INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Towards a Unified Aggregation Framework for Preferences and Judgements

Luigi Marengo [a]
Simona Settepanella [b]
Yan X. Zhang [c]

[a] LUISS, Rome, Italy
[b] Hokkaido University, Sapporo, Japan
[c] San José University, CA, USA

2018/28 September 2018
ISSN(ONLINE) 2284-0400

Towards a Unified Aggregation Framework for Preferences and Judgements

Luigi Marengo[1], Simona Settepanella[2], and Yan X. Zhang[3]

[1]*Dept. of Management, LUISS, Roma, Italy, l.marengo@luiss.it*
[2]*Dept. of Mathematics, Hokkaido University, Kita-ku, Sapporo, Japan, s.settepanella@math.sci.hokudai.ac.jp*
[3]*Dept. of Mathematics, San Jose State University, San Jose, CA, USA, yan.x.zhang@math.sjsu.edu*

September 2018

Abstract

The “doctrinal paradox”, also called “discursive dilemma”, shows that the aggregation of judgements held by different individuals is problematic and can lead to group-level inconsistencies, although each individual is consistent. This aggregation problem has intuitive similarities with the Condorcet paradox in the aggregation of preferences. Indeed, List and Pettit (2002) proved an impossibility theorem in the framework of judgement aggregation, analogous to Arrow’s Theorem from the framework of preference aggregation. However, List and Pettit (2004) claim that the judgement aggregation framework is “more expressive” than the classical social choice framework, in the sense that while the framework of preference aggregation can be mapped into the framework of judgement aggregation, there exists no obvious reverse mapping. In this paper we show instead that the social choice framework has enough power to express the judgement aggregation framework. To do so, we present a *graph-theoretic* version of the social choice framework and show that it is sufficient to embed the judgement aggregation framework. As an application of this framework, we show that the doctrinal paradox and Condorcet’s paradox (both under the majority aggregation rule) arise for essentially the same reason.

1 Introduction

The classical social choice literature studies the aggregation of individual preferences (usually represented as partial orders or sets of ordered pairs) into collective "social" preferences. The judgement aggregation literature studies the aggregation of individual judgements (assignments of True or False to sets of logical propositions) into a collective (e.g. of a jury or a committee) judgement.

The "paradox" that aggregating "rational" preferences can lead to "irrational" outcomes has been known at least since the 17th century (de Caritat marquis de Condorcet, 1785). Condorcet considers a case in which individual voters have transitive preferences over outcomes and shows that aggregate preferences constructed via majority voting may fail to comply with transitivity. A general formal framework has been provided by Kenneth Arrow, whose impossibility theorem proves that, in general, no aggregation mechanism (except, trivially, dictatorship by one individual) exists which ensures transitivity of social preferences even when all individual agents comply with transitivity (Arrow, 1951).

Something similar happens in judgement aggregation. The "doctrinal paradox", first formulated by Kornhauser and Sager (1993) (though it is possible to find some partial antecedents) concerns instead the problem of aggregation of judgements. Individual agents have heterogeneous believes on "atomic" propositions and are correctly computing some logical operations (and, or, not, logical implication, etc.) on such propositions. The "paradox" happens when the aggregation on the atomic propositions and their logical expressions produces "illogical" aggregate results.

List and Pettit (2002) formalized this paradox and proved an impossibility result somewhat similar to Arrow's theorem. In a subsequent paper which compares the two theorems (List and Pettit, 2004), List and Pettit made the strong claim that the judgement aggregation framework is "more powerful" than the one of preference aggregation, in that their model of judgement aggregation includes the one of preference aggregation, but, in general, the reverse is not true. They also compare the doctrinal paradox and the Condorcet paradox, arguing that the two paradoxes are fundamentally different.

In this paper we try to challenge these conclusions by using a different formalization of the social choice problem of preference aggregation. We use directed graphs to model preferences and we show that, within this framework, judgement aggregation can indeed be represented as a special case and that the Condorcet and doctrinal paradoxes are strongly intertwined.

The paper is organized as follows: in Section 2, we present the preference aggregation (or social choice) framework with its Condorcet paradox and the judgement aggregation framework with its doctrinal paradox. In Section 3 we illustrate the main motivation of this paper. In Section 4, we introduce our *graph-theoretic framework* and show that we can incorporate the judgement aggregation framework inside the graph-theoretic framework. Finally, we show in Section 5.4 that the Condorcet and doctrinal paradoxes happen for essentially the same reason. We conclude with some discussion and directions for further research in Section 6.

We emphasize that we do not wish to diminish the importance of the many valuable contributions to the judgement aggregation literature: (List and Pettit, 2002, 2004; Dietrich and List, 2007a,b, and many others). Our main goal is to show that the classical social choice framework is not too different from the judgement aggregation framework in terms of expressive power and that many classical results in the social choice literature can be extended to the judgement aggregation framework (and vice-versa). Last but not least, we submit that our formalization is a particularly powerful alternative to the already existing ones.

2 Preliminaries

2.1 The preference aggregation framework

There are n *voters* $\{1, 2, \ldots, n\}$ voting on a set X of *alternatives* according to individual preferences. We assume that such individual preferences are antisymmetric and transitive. Let $\mathcal{O}_>$ be the set of *total orders* on X and $\mathcal{O}_\geq$ be the set of *partial orders* on X. Hence if m is the cardinality of X, an element o of $\mathcal{O}_\geq$ is equivalent to $\binom{m}{2}$ *pairwise preferences* $o(i, j)$, with $i \neq j \in X$, where:

- $o(i, j) \in \{1, -1, 0\}$;
- $o(i, j) = 1, -1$, or 0 respectively if o prefers i to j, j to i, or is indifferent between i and j respectively.
- (Anti-symmetric) $o(i, j) = -o(j, i)$ for all $i \neq j$;
- (Transitive) If $o(i, j) = o(j, k) = 1$, then $o(i, k) = 1$.

If $o(i, j) = 1$ or -1, we say that o has a *strict* pairwise preference. If all the $o(i, j)$ are strict, we say that o is an element of $\mathcal{O}_>$, the *total orders* on X. In this case $o(i, j) \in \{1, -1\}$, since indifference is ruled out. Total orders can also be considered as permutations and partial orders can also be considered as ordered set partitions. We now assume that each voter has a *preference* in $\mathcal{O}_\geq$, where the conditions of being a total order capture the idea that the preferences are "rational".

The social choice literature studies *social preference functions*, i.e. functions $f\colon \mathcal{O}_\geq^n \to \mathcal{O}_\geq$ which aggregate n preferences into a single collective preference. We remark that there are different variations of this set-up which use different domains and codomains (for example, the celebrated Gibbard-Satterwaite Theorem looks at *social choice functions* $f\colon \mathcal{O}_>^n \to X$ instead, and there are versions of Arrow's Theorem which examine functions such as $f\colon \mathcal{O}_>^n \to \mathcal{O}_>$), but our definition of social preference functions is a quintessential model.

