Title
Arrow's Decisive Coalitions

Authors
Holliday, Wesley Halcrow
Pacuit, Eric

Publication Date
2018-10-14

Arrow's Decisive Coalitions*

Wesley H. Holliday† and Eric Pacuit‡

† University of California, Berkeley

‡ University of Maryland, College Park

Preprint of November 21, 2018

Forthcoming in the special issue of *Social Choice and Welfare* in Memory of Kenneth J. Arrow, edited by Eric Maskin, Amartya Sen, and Kotaro Suzumura.

Abstract

In his classic monograph, *Social Choice and Individual Values*, Arrow introduced the notion of a decisive coalition of voters as part of his mathematical framework for social choice theory. The subsequent literature on Arrow's Impossibility Theorem has shown the importance for social choice theory of reasoning about coalitions of voters with different grades of decisiveness. The goal of this paper is a fine-grained analysis of reasoning about decisive coalitions, formalizing how the concept of a decisive coalition gives rise to a social choice theoretic language and logic all of its own. We show that given Arrow's axioms of the Independence of Irrelevant Alternatives and Universal Domain, rationality postulates for social preference correspond to strong axioms about decisive coalitions. We demonstrate this correspondence with results of a kind familiar in economics—representation theorems—as well as results of a kind coming from mathematical logic—completeness theorems. We present a complete logic for reasoning about decisive coalitions, along with formal proofs of Arrow's and Wilson's theorems. In addition, we prove the correctness of an algorithm for calculating, given any social rationality postulate of a certain form in the language of binary preference, the corresponding axiom in the language of decisive coalitions. These results suggest for social choice theory new perspectives and tools from logic.

1 Introduction

Kenneth Arrow's monograph *Social Choice and Individual Values* [Arrow, 2012], originally published in 1951, revolutionized the study of social decision making. The centerpiece of the book is what is now called Arrow's Impossibility Theorem, which shows that the only group decision method satisfying several plausible principles is a *dictatorship*. This ground-breaking theorem has been studied from a variety of perspectives,[1] and many different proofs

*We wish to thank the referees for *Social Choice and Welfare*, Mikayla Kelley, and Kotaro Suzumura for helpful comments.

[1] For instance, see Saari 2008 for a geometric perspective, Baigent 2010 for a topological perspective, Abramsky 2015 for a category theoretic perspective, Lauwers and Van Liedekerke 1995 and Herzberg and Eckert 2012 for model-theoretic perspectives, Fishburn and Rubinstein 1986 and Shelah 2005 for algebraic perspectives, and Bao and Halpern 2017 for a quantum perspective.

can be found in the social choice literature (e.g., Fishburn 1970, Barberá 1980, Reny 2001, Geanakoplos 2005, Yu 2012, Sen 2014). The ramifications of Arrow's result are far-reaching: the Impossibility Theorem and related results have influenced discussions in many areas of economics [Buchanan, 1954, Hammond, 1976, Campbell, 1992, Fleurbaey and Mongin, 2005, Le Breton and Weymark, 2010, Sen, 2017], political science [Riker, 1982, Dryzek and List, 2003, Mackie, 2003, Patty and Penn, 2014] and philosophy [Hurley, 1985, Okasha, 2011, Kroedel and Huber, 2013, Stegenga, 2013, Morreau, 2015, MacAskill, 2016]. It is a testament to Arrow's insight that his theorem continues to inspire generations of researchers.

One lesson from Arrow's own proof of his Impossibility Theorem, as well as the subsequent literature, is the importance for social choice theory of reasoning about decisive coalitions of voters.[2] To prove impossibility theorems, social choice theorists often proceed by laying down axioms about group decision methods and then deriving consequences about decisive coalitions. One might have different views about the acceptability of such consequences depending on the group decision problem being studied. For example, in a voting context of "one person, one vote," one might think that only majorities should be decisive. In a welfare context, one might think that a group of the least well-off members of society could be decisive, even though they do not constitute a majority. Or in a context of aggregating expert opinions, one might think that some non-majority group of experts may be authoritative on the question of whether policy x is better than policy y, although they are not authoritative on whether x' is better than y'. A powerful feature of Arrow's framework is that one can ask questions about decisiveness in all of these contexts.

The goal of this paper is a fine-grained analysis of reasoning about decisive coalitions, formalizing how the concept of a decisive coalition gives rise to a social choice theoretic language and logic all of its own. We show that given two of Arrow's axioms about group decision methods, namely the Independence of Irrelevant Alternatives (IIA) and Universal Domain (UD), rationality postulates for group preference correspond to strong axioms about decisive coalitions. We demonstrate this correspondence with two types of results. The first type is familiar in economics: representation theorems. The second type of result, building on the representation theorems, comes from mathematical logic: completeness theorems. Our completeness theorems show that the axioms about decisive coalitions that we identify capture *every consequence* of IIA, UD, and the rationality postulates that can be stated in the language of decisive coalitions, in the sense that every such consequence is derivable from our axioms, using some simple rules of inference. Without such a theorem, there would be no guarantee that our set of axioms about decisive coalitions does not miss some important principles about decisive coalitions that follow from Arrow's axioms about group decision methods. With the completeness theorem, we have a guarantee. What is important for social choice theory is not having a completeness theorem per se, but rather having a set of axioms about decisive coalitions that comes with such a guarantee.

This work continues a strong tradition of using ideas from mathematical logic in social choice theory. A number of social choice theorists, including Arrow and Amartya Sen, have noted the influence of mathematical logic on their thinking.[3] An early use of results from

[2]Reasoning about decisive coalitions is also central in the theory of judgment aggregation (see, e.g., List and Polak 2010).

[3]In a chapter entitled "Origins of the Impossibility Theorem" [Arrow, 2014], Arrow recalls: "As it happens,

mathematical logic in social choice theory is Rubinstein's [1984] proof of the equivalence between multi-profile approaches and single-profile approaches to social choice.[4] There is now a rich literature developing logical systems that can formalize social choice results [Wiedijk, 2007, Agotnes et al., 2009, Nipkow, 2009, Tang and Lin, 2009, Troquard et al., 2011, Endriss, 2011, Grandi and Endriss, 2013, Ciná and Endriss, 2016, Pacuit and Yang, 2016, Salles, 2017]. We do not attempt here a comparison between our logical approach and others found in this literature. In short, our goal is different: rather than adapting existing logical systems to formalize social choice results, we aim to develop a new logical system that captures the style of reasoning about decisive coalitions familiar to social choice theorists.

Indeed, we show that there is a simple logic of decisive coalitions that unifies many impossibility results in the social choice literature. Using this logic, we give formal proofs of Arrow's theorem, Wilson's [1972] theorem, and Gibbard's [2014] Oligarchy Theorem.

Following Monjardet [1967, 1978, 1983], Kirman and Sondermann [1972], and Hansson [1976], we can view Arrow's theorem as resulting from two facts. The first is that Arrow's axioms imply that the collection of decisive coalitions is an *ultrafilter*. An ultrafilter on a nonempty set V is a nonempty set $\mathcal{U}$ of subsets of V such that (i) $\varnothing \notin \mathcal{U}$, (ii) for every $A \subseteq V$, either $A \in \mathcal{U}$ or $A^c \in \mathcal{U}$, (iii) the set is closed under supersets (for all $A, B \subseteq V$, if $A \in \mathcal{U}$ and $A \subseteq B$, then $B \in \mathcal{U}$), and (iv) the set is closed under intersection (for all $A, B \subseteq V$, if $A \in \mathcal{U}$ and $B \in \mathcal{U}$, then $A \cap B \in \mathcal{U}$). The second fact is that any ultrafilter on a finite set is principal, meaning that it contains a set with exactly one element. Thus, if the set of voters is finite, then there is a decisive coalition with exactly one member, which is the definition of a dictatorship. From this perspective, Arrow's theorem may be seen as resulting from the combination of a social choice theoretic fact about decisive coalitions—under Arrow's axioms, they form an ultrafilter—and a purely combinatorial fact—any ultrafilter on a finite set is principal.[5] In this paper, what we mean by a proof of Arrow's theorem is a proof that the decisive coalitions form an ultrafilter. Before one even reaches the conclusion of dictatorship, the conclusion that the set of decisive coalitions is closed under intersection seems highly undesirable (see Hansson 1976 for a similar observation). While this holds trivially when the group decision method is unanimity, in which case the only decisive coalition is the set of all voters, closure under intersection makes little sense when decisiveness depends in part on a coalition being *large enough*—for example, it is obviously violated by the method of majority rule.

The rest of the paper is organized as follows. Section 2 provides the familiar definitions from the social choice framework, including the two main notions of decisiveness that we will study. In Section 3, we prove our main representation theorems for decisive coalitions. Building on this representation theory, Section 4 introduces a family of logics for reasoning about decisive coalitions. We explain the key property of completeness of these logics, which we prove in the Appendix. In Sections 5 and 6, we show how to use the logics from Section

during my college years, I was fascinated by mathematical logic, a subject I read on my own until, by a curious set of chances, the great Polish logician, Alfred Tarski, taught one year at The City College (in New York), where I was a senior. He chose to give a course on the calculus of relations, and I was introduced to such topics as transitivity and orderings" (p. 154). Sen [2017, p. xvi] even reports having taught mathematical logic at Delhi University. For his views on the relation between logic and economics, see Sen 2017, p. 108.

[4] Earlier Murakami [1968] applied results about three-valued logic to the analysis of voting rules.

[5] This separation of the social choice theoretic content and the combinatorial content of Arrow's theorem is also a theme in Makinson 1996.

4 to prove Arrow's theorem and variations thereof. In Section 7, we show how to extend our framework to reason about the relations between the main notions of decisiveness, as well as other notions of decisiveness in the social choice literature. Section 8 takes a "meta"-perspective on the approach of this paper, by showing that there is an algorithm for translating rationality postulates for social preferences into axioms about decisive coalitions. We conclude in Section 9 with a discussion of future work.

2 Preliminaries

Recall that a binary relation P on a set X is a *strict weak order* if and only if P satisfies the following conditions for all $x, y, z \in X$:

asymmetry: if xPy, then *not* yPx;

negative transitivity: if xPy, then xPz or zPy.

Negative transitivity is equivalent to the condition that if *not* xPz and *not* zPy, then *not* xPy, which explains the name. Together negative transitivity and asymmetry imply that P is transitive. In the social choice literature, 'transitivity' usually refers to the transitivity of a weak preference relation R, while 'quasi-transitivity' refers to the transitivity of the corresponding strict preference relation P. However, since we take P as primitive, we use 'transitivity' for the transitivity of P:

transitivity: if xPy and yPz, then xPz.

Let xNy if and only if neither xPy nor yPx. We call N the relation of *P-noncomparability*. If P is a strict weak order, then N satisfies the following for all $x, y, z \in X$:

transitivity of P-noncomparability: if xNy and yNz, then xNz.

The relation P is a *strict linear order* if and only if it satisfies asymmetry, transitivity, and weak completeness: for all $x, y \in X$, if $x \neq y$, then xPy or yPx.

We adopt the following notation for classes of binary relations on X:

$P(X)$ is the set of all asymmetric binary relations on X;

$O(X)$ is the set of all strict weak orders on X;

$L(X)$ is the set of all strict linear orders on X.

For any nonempty set X, whose members we call *alternatives*, and any nonempty set V, whose members we call *voters*, a *profile* $\mathbf{P}$ *for* $\langle X, V\rangle$ is an element of $O(X)^V$, i.e., a function assigning to each $i \in V$ a relation $P_i \in O(X)$, which we call *i's strict preference relation*. For $x, y \in X$, let:

$$\begin{aligned} \mathbf{P}(x,y) &= \{i \in V \mid xP_iy\}; \\ \mathbf{P}_{|\{x,y\}} &= \text{the function assigning to each } i \in V \text{ the relation } P_i \cap \{x,y\}^2. \end{aligned}$$

Adapting the terminology of Sen [2017, Ch. 2*], we say that a *collective choice rule* (CCR) for $\langle X, V\rangle$ is a function f from a subset of $O(X)^V$ to $P(X)$.[6] By $xf(\mathbf{P})y$, we mean $\langle x, y\rangle \in f(\mathbf{P})$. We adopt the following terminology and abbreviations for standard conditions on a CCR f:

Domain Conditions

- universal domain (UD): $\mathrm{dom}(f) = O(X)^V$.[7]
- linear domain (LD): $\mathrm{dom}(f) = L(X)^V$.

Codomain Conditions ("Rationality Postulates")

- transitive rationality (TR): for all $\mathbf{P} \in \mathrm{dom}(f)$, $f(\mathbf{P})$ is transitive.
- full rationality (FR): for all $\mathbf{P} \in \mathrm{dom}(f)$, $f(\mathbf{P})$ is a strict weak order.

