

Kornhauser, L.A. (1992), Modelling Collegial Courts II: Legal Doctrine, *Journal of Law, Economics and Organization*, 8, 441-470.

Kornhauser, L.A., Sager L.G. (1986), Unpacking the Court, *The Yale Law Journal*, 96, 82-117.

Kornhauser, L.A., Sager L.G. (1993), The One and the Many: Adjudication in Collegial Courts, *California Law Review*, 81, 1-59.

Kornhauser, L.A, Sager, L.G. (2004), The Many as One: Integrity and Group Choice in Paradoxical Cases, *Philosophy and Public Affairs*, 32, 249-276.

Lang, J., Pigozzi, G., Slavkovik, M., van der Torre, L. (2011), Judgment Aggregation Rules Based on Minimization, in *Proceedings of the 13th*

Conference on the Theoretical Aspects of Rationality and Knowledge (TARK XIII), 238-246. Extended version, LAMSADE WP, Université Paris Dauphine, 2011 (www.pigozzi.org/Publications.html).

Le Breton, M., Weymark, J. (2011), Arrovian Social Choice Theory on Economic Domains, in *Handbook of Social Choice and Welfare*, ed. by Arrow, K.J., Sen, A.K., Suzumura, K., Amsterdam, North Holland, II, 191-299.

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

Lehtinen, A. (2011), A Welfarist Critique of Social Choice Theory, *Journal of Theoretical Politics*, 23, 359-381.

Lewis, D. (1973), *Counterfactuals*, Oxford, Blackwell

Lismont, L., Mongin, P. (2003), Strong Completeness for Weak Logics of Common Belief, *Journal of Philosophical Logic*, 32, 115-137.

List, C. (2003), A Possibility Theorem on Aggregation over Multiple Interconnected Propositions, *Mathematical Social Sciences,* 45, 1-13. (Corrigendum, *Mathematical Social Sciences* 52, 109-110.)

List, C. (2004), A Model of Path-Dependence in Decisions Over Multiple Propositions, *American Political Science Review*, 98, 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., Petit, P. (2005), On the Many as One: A Reply to Kornhauser and Sager, *Philosophy and Public Affairs*, 33, 377-390.

List, C., Polak, B. (2010), Introduction to Judgment Aggregation, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 441–466.

List, C., Puppe, C. (2009), Judgment Aggregation: A Survey, in *Oxford Handbook of Rational and Social Choice*, ed. by Anand, P., Puppe, C., Pattanaik, P., Oxford, Oxford University Press.

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

Makinson, D. (2005), *Bridges from Classical Logic to Nonmonotonic Logic*, Texts in Computer Science Series, King's College London.

Miller, M.K., Osherson, D. (2008), Methods for Distance-based Judgment Aggregation, *Social Choice and Welfare*, 32, 575-601.

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

Mongin, P. (1997), Spurious Unanimity and the Pareto Principle, THEMA WP, Université de Cergy-Pontoise.

Mongin, P. (2003), L'axiomatisation et les théories économiques, *Revue économique*, 54, 99-138.

Mongin, P. (2008), Factoring Out the Impossibility of Logical Aggregation, *Journal of Economic Theory*, 141, 100-113.

Mongin, P. (2012), Un précurseur méconnu de la théorie de l'agrégation logique, *Revue économique*, forthcoming.

Monjardet, B. (2003), De Condorcet à Arrow via Guilbaud, Nakamura et les "jeux simples", *Mathématiques et sciences humaines / Mathematics and Social Sciences*, n° 163, 5-32.

Nash, J.R. (2003), A Context-Sensitive Voting Protocol Paradigm for Multimembers Courts, *Stanford Law Review*, 56, 75-159.

Nehring, K. (2003), Arrow's Theorem as a Corollary, *Economics Letters*, 80, 379-382.

Nehring, K. (2005), The (Im)Possibility of a Paretian Rational, Economics WP 0068, Institute for Advanced Study, School of Social Science.

Nehring, K., Puppe, C. (2002), Strategy-Proof Social Choice on Single-Peaked Domains: Possibility, Impossibility and the Space Between, Department of Economics WP, University of California, Davies.

Nehring, K., Puppe, C. (2005), The Structure of Strategy-Proof Social Choice. Part II: Non-dictatorship, Anonymity, and Neutrality, Department of Economics WP, University of Karlsruhe.

Nehring, K., Puppe, C. (2008), Consistent Judgement Aggregation: The Truth-Functional Case, *Social Choice and Welfare*, 31, 41-57.

Nehring, K., Puppe, C. (2010), Abstract Arrowian Aggregation, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 467-494.

Nehring, K., Puppe, C. (2010b), Justifiable Group Choice, *Journal of Economic Theory*, 145, Symposium on Judgment Aggregation, 583-602.

Nute, D., Cross, C.B. (2001), Conditional Logic, in *Handbook of Philosophical Logic*, ed. by Gabbay, D., Guenthner, F., Dordrecht, Kluwer, IV.

Pauly, M. (2007), Axiomatizing Collective Judgment Sets in a Minimal Logical Language, *Synthese*, 158, 233-250.

Pauly, M. (2008), On the Role of Language in Social Choice Theory, *Synthese*, 163, 227-243.

Pauly, M., 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.

Pigozzi, G., Hartmann, S. (2007), Judgment Aggregation and the Problem of Truth-tracking, in *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge* (TARK XI), 248-252.

Priest, G. (2002), Paraconsistent Logic, in *Handbook of Philosophical Logic* (2nd ed.) ed. by Gabbay, D., Guenthner, F., Dordrecht, Kluwer, VI, 287-393.

Priest, G., Routley, R., Norman, J. (ed. by) (1989), *Paraconsistent Logic: Essays on the Inconsistent*, München, Philosophia Verlag.

Sen, A. (1970), *Collective Choice and Social Welfare*, Amsterdam, North Holland.

Sen, A. (1986), Social Choice Theory, in *Handbook of Mathematical Economics*, ed. by Arrow, K.J., Intriligator, M.D., Amsterdam, North Holland, III, 63-77.

Stalnaker, R. (1968), A Theory of Conditionals, in *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, 2, 98-112.

van Hees, M. (2007), The Limits of Epistemic Democracy, *Social Choice and Welfare,* 28, 649-666.

Vickers, J. (1989), An Agenda for Subjectivism, *Erkenntnis*, 31, 397-416.

Wilson, R. (1972), Social Choice Theory without the Pareto Principle, *Journal of Economic Theory*, 5, 478-486.

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