The social choice literature requires individual and social preferences to satisfy some additional properties. The most common ones include:

- (Pareto) A social preference function f is *Pareto* if for all $1 \leq i \leq n$ and some x, y we have $o_i(x, y) = 1$, we must also have $o(x, y) = 1$ where $o = f(o_1, \ldots, o_n)$.
- (IIA / Indifference of Irrelevant Alternatives) A social preference function f has *IIA* if whenever if $(o_1, \ldots, o_n)$ and $(o_1^*, \ldots, o_n^*)$ are two lists of preferences in the domain of f and there exists x and y such that for all i, $o_i(x, y) = o_i^*(x, y)$, then if $o = f(o_1, \ldots, o_n)$ and $o^* = f(o_1^*, \ldots, o_n^*)$, we must have $o(x, y) = o^*(x, y)$.

A common, more visually intuitive reformulation of the social choice framework is to represent each preference $o \in \mathcal{O}_{\geq}$ as a directed graph on $|X| = m$ vertices corresponding to the alternatives, where we draw a directed edge $y \to x$ if and only if $o(x, y) = 1$. Not all directed graphs of m alternatives appear as potential preferences (because for instance they may violate transitivity). Thus, the graphs that do correspond to preferences form a strict "rational" subset of the graphs on m vertices. In particular, these "rational" graphs cannot have cycles: any sequence of directed edges of the form

$$v_1 \leftarrow v_2 \leftarrow \cdots \leftarrow v_k \leftarrow v_1$$

corresponds to a sequence of pairwise preferences

$$o(v_1, v_2) = \cdots = o(v_{k-1}, v_k) = o(v_k, v_1) = 1$$

which violates the transitive property. Thus, social preference functions aggregate n "rational" graphs (preferences) into another "rational" graph. Our proposed generalization in Section 4 maintains this formulation of the social choice framework.

2.2 The Condorcet paradox

One particular candidate for a social preference function is the following *majority rule* function: given preferences $o_1, \ldots, o_n$, for all $(j, k) \in X \times X$, define $o(j, k) \in \mathcal{O}_{\geq}$ to be the majority of $o_i(j, k)$ over all i. Then define o to be our aggregate social preference. Note that we have to allow elements in $\mathcal{O}_{\geq}$ in case there are ties. Thus, $o(j, k)$ are "majority rule" pairwise preferences between alternatives j and k.

The Condorcet "paradox" (de Caritat marquis de Condorcet, 1785), states that under majority rule $o(j, k)$ may have cycles when $|X| > 2$, which means it cannot come from a partial order $o \in \mathcal{O}_{\geq}$. Formally, this means that the "majority rule" is not, in general, a valid social preference function, always capable of producing a rationally consistent "consensus preference".

Example 2.1. Table 1 shows an example of the Condorcet paradox on 3 alternatives $\{A, B, C\}$ with 3 voters; note that it suffices to describe 3 out of 6 pairwise preferences for each voter because the others are determined by anti-symmetry. The "majority rule" aggregation then fails transitivity, because $o(A, C)$ should equal 1 from $o(A, B) = O(B, C) = 1$, but $o(C, A) = 1$, a contradiction. We can transform the above into the graphs in Figure 1.

	$o(A,B)$	$o(B,C)$	$o(C,A)$
Voter 1	-1	1	1
Voter 2	1	-1	1
Voter 3	1	1	-1
Majority	1	1	1

Table 1: Table of preferences for the Condorcet paradox.

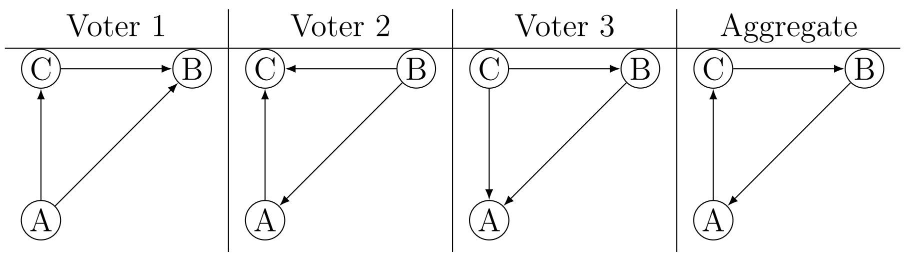

Figure 1: Graphical representation of Table 1. Note the aggregate graph contains a cycle (Condorcet paradox).

2.3 The judgement aggregation framework

In the judgement aggregation framework, we have n *judges* $\mathcal{J} = \{j_1, j_2, \ldots, j_n\}$ who are called to express judgements on a set $\Xi = \Xi_a \cup \Xi_c$ of $(m+s)$ logical propositions, which is the (disjoint) union of the set $\Xi_a = \{P_1, \ldots, P_m\}$ of *atomic* propositions and the set $\Xi_c = \{C_1, \ldots, C_s\}$ of *compound* propositions. A *compound* proposition is something in the boolean algebra generated by logical operations on atomic propositions, such as $P \wedge Q$, $P \vee Q$, $\neg P \rightarrow Q$, etc. We assume that such compound propositions are *nontrivial*; that is, they cannot be constant but must be surjective functions onto $\{0, 1\}$, where 0 stands for "False" and 1 stands for "True".

Atomic propositions generally represent basic facts (i.e. P*: the suspect is guilty of "breaking"*; Q*: the suspect is guilty of "entering"*), compound propositions represent laws or rules (e.g. $P \wedge Q$*: the suspect is guilty of "breaking and entering" when (the suspect is guilty of "breaking") and (the suspect is guilty of "entering")*). A judgement is *logical* if and only if given the judgement's atomic propositions, the value of the judgement's compound propositions agrees with the respective rules.

Each judge judges each proposition to be 1 (True) or 0 (False). Formally, each judge's *judgement* is a function[1] $\Xi \rightarrow \{0, 1\}$ that assigns a truth value to each proposition. Equivalently, we can think of

$$J = (J_a || J_c) = (J(P_1), \ldots, J(P_m), J(C_1), \ldots, J(C_s))$$

[1]List and Pettit encode judgements using sets of propositions while we use a functions to preserve coherence with the rest of the paper. The two encodings are equally expressive.

(the $||$ denotes concatenation) as a vector of length $(m+s)$, with the first m coordinates forming a vector J_a corresponding to atomic propositions and the last s coordinates forming a vector J_c corresponding to compound propositions.

Similarly to voters aggregating individual "rational" preferences into social choices, we want judges to aggregate "logical" judgements into a jury decision. Given a set Ξ of propositions, let $\mathbb{U}_\Xi$, the *logical judgements*, be the set of judgements J where for each $1 \leq i \leq s$, if $C_i \in \Xi_c$ is determined by the function $\hat{C}(P_1, \ldots, P_m)$, then $J(C_i) = \hat{C}(J(P_1), J(P_2), \ldots, J(P_m))$ where $J(P_i) \in \{0,1\}$ denote the value of J on proposition P_i. List and Pettit call this property "deductive closure" as it ensures that the judge's opinions are internally consistent.

As an example, suppose we have $m = 2$ atomic propositions $P_1 = P$ and $P_2 = Q$ and $c = 1$ compound proposition $C_1 = P \wedge Q$. Let a judge j_i have judgement J_i. In order to simplify notation, we will denote by $J_{a,i} = (J_i)_a$ and $J_{c,i} = (J_i)_c$. Suppose $J_{a,i} = (0,0)$ and J_i is logical, then its value on the unique compound proposition C_1 assigned by j_i must be $J_{c,i} = (0 \wedge 0) = (0)$. Thus, the final logical judgement is $J_i = (0,0,0)$.