Interprofile Conditions

- independence of irrelevant alternatives (IIA): for all $\mathbf{P}, \mathbf{P}' \in \mathrm{dom}(f)$ and $x, y \in X$, if $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$, then $xf(\mathbf{P})y$ if and only if $xf(\mathbf{P}')y$.
- positive association of social and individual values (PA): for all $\mathbf{P}, \mathbf{P}' \in \mathrm{dom}(f)$, if for all $x, y \in X \setminus \{x^\star\}$, $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$, $\mathbf{P}(x^\star, y) \subseteq \mathbf{P}'(x^\star, y)$, and $\mathbf{P}(y, x^\star) \supseteq \mathbf{P}'(y, x^\star)$, then for all $y \in X$, $x^\star f(\mathbf{P})y$ implies $x^\star f(\mathbf{P}')y$.
- semi-positive association of social and individual values (SPA): for all $\mathbf{P}, \mathbf{P}' \in \mathrm{dom}(f)$, if for all $x, y \in X \setminus \{x^\star\}$, $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$, $\mathbf{P}(x^\star, y) = \mathbf{P}'(x^\star, y)$, and $\mathbf{P}(y, x^\star) \supseteq \mathbf{P}'(y, x^\star)$, then for all $y \in X$, $x^\star f(\mathbf{P})y$ implies $x^\star f(\mathbf{P}')y$.[8]

Decisiveness Conditions

- Pareto (P): for all $\mathbf{P} \in \mathrm{dom}(f)$ and $x, y \in X$, if $\mathbf{P}(x, y) = V$, then $xf(\mathbf{P})y$.
- dictatorship: there is an $i \in V$ such that for all $\mathbf{P} \in \mathrm{dom}(f)$ and $x, y \in X$, if xP_iy, then $xf(\mathbf{P})y$.

Some implications between these conditions are immediate. For example, FR implies TR, and LD implies SPA.[9] The implication in Arrow's theorem, of course, is far from obvious.

Theorem 2.1 (Arrow)**.** Assume that $|X| \geq 3$ and V is finite. Then any CCR for $\langle X, V\rangle$ satisfying UD, IIA, FR, and P is a dictatorship.

Next we recall the key notions of decisiveness that are the focus of this paper.

[6] Strictly speaking, as Sen—like Arrow—takes the weak preference relation R as primitive, Sen's CCRs send a vector of reflexive, transitive, and complete relations on X to a binary relation on X.

[7] In this paper, we do not discuss domain restrictions, other than LD. However, we do note that our results continue to hold on any "saturated" domain that is rich enough for Arrow's theorem to hold [Kalai et al., 1979, Le Breton and Weymark, 2010].

[8] SPA is not as standard as the other conditions, but it has a natural connection to our main topic of decisive coalitions, explained at the end of this section.

[9] For the latter implication, if $\mathbf{P}, \mathbf{P}' \in L(X)$, then $\mathbf{P}(x^\star, y) = \mathbf{P}'(x^\star, y)$ implies $\mathbf{P}(y, x^\star) = \mathbf{P}'(y, x^\star)$ and hence $\mathbf{P}_{|\{x^\star,y\}} = \mathbf{P}'_{|\{x^\star,y\}}$. So if $\mathrm{dom}(f) = L(X)$, then the assumption of SPA implies that for all $x, y \in X \setminus \{x^\star\}$, $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$ and $\mathbf{P}_{|\{x^\star,y\}} = \mathbf{P}'_{|\{x^\star,y\}}$, which implies $\mathbf{P} = \mathbf{P}'$, so $f(\mathbf{P}) = f(\mathbf{P}')$.

Definition 2.2. Let f be a CCR for $\langle X, V\rangle$. For any $x, y \in X$ and $A \subseteq V$:

1. A is *almost decisive for x over y according to f* if and only if for all $\mathbf{P} \in \text{dom}(f)$, if $A = \mathbf{P}(x, y)$ and $A^c = \mathbf{P}(y, x)$, then $x f(\mathbf{P}) y$;

2. A is *almost decisive according to f* if and only if for every $x, y \in X$, A is almost decisive for x over y according to f;

3. A is *decisive for x over y according to f* if and only if for all $\mathbf{P} \in \text{dom}(f)$, if $A \subseteq \mathbf{P}(x, y)$, then $x f(\mathbf{P}) y$;

4. A is *decisive according to f* if and only if for every distinct $x, y \in X$, A is decisive for x over y according to f.

We can keep track of the decisiveness data for f using the two functions in Definition 2.3. For any set S, $\wp(S)$ is the set of all subsets of S.

Definition 2.3. Let f be a CCR for $\langle X, V\rangle$. We define functions $\widehat{D}_f : X^2 \to \wp(\wp(V))$ and $\overline{D}_f : X^2 \to \wp(\wp(V))$ as follows:

1. $\widehat{D}_f(x, y)$ is the set of all $A \subseteq V$ that are almost decisive for x over y according to f;

2. $\overline{D}_f(x, y)$ is the set of all $A \subseteq V$ that are decisive for x over y according to f.

Remark 2.4. In the literature, the letter D is often used for almost decisiveness, but we wish to use D for either $\widehat{D}$ or $\overline{D}$ depending on the context.

Clearly $\overline{D}_f(x, y) \subseteq \widehat{D}_f(x, y)$. Also observe that PA together with UD implies $\overline{D}_f(x, y) = \widehat{D}_f(x, y)$. Under SPA, there is still a distinction between $\overline{D}$ and $\widehat{D}$, but SPA together with UD implies that $\overline{D}$ and $\widehat{D}$ are nicely related as follows for $x \neq y$:

$$A \in \overline{D}(x, y) \text{ if and only if for all } B \supseteq A, \text{ we have } B \in \widehat{D}(x, y).$$

The left to right implication is obvious. From right to left, suppose $A \notin \overline{D}(x, y)$, so there is some profile $\mathbf{P}$ in which $A \subseteq \mathbf{P}(x, y)$ but *not* $x f(\mathbf{P}) y$. It follows by UD that there is a profile $\mathbf{P}'$ such that $\mathbf{P}(x, y) = \mathbf{P}'(x, y)$ and $\mathbf{P}'(x, y)^c = \mathbf{P}'(y, x)$. Then by SPA, *not* $x f(\mathbf{P}) y$ implies *not* $x f(\mathbf{P}') y$. Thus, $\mathbf{P}(x, y) \notin \widehat{D}(x, y)$, so we have found a $B \supseteq A$ such that $B \notin \widehat{D}(x, y)$.

3 Representation

One of the most important kinds of results in mathematics in general and in axiomatic economics in particular is the *representation theorem*, showing that any abstract structure satisfying certain axioms can be represented as arising from a more concrete structure. Examples include: a binary preference relation $\succeq$ on lotteries can be represented by a linear utility function, in the sense that for any lotteries L and M, $L \succeq M \Leftrightarrow$ the expected utility of L is at least that of M, if and only if $\succeq$ satisfies the axioms in von Neumann and Morgenstern 1947; a choice function c on a finite set S can be represented by a weak ordering R on S, in the sense that for any $A \subseteq S$, $c(A) = \{x \in A \mid xRy \text{ for all } y \in A\}$, if and only if c satisfies the axioms in Arrow 1959; a binary relation $\geqslant$ on the powerset of a finite set S can be represented by a probability function μ on $\wp(S)$, in the sense that for any $A, B \subseteq S$, $A \geqslant B \Leftrightarrow \mu(A) \geq \mu(B)$, if and only if $\geqslant$ satisfies the axioms in Kraft et al. 1959.

For our purposes in this paper, the key notion of representation is the following.

Definition 3.1. Let K be a class of CCRs for $\langle X, V\rangle$.

1. A function $D : X^2 \to \wp(\wp(V))$ is *almost-decisively representable in* K if and only if there is an $f \in K$ such that $D = \widehat{D}_f$, in which case we say that f *almost-decisively represents* D.

2. A function $D : X^2 \to \wp(\wp(V))$ is *decisively representable in* K if and only if there is an $f \in K$ such that $D = \overline{D}_f$, in which case we say that f *decisively represents* D.

In Sections 3.1–3.2, we prove representation theorems for both notions of representability. We start with the proofs for almost-decisive representability, as they are somewhat simpler.

Remark 3.2. A *simple game* (as in, e.g., Taylor and Zwicker 1992) is a pair $\langle N, W\rangle$ where N is a finite set and W is a family of subsets of N, thought of as the "winning coalitions" of voters.[10] If we assume that our set V of voters is finite, then a function $D : X^2 \to \wp(\wp(V))$ as in Definition 3.1 yields, for each $\langle x, y\rangle \in X^2$, a simple game $\langle V, D(x, y)\rangle$.

3.1 Almost Decisive Coalitions

First, we will give necessary and sufficient conditions for almost-decisively representability in the class of CCRs satisfying LD, IIA, and TR (resp. FR). Then we will show how the same conditions are necessary and sufficient for almost-decisively representability in the class of CCRs satisfying UD, IIA, and TR (resp. FR).

Theorem 3.3. Let X and V be nonempty sets with $|X| \geq 3$. A function $D : X^2 \to \wp(\wp(V))$ is almost-decisively representable in the class of CCRs for $\langle X, V\rangle$ satisfying LD, IIA, and TR (resp. FR) if and only if for all $A, B, C \subseteq V$ and $x, y, z \in X$ with $x \neq y$, $y \neq z$, and $x \neq z$:

1. $A \in D(x, x)$;

2. if $A \in D(x, y)$, then $A^c \notin D(y, x)$;

3. for TR: if $A \in D(x, y)$, $B \in D(y, z)$, and $A \cap B \subseteq C \subseteq A \cup B$, then $C \in D(x, z)$;

4. for FR: if $A \in D(x, y)$ and $B \cap C \subseteq A \subseteq B \cup C$, then $B \in D(x, z)$ or $C \in D(z, y)$.

Proof. Suppose D is almost-decisively represented by a CCR f for $\langle X, V\rangle$ satisfying LD, IIA, and TR (resp. FR).

For property 1, there is no profile $\mathbf{P}$ such that $A = \mathbf{P}(x, x)$ and $A^c = \mathbf{P}(x, x)$, so trivially $A \in D(x, x)$. Property 2 is also immediate from the definition of almost decisiveness.

For property 3 under TR, to show that $C \in D(x, z)$, consider any profile $\mathbf{P}$ in which $C = \mathbf{P}(x, z)$ and $C^c = \mathbf{P}(z, x)$. We must show $xf(\mathbf{P})z$. First, we claim that there is a profile $\mathbf{P}' \in L(X)^V$ such that:

$$A = \mathbf{P}'(x, y),\ A^c = \mathbf{P}'(y, x) \qquad B = \mathbf{P}'(y, z),\ B^c = \mathbf{P}'(z, y) \qquad C = \mathbf{P}'(x, z),\ C^c = \mathbf{P}'(z, x).$$

[10] In Taylor and Zwicker 1999, the definition of a 'simple game' requires that W be closed under supersets, while the term 'hypergraph' is used for the more general structures.

To see that there is such a profile, define for each $i \in V$ a binary relation P_i^0 on $\{x, y, z\}$ according to the cell to which i belongs in the partition in Figure 1: e.g., if $i \in A \cap B^c \cap C$, define P_i^0 as the strict linear order with $xP_i^0 zP_i^0 y$. Since $A \cap B \subseteq C \subseteq A \cup B$, we have $A \cap B \cap C^c = \varnothing$ and $A^c \cap B^c \cap C = \varnothing$. If $i \in V$ belongs to any of the nonempty cells of the partition, then P_i^0 is a strict linear order on $\{x, y, z\}$; hence it can be extended to a strict linear order P_i' on all of X. The resulting profile $\mathbf{P}' \in L(X)^V$ can be seen to satisfy the bulleted equations above by checking the table in Figure 1. Then since $A \in D(x, y)$ and $B \in D(y, z)$, we have $xf(\mathbf{P}')y$ and $yf(\mathbf{P}')z$, which together imply $xf(\mathbf{P}')z$ by the transitivity condition of TR. Then since $\mathbf{P}_{|\{x,z\}} = \mathbf{P}'_{|\{x,z\}}$, we have $xf(\mathbf{P})z$ by IIA.

cell	nonempty?	strict linear order
$A \cap B \cap C$	maybe	$xP_i^0 yP_i^0 z$
$A \cap B \cap C^c$	no	
$A \cap B^c \cap C$	maybe	$xP_i^0 zP_i^0 y$
$A \cap B^c \cap C^c$	maybe	$zP_i^0 xP_i^0 y$
$A^c \cap B \cap C$	maybe	$yP_i^0 xP_i^0 z$
$A^c \cap B \cap C^c$	maybe	$yP_i^0 zP_i^0 x$
$A^c \cap B^c \cap C$	no	
$A^c \cap B^c \cap C^c$	maybe	$zP_i^0 yP_i^0 x$

Figure 1: cells and associated orders for the proof of property 3 in Theorem 3.3

For property 4 under FR, suppose $B \notin D(x, z)$, so there is a profile $\mathbf{P}$ in which $B = \mathbf{P}(x, z)$, $B^c = \mathbf{P}(z, x)$, and *not* $xf(\mathbf{P})z$. First, we claim that there is a profile $\mathbf{P}' \in L(X)^V$ such that:

$$A = \mathbf{P}'(x, y),\ A^c = \mathbf{P}'(y, x) \qquad B = \mathbf{P}'(x, z),\ B^c = \mathbf{P}'(z, x) \qquad C = \mathbf{P}'(z, y),\ C^c = \mathbf{P}'(y, z).$$