We remark that, since for any $J_a \in \{0,1\}^m$ there exists one and only one element $J_c \in \mathbb{U}_{\Xi_c}$ such that $J = J_a || J_c \in \mathbb{U}_\Xi$, this implies that $\mathbb{U}_\Xi$ is isomorphic to $\{0,1\}^m$ as a set, and can be thought of as vertices of the m-dimensional Hamming cube in $\mathbb{R}^m$.

We now define *judgement aggregation functions* to be functions $f \colon \mathbb{U}_\Xi^n \to \mathbb{U}_\Xi$. In other words, judgement aggregation functions aggregate n logical judgements into a single logical judgement and are analogous to social preference functions in the social choice framework.

Like social preference functions, also judgement aggregation functions are usually required to satisfy some additional properties:

- (Unanimity) A judgement aggregation function f is unanimous if whenever every judge holds the same judgement, then the latter is also be the aggregate judgement. Formally: if $J_i(P) = x$ for all J_i with $x \in \{0,1\}$, then $J(P) = x$. Notice that unanimity implies $f(J, \ldots, J) = J$.

- (Propositionwise Independence) As defined in Dietrich and List (2013), a judgement aggregation function f is propositionwise independent if there exist $(m+s)$ functions $f_i \colon \{0,1\}^n \to \{0,1\}$, $1 \leq i \leq (m+s)$, such that whenever $f(J_1, \ldots, J_n) = J$, for each i we have

$$(J)_i = f_i((J_1)_i, \ldots, (J_n)_i).$$

  This property captures the notion that if judges want to judge whether a defendant committed a specific crime, it suffices to ask all the judges whether the defendant committed that specific crime (as opposed to also needing judgement regarding other crimes).

- (Systematicity) As defined in List and Pettit (2002), a judgement aggregation function f satisfies the property of systematicity if there exists a single function

$\tilde{f}\colon \{0,1\}^n \to \{0,1\}$ such that whenever $f(J_1,\ldots,J_n) = J$, for each i we have

$$(J)_i = \tilde{f}((J_1)_i,\ldots,(J_n)_i).$$

In particular, note that systematicity is a special case of propositionwise independence.

Notice that propositionwise independence can also be given a different interpretation: f is propositionwise independent if and only if the table of propositions (e.g Table 2) is "commutative" in the sense that we get the same result if:

- we first compute the value of a compound proposition C_i for each judge and then we aggregate (i.e. we first move horizontally in the table of propositions from atomic to compound propositions and then aggregate vertically), or

- we first aggregate the judgements on each atomic proposition and then we compute the value of C_i on these aggregated values (i.e. we first move vertically on the table of preferences aggregating preferences of each judge and then compute the compound proposition horizontally).

In Section 4 we will discuss how these conditions relate to the analogous conditions imposed on the preference aggregation functions in social choice.

2.4 The doctrinal paradox

Like in social choice, a majority rule seems a natural candidate for aggregation also in the domain of judgements. *Majority judgement* simply states that the jury's judgement on a proposition (either atomic or compound) corresponds to the judgement held by the majority of judges (juries with an odd number of judges are often used to avoid ties). The doctrinal paradox was presented in Kornhauser and Sager (1986) and has some important real life examples such as the famous US Supreme Court case *Arizona vs. Fulminante.* It describes the possibility that majority aggregation on either atomic or derived compound propositions may lead to different results.

Example 2.2. Consider the following simple example (List, 2012): suppose there are three judges, $N = \{1,2,3\}$, and three propositions, $\Xi = \{P,Q,R\}$, where P stands for "the defendant was contractually obliged not to do action W", Q for "the defendant did action W", and R for "the defendant is liable for breach of contract". Assume legal doctrine requires that the premises P and Q are jointly necessary and sufficient for the conclusion R. Suppose that the individual judgements are given by Table 2. We can immediately notice that majority aggregation on P and Q produces True for both, and therefore would lead logically to the conclusion that also R is True. But majority aggregation on R produces False. There is therefore an inconsistency between the **premise-based aggregation** and the **conclusion-based aggregation**, which is why this phenomenon is called the **doctrinal paradox**.

	P (obligation)	Q (action)	R (liability)($:\iff (P \wedge Q)$)
Judge 1	True	True	True
Judge 2	False	True	False
Judge 3	True	False	False
Majority	True	True	False

Table 2: Judgements leading to a doctrinal paradox.

3 A common framework for preference and judgement aggregation

In section 4 of their 2004 paper, List and Pettit address the question of whether the preference aggregation and the judgement aggregation frameworks can be mapped into each other (List and Pettit, 2004, pp. 215–220).

They answer negatively by showing counter-examples in which either we have a doctrinal but not a Condorcet paradox or vice versa. In this and the following section we argue that by adding to the judgement aggregation framework the plausible assumption of proposition-wise consistency (which List and Pettit discard because it may lead to incomplete preferences) and by using our graph theoretic model, preference and judgement aggregations can indeed be mapped into each other.

3.1 List and Pettit's example

List and Pettit consider again the judgement aggregation problem represented in Table 2 and assume further that the preferences orderings of judges $1, 2$ and 3 over the possible pairs of judgements on the questions p and q are as follows:

1. **Judge 1**: True-True > False-True >True-False >False-False
2. **Judge 2**: False-True > True-True >True-False >False-False
3. **Judge 3**: True-False > False-True >True-True >False-False

then they show that such a preference structure does not generate a Condorcet paradox while indeed it leads to a discursive dilemma.[2]

[2]If on the contrary judges had for instance the following preferences there would be a Condorcet paradox, but not a discursive dilemma:

1. **Judge 1**: False-False > True-False > False-True > True-True
2. **Judge 2**: False-True > False-False >True-False > True-True
3. **Judge 3**: True-False > False-True > False-False > True-True

The problem with this approach is that Table 2 only provides the best preferred alternative for each judge, but does not say anything on the remaining preferences. In the following subsection we will show how, starting from example in Table 2, it is possible to retrieve a partial order on the set of couples {True-True, False-False, False-True, True-False} $= \{0,1\}^2$ for True$=1$ and False$=0$ sufficient to aggregate preferences via majority rule in such a way that discursive dilemma arises if and only if Condorcet paradox arises.

3.2 Example of our main result: Graphical judgement aggregation

Before recasting List and Pettit's example in our framework, let us remark that what they outline is a preference structure over objects made of two elements, i.e. the beliefs on propositions p and q. Social choice over multidimensional objects has been studied, among the others, in Marengo and Pasquali (2011) and Marengo and Settepanella (2014). Take, for example, the textbook case of a group of friends deciding "what to do tonight". The textbook example would give a list of alternatives, such as "restaurant, movies, concert", but in reality these are "complex" objects made of several interacting dimensions. For instance the object "going to the restaurant" includes type of food, price range, at what time, with whom, etc. Such elements typically involve non trivial interdependencies: for instance we have friends with whom we like to go to the movies but not to the restaurant or vice versa (e.g. because our cinematographic tastes are aligned while the culinary ones are incompatible).

Consider for example a choice on only two items: at what time to go out and where to go. If each one of these two variables can assume two values, e.g. "go out at 8pm" or "go out at 10pm","go to restaurant" and "go to the movies", then we are left with 4 possible choices representing all combinations. Then we have 2 propositions $p =$"where to go" $q =$"when to go" that can assume each one two values that we can call, for simplicity, $0, 1$. If we assume that they exclude the possibility to "go to the restaurant at 10pm", we are left with 3 alternatives that we can call $A = 00 =$"go to the restaurant at 8am", $C = 10 =$"go to watch a movie at 8pm" and $B = 11 =$"go to watch a movie at 10am". If we keep preferences according to Example 2.1 then we get Table 3 which is graphically represented in Figure 2. Notice that horizontal edges have the meaning of "preferences on what time to go" while vertical edges represents preferences on "where to go".

3.3 Example of our main result: graphical judgement aggregation

In order to build a partial order on the set of preferences $\{0,1\}^2$ starting from judgements in Table 2 we assume the judge has a *proposition-wise consistency* when his/her judgement is converted to preferences, i.e. we require that each judge to consistently prefer his/her own choice on each proposition (either atomic or compound), when

	$o(A,B)$	$o(B,C)$	$o(C,A)$
Voter 1	11	11	10
Voter 2	00	10	10
Voter 3	00	11	00
Majority	11	11	10

Table 3: Table of preferences for the Condorcet paradox in a multi-dimensional decision.

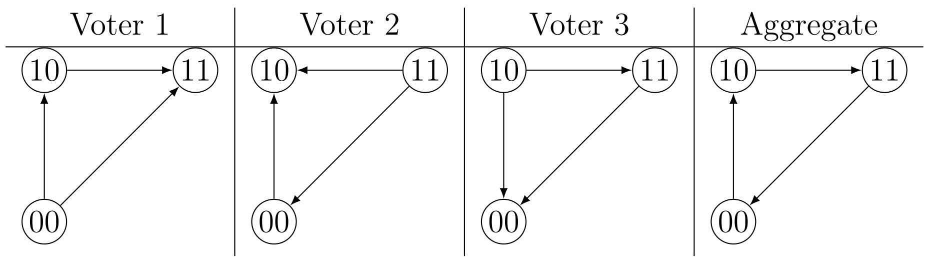

Figure 2: Graphical representation of Table 3. Note the similarity to Figure 1.

the value of other propositions is fixed. Formally, if J is the set of judgements of a judge on the m atomic propositions and the judge is called to decide between two possibilities $v, v' \in \{0,1\}^m$, such that the i-th entries $v_i = v'_i$ for any $i \neq t$ and $v_t = J_t$ and $v'_t \neq J_t$ then the judge will prefer v to v'.

For example, if the judge's beliefs on the atomic propositions (p, q) are True-True, then, when choosing between two judgements which differ only by a single atomic proposition, she will always prefer the judgement which assigns True to such a proposition. Thus her preferences will be: True-True>True-False>False-False and True-True>False-True>False-False. Whereas True-False and False-True will not be comparable.

Moreover, given the value of the compound proposition $r = p \wedge q$, then the judge will be consistent with this result preferring values of couples p and q determining the same value of r of her own preferred choice. For example if her belief is True-True for the atomic propositions p and q, then her preferred value of r is True and, since value of r is False in any other couple different from True-True, we get True-True > True- False, True-True > False - False and True-True > False-True.

One of List and Pettit's original objections to the idea of embedding the judgement framework into the social choice framework is that they found it impossible to get a total order from the judgement framework. While we do agree on this, the key point of our objection is that a total order is actually not necessary to get a Condorcet paradox. We will use a graph-theoretic framework to show this.

Now, if we assume proposition-wise consistency, the judgements of judges $1, 2$ and 3 in Table 2 corresponds to graphs in Figure 3. Aggregation by majority rule

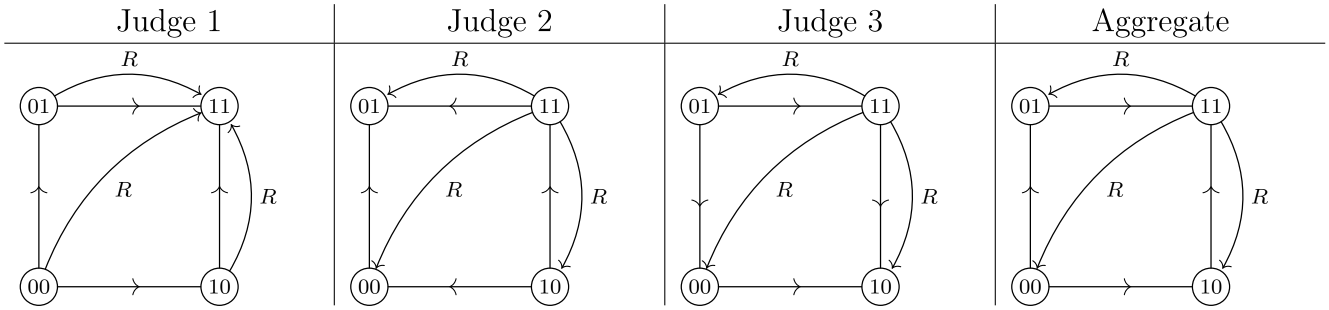

Figure 3: The graphs G_i of preferences of judges $j_i, i = 1, 2, 3$ in Table 2 and their aggregated graph (under majority rule). Curved edges labelled R represent preferences on the compound proposition $R = P \wedge Q$.

gives rise to a cycle involving vertices $(0,0), (0,1)$ and $(1,1)$ corresponding to a cycle between False-False, False-True and True-True. This is a Condorcet cycle which corresponds to the discursive dilemma in Table 2. In Section 5.4 we will give a general proof of the equivalence between the two paradoxes beyond this particular example.

4 The graph-theoretic framework

We are now ready to spell out the details of our graph-theoretic framework which allows us to model both preferences and judgements as an s-tuple of directed graphs. In the case of judgements each of the s directed graphs corresponds to a different compound proposition. The case of preferences corresponds to the special case where $s = 1$. Thus, the graph-theoretic framework is a slight generalization of the social choice framework that allows $s > 1$. We now show that this is enough to encompass the judgement aggregation framework as well.

We have n individuals (e.g. voters or judges), a set of N alternatives, where each alternative is an m-dimensional object, and a set of s *labels* for s-tuple of directed graphs. Each individual is characterized by an s-tuple of *preference graphs*, each one with N vertices corresponding to the alternatives. In the case of preferences $s = 1$ and the s-tuple of graphs will contain only one element.

We will consider some subset of preference graphs as *rational*. We define *rational aggregation functions* to be functions that aggregate rational preference graphs to rational preference graphs. Then the preferences (seen as graphs) from the social choice framework are exactly the rational graphs, and the social preference functions are exactly the the rational aggregation functions.

Recall that Ξ contains atomic propositions P_j, $j = 1, \ldots, m$ and compound propositions C_j, $j = 1, \ldots, s$. Given a judgement $J = (J_a || J_c) \in \{0,1\}^{m+s}$, we build s associated graphs $G_i(J) = (V(J), E_i(J))$, each of them corresponding to each compound proposition C_i, as follows:

1. Each $G_i(J)$ has the same set of vertices $V(J)$, which consists of the 2^m vertices $v \in \{0,1\}^m$ of the m-dimensional cube in $\mathbb{R}^m$. As discussed in Section 2.3, these correspond to the 2^m potential choices of J_a.