The proof that there is such a profile follows the same partitioning strategy as that for property 3 above, but now using the table in Figure 2 and the fact that from $B \cap C \subseteq A \subseteq B \cup C$, we have $A^c \cap B \cap C = \varnothing$ and $A \cap B^c \cap C^c = \varnothing$. Then since $\mathbf{P}_{|\{x,z\}} = \mathbf{P}'_{|\{x,z\}}$, from *not* $xf(\mathbf{P})z$ we have *not* $xf(\mathbf{P}')z$ by IIA, and since $A \in D(x, y)$, we have $xf(\mathbf{P}')y$. Therefore, $zf(\mathbf{P}')y$ by the negative transitivity condition of FR. Now we claim that $C \in D(z, y)$. Consider any profile $\mathbf{P}''$ in which $C = \mathbf{P}''(z, y)$ and $C^c = \mathbf{P}''(y, z)$. Then $\mathbf{P}'_{|\{y,z\}} = \mathbf{P}''_{|\{y,z\}}$, so by IIA, $zf(\mathbf{P}')y$ implies $zf(\mathbf{P}'')y$. Thus, $C \in D(z, y)$.

cell	nonempty?	strict linear order
$A \cap B \cap C$	maybe	$xP_i^0 zP_i^0 y$
$A \cap B \cap C^c$	maybe	$xP_i^0 yP_i^0 z$
$A \cap B^c \cap C$	maybe	$zP_i^0 xP_i^0 y$
$A \cap B^c \cap C^c$	no	
$A^c \cap B \cap C$	no	
$A^c \cap B \cap C^c$	maybe	$yP_i^0 xP_i^0 z$
$A^c \cap B^c \cap C$	maybe	$zP_i^0 yP_i^0 x$
$A^c \cap B^c \cap C^c$	maybe	$yP_i^0 zP_i^0 x$

Figure 2: cells and associated orders for the proof of property 4 in Theorem 3.3

Conversely, suppose properties 1, 2, and 3 (resp. 1, 2, and 4) hold. We define f as follows: for any $\mathbf{P} \in L(X)^V$, let $xf(\mathbf{P})y$ if and only if $x \neq y$ and $\mathbf{P}(x,y) \in D(x,y)$. First, we claim that f satisfies TR (resp. FR):

- asymmetry: if $xf(\mathbf{P})y$, then *not* $yf(\mathbf{P})x$. Suppose $xf(\mathbf{P})y$, so $\mathbf{P}(x,y) \in D(x,y)$, which implies $\mathbf{P}(x,y)^c \notin D(y,x)$ by property 2. Since $\mathbf{P} \in L(X)^V$, we have $\mathbf{P}(x,y)^c = \mathbf{P}(y,x)$, so $\mathbf{P}(x,y)^c \notin D(y,x)$ implies $\mathbf{P}(y,x) \notin D(y,x)$ and hence *not* $yf(\mathbf{P})x$.

  Given asymmetry, we may assume without loss of generality in the following proofs of transitivity and negative transitivity that x, y, and z are pairwise distinct.

- transitivity: if $xf(\mathbf{P})y$ and $yf(\mathbf{P})z$, then $xf(\mathbf{P})z$. Suppose $xf(\mathbf{P})y$ and $yf(\mathbf{P})z$, so $\mathbf{P}(x,y) \in D(x,y)$ and $\mathbf{P}(y,z) \in D(y,z)$. Then since $\mathbf{P}(x,y) \cap \mathbf{P}(y,z) \subseteq \mathbf{P}(x,z) \subseteq \mathbf{P}(x,y) \cup \mathbf{P}(y,z)$, it follows by property 3 that $\mathbf{P}(x,z) \in D(x,z)$ and hence $xf(\mathbf{P})z$.

- negative transitivity (assuming property 4): if $xf(\mathbf{P})y$, then $xf(\mathbf{P})z$ or $zf(\mathbf{P})y$. Suppose $xf(\mathbf{P})y$, so $\mathbf{P}(x,y) \in D(x,y)$. Then since $\mathbf{P}(x,z) \cap \mathbf{P}(z,y) \subseteq \mathbf{P}(x,y) \subseteq \mathbf{P}(x,z) \cup \mathbf{P}(z,y)$, it follows by property 4 that $\mathbf{P}(x,z) \in D(x,z)$ or $\mathbf{P}(z,y) \in D(z,y)$, which implies $xf(\mathbf{P})z$ or $zf(\mathbf{P})y$.

Next observe that f satisfies IIA. Suppose $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$. Then $\mathbf{P}(x,y) = \mathbf{P}'(x,y)$ and $\mathbf{P}(y,x) = \mathbf{P}'(y,x)$. It follows by the definition of f that $f(\mathbf{P})$ and $f(\mathbf{P}')$ agree with respect to x and y.

Finally, we claim that D is almost-decisively represented by f. Suppose $A \in D(x,y)$. Then by definition of f, for any profile $\mathbf{P} \in L(X)^V$, if $A = \mathbf{P}(x,y)$ and $A^c = \mathbf{P}(y,x)$ (which together imply $x \neq y$), then $xf(\mathbf{P})y$. Hence $A \in \widehat{D}_f(x,y)$. Conversely, suppose $A \notin D(x,y)$, so by property 1, $x \neq y$. Consider a profile $\mathbf{P} \in L(X)^V$ in which $A = \mathbf{P}(x,y)$. Then $A = \mathbf{P}(x,y) \notin D(x,y)$, so by definition of f we have *not* $xf(\mathbf{P})y$. Therefore, $A \notin \widehat{D}_f(x,y)$. □

We will now show that representability in the class of CCRs satisfying LD, IIA, and TR (resp. FR) is sufficient for representability in the class of CCRs satisfying UD, IIA, and TR (resp. FR), using the following lemma.

Lemma 3.4. If $f : L(X)^V \to P(X)$, then there is an $f^+ : O(X)^V \to P(X)$ such that:

1. for all $\mathbf{P} \in L(X)^V$, $f(\mathbf{P}) = f^+(\mathbf{P})$;
2. $\mathrm{ran}(f) = \mathrm{ran}(f^+)$;
3. if f satisfies IIA, so does f^+;
4. for all $x, y \in X$, $\widehat{D}_f(x,y) = \widehat{D}_{f^+}(x,y)$ and $\overline{D}_f(x,y) = \overline{D}_{f^+}(x,y)$.

Proof. Fix any strict linear order S on X. Given a profile $\mathbf{P} \in O(X)^V$, we define $\mathbf{P}^+ \in L(X)^V$ as follows. If *not* xN_iy, then P_i^+ matches P_i on x, y. If xN_iy, then P_i^+ matches S on x, y. We claim that P^+ is a strict linear order:

- For weak completeness, suppose $x \neq y$. If *not* xN_iy, then P_i^+ matches P_i on x, y, so *not* xN_i^+y, which implies xP_i^+y or yP_i^+x. On the other hand, if xN_iy, then P_i^+ matches the strict linear order S on x, y, so again *not* xN_i^+y given $x \neq y$.

- For asymmetry, suppose xP_i^+y. If *not* xN_iy, then P_i^+ matches P_i on x, y, so the asymmetry of P_i gives us *not* yP_i^+x. On the other hand, if xN_iy, then P_i^+ matches S on x, y, so the asymmetry of S gives us *not* yP_i^+x.

- For transitivity, suppose xP_i^+y and yP_i^+z. Case 1: suppose xN_iy and yN_iz. Then by the transitivity of P_i-noncomparability, xN_iz. In this case, P_i^+ matches S on x, y, on y, z, and on x, z, so xP_i^+y and yP_i^+z together imply xP_i^+z by the transitivity of S. Case 2: suppose *not* xN_iy and *not* yN_iz. In this case, P_i^+ matches P_i on x, y and on y, z, so xP_i^+y and yP_i^+z together imply xP_iy and yP_iz, which in turn imply xP_iz by the transitivity of P_i. Thus, we have *not* xN_iz, so that P_i^+ matches P_i on x, z. Therefore, from xP_iz we have xP_i^+z. Case 3: *not* xN_iy, but yN_iz. Then xP_i^+y implies xP_iy. By the negative transitivity of P_i, we have xP_iz or zP_iy. But yN_iz, so we conclude that xP_iz and hence xP_i^+z. Case 4: xN_iy, but *not* yN_iz. Then yP_i^+z implies yP_iz. By the negative transitivity of P_i, we have yP_ix or xP_iz. But xN_iy, so we conclude that xP_iz and hence xP_i^+z.

Now given $f : L(X)^V \to P(X)$, we define $f^+ : O(X)^V \to P(X)$ by $f^+(\mathbf{P}) = f(\mathbf{P}^+)$. This definition immediately implies properties 1 and 2 in the statement of the lemma.

The definition of $\mathbf{P}^+$, using the fixed S, ensures that for any $\mathbf{P}, \mathbf{P}' \in O(X)^V$, we have $\mathbf{P}_{|\{x,y\}} = \mathbf{P}'_{|\{x,y\}}$ if and only if $\mathbf{P}^+_{|\{x,y\}} = \mathbf{P}'^+_{|\{x,y\}}$. Thus, IIA for f implies IIA for f^+.

Next, we show that $\widehat{D}_f(x, y) = \widehat{D}_{f^+}(x, y)$. The inclusion $\widehat{D}_{f^+}(x, y) \subseteq \widehat{D}_f(x, y)$ is obvious. For $\widehat{D}_f(x, y) \subseteq \widehat{D}_{f^+}(x, y)$, suppose $A \in \widehat{D}_f(x, y)$ and $\mathbf{P} \in O(X)^V$ is such that $A = \mathbf{P}(x, y)$ and $A^c = \mathbf{P}(y, x)$. It follows that $A = \mathbf{P}^+(x, y)$ and $A^c = \mathbf{P}^+(y, x)$, which with $A \in \widehat{D}_f(x, y)$ implies $xf(\mathbf{P}^+)y$ and hence $xf^+(\mathbf{P})y$. Thus, $A \in \widehat{D}_{f^+}(x, y)$.

The proof that $\overline{D}_f(x, y) = \overline{D}_{f^+}(x, y)$ is essentially the same. The inclusion $\overline{D}_{f^+}(x, y) \subseteq \overline{D}_f(x, y)$ is again obvious. For $\overline{D}_f(x, y) \subseteq \overline{D}_{f^+}(x, y)$, suppose $A \in \overline{D}_f(x, y)$ and $\mathbf{P} \in O(X)^V$ is such that $A \subseteq \mathbf{P}(x, y)$. It follows that $A \subseteq \mathbf{P}^+(x, y)$, which with $A \in \overline{D}_f(x, y)$ implies $xf(\mathbf{P}^+)y$ and hence $xf^+(\mathbf{P})y$. Thus, $A \in \overline{D}_{f^+}(x, y)$. □

We can now prove our desired representation theorem for UD instead of LD.

Theorem 3.5. Let X and V be nonempty sets with $|X| \geq 3$. A function $D : X^2 \to \wp(\wp(V))$ is almost-decisively representable in the class of CCRs for $\langle X, V \rangle$ satisfying UD, IIA, and TR (resp. FR) if and only if D satisfies properties 1, 2, and 3 (resp. 1, 2, and 4) of Theorem 3.3.

Proof. Inspection of the proof of Theorem 3.3 shows that the proof of the properties did not use the assumption that the domain of f is $L(X)^V$. The proofs work equally well under the assumption that the domain of f is $O(X)^V$.

Conversely, if f satisfies the properties, then its representability in the class of CCRs satisfying UD, IIA, and TR (resp. FR) follows from Theorem 3.3 plus Lemma 3.4. □

3.2 Decisive Coalitions

The representation theorem for decisive coalitions is almost the same—only the first two properties slightly change, and we add one property—but now the verification of cases in the proof is more involved and the proof for property 4 requires more care.

Theorem 3.6. Let X and V be nonempty sets with $|X| \geq 3$. A function $D : X^2 \to \wp(\wp(V))$ is decisively representable in the class of CCRs for $\langle X, V\rangle$ satisfying UD, IIA, and TR (resp. FR) if and only if for all $A, B, C \subseteq V$ and $x, y, z \in X$ with $x \neq y$, $y \neq z$, and $x \neq z$:

1. $A \in D(x, x)$ if and only if $A \neq \varnothing$;
2. if $A \in D(x, y)$ and $A \cap B = \varnothing$, then $B \notin D(y, x)$;
3. for TR: if $A \in D(x, y)$, $B \in D(y, z)$, and $A \cap B \subseteq C \subseteq A \cup B$, then $C \in D(x, z)$;
4. for FR: if $A \in D(x, y)$ and $B \cap C \subseteq A \subseteq B \cup C$, then $B \in D(x, z)$ or $C \in D(z, y)$;
5. if $A \in D(x, y)$ and $A \subseteq B$, then $B \in D(x, y)$.