2. There is an edge (v, w) from v to w,$v \neq w$, is in $E_i(J)$ if and only if one of the following occurs:

   i) they differ only for the value of one entry t, and the value $(J)_t$ of the t entry of J equals the t-entry $(w)_t$ of w (and hence $(v)_t = \neg(J)_t$). This captures the idea of *judge systematicity* from Section 3.3. We call such an edge an *atomic* edge.

   ii) $C_i(v) \neq C_i(w)$, $C_i(w) = (J_c)_i$, $C_i(v) \neq (J_c)_i$. We call such an edge a *compound* edge.

Note that all the graphs G_i have the same vertices and atomic edges. We call the set of all graphs on 2^m vertices $Gr(m)$. Let $G : \mathbb{U}_\Xi \to (Gr(m))^s$ be the map that sends a logical judgement J to the s-tuple of such graphs $(G_1, \ldots, G_s)$.

Remark 4.1. Because each judge j_i has a logical judgement $J_i \in \mathbb{U}_\Xi$, the above correspondence associates to each judge preference J_i its own s-tuple of graphs $(G_1(J_i), \ldots, G_s(J_i))$. The following considerations apply:

1. in order to draw atomic edges we assume the *proposition-wise consistency* defined in Section 3.3 in the judge's preferences, i.e. we assume that the preferred value of the judge j_i on proposition P_j is independent from the values of other propositions. That is, if the value of any proposition P_k, $k \neq j$, is left unchanged (the Hamming distance between vertices representing those preferences is 1) then the judge will always prefer the vertex in which the proposition P_j has his or her preferred value;

2. point ii) states that the judge is consistent with the value of each compound proposition obtained by its atomic propositions, i.e. the judge will always prefer the vertex v such that the value of compound proposition $C_i(v)$ agrees with his/her judgement of C_i. Notice that no edge is drawn between v and w if they have the same aggregate value in compound proposition C_i.

Example 4.2. We first give a simple example showing we indeed generalized the social choice framework. Reconsider Example 2.2 from Table 2. As we only have one compound proposition, $s = 1$, so each judge j_i has only one associated graph G_i. We again get G_1, G_2, G_3 in Figure 3.

Example 4.3. Consider now the case where we have two atomic propositions P and Q and two compound propositions, $C_1 = P \wedge Q$ and $C_2 = P \oplus Q$ (recall that $P \oplus Q$ is XOR/exclusive-OR; that is, it is true if and only if P or Q are true, but not both). We still use majority judgement as aggregation, which gives us Table 4. We see that

	P	Q	$C_1 = P \wedge Q$	$C_2 = P \oplus Q$
Judge 1	True	True	True	False
Judge 2	False	True	False	True
Judge 3	True	False	False	True
Majority	True	True	False	True

Table 4: Table of judgements for Example 4.3.

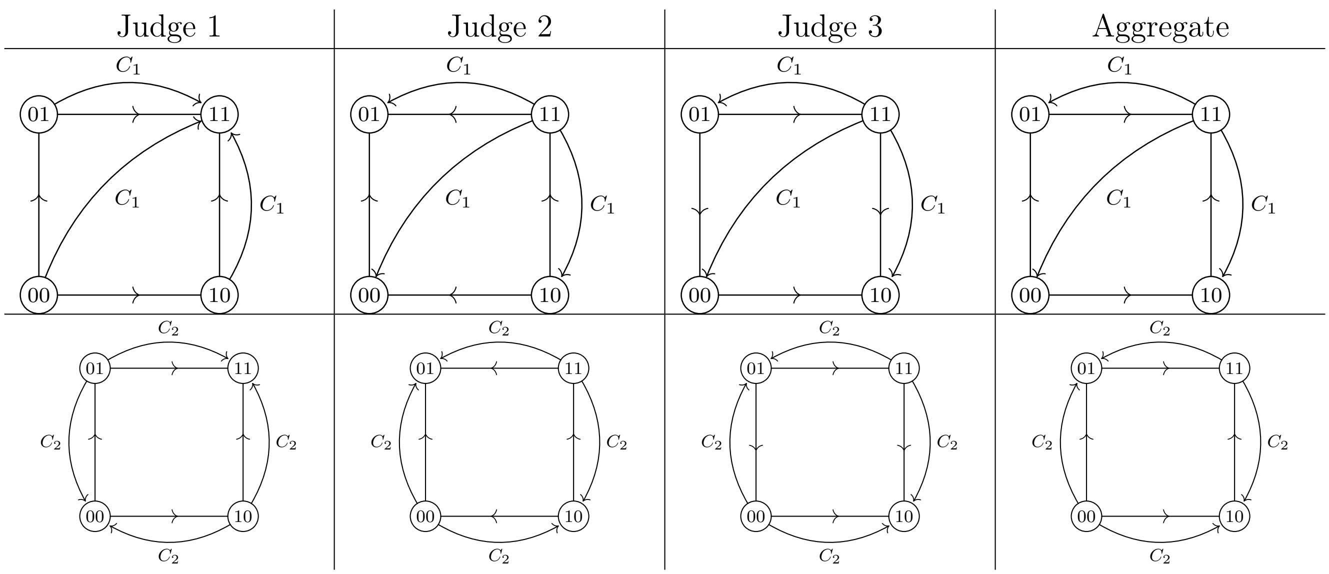

Figure 4: The k-th row contains the graphs $G_{i,k}$ of preferences of judges $j_i, i = 1, 2, 3$, in Table 4 and their aggregation (under majority rule). The graphs $G_{i,1}$ are exactly the G_i in Figure 3. Curved edges labelled by C_i represent preferences on the compound propositions $C_i = P \wedge Q$ and $C_2 = P \oplus Q$.

the aggregated judgement is again not logical (for both propositions). Now, each judge j_i's judgement J_i is translated into two graphs $G(J_i) = (G_{i,1}, G_{i,2})$ because we have 2 compound propositions, and the aggregated judgement J is also translated into two graphs $G(J) = (G_1, G_2)$. The graphs $G_{1,1}, G_{2,1}, G_{3,1}$ are exactly the graphs in Figure 3 because they correspond to $P \wedge Q$. The new graphs $G_{1,2}, G_{2,2}, G_{3,2}$ and their aggregation can be seen in Figure 4.

In Section 5.1, we will prove that logical judgements must have a "most-preferred choice" (which we will call *global optimum*) with no edges going out. It is easy to check that the aggregated graphs in Figure 4 have no such vertex.

We remark that we could have detected the logical inconsistency of the aggregated judgement by looking at each compound proposition individually. Indeed, in order for a judgement to be consistent, it has to be consistent in each compound proposition. This will be elucidated in the following section.