Proof. Suppose D is decisively represented by a CCR f satisfying UD, IIA, and TR (resp. FR).

For property 1, for any profile $\mathbf{P}$, we have $\mathbf{P}(x, x) = \varnothing$. So if $A \neq \varnothing$, then we trivially have $A \in D(x, x)$. On the other hand, if $A = \varnothing$, then since *not* $xf(\mathbf{P})x$, we have $A \notin D(x, x)$.

For property 2, suppose $A \in D(x, y)$ and $A \cap B = \varnothing$. Then by UD, there is a profile $\mathbf{P}$ in which $\mathbf{P}(x, y) = A$ and $\mathbf{P}(y, x) = B$. Since $A \in D(x, y)$, it follows that $xf(\mathbf{P})y$, which with asymmetry implies *not* $yf(\mathbf{P})x$. Therefore, $B \notin D(y, x)$.

cell	nonempty?	strict weak order
$A \cap B \cap C$	maybe	linear: $xP_i^0 yP_i^0 z$
$A \cap B \cap C^c$	no	
$A \cap B^c \cap C$	maybe	linear: $xP_i^0 zP_i^0 y$
$A \cap B^c \cap C^c$	maybe	$xP_i^0 y$, $zP_i^0 y$, and P_i^0 relates x and z however P_i does
$A^c \cap B \cap C$	maybe	linear: $yP_i^0 xP_i^0 z$
$A^c \cap B \cap C^c$	maybe	$yP_i^0 x$, $yP_i^0 z$, and P_i^0 relates x and z however P_i does
$A^c \cap B^c \cap C$	no	
$A^c \cap B^c \cap C^c$	maybe	$\begin{cases} \text{linear} : xP_i^0 yP_i^0 z & \text{if } xP_i z \\ \text{linear} : zP_i^0 yP_i^0 x & \text{if } zP_i x \\ x, y, z \text{ noncomparable in } P_i^0 & \text{if } xN_i z \end{cases}$

Figure 3: cells and associated orders for the proof of property 3 in Theorem 3.6

For property 3 under TR, to show that $C \in D(x, z)$, consider any profile $\mathbf{P}$ in which $C \subseteq \mathbf{P}(x, z)$. We must show $xf(\mathbf{P})z$. First, we claim that there is a profile $\mathbf{P}'$ such that:

$$A \subseteq \mathbf{P}'(x, y) \qquad B \subseteq \mathbf{P}'(y, z) \qquad \mathbf{P}_{|\{x,z\}} = \mathbf{P}'_{|\{x,z\}}.$$

As in the proof of Theorem 3.3, to see that there is such a profile, define for each $i \in V$ a binary relation P_i^0 on $\{x, y, z\}$ according to the cell to which i belongs in the partition in Figure 3. Since $A \cap B \subseteq C \subseteq A \cup B$, we have $A \cap B \cap C^c = \varnothing$ and $A^c \cap B^c \cap C = \varnothing$. One can then check by cases that if $i \in V$ belongs to any of the nonempty cells of the partition, then P_i^0 is a strict weak order on $\{x, y, z\}$; hence it can be extended to a strict weak order P_i' on all of X. The resulting profile $\mathbf{P}'$ can be seen to satisfy the bulleted equations above by checking the table in Figure 3. Then since $A \in D(x, y)$ and $B \in D(y, z)$, we have $xf(\mathbf{P}')y$ and $yf(\mathbf{P}')z$, which together imply $xf(\mathbf{P}')z$ by the transitivity condition of TR. Then since $\mathbf{P}_{|\{x,z\}} = \mathbf{P}'_{|\{x,z\}}$, we have $xf(\mathbf{P})z$ by IIA.

For property 4 under FR, suppose $B \notin D(x, z)$, so there is a profile $\mathbf{P}$ in which $B \subseteq \mathbf{P}(x, z)$ and *not* $xf(\mathbf{P})z$. To show that $C \in D(z, y)$, consider a profile $\mathbf{P}^\star$ such that $C \subseteq \mathbf{P}^\star(z, y)$. We claim that there is a profile $\mathbf{P}'$ such that:

$$A \subseteq \mathbf{P}'(x, y) \qquad \mathbf{P}'_{|\{y,z\}} = \mathbf{P}^\star_{|\{y,z\}} \qquad \mathbf{P}'_{|\{x,z\}} = \mathbf{P}_{|\{x,z\}}.$$

The proof that there is such a profile follows the same partitioning strategy as that for property 3 above, only now using the table in Figure 4 and the fact that from $B \cap C \subseteq A \subseteq B \cup C$, we have $A^c \cap B \cap C = \varnothing$ and $A \cap B^c \cap C^c = \varnothing$. Then since $\mathbf{P}_{|\{x,z\}} = \mathbf{P}'_{|\{x,z\}}$, from *not* $xf(\mathbf{P})z$ we have *not* $xf(\mathbf{P}')z$ by IIA, and since $A \in D(x, y)$, we have $xf(\mathbf{P}')y$. Therefore, $zf(\mathbf{P}')y$ by the negative transitivity condition of FR. Then since $\mathbf{P}'_{|\{y,z\}} = \mathbf{P}^\star_{|\{y,z\}}$, we have $zf(\mathbf{P}^\star)y$ by IIA. Thus, $C \in D(z, y)$.

cell	nonempty?	strict weak order
$A \cap B \cap C$	maybe	linear: $xP^0_i zP^0_i y$
$A \cap B \cap C^c$	maybe	$xP^0_i y$, $xP^0_i z$, and P^0_i relates y and z however $P^\star_i$ does
$A \cap B^c \cap C$	maybe	$xP^0_i y$, $zP^0_i y$, and P^0_i relates x and z however P_i does
$A \cap B^c \cap C^c$	no	
$A^c \cap B \cap C$	no	
$A^c \cap B \cap C^c$	maybe	$\begin{cases} \text{linear} : xP^0_i yP^0_i z & \text{if } yP^\star_i z \\ \text{linear} : xP^0_i zP^0_i y & \text{if } zP^\star_i y \\ xP^0_i y, xP^0_i z, yN^0_i z & \text{if } yN^\star_i z \end{cases}$
$A^c \cap B^c \cap C$	maybe	$xP^0_i y$, $zP^0_i y$, and P^0_i relates x and z however P_i does
$A^c \cap B^c \cap C^c$	maybe	$\begin{cases} \text{linear} : xP^0_i zP^0_i y & \text{if } xP_i z, zP^\star_i y \\ \text{linear} : xP^0_i yP^0_i z & \text{if } xP_i z, yP^\star_i z \\ xP^0_i z, xP^0_i y, zN^0_i y & \text{if } xP_i z, zN^\star_i y \\ \text{linear} : zP^0_i xP^0_i y & \text{if } zP_i x, zP^\star_i y \\ \text{linear} : yP^0_i zP^0_i x & \text{if } zP_i x, yP^\star_i z \\ yP^0_i x, zP^0_i x, zN^0_i y & \text{if } zP_i x, zN^\star_i y \\ xP^0_i y, zP^0_i y, xN^0_i z & \text{if } xN_i z, zP^\star_i y \\ yP^0_i x, yP^0_i z, xN^0_i z & \text{if } xN_i z, yP^\star_i z \\ x, y, z \text{ noncomparable in } P^0_i & \text{if } xN_i z, zN^\star_i y \end{cases}$

Figure 4: cells and associated orders for the proof of property 4 in Theorem 3.6

Finally, that property 5 holds is immediate from the definition of decisiveness.

Conversely, suppose properties 1–3 and 5 (resp. properties 1, 2, 4, and 5) hold. We define f as follows: for any $\mathbf{P} \in O(X)^V$, let $xf(\mathbf{P})y$ if and only if $\mathbf{P}(x, y) \in D(x, y)$. First, we claim that f satisfies TR (resp. FR):

- asymmetry: if $xf(\mathbf{P})y$, then *not* $yf(\mathbf{P})x$. Suppose $xf(\mathbf{P})y$, so $\mathbf{P}(x, y) \in D(x, y)$. Since $\mathbf{P}(x, y) \cap \mathbf{P}(y, x) = \varnothing$, property 2 implies $\mathbf{P}(y, x) \notin D(y, x)$ and hence *not* $yf(\mathbf{P})x$.

  Given asymmetry, we may assume without loss of generality in the proofs of transitivity and negative transitivity that x, y, and z are pairwise distinct.

- transitivity: same as in the proof of Theorem 3.3.

- negative transitivity (assuming property 4): same as in the proof of Theorem 3.3.

The proof that f satisfies IIA is also the same as in the proof of Theorem 3.3.

Finally, we claim that D is decisively represented by f. Suppose $A \in D(x,y)$ and consider any profile $\mathbf{P} \in O(X)^V$ in which $A \subseteq \mathbf{P}(x,y)$. Then by property 5, $A \in D(x,y)$ implies $\mathbf{P}(x,y) \in D(x,y)$, which implies $xf(\mathbf{P})y$ by the definition of f. Hence $A \in \overline{D}_f(x,y)$. Conversely, suppose $A \notin D(x,y)$, so by property 1, we have $A = \varnothing$ if $x = y$. Consider a profile $\mathbf{P} \in O(X)^V$ in which $A = \mathbf{P}(x,y)$. Then $A = \mathbf{P}(x,y) \notin D(x,y)$, so by definition of f we have *not* $xf(\mathbf{P})y$. Therefore, $A \notin \overline{D}_f(x,y)$. □

4 Logics

In this section, we show how to parlay the representation results of Section 3 into logical calculi in which we can give purely formal derivations of theorems about decisive coalitions.

For the rest of the paper, fix a finite set X with $|X| \geq 3$.[11]

First, we must specify the *formal language* in which we will give our derivations.

Definition 4.1. Let Coal be a nonempty set, called the set of *coalition labels*. The set Term of *coalition terms* is generated by the following grammar:

$$t ::= a \mid 0 \mid 1 \mid -t \mid (t \sqcap t) \mid (t \sqcup t)$$

where $a \in \mathsf{Coal}$. This means that each $a \in \mathsf{Coal}$ is a term; that '0' and '1' are terms; that the result of prefixing the symbol '$-$' to any term is a term; that the result of putting the symbol '$\sqcap$' (resp. '$\sqcup$') between any two terms and then adding parentheses is also a term; and nothing is a term unless it can be constructed using these grammatical rules.

Let Alt be a set with $|\mathsf{Alt}| = |X|$, called the set of *alternative labels*. The set Form of *formulas* is generated by the following grammar[12]:

$$\varphi ::= t \equiv t \mid D_{x>y}(t) \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid (\varphi \to \varphi)$$

where $t \in \mathsf{Term}$ and $x, y \in \mathsf{Alt}$. This is to be interpreted in the same way as the grammar for terms above. In addition, we define some key abbreviations[13]:

$$(s \sqsubseteq t) := s \sqcap t \equiv s \qquad\qquad D(t) := \bigwedge_{x,y \in \mathsf{Alt},\, x \neq y} D_{x>y}(t),$$

where the second stands for the formula constructed recursively by $\varphi_0 := D_{x_0>y_0}(t)$, $\varphi_{n+1} := (\varphi_n \wedge D_{x_{n+1}>y_{n+1}}(t))$, where $\langle x_0, y_0\rangle, \dots, \langle x_m, y_m\rangle$ are the pairs of distinct elements from Alt.

We may add parentheses for readability and drop parentheses when no ambiguity arises.

So far, the terms and formulas of Definition 4.1 are just meaningless strings of symbols. We endow them with a precise mathematical meaning by defining a *formal semantics*.

First, we give a semantic interpretation of the coalition terms and alternative labels.

[11] The reason for this finiteness assumption is given in footnote 18.

[12] Note that '$\equiv$' is a symbol in our formal language, which will be interpreted using the equality relation (Definition 4.4), which we denote by '$=$'.

[13] We use '$:=$' to mean "is by definition an abbreviation for."

Definition 4.2. Given a nonempty set V, a *coalition labeling* for V is a function $\alpha : \mathsf{Coal} \to \wp(V)$. We extend α to a function $\dot{\alpha} : \mathsf{Term} \to \wp(V)$ as follows:

$$\begin{array}{ll} \dot{\alpha}(a) = \alpha(a) \text{ for } a \in \mathsf{Coal} & \dot{\alpha}(-t) = \alpha(t)^c \\ \dot{\alpha}(0) = \varnothing & \dot{\alpha}(s \sqcap t) = \dot{\alpha}(s) \cap \dot{\alpha}(t) \\ \dot{\alpha}(1) = V & \dot{\alpha}(s \sqcup t) = \dot{\alpha}(s) \cup \dot{\alpha}(t). \end{array}$$

Note that we allow distinct coalition labels to stand for the same coalition. By contrast, we require distinct alternative labels to stand for distinct alternatives.[14]

Definition 4.3. An *alternative labeling* is a bijection $\beta : \mathsf{Alt} \to X$.

We can now give the formulas of Definition 4.1 a precise mathematical "meaning" by defining when they are *true* of a CCR f. There are two meanings we can assign: interpret D as *decisive* or interpret D as *almost decisive*. We call the former the "decisiveness semantics" for our language and the latter the "almost decisiveness semantics."

Definition 4.4. Let V be any nonempty set, f a CCR for $\langle X, V\rangle$, α a coalition labeling for V, and β an alternative labeling. We inductively define the notion of a formula φ being *true of f relative to α, β according to decisiveness semantics* (notation: $f \models_{\alpha,\beta} \varphi$) as follows:

1. $f \models_{\alpha,\beta} s \equiv t$ if and only if $\dot{\alpha}(s) = \dot{\alpha}(t)$;
2. $f \models_{\alpha,\beta} D_{x>y}(t)$ if and only if $\dot{\alpha}(t) \in \overline{D}_f(\beta(x), \beta(y))$;
3. $f \models_{\alpha,\beta} \neg\varphi$ if and only if $f \not\models_{\alpha} \varphi$;
4. $f \models_{\alpha,\beta} \varphi \wedge \psi$ if and only if $f \models_{\alpha,\beta} \varphi$ and $f \models_{\alpha,\beta} \psi$;
5. $f \models_{\alpha,\beta} \varphi \vee \psi$ if and only if $f \models_{\alpha,\beta} \varphi$ or $f \models_{\alpha,\beta} \psi$;
6. $f \models_{\alpha,\beta} \varphi \to \psi$ if and only if $f \not\models_{\alpha,\beta} \varphi$ or $f \models_{\alpha,\beta} \psi$.

A formula φ is simply *true of f according to decisiveness semantics* if and only if φ is true of f relative to every coalition labeling and alternative labeling. If φ is not true of f, then we say that f *refutes* φ.

We define truth according to *almost decisiveness semantics* in the same way, except with a modified clause for D:

2.′ $f \models_{\alpha,\beta} D_{x>y}(t)$ if and only if $\dot{\alpha}(t) \in \widehat{D}_f(\beta(x), \beta(y))$.

According to Definition 4.4, the formula $s \equiv t$ (resp. $s \sqsubseteq t$) "says" that the coalition denoted by s is the same as (resp. a subset of) the coalition denoted by t; the formula $D_{x>y}(t)$ "says" that the coalition denoted by t is (almost) decisive for the alternative denoted by x over the alternative denoted by y; the formula $\neg\varphi$ "says" that φ is not true; the formula $\varphi \wedge \psi$ "says" that both φ and ψ are true; and so on.

Since Definition 4.4 supplies a notion of a formula φ being true of a CCR f, for any class K of CCRs we can ask the following key logical question:

[14] A different approach would allow distinct alternative labels to stand for the same alternative, but would then add identity formulas $x = y$ to our language, so that we could express a requirement of distinctness when needed by a formula of the form $\neg\, x = y$.

Is there a finite formal calculus for deriving all and only the formulas that are true of all CCR*s in* K?

We will show that the answer is *yes* where K is the class of CCRs satisfying UD, IIA, and TR (resp. FR).[15] Since one may wish to vary the class of CCRs at issue, we first define a generic notion of a logic for our language—and then define several specific logics thereafter. For this purpose, we say that a formula $s \equiv t$ is a *valid equation* if and only if $\dot{\alpha}(s) = \dot{\alpha}(t)$ for any coalition labeling α. Examples of valid equations are $-(s \sqcap t) \equiv (-s \sqcup -t)$ and $(s \sqcap (t \sqcup u)) \equiv ((s \sqcap t) \sqcup (s \sqcap u))$. It is a well-known fact of Boolean algebra that there is an algorithm for deciding whether a given formula $s \equiv t$ is a valid equation in this sense, and one can give a finite list of valid equations from which all others are derivable.

Definition 4.5. A *decisiveness logic* is any set $\mathbf{L}$ of formulas—called the *theorems* of $\mathbf{L}$—that contains all instances of the following axioms 1–3 and is closed under rules 4 and 5:

1. all valid equations $s \equiv t$ and $\neg(0 \equiv 1)$;[16]
2. Leibniz's law $s \equiv t \to (\varphi[s/u] \leftrightarrow \varphi[t/u])$, where $\psi[u'/u]$ is the result of replacing all occurrences in ψ of the term u by the term u';
3. all tautologies of propositional logic;
4. if φ and $\varphi \to \psi$ are theorems of $\mathbf{L}$, then ψ is a theorem of $\mathbf{L}$;
5. if φ is a theorem of $\mathbf{L}$, then so is any formula obtained from φ by substituting a coalition term for all occurrences of a coalition label in φ or by substituting an alternative label that does not occur in φ for all occurrences of an alternative label in φ.

We write $\vdash_{\mathbf{L}} \varphi$ to indicate that φ is a theorem of $\mathbf{L}$.

First, we define the logics appropriate for *almost decisiveness*.

Definition 4.6. Let $\widehat{\mathbf{T}}$ be the smallest decisiveness logic that contains the following axioms for $a, b, c \in \mathsf{Coal}$ and $x, y, z \in \mathsf{Alt}$ such that $x \neq y$, $x \neq z$, and $y \neq z$:

1. $D_{x>x}(a)$;
2. $D_{x>y}(a) \to \neg D_{y>x}(-a)$;
3. transitivity axiom: $\big(D_{x>y}(a) \wedge D_{y>z}(b) \wedge (a \sqcap b \sqsubseteq c) \wedge (c \sqsubseteq a \sqcup b)\big) \to D_{x>z}(c)$.

Let $\widehat{\mathbf{W}}$ be the smallest decisiveness logic that contains the axioms of $\widehat{\mathbf{T}}$ as well as the following for $a, b, c \in \mathsf{Coal}$ and $x, y, z \in \mathsf{Alt}$ such that $x \neq y$, $x \neq z$, and $y \neq z$:

4. negative transitivity axiom: $\big(D_{x>y}(a) \wedge (b \sqcap c \sqsubseteq a) \wedge (a \sqsubseteq b \sqcup c)\big) \to \big(D_{x>z}(b) \vee D_{z>y}(c)\big)$.

Next, we define the logics appropriate for *decisiveness*.

Definition 4.7. Let $\overline{\mathbf{T}}$ be the smallest decisiveness logic that contains the following axioms for $a, b, c \in \mathsf{Coal}$ and $x, y, z \in \mathsf{Alt}$ such that $x \neq y$, $x \neq z$, and $y \neq z$:

[15] In fact, we also have all the resources to show that the answer is *yes* for the class of CCRs satisfying LD, IIA, and TR (resp. FR), but we do not want to overload the reader with variations on our results.

[16] The axiom $\neg(0 \equiv 1)$ reflects our assumption that the set V of all voters is nonempty.

1. $D_{x>x}(a) \leftrightarrow \neg(a \equiv 0)$;
2. $(D_{x>y}(a) \wedge ((a \sqcap b) \equiv 0)) \rightarrow \neg D_{y>x}(b)$;
3. $(D_{x>y}(a) \wedge (a \sqsubseteq b)) \rightarrow D_{x>y}(b)$;
4. the transitivity axiom as in Definition 4.6.

Let $\overline{\mathbf{W}}$ be the smallest decisiveness logic that contains the axioms of $\overline{\mathbf{T}}$ as well as the following for $a, b, c \in \mathsf{Coal}$ and $x, y, z \in \mathsf{Alt}$ such that $x \neq y$, $x \neq z$, and $y \neq z$:

5. the negative transitivity axiom as in Definition 4.6.

We can now state the result that our logics can derive as theorems all and only the formulas that are true of all CCRs in the classes we care about. Logicians call the 'all' part *completeness* and the 'only' part *soundness*. In Theorem 4.8 below, note that for soundness we allow V to be of arbitrary cardinality, which makes for a stronger result than restricting to finite V (soundness with respect to a larger class of CCRs implies soundness with respect to a smaller class of CCRs); yet for completeness we restrict to finite V, which makes for a stronger result than allowing V to be of arbitrary cardinality (completeness with respect to a smaller class of CCRs implies completeness with respect to a larger class of CCRs).

Theorem 4.8.

1. Soundness: if φ is a theorem of $\widehat{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$), then for any nonempty set V, φ is true of all CCRs for $\langle X, V\rangle$ satisfying UD, IIA, and TR (resp. FR), according to the almost decisiveness semantics.
2. Completeness: if for any finite nonempty set V, φ is true of all CCRs for $\langle X, V\rangle$ satisfying UD, IIA, and TR (resp. FR), according to the almost decisiveness semantics, then φ is a theorem of $\widehat{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$).
3. Soundness: if φ is a theorem of $\overline{\mathbf{T}}$ (resp. $\overline{\mathbf{W}}$), then for any nonempty set V, φ is true of all CCRs satisfying UD, IIA, and TR (resp. FR), according to the decisiveness semantics.
4. Completeness: if for any finite nonempty set V, φ is true of all CCRs satisfying UD, IIA, and TR (resp. FR), according to the decisiveness semantics, then φ is a theorem of $\overline{\mathbf{T}}$ (resp. $\overline{\mathbf{W}}$).

Proof. Soundness is straightforward. First, simply check that each axiom of the logic is true of all CCRs in the relevant class. Second, check our two rules: if φ and $\varphi \rightarrow \psi$ are both true of all CCRs in the class, then so is ψ; and if φ is true of all CCRs in the class, then so is any formula obtained from φ by the substitution rule. Since the logic is defined inductively as the smallest set of formulas that contains the axioms and is closed under the rules, it follows that all theorems of the logic are true of all CCRs in the relevant class.

Completeness is more involved—we give a sketch of the proof in the Appendix. □

The existence of a sound and complete formal logic for reasoning about (almost) decisive coalitions raises the possibility that such reasoning might be carried out automatically by a computer. A set Σ of formulas in our formal language is *decidable* if and only if there is an *algorithm* which, given any formula of the language, will decide in a finite amount of time whether the formula belongs to the set Σ. The obvious question to ask is whether the set of theorems of our sound and complete logic is decidable. The answer is affirmative.

Theorem 4.9. The set of theorems of $\widehat{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$, $\overline{\mathbf{T}}$, $\overline{\mathbf{W}}$) is decidable.

Proof. Decidability follows from our completeness proof in the Appendix. Inspection of the proof shows that for any formula φ, we can compute an integer $n(\varphi)$ such that if φ is not a theorem, then there is a finite nonempty set V with $|V| \leq n(\varphi)$ such that some CCR f for $\langle X, V\rangle$ satisfying the relevant properties (e.g., UD, IIA, and TR) refutes φ. For every such V, of which there are only finitely many up to isomorphism, we can check every such CCR for $\langle X, V\rangle$, of which there are only finitely many. Then since φ is a theorem if and only if it is not refuted by any of these CCRs, we have an algorithm for deciding theoremhood. $\square$

A further question, which we leave for future work, is what is the *computational complexity* of deciding whether a formula is a theorem of $\widehat{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$, $\overline{\mathbf{T}}$, $\overline{\mathbf{W}}$).[17]

5 Derivations

Our reason for introducing the logics of the previous section is so that we may give simple formal derivations of important social-choice theoretic facts about (almost) decisive coalitions. Such derivations allow us to easily see which axioms and further assumptions are used. In this section, we give derivations of a series of lemmas that lead up to Arrow's theorem. In Section 6, we derive some variations: the Oligarchy Theorem and Wilson's theorem.

A nice feature of the lemmas to follow is that they almost all go through in *both* the logic of decisive coalitions and the logic of almost decisive coalitions, since the lemmas only use axioms common to both logics. To make this clear, we will use the following notation:

'$\vdash_{\mathbf{T}} \varphi$' means that φ is a theorem of *both* $\overline{\mathbf{T}}$ and $\widehat{\mathbf{T}}$;

'$\vdash_{\mathbf{W}} \varphi$' means that φ is a theorem of *both* $\overline{\mathbf{W}}$ and $\widehat{\mathbf{W}}$.

In what follows, whenever we write '(almost) decisive' in some statement, we intend that the reader may read the statement with 'decisive' or with 'almost decisive'.

5.1 Existential Assumptions

We will prove our series of lemmas conditional on certain assumptions about the existence of (almost) decisive coalitions. The strongest of such assumptions is the Pareto principle stating that the set of all voters is (almost) decisive, which we express as $D(1)$. Yet we are also interested in weaker assumptions (implied by Pareto, assuming UD).

[17] A lower bound is easy: since our logic includes propositional logic, deciding theoremhood is a co-NP-hard problem. The question is: what is an upper bound?