We denote by $Gr(m)^G = G(\{0,1\}^{m+s}) \subset Gr(m)^s$ the set of s-tuples of graphs $(G_1, \ldots, G_s)$ obtainable from any judgement (even non-logical ones) in $\{0,1\}^{m+s}$. Our main result is:

Theorem 4.4. *The function G that associates to each element $J \in \{0,1\}^{m+s}$ its s-tuple of graphs $(G_1(J), \ldots, G_s(J))$ gives a bijection between $\{0,1\}^{m+s}$ and $Gr(m)^G$. G naturally induces a bijection $\tilde{G}$ between sets of functions*

$$\{f | f\colon (\{0,1\}^{m+s})^n \to \{0,1\}^{m+s}\} \simeq \{f_{gr} | f_{gr}\colon (Gr(m)^G)^n \to (Gr(m)^G)\}.$$

Proof. Suppose $G(J_a||J_c) = G^1$ and $G(J'_a||J'_c) = G^2$. Suppose $J_a \neq J'_a$, then it is clear that the set of atomic edges in the graphs are also different. If $J_c \neq J'_c$, then there exists some i such that $(J_c)_i \neq (J'_c)_i$. Because C_i is nontrivial, we also know there exist $v \neq w \in \{0,1\}^m$ such that $C_i(v) = (J_c)_i$ and $C(w) = (J'_c)_i$. Consider vertices labelled with v and w in the two graphs G^1_i and G^2_i. By the definition of G, in G^1_i we have a compound edge $w \to v$ and in G^2_i we have a compound edge $v \to w$. Thus the two tuples of graphs must be different if $J_c \neq J'_c$. We now conclude that if $G^1 = G^2$, we must have $J_a = J'_a$ and $J_c = J'_c$. Equivalently, G is one-to-one. Thus, G gives a bijection from the domain $\{0,1\}^{m+s}$ to the range $Gr(m)^G$.

Because of this bijection, when given input $G^1, \ldots, G^n$, each belonging to $Gr(m)^G$, each G^i must be equal to some $G(J_i)$ for a unique judgement $J_i \in \{0,1\}^{m+s}$. Thus,

$$\begin{aligned} f_{gr}(G^1, \ldots, G^n) &= f_{gr}(G(J_1), G(J_2), \ldots, G(J_n)) \\ &= G(f(G^{-1}(G(J_1)), G^{-1}(G(J_2)), \ldots, G^{-1}(G(J_n)))) \\ &= G(f(J_1, J_2, \ldots, J_n)) \end{aligned}$$

is a well-defined function that has codomain in $\{0,1\}^{m+s}$. □

The main strength of Theorem 4.4 is that we can completely package the function f into the language of directed graphs on 2^m vertices. This addresses our goal of showing that we can indeed embed the original definition of judgement aggregation into a generalization of the social choice framework. In particular, we can define the rational graphs in our model to be the total graphs that come from logical judgements. Because of Theorem 4.4, we know that the judgement aggregation functions from the judgement aggregation framework are simply the rational aggregation functions in this instance of the graph-theoretic framework.

5 Applications

In Section 4 we showed that we can embed the judgement aggregation framework into the graph-theoretic framework. In this section we discuss some straightforward applications.

5.1 Detecting logical judgements with global optima

A first interesting application of our graph-theoretic framework is that it allows to extend to judgement aggregations problems results already known in graph/theoretic models of social choice. In particular, some recent results concerning local and global optima in social choice among multi-dimensional alternatives (Marengo and Settepanella, 2014; Amendola and Settepanella, 2012; Amendola et al., 2015) have a direct equivalent in judgement aggregation.

We call a *global optimum* of a graph $G = (V, E)$ a vertex $v \in V$ such that:

i) for all $w \in V$ it exists a path $w \to v$, i.e. a sequence of edges $(w, v_1), (v_1, v_2), \ldots, (v_n, v) \in E$, and

ii) for all $w \in V$, $(v, w) \notin E$.

Note that property ii) implies that a global optimum, if it exists, is unique.

Now, it turns out that this intuitive notion of global optimum (which can be visually checked once the graph has been drawn) corresponds to logical consistency of judgements, as shown by the following theorem:

Theorem 5.1. *A judgement $J = (J_a||J_c) \in \{0, 1\}^{m+s}$ is logically consistent if and only if each of its associated graphs $G_i(J)$ has a global optimum. Moreover, when all the graphs $G_i(J)$ have a global optimum, the global optimum is the same vertex for all of them and it is exactly the vertex J_a.*

Proof. Assume J is logically consistent. For any i, let $v = J_a$ be the vertex corresponding to the atomic judgement part J_a of $J = (J_a, J_c)$. For any other vertex w, by the construction of the graph, if we have an edge from v into w, by the rules of the graph construction we must have w disagree with v (resp. $C_i(w)$ disagree with $C_i(v)$) on an atomic (resp. compound) proposition on which w agrees with J_a (resp. $C_i(w)$ agrees with J_c), which is impossible because v corresponds to J_a (resp. $C_i(v)$

agrees with $(J_c)_i$ because J is logical). Furthermore, at least one path will always exist from any $w \neq v$ to v via atomic edges: simply change the atomic propositions one at a time. Thus, $v = J_a$ is a global optimum for every G_i.

Now, suppose J is not logical. Because J is not logical, there is some i such that $C_i(J_a) \neq (J_c)_i$. We look at the graph G_i. Suppose it has a global optimum v. By the atomic edges, we must have $v = J_a$. This means $C_i(v) \neq (J_c)_i$. Because C_i is nontrivial, there must be some $w \in \{0,1\}^m$ such that $C_i(w) = (J_c)_i$. By definition, we have a compound edge from v to w, which gives a contradiction. Thus, G_i has no global optimum. □

Thus, by theorem 4.4, we know we can do judgement aggregation over graphs instead of tables of True/False values. Theorem 5.1 makes such a perspective even more useful: we can now translate the idea of logical consistency in judgement aggregation to the graph-theoretic idea of global optima.

5.2 The total graph

In the graph-theoretic framework, we characterize each individual by an s-tuple of preference graphs, with N vertices corresponding to the alternatives. Instead of drawing s graphs $(G_1, \ldots, G_s)$, we can also define a single graph, the *total graph* G_T, by just including all the edges from all the $E(G_i)$ and labelling the edges to avoid ambiguity. Then we can also consider each individual to be characterized by a single preference graph with N vertices and, possibly, labelled edges.

In the case of preferences we assume that all edges are equally labelled, while in the case of judgements each edge of the directed graph G_T is labelled by the proposition (either atomic or compound) it refers to. If $P_j, j = 1, \ldots, m$, are atomic propositions and $C_j, j = 1, \ldots, s$, are compound propositions, we can get G_T from the collection of all G_i labelling by C_i the compound edges in G_i and by P_j the atomic edges corresponding to the atomic proposition P_j (recall that the atomic edges are common to all $G_i, i = 1, \ldots, s$). As an example, the total graphs $G_T(J_i)$ and $G_T(J)$ associated, respectively, to the judgement of judges j_i and their majority rule aggregation in Example 4.3 are given in Figures 5.

We can get G_T from the collection of all G_i or vice-versa, so G_T contains the same information as the tuple of graphs $(G_1, \ldots, G_s)$. The total graph G_T stores the information more efficiently and is more mathematically natural for certain statements and proofs, especially those where the roles of atomic and compound edges are not important (because the total graph treats the two types of edges equally). An example is the results in the next section comparing judgement and preference aggregation concepts. On the other hand, the s-tuple of graphs can be a better model when studying judgement aggregation. Indeed, for example, Theorem 5.1 tells us that it is sufficient that only one of the G_i does not admit a global optimum to state that the judgement is not logical. Also, when we want to think of the atomic and compound propositions as different, the s-tuple of graphs makes more sense.