Let the *existential assumption* (EA) be the assumption that for every pair of alternatives x, y, there is some coalition $c_{x,y}$ that is (almost) decisive for x over y:[18]

$$EA := \bigwedge_{x,y \in \mathsf{Alt}} D_{x>y}(c_{x,y}).$$

Let the *weak existential assumption* (WEA) be the assumption that for every pair of distinct alternatives x, y, there is some coalition $c_{x,y}$ that is *not* (almost) decisive for x over y:

$$WEA := \bigwedge_{x,y \in \mathsf{Alt},\, x \neq y} \neg D_{x>y}(c_{x,y}).$$

If we interpret D as decisiveness, then WEA is equivalent to the well-known condition of

> non-imposition (NI): a CCR f satisfies non-imposition if and only if for every $x, y \in X$, there exists a profile $\mathbf{P} \in \text{dom}(f)$ such that *not* $xf(\mathbf{P})y$.

Finally, let *non-emptiness* (NE) be the assumption that there is some pair of distinct alternatives s, t and some coalition e such that e is (almost) decisive for s over t:

$$NE := D_{s>t}(e).$$

Note that EA implies WEA and NE, but not vice versa.

Where $\boldsymbol{c}$ is a tuple of $c_{x,y}$'s, we may write $EA(\boldsymbol{c})$ (resp. $WEA(\boldsymbol{c})$) to make clear that we obtain a different formula for each different choice of the coalition labels $c_{x,y}$. But by proving that an implication $EA \to \varphi$ (resp. $WEA \to \varphi$) holds for *any* choice of the labels $c_{x,y}$, including $c_{x,y}$'s all of which do not occur in φ, we will thereby prove that the *existence of some coalitions or other* satisfying EA (resp. WEA) is sufficient to imply φ. In this way, we can capture existential assumptions without having explicit quantification over coalitions in our formal language. Similarly, we may write $NE(e)$ or $NE(e; s, t)$ to make clear the dependence on the choice of e, s, and t, but the same points apply.

To be precise, throughout we will take advantage of the following familiar mathematical point, formalized in first-order logic: if we have proved an implication $\psi \to \varphi$, where all of the variables of ψ are among $v_1, \dots, v_n$, all of the variables of φ are among $v'_1, \dots, v'_m$, and $\{v_1, \dots, v_n\} \cap \{v'_1, \dots, v'_m\} = \varnothing$, then we may infer $(\exists v_1 \dots \exists v_1 \psi) \to (\forall v'_1 \dots \forall v'_m \varphi)$, where $\exists$ and $\forall$ are the existential ("there exists") and universal ("for all") quantifiers, respectively.

5.2 Contagion Lemma from EA

Our first set of formal derivations establish the "contagion" lemma that (almost) decisiveness over one pair of alternatives spreads to (almost) decisiveness over all pairs of alternatives. This is often proved using the Pareto principle, but the weaker principle EA suffices.

[18] Note that here we use the assumption that Alt is finite (which follows from our assumptions that X is finite and that $|\mathsf{Alt}| = |X|$), so that EA can be written as a finite conjunction in our language.

Lemma 5.1. Assume $x \neq y$, $x \neq v$, $w \neq y$.[19]

1. $\vdash_{\mathbf{T}} D_{y>v}(c) \to \big(D_{x>y}(a) \to D_{x>v}(a)\big)$ 2. $\vdash_{\mathbf{T}} D_{w>x}(c) \to \big(D_{x>y}(a) \to D_{w>y}(a)\big)$.

Proof. For part 1, if $y = v$, the principle is a tautology of propositional logic. Otherwise, we derive the principle as follows:

(1) $\big(D_{x>y}(a) \wedge D_{y>v}(c) \wedge (a \sqcap c \sqsubseteq a) \wedge (a \sqsubseteq a \sqcup c)\big) \to D_{x>v}(a)$, instance of transitivity axiom

(2) $(a \sqcap c \sqsubseteq a) \wedge (a \sqsubseteq a \sqcup c)$, valid Boolean inequalities

(3) $\big(D_{x>y}(a) \wedge D_{y>v}(c)\big) \to D_{x>v}(a)$ from (1) and (2) by propositional logic

(4) $D_{y>v}(c) \to \big(D_{x>y}(a) \to D_{x>v}(a)\big)$ from (3) by propositional logic.

The proof for part 2 is analogous, using:

(1′) $\big(D_{w>x}(c) \wedge D_{x>y}(a) \wedge (c \sqcap a \sqsubseteq a) \wedge (a \sqsubseteq c \sqcup a)\big) \to D_{w>y}(a)$, instance of transitivity axiom. □

Lemma 5.2. Assume $x \neq y$, $x \neq z$, $w \neq z$, $w \neq v$.

$$\vdash_{\mathbf{T}} \big(D_{y>z}(c) \wedge D_{w>x}(d) \wedge D_{z>v}(e)\big) \to \big(D_{x>y}(a) \to D_{w>v}(a)\big).$$

Proof.

(1) $D_{y>z}(c) \to \big(D_{x>y}(a) \to D_{x>z}(a)\big)$ by Lemma 5.1.1

(2) $D_{w>x}(d) \to \big(D_{x>z}(a) \to D_{w>z}(a)\big)$ by Lemma 5.1.2

(3) $D_{z>v}(e) \to \big(D_{w>z}(a) \to D_{w>v}(a)\big)$ by Lemma 5.1.1

(4) $\big(D_{y>z}(c) \wedge D_{w>x}(d) \wedge D_{z>v}(e)\big) \to \big(D_{x>y}(a) \to D_{w>v}(a)\big)$ from (1)–(3) by propositional logic. □

Lemma 5.3. Assume $x \neq y$.

1. $\vdash_{\mathbf{T}} EA \to (D_{x>y}(a) \to D(a))$ 2. $\vdash_{\mathbf{T}} D(1) \to (D_{x>y}(a) \to D(a))$.

Proof. Part 1 is by repeated application of Lemma 5.2. Part 2 is immediate from part 1. □

5.3 Contagion Lemma from WEA

We can also prove the contagion lemma from the weaker assumption of WEA, rather than EA, if we strengthen the logic from $\widehat{\mathbf{T}}$ (resp. $\overline{\mathbf{T}}$) to $\widehat{\mathbf{W}}$ (resp. $\overline{\mathbf{W}}$).

Lemma 5.4. Assume $x \neq y$, $x \neq v$, $w \neq y$.

1. $\vdash_{\mathbf{W}} \neg D_{v>y}(c) \to \big(D_{x>y}(a) \to D_{x>v}(a)\big)$ 2. $\vdash_{\mathbf{W}} \neg D_{x>w}(c) \to \big(D_{x>y}(a) \to D_{w>y}(a)\big)$.

[19] For $\widehat{\mathbf{T}}$, we can drop the assumption that $x \neq v$ for part 1 and the assumption that $w \neq y$ for part 2; and we can drop the same assumptions for $\overline{\mathbf{T}}$ if we add the assumption $\neg(a \equiv 0)$ to the conditionals.

Proof. For part 1, if $y = v$, the principle is a tautology of propositional logic. Otherwise, we derive the principle as follows:

(1) $\big(D_{x>y}(a) \wedge (a \sqcap c \sqsubseteq a) \wedge (a \sqsubseteq a \sqcup c)\big) \to \big(D_{x>v}(a) \vee D_{v>y}(c)\big)$, instance of negative transitivity axiom

(2) $(a \sqcap c \sqsubseteq a) \wedge (a \sqsubseteq a \sqcup c)$, valid Boolean inequalities

(3) $D_{x>y}(a) \to \big(D_{x>v}(a) \vee D_{v>y}(c)\big)$ from (1) and (2)

(4) $\neg D_{v>y}(c) \to (D_{x>y}(a) \to D_{x>v}(a))$ from (3) by propositional logic.

For part 2, the proof is analogous, using:

(1′) $\big(D_{x>y}(a) \wedge (c \sqcap a \sqsubseteq c) \wedge (c \sqsubseteq c \sqcup a)\big) \to \big(D_{x>w}(c) \vee D_{w>y}(a)\big)$, instance of negative transitivity axiom. □

Lemma 5.5. Assume $x \neq y$, $x \neq z$, $w \neq z$, $w \neq v$.

$$\vdash_{\mathbf{W}} \big(\neg D_{z>y}(c) \wedge \neg D_{x>w}(d) \wedge \neg D_{v>z}(e)\big) \to \big(D_{x>y}(a) \to D_{w>v}(a)\big).$$

Proof.

(1) $\neg D_{z>y}(c) \to \big(D_{x>y}(a) \to D_{x>z}(a)\big)$ by Lemma 5.4.1

(2) $\neg D_{x>w}(d) \to \big(D_{x>z}(a) \to D_{w>z}(a)\big)$ by Lemma 5.4.2

(3) $\neg D_{v>z}(e) \to \big(D_{w>z}(a) \to D_{w>v}(a)\big)$ by Lemma 5.4.1

(4) $\big(\neg D_{z>y}(c) \wedge \neg D_{x>w}(d) \wedge \neg D_{v>z}(e)\big) \to \big(D_{x>y}(a) \to D_{w>v}(a)\big)$ from (1)–(3) by propositional logic. □

Lemma 5.6. Assume $x \neq y$.

$$\vdash_{\mathbf{W}} WEA \to (D_{x>y}(a) \to D(a)).$$

Proof. By repeated application of Lemma 5.5. □

5.4 Intersection Lemma

Our next set of derivations establishes what could be called Hansson's lemma (see Hansson 1976, p. 92) for (almost) decisive coalitions: UD, IIA, and TR together imply that the set of (almost) decisive coalitions is closed under intersections.

Lemma 5.7. Assume $x \neq y$, $y \neq z$, and $x \neq z$.

$$\vdash_{\mathbf{T}} \big(D_{x>z}(a) \wedge D_{z>y}(b)\big) \to D_{x>y}(a \sqcap b).$$

Proof.

(1) $\big(D_{x>z}(a) \wedge D_{z>y}(b) \wedge (a \sqcap b \sqsubseteq a \sqcap b) \wedge (a \sqcap b \sqsubseteq a \sqcup b)\big) \to D_{x>y}(a \sqcap b)$, instance of transitivity axiom

(2) $(a \sqcap b \sqsubseteq a \sqcap b) \wedge (a \sqcap b \sqsubseteq a \sqcup b)$, valid Boolean inequalities

(3) $\big(D_{x>z}(a) \wedge D_{z>y}(b)\big) \to D_{x>y}(a \sqcap b)$ from (1) and (2) by propositional logic. □

Lemma 5.8. Assume $x \neq y$, $y \neq z$, and $x \neq z$.

$$\vdash_{\mathbf{T}} (D_{y>z}(c) \wedge D_{z>x}(d)) \to \big((D_{x>y}(a) \wedge D_{x>y}(b)) \to D_{x>y}(a \sqcap b)\big).$$

Proof.

(1) $\big(D_{x>z}(a) \wedge D_{z>y}(b)\big) \to D_{x>y}(a \sqcap b)$ by Lemma 5.7

(2) $D_{y>z}(c) \to \big(D_{x>y}(a) \to D_{x>z}(a)\big)$ by Lemma 5.1.1

(3) $D_{z>x}(d) \to \big(D_{x>y}(b) \to D_{z>y}(b)\big)$ by Lemma 5.1.2

(4) $(D_{y>z}(c) \wedge D_{z>x}(d)) \to \big((D_{x>y}(a) \wedge D_{x>y}(b)) \to (D_{x>z}(a) \wedge D_{z>y}(b))\big)$ from (2) and (3) by propositional logic

(5) $(D_{y>z}(c) \wedge D_{z>x}(d)) \to \big((D_{x>y}(a) \wedge D_{x>y}(b)) \to D_{x>y}(a \sqcap b)\big)$ from (1) and (4) by propositional logic. □

Lemma 5.9. $\vdash_{\mathbf{T}} (D(a) \wedge D(b)) \to D(a \sqcap b))$.

Proof. By Lemma 5.8, $\vdash_{\mathbf{T}} EA \to ((D(a) \wedge D(b)) \to D(a \sqcap b))$ for any formulation of EA with any tuple of coalition labels, i.e., any $EA(\mathbf{c})$. Then since $\vdash_{\mathbf{T}} D(a) \to EA(\mathbf{a})$, by propositional logic we obtain $\vdash_{\mathbf{T}} (D(a) \wedge D(b)) \to D(a \sqcap b))$. □

In Section 6.1, we show how Lemma 5.9 can be combined with one other lemma to yield a version of Gibbard's [2014] Oligarchy Theorem.

5.5 Filter Lemmas

Our next set of lemmas shows that UD, IIA, TR, and P together imply that the set of almost decisive sets is not only closed under intersection, but also closed under supersets. Thus, the set of almost decisive sets is a *filter*. For decisive sets, closure under supersets is immediate, as $(D_{x>y}(a) \wedge (a \sqsubseteq b)) \to D_{x>y}(b)$ is an axiom of $\overline{\mathbf{T}}$.

Lemma 5.10. Assume $x \neq y$, $y \neq z$, and $x \neq z$.

$$\vdash_{\widehat{\mathbf{T}}} (D_{y>z}(c) \wedge D_{z>y}(1)) \to \big((D_{x>y}(a) \wedge (a \sqsubseteq b)) \to D_{x>y}(b)\big).$$

Proof.

(1) $D_{y>z}(c) \to (D_{x>y}(a) \to D_{x>z}(a))$ by Lemma 5.1.1

(2) $\big(D_{x>z}(a) \wedge D_{z>y}(1) \wedge (a \sqcap 1 \sqsubseteq b) \wedge (b \sqsubseteq a \sqcup 1)\big) \to D_{x>y}(b)$, instance of transitivity axiom

(3) $(a \sqsubseteq b) \to ((a \sqcap 1 \sqsubseteq b) \wedge (b \sqsubseteq a \sqcup 1))$ by Boolean reasoning[20]

(4) $(a \sqsubseteq b) \to \big((D_{x>z}(a) \wedge D_{z>y}(1)) \to D_{x>y}(b)\big)$ from (2) and (3) by propositional logic

[20] By 'Boolean reasoning', we mean a combined use of valid Boolean equations, Leibniz's law, and propositional logic.

(5) $(D_{y>z}(c) \wedge D_{z>y}(1)) \to \big((D_{x>y}(a) \wedge (a \sqsubseteq b)) \to D_{x>y}(b)\big)$ from (1) and (4) by propositional logic. $\square$

Lemma 5.11.

$$1.\ \vdash_{\overline{\mathbf{T}}} (D(a) \wedge a \sqsubseteq b) \to D(b) \qquad 2.\ \vdash_{\widehat{\mathbf{T}}} D(1) \to \big((D(a) \wedge a \sqsubseteq b) \to D(b)\big).$$

Proof. Part 1 uses the axiom $(D_{x>y}(a) \wedge (a \sqsubseteq b)) \to D_{x>y}(b)$ of $\overline{\mathbf{T}}$ repeatedly and the definition of $D(t)$. Part 2 uses Lemma 5.10 repeatedly and the definition of $D(t)$. $\square$

Lemma 5.12.

$$1.\ \vdash_{\overline{\mathbf{T}}} D(a \sqcap b) \to (D(a) \wedge D(b)) \qquad 2.\ \vdash_{\widehat{\mathbf{T}}} D(1) \to \big(D(a \sqcap b) \to (D(a) \wedge D(b))\big).$$

Proof. Immediate from Lemma 5.11 using the Boolean inequalities $a \sqcap b \sqsubseteq a$ and $a \sqcap b \sqsubseteq b$. $\square$

In addition, the set of (almost) decisive sets is a *proper* filter (assuming P for almost decisiveness), meaning that the empty set does not belong to the filter.

Lemma 5.13.

$$1.\ \vdash_{\overline{\mathbf{T}}} \neg D(0) \qquad 2.\ \vdash_{\widehat{\mathbf{T}}} D(1) \to \neg D(0).$$

Proof. Part 1 uses the axiom $(D_{x>y}(0) \wedge ((0 \sqcap 0) \equiv 0)) \to \neg D_{y>x}(0)$, the valid equation $(0 \sqcap 0) \equiv 0$, the definition of $D(0)$, and propositional logic. For part 2, by the axiom $D_{w>v}(1) \to \neg D_{v>w}(-1)$ and Boolean reasoning, we have $\vdash D_{w>v}(1) \to \neg D_{v>w}(0)$, which clearly implies $\vdash D(1) \to \neg D(0)$ given the definitions of $D(1)$ and $D(0)$. $\square$

5.6 Ultrafilter Lemma

Our final lemma shows, as Kirman and Sondermann [1972] and Hansson [1976] observed, that Arrow's assumptions imply that the set of (almost) decisive coalitions is not only a proper filter but in fact an *ultrafilter*. For this result, we use FR instead of TR.

Lemma 5.14. $\vdash_{\mathbf{W}} (WEA \wedge NE) \to (D(a) \vee D(-a))$.

Proof. By propositional logic, it suffices to prove $\vdash (WEA \wedge NE) \to (\neg D(a) \to D(-a))$. Since $\neg D(a)$ is equivalent to $\bigvee_{x,y \in \mathsf{Alt},\, x \neq y} \neg D_{x>y}(a)$, by propositional logic it suffices to show that for each distinct $u, w \in \mathsf{Alt}$, we have $\vdash (WEA \wedge NE) \to (\neg D_{u>w}(a) \to D(-a))$:

(1) $WEA \to (D_{s>t}(e) \to D_{u>v}(e))$ by Lemma 5.5

(2) $(WEA \wedge NE) \to D_{u>v}(e)$ by (1), since $NE := D_{s>t}(e)$

(3) $\big(D_{u>v}(e) \wedge (a \sqcap -a \sqsubseteq e) \wedge (e \sqsubseteq a \sqcup -a)\big) \to \big(D_{u>w}(a) \vee D_{w>v}(-a)\big)$, instance of negative transitivity axiom

(4) $(a \sqcap -a \sqsubseteq e) \wedge (e \sqsubseteq a \sqcup -a)$, valid Boolean inequalities

(5) $D_{u>v}(e) \to (D_{u>w}(a) \vee D_{w>v}(-a))$ from (2) and (3) by propositional logic

(6) $D_{u>v}(e) \to (\neg D_{u>w}(a) \to D_{w>v}(-a))$ from (4) by propositional logic

(7) $WEA \to (D_{w>v}(-a) \to D(-a))$ by Lemma 5.6

(8) $(WEA \wedge NE) \to (\neg D_{u>w}(a) \to D(-a))$ by (1), (5), and (6). $\square$

5.7 Arrow's Theorem

We can now prove Arrow's theorem for both decisive coalitions and almost decisive coalitions. Dictatorship is equivalent to there being a singleton coalition $\{i\}$ that is decisive for every pair of alternatives. Almost dictatorship is the weaker condition that there is a singleton coalition $\{i\}$ that is almost decisive for every pair of alternatives:

> almost dictatorship: there is an $i \in V$ such that for every $\mathbf{P} \in \mathrm{dom}(f)$ and $x, y \in X$, if xP_iy and yP_jx for all $j \in V \setminus \{i\}$, then $xf(\mathbf{P})y$.

Assuming LD or UD and PA, almost dictatorship is equivalent to dictatorship.

As in Section 1, to prove that there is an (almost) dictatorship in the case where V is finite, it suffices to show that the set of (almost) decisive coalitions is an ultrafilter. Thus, in our formalism, Arrow's theorem for (almost) decisive sets is the following.

Theorem 5.15 (Arrow's Theorem)**.** In the logic $\mathbf{W}$, we have:

1. $\vdash_{\mathbf{W}} D(1) \to \neg D(0)$;
2. $\vdash_{\mathbf{W}} D(1) \to (D(a) \vee D(-a))$;
3. $\vdash_{\mathbf{W}} D(1) \to \big((D(a) \wedge D(b)) \leftrightarrow D(a \sqcap b)\big)$.

Combining the lemmas of Section 5.6, we obtain the following stronger result.

Theorem 5.16 (Stronger Arrow's Theorem)**.** In the logics $\mathbf{T}$ and $\mathbf{W}$, we have:

1. $\vdash_{\overline{\mathbf{T}}} \neg D(0)$ and $\vdash_{\widehat{\mathbf{T}}} D(1) \to \neg D(0)$;
2. $\vdash_{\mathbf{W}} (WEA \wedge NE) \to (D(a) \vee D(-a))$;
3. $\vdash_{\mathbf{T}} ((D(a) \wedge D(b)) \to D(a \sqcap b)$;
4. $\vdash_{\overline{\mathbf{T}}} D(a \sqcap b) \to (D(a) \wedge D(b))$ and $\vdash_{\widehat{\mathbf{T}}} D(1) \to \big(D(a \sqcap b) \to (D(a) \wedge D(b))\big)$.

6 Variations

In this section, we briefly show how some of the lemmas of Section 5 can be combined with additional lemmas to give formal derivations of variations on the Arrovian theme.

6.1 Almost Oligarchy Theorem

Our first example is a version of Gibbard's [2014] Oligarchy Theorem (cf. Guha 1972, Mas-Colell and Sonnenschein 1972, and Weymark 1984). For its statement, we adopt the term 'semi-decisive' from Sen [2017, p. 125]. Others instead say that a coalition is "blocking" [Suzumura, 1983, p. 71] or has "veto power" [Campbell and Kelly, 2002, p. 59].

Definition 6.1. Let f be a CCR for $\langle X, V\rangle$. For any $x, y \in X$ and $A \subseteq V$:

1. A is *almost semi-decisive for x over y according to* f if and only if for all $\mathbf{P} \in \mathrm{dom}(f)$, if $A = \mathbf{P}(x, y)$ and $A^c = \mathbf{P}(y, x)$, then *not* $yf(\mathbf{P})x$;

2. A is *almost semi-decisive according to* f if for all $x, y \in X$, A is almost semi-decisive for x over y according to f.

Clearly A is almost semi-decisive for x over y if and only if A^c is not almost decisive for y over x. Thus, we can define almost semi-decisiveness predicates in the language of Section 4, assuming the *almost decisiveness semantics* for D (recall Definition 4.4):

$$S_{x>y}(a) := \neg D_{y>x}(-a) \qquad\qquad S(a) := \bigwedge_{x,y\in\mathsf{Alt}} S_{x>y}(a).$$

Definition 6.2. Let f be a CCR and $A \subseteq V$.

1. A is an *almost oligarchy according to* f if and only if A is almost decisive according to f and for each $i \in A$, $\{i\}$ is almost semi-decisive according to f.

2. A is a *strong almost oligarchy according to* f if and only if A is almost decisive according to f and for each nonempty $B \subseteq A$, B is almost semi-decisive according to f.

We now obtain what could be called the "almost oligarchy theorem."

Theorem 6.3. Assume that $|X| \geq 3$ and V is finite. If a CCR f satisfies UD, IIA, EA, and TR, then there exists a strong almost oligarchy according to f.

The two key lemmas for the proof of Theorem 6.3 are (i) that the set of almost decisive coalitions is non-empty and closed under intersection, and (ii) that if a coalition A is not almost semi-decisive, then A^c is almost decisive. For by (i) and the finiteness of V, the intersection B of all almost decisive coalitions is itself almost decisive, and then by (ii), any nonempty subset $A \subseteq B$ is almost semi-decisive; for if not, then A^c is almost decisive, which implies $B \subseteq A^c$ and hence $A \subseteq B \subseteq A^c$, so $A = \varnothing$, a contradiction. We already proved a formal version of (i) (non-emptiness follows from Lemma 5.3 and closure under intersection from Lemma 5.9), so it only remains to prove the following formal version of (ii).

Lemma 6.4. $\vdash_{\mathbf{T}} EA \to (\neg S(a) \to D(-a))$.

Proof. By propositional logic, $\neg S(a)$ is equivalent to $\bigvee_{x,y\in\mathsf{Alt}} \neg S_{x>y}(a)$, which by definition is $\bigvee_{x,y\in\mathsf{Alt}} D_{x>y}(-a)$. Thus, it suffices to show that for any distinct x, y, we have $\vdash_{\mathbf{T}} EA \to (D_{x>y}(-a) \to D(-a))$, which we already established in Lemma 5.3. $\square$

We cannot define the notion of semi-decisiveness[21] in terms of either almost decisiveness or decisiveness, so if we want to prove the original oligarchy theorem, rather than the "almost oligarchy theorem," we must extend our language. For example, one could add a new primitive predicate $\overline{S}_{x>y}$ for semi-decisiveness to our language. One could then study the joint logic of all these decisiveness notions, which we leave for future work.

[21] A is *semi-decisive for x over y according to* f if and only if for any $\mathbf{P} \in \mathrm{dom}(f)$, if $A \subseteq \mathbf{P}(x, y)$, then *not* $yf(\mathbf{P})x$.

6.2 Almost Wilson's Theorem

Our second example is an "almost" version of Wilson's [1972] theorem, which depends on the following key notions.

Definition 6.5. Let f be a CCR for $\langle X, V\rangle$. For any $x, y \in X$ and $A \subseteq V$:

1. A is *almost inversely decisive for x over y according to f* if and only if for all $\mathbf{P} \in \mathrm{dom}(f)$, if $A = \mathbf{P}(x, y)$ and $A^c = \mathbf{P}(y, x)$, then $yf(\mathbf{P})x$;
2. A is *almost inversely decisive according to f* if and only if for every $x, y \in X$, A is almost inversely decisive for x over y according to f;
3. A is *inversely decisive for x over y according to f* if and only if for all $\mathbf{P} \in \mathrm{dom}(f)$, if $A \subseteq \mathbf{P}(x, y)$, then $yf(\mathbf{P})x$;
4. A is *inversely decisive according to f* if and only if for every distinct $x, y \in X$, A is inversely decisive for x over y according to f.

We can define almost inverse decisiveness predicates in the language of Section 4, assuming the *almost decisiveness semantics* for D (recall Definition 4.4):

$$I_{x>y}(a) := D_{y>x}(-a) \qquad\qquad I(a) := D(-a).$$

Just as the Pareto assumption $D(1)$ gave us that the set of almost decisive coalitions is closed under *supersets*, the *inverse* Pareto assumption $D(0)$ gives us that the set of almost decisive coalitions is closed under *subsets*.