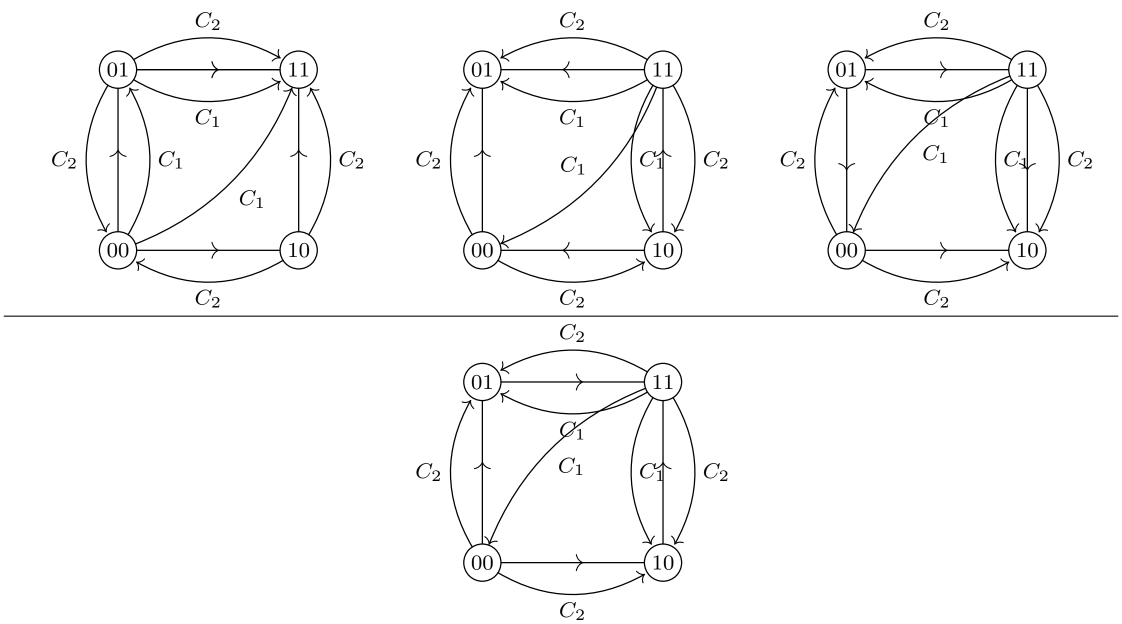

Figure 5: The first row contains the total graphs $G_T(J_i)$ of preferences of judges $j_i, i = 1, 2, 3$ in Table 4. The second row contains their aggregation under majority rule. Note a pair of vertices may have more than 1 compound edge between them. For the sake of simplicity the labels of the atomic propositions are omitted, even though they are technically labelled as well.

5.3 Comparing judgement and preference aggregation concepts

The following corollary shows that concepts from judgement and preference aggregations are actually equivalent within our graph-theoretic framework.

Corollary 5.2. *In each row in the following table, the left property (from the social choice framework) and the right property (from the judgement aggregation framework) correspond to the same property under the graph-theoretic framework. Recall that the definitions for the properties can be found in Section 2.*

social choice	*judgement aggregation*
Pareto	*unanimity*
IIA	*propositionwise-independence*
majority rule	*majority judgement*

Proof. We omit the proofs since the correspondences are rather trivial. Furthermore, the majority rule vs. majority judgement correspondence will be made more explicit in Section 5.4. □

For a label L and an edge-labelled graph G (such as a total graph G_T), let $l(G, L)$ be the subgraph of G consisting of edges of label L. Also, because a total graph G_T contains the same information as a s-tuple of graphs, we can assume a graph aggregation function f_{gr} aggregates a n-tuple of total graphs into a single total graph. We then obtain two more corollaries that translate the judgement aggregation concepts of *propositionwise independence* and *systematiciy*.

Corollary 5.3. *Let $\tilde{G}(f) = f_{gr}$. Then f is propositionwise independent if and only if there exist functions $\{f_L\}$, one for each label L, such that whenever $f_{gr}(G_T^1, \ldots, G_T^n) = G_T$ for total graphs $G_T^1, \ldots G_T^n$, for each label L,*

$$f_L(l(G_T^1, L), \ldots, l(G_T^n, L)) = l(G_T, L).$$

Corollary 5.4. *Let $\tilde{G}(f) = f_{gr}$. Then f is systematic if and only if is propositionwise independent and there exists a function $\tilde{f}$ such that $f_L = \tilde{f}$ for any label L.*

The proofs are rather trivial when considering the total graph G_T defined in Section 5.2 (while the demonstration with the s-tuples would be much more cumbersome) so we omit them.

5.4 Doctrinal vs Condorcet paradoxes

We finish by comparing the Condorcet and the doctrinal paradoxes. List and Pettit (2004) contrast the two paradoxes. They refer to May's Theorem (May, 1952), which states that simple majority voting is the only social preference function satisfying certain conditions on two individuals and present both the Condorcet and the doctrinal paradoxes as violations of May's Theorem when extended beyond its assumptions. In particular, they present the Condorcet paradox as deriving from violations of May's Theorem when we allow for more than two alternatives, and the doctrinal paradox when we allow for compound propositions.

While we think their argument has merit, we would like to offer a complementary perspective by focusing on the similarity between the two paradoxes: we claim that the Condorcet and doctrinal paradoxes derive from the **same aggregation problem under the graph-theoretic framework**. Formally, we define the *majority graph-aggregation rule* as the following function on n-tuples of total graphs: $f(G_T^1, \ldots, G_T^1) = G_T$, where:

- $V(G_T) = V(G_T^1) = \cdots = V(G_T^n)$;
- for every pair of vertices (v, w) and label L, count the number of times x that (v, w) appears as an L-labelled edge among the n graphs G_i and the number of times y that (w, v) appears as an L-labelled edge among the G_i. Then place an L-labelled edge (v, w) in G if $x > y$, (w, v) if $x < y$, and no edge otherwise.

We now show that this rule amounts to the majority aggregation functions in both frameworks:

- for majority rule in the social choice framework, recall that it just the special case of the graph-theoretic framework where graphs come from total orders and there is only one possible label. Thus, the majority rule simply checks if there are at least $n/2$ directed edges (v, w), which makes it a special case of the majority graph-aggregation rule;
- for majority rule in the judgement aggregation framework, when translated to total graphs as explained in Section 4, we see that for each (either atomic or compound) proposition P, there are exactly two cases for all the P-labelled edges, with all the edges pointing in one direction if the judgement is $P = True$ and in the other direction if the judgement is $P = False$. The majority judgement then labels all such edges in the direction which appears more than $n/2$ times. This is exactly what the majority graph-aggregation rule would do.

We finish by showing that both paradoxes indeed happen through the majority graph-aggregation rule. First, in the social choice framework, we revisit Example 2.1, where 3 preferences form a cycle when aggregated under majority rule. We know that preference graphs, being "rational," cannot contain cycles, so the Condorcet paradox happens because the majority graph-aggregation rule creates a non-rational graph.

As to the judgement aggregation framework, let us reconsider Example 2.2, recast in our graph-theoretic framework developed in Section 4. Figure 6 visually compares the two examples and provides the basic intuition. If we observe that the atomic edge between 00 and 01, the atomic edge between 01 and 11, and the compound edge

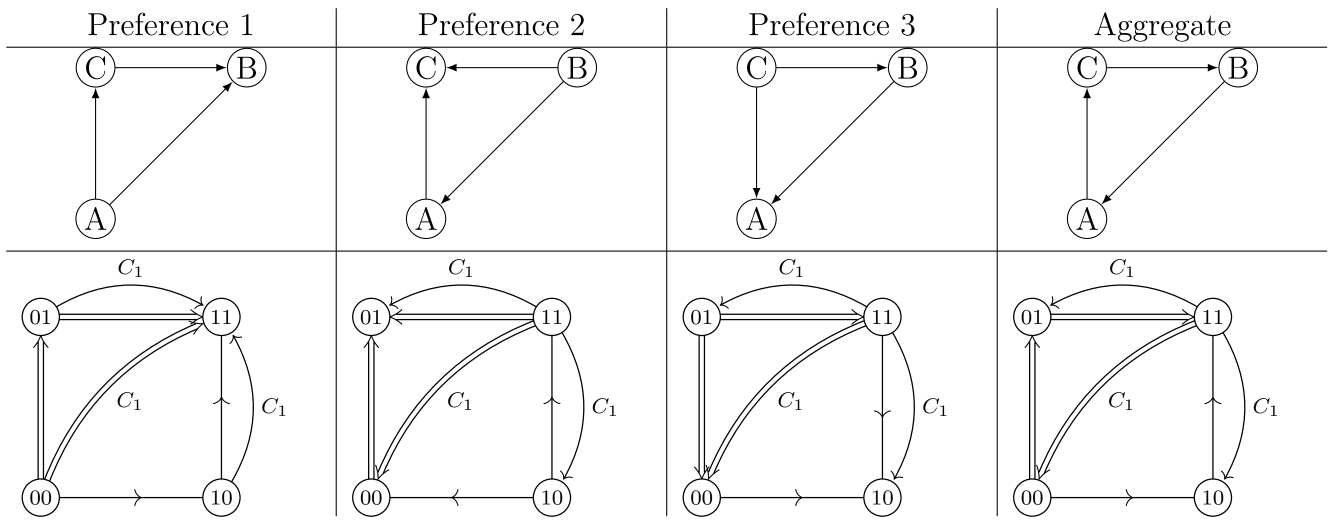

Figure 6: Combining the graphs from Example 2.1 (above) and Example 2.2 (below). Contrast the edges between A, B, C in the top graphs and the double-arrowed edges between $00, 11, 01$ in the bottom graphs. To make the graph more readable, we omitted all labels except those of compound propositions.

between 00 and 11 from the judgement-aggregation example have the same orientations as the edges, respectively, between A and C, between C and B, and between A and B from the social choice example, it is now clear that the two paradoxes have actually the same nature. This happens because creating a cycle among 00, 01, and 11 in our graph implies that none of them can be a global optimum. In addition, also 10, 00 and 11 form a cycle and therefore neither 10 can be a global optimum. Thus, as showed by Theorem 5.1 the aggregate graph is not "rational."

6 Conclusion

In this work, we have presented a unified graph-theoretic framework which, by providing a simple generalization of the classical social choice framework, allows to present the judgement aggregation problem as a particular case.

As an application of this generalization, we showed that the doctrinal and the Condorcet paradoxes can be studied as two similar problems originating from the application of the same aggregation rule, the *graph-theoretic majority aggregation rule*. Furthermore, the two "paradoxes" arise for basically the same reason.

There are many questions that we leave to future research. Both the social choice and judgement aggregation frameworks have been developing largely separately, and it would be useful to see whether this unified framework can help import and extend in one framework the results already available in the other. Some possible directions include:

- Impossibility Theorems in Judgement Aggregation: following up on List and Pettit's impossibility theorem (List and Pettit, 2002), a series of work by Dietrich and List (Dietrich and List, 2007a,b, 2013) have proved an increasingly strong list of impossibility theorems about propositionwise independent judgement aggregation. Are there obvious generalizations to the graph-theoretic framework? If so, what do they mean for the classical social choice framework?

- Arrow's Theorem: List and Pettit compare and contrast their impossibility theorem in the judgement aggregation framework with Arrow's Theorem in social choice (List and Pettit, 2004). The third author of this paper has an upcoming work (Zhang, 2018) on an analogue of Arrow's Theorem under the doctrinal framework. Possibly the correct generalization of both theorems would be a visually intuitive one in the graph-theoretic framework.

- Social choice on complex multidimensional objects: as already mentioned, one of the interesting properties of our graph-theoretic framework is that it naturally deals with choices on multidimensional alternatives and this enables to encompass the judgement aggregation problem which is inherently multidimensional. Multidimensional social choice has been studied both in many contexts and with different methods (Kramer, 1972) and recently has been analyzed with a graph-theoretic model similar to the one presented in this paper (Amendola and Settepanella, 2012; Marengo and Settepanella, 2014; Amendola et al., 2015), which has produced a series of analytical results on the properties of aggregation and, in particular, on the likelihood to find (and avoid) cycles and on the existence of multiple or unique equilibria. The extension of such results to judgement aggregation seems particularly promising.

Acknowledgments

We wish to thank Yoshiki Nakajima for his initial exploration of this work with the first two authors.

References

G. Amendola and S. Settepanella. Optimality in social choice. *The Journal of Mathematical Sociology*, 36:44–77, 2012.

G. Amendola, L. Marengo, D. Pirino, S. Settepanella, and A. Takemura. Decidability in complex social choices. *Evolutionary and Institutional Economics Review*, 12: 141–168, 2015.

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

J.-A.-N. de Caritat marquis de Condorcet. *Essai sur l'Application de l'Analyse aux Probabilités de Decision Rendue à la Pluralité des Voix*. Imprimerie Royale, Paris, 1785.

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

F. Dietrich and C. List. Judgment aggregation without full rationality. *Social Choice and Welfare*, 31:15–39, 2007b.

F. Dietrich and C. List. Propositionwise judgment aggregation: the general case. *Social Choice and Welfare*, 40:1067–1095, 2013.

L. A. Kornhauser and L. G. Sager. The one and the many: adjudication in collegial courts. *California Law Review*, 81:1–59, 1993.

G. H. Kramer. Sophisticated voting over multidimensional choice spaces. *Journal of Mathematical Sociology*, 2:165–180, 1972.

C. List. Social choice theory. In E. Zalta, editor, *Stanford Encyclopedia of Philosophy*. The Metaphysics Research Lab, Stanford, 2012.

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

C. List and P. Pettit. Aggregating sets of judgements: Two impossibility results compared. *Synthese*, 140:208–235, 2004.

L. Marengo and C. Pasquali. The construction of choice. a computational voting model. *Journal of Economic Interaction and Coordination*, 6:139–156, 2011.

L. Marengo and S. Settepanella. Social choice among complex objects. *Annali della Scuola Normale Superiore Pisa Classe di Scienze*, 13:1–26, 2014.

K. May. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica*, 20:680–684, 1952.

Y. X. Zhang. Applying Fourier Analysis to Judgment Aggregation, 2018.