Lemma 6.6. Assume $x \neq y$, $y \neq z$, and $x \neq z$.

$$\vdash_{\mathbf{T}} (D_{y>z}(c) \wedge D_{z>y}(0)) \to \big((D_{x>y}(a) \wedge (b \sqsubseteq a)) \to D_{x>y}(b)\big).$$

Proof. The proof is the same as that of Lemma 5.10, with two exceptions: all occurrences of 1 in the proof Lemma 5.10 are replaced by 0, and all occurrences of $(a \sqsubseteq b)$ in the proof of Lemma 5.10 are replaced by $(b \sqsubseteq a)$. The key point is that

$$\big(D_{x>z}(a) \wedge D_{z>y}(0) \wedge (a \sqcap 0 \sqsubseteq b) \wedge (b \sqsubseteq a \sqcup 0)\big) \to D_{x>y}(b)$$

is an instance of the transitivity axiom, and $(b \sqsubseteq a) \to ((a \sqcap 0 \sqsubseteq b) \wedge (b \sqsubseteq a \sqcup 0))$ is provable by Boolean reasoning. □

Lemma 6.7. $\vdash_{\mathbf{T}} D(0) \to \big((D(a) \wedge (b \sqsubseteq a)) \to D(b)\big)$.

Proof. By repeated application of Lemma 6.6. □

Next, just as the almost decisive coalitions are closed under intersection, the almost inversely decisive coalitions are closed under union.

Lemma 6.8. $\vdash_{\mathbf{T}} EA \to ((D(a) \wedge D(b)) \to D(a \sqcup b))$.

Proof. The proof is the same as that of Lemma 5.9, which uses the proof of Lemma 5.8, which in turn uses the proof of Lemma 5.7, with two exceptions. First, the first two lines of the proof of Lemma 5.7 change to:

(1′) $\big(D_{x>z}(a) \wedge D_{z>y}(b) \wedge (a \sqcap b \sqsubseteq a \sqcup b) \wedge (a \sqcup b \sqsubseteq a \sqcup b)\big) \to D_{x>y}(a \sqcup b)$, instance of transitivity axiom

(2′) $(a \sqcap b \sqsubseteq a \sqcup b) \wedge (a \sqcup b \sqsubseteq a \sqcup b)$, valid Boolean inequalities.

Second, in the later lines of the proofs, all instances of $a \sqcap b$ are replaced by $a \sqcup b$. $\Box$

We can now state and prove the version of Wilson's theorem for almost decisive and almost inversely decisive sets.

Theorem 6.9 (Almost Wilson's Theorem)**.** In the logics $\mathbf{T}$ and $\mathbf{W}$, we have:

1. $\vdash_{\mathbf{W}} (WEA \wedge NE) \to (D(1) \vee I(1))$
2. $\vdash_{\mathbf{W}} D(1) \to (D(a) \vee D(-a))$
3. $\vdash_{\mathbf{T}} D(1) \to \big(D(a \sqcap b) \leftrightarrow (D(a) \wedge D(b))\big)$
4. $\vdash_{\mathbf{W}} I(1) \to (I(a) \vee I(-a))$
5. $\vdash_{\mathbf{T}} I(1) \to \big(I(a \sqcap b) \leftrightarrow (I(a) \wedge I(b))\big)$.

Proof. Part 1 follows from Theorem 5.16.2 with $a := 1$ and the definition $I(1) := D(-1)$.

Parts 2–3 follow from Theorem 5.15.2-3.

For part 4, $I(1) \to (I(a) \vee I(-a))$ is an abbreviation for $D(-1) \to (D(-1) \vee D(--1))$. Given the axiom $D_{x>y}(-1) \to \neg D_{y>x}(1)$, we have $\vdash D(-1) \to WEA(\mathbf{1})$.[22] Clearly we also have $\vdash D(-1) \to NE(-1)$. Then by Lemma 5.14, we have $\vdash D(-1) \to (D(a) \vee D(-a))$ and hence $\vdash D(-1) \to (D(-1) \vee D(--1))$.

For part 5, we must show $\vdash_{\mathbf{T}} D(-1) \to (D(-(a \sqcap b)) \leftrightarrow (D(-a) \wedge D(-b))$. By Boolean reasoning, it suffices to show that $\vdash_{\mathbf{T}} D(0) \to (D(-a \sqcup -b) \leftrightarrow (D(-a) \wedge D(-b))$, which clearly follows from Lemmas 6.7 and 6.8. $\Box$

We cannot define the notion of inverse decisiveness in terms of either almost decisiveness or decisiveness, so if we want to prove the original version of Wilson's theorem, rather than the "almost" version, we must extend our language to one with predicates for both decisiveness and inverse decisiveness, as in Section 7.4.

7 Relating Kinds of Decisiveness

In Section 4, we presented logical calculi for reasoning about decisive coalitions or about almost decisive coalitions—but not for reasoning about the relations between decisive and almost decisive coalitions at the same time. In this brief section, we explain how to apply our methodology to obtain combined logics for decisive and almost decisive coalition, as well as inversely decisive and almost inversely decisive coalitions.

7.1 Decisiveness Regimes

Our first step is to extend the representation theory of Section 3 to cover almost decisive coalitions and decisive coalitions at the same time.

Definition 7.1. A *decisiveness regime* for $\langle X, V\rangle$ is a pair $\langle \widehat{D}, \overline{D}\rangle$ where $\widehat{D} : X^2 \to \wp(\wp(V))$ and $\overline{D} : X^2 \to \wp(\wp(V))$.

[22] By $WEA(\mathbf{1})$, we mean the WEA formula of Section 5.1 in which each $c_{x,y}$ is 1.

Definition 7.2. Let K be a class of CCRs for $\langle X, V\rangle$. A decisiveness regime $\langle \widehat{D}, \overline{D}\rangle$ for $\langle X, V\rangle$ is *representable in* K if and only if there is an $f \in K$ such that $\widehat{D} = \widehat{D}_f$ and $\overline{D} = \overline{D}_f$, in which case we say that f *represents* $\langle \widehat{D}, \overline{D}\rangle$.

We extend Theorem 3.3 to decisiveness regimes as follows.

Theorem 7.3. Let X and V be nonempty sets with $|X| \geq 3$. A decisiveness regime $\langle \widehat{D}, \overline{D}\rangle$ for $\langle X, V\rangle$ is representable in the class of CCRs for $\langle X, V\rangle$ satisfying LD, IIA, and TR (resp. FR) if and only if for all $A, B, C \subseteq V$ and $x, y, z \in X$ with $x \neq y$, $y \neq z$, and $x \neq z$:

1. $A \in \widehat{D}(x, x)$;
2. if $A \in \widehat{D}(x, y)$, then $A^c \notin \widehat{D}(y, x)$;
3. for TR: if $A \in \widehat{D}(x, y)$, $B \in \widehat{D}(y, z)$, and $A \cap B \subseteq C \subseteq A \cup B$, then $C \in \widehat{D}(x, z)$;
4. for FR: if $A \in \widehat{D}(x, y)$ and $B \cap C \subseteq A \subseteq B \cup C$, then $B \in \widehat{D}(x, z)$ or $C \in \widehat{D}(z, y)$;
5. $A \in \overline{D}(x, x)$ if and only if $A \neq \varnothing$;
6. $A \in \overline{D}(x, y)$ if and only if for all $B \supseteq A$, we have $B \in \widehat{D}(x, y)$.

Proof. Suppose $\langle \widehat{D}, \overline{D}\rangle$ is represented by a CCR f satisfying LD, IIA, and TR (resp. FR). The proof that $\widehat{D}$ satisfies the properties was already given in the proof of Theorem 3.3. The proof that $\overline{D}$ satisfies property 5 is the same as in the proof of Theorem 3.6. The proof that $\langle \widehat{D}, \overline{D}\rangle$ satisfies property 6 is essentially the same as at the end of Section 2, only using LD instead of UD and SPA.

For the converse, we define the CCR f exactly as in the proof of Theorem 3.3. There we already showed that $\widehat{D} = \widehat{D}_f$, so it only remains to show $\overline{D} = \overline{D}_f$. Suppose $A \in \overline{D}(x, y)$. To show that $A \in \overline{D}_f(x, y)$, consider a profile $\mathbf{P}$ in which $A \subseteq \mathbf{P}(x, y)$. By the left-to-right direction of property 6, since $A \in \overline{D}(x, y)$ and $A \subseteq \mathbf{P}(x, y)$, we have $\mathbf{P}(x, y) \in \widehat{D}(x, y)$, so $x f(\mathbf{P}) y$. Hence $A \in \overline{D}_f(\mathbf{P})$. Conversely, suppose $A \in \overline{D}_f(\mathbf{P})$. Let $B \supseteq A$ and consider any profile $\mathbf{P}$ in which $B = \mathbf{P}(x, y)$ and $B^c = \mathbf{P}(y, x)$. Then since $A \in \overline{D}_f(\mathbf{P})$ and $A \subseteq \mathbf{P}(x, y)$, we have $x f(\mathbf{P}) y$. This shows that $B \in \widehat{D}(x, y)$. Since B was arbitrary, it follows that $B \in \widehat{D}(x, y)$ for all $B \supseteq A$. Hence $A \in \overline{D}(x, y)$ by the right-to-left direction of property 6. □

We leave for future work the problem of obtaining a version of Theorem 7.3 with LD replaced by the conjunction of UD and SPA. It follows from the observations at the end of Section 2 that if $\langle \widehat{D}, \overline{D}\rangle$ is representable in the class of CCRs satisfying UD, SPA, and any other properties, then $\langle \widehat{D}, \overline{D}\rangle$ satisfies property 6 of Theorem 7.3. The open question is how to show that if a decisiveness regime satisfies properties 1–6 of Theorem 7.3, then it is representable in the class of CCRs satisfying UD, SPA, IIA, and TR (resp. FR).

7.2 Combined Logics

Our next step is to combine and extend the logics of Section 4 to logics for reasoning about decisive coalitions and almost decisive coalition in the same system.

We define a formal language as in Definition 4.1 but with formulas $\widehat{D}_{x>y}(t)$ and $\overline{D}_{x>y}(t)$ for every $t \in \mathsf{Term}$ and $x, y \in \mathsf{Alt}$, instead of $D_{x>y}(t)$. We then modify the semantics of Definition 4.4 in the obvious way with the following clauses:

- $f \models_{\alpha,\beta} \widehat{D}_{x>y}(t)$ if and only if $\dot{\alpha}(t) \in \widehat{D}_f(\beta(x), \beta(y))$;
- $f \models_{\alpha,\beta} \overline{D}_{x>y}(t)$ if and only if $\dot{\alpha}(t) \in \overline{D}_f(\beta(x), \beta(y))$.

Finally, we define the logic $\overline{\widehat{\mathbf{T}}}$ (resp. $\overline{\widehat{\mathbf{W}}}$) as follows:

- all axioms and rules from Definition 4.5;
- the axioms and rules of $\widehat{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$) from Definition 4.6, but for $\widehat{D}$ instead of D;
- $\overline{D}_{x>x}(a) \leftrightarrow \neg(a \equiv 0)$;
- $(\overline{D}_{x>y}(a) \wedge (a \sqsubseteq b)) \to \widehat{D}_{x>y}(b)$;
- $\overline{D}$-generalization rule: if $\varphi \to ((t \sqsubseteq b) \to \widehat{D}_{x>y}(b))$ is a theorem and $b \in \mathsf{Coal}$ does not occur in φ or t, then $\varphi \to \overline{D}_{x>y}(t)$ is a theorem.

The axioms $\overline{D}_{x>x}(a) \leftrightarrow \neg(a \equiv 0)$ and $(\overline{D}_{x>y}(a) \wedge (a \sqsubseteq b)) \to \widehat{D}_{x>y}(b)$ match property 5 and the left-to-right direction of property 6 of Theorem 7.3, respectively. The $\overline{D}$-generalization rule is designed to match the right-to-left direction of property 6 of Theorem 7.3. Roughly, the effect of the assumption that b does not occur in φ or t is that b denotes an *arbitrary* superset of the coalition that t denotes; then the fact that φ implies $\widehat{D}_{x>y}(b)$ means that φ implies that *every* superset of the coalition that t denotes is almost decisive; that in turn implies—assuming LD (or UD and SPA)—that the coalition that t denotes is decisive. For a more precise explanation, see the proof of Theorem 7.4.

Theorem 7.4. Let φ be any formula of the expanded language with $\widehat{D}$ and $\overline{D}$.

1. Soundness: if φ is a theorem of $\overline{\widehat{\mathbf{T}}}$ (resp. $\overline{\widehat{\mathbf{W}}}$), then for any nonempty set V, φ is true of all CCRs for $\langle X, V\rangle$ satisfying LD (or UD and SPA), IIA, and TR (resp. FR).
2. Completeness: if for any nonempty finite set V, φ is true of all CCRs for $\langle X, V\rangle$ satisfying LD, IIA, and TR (resp. FR), then φ is a theorem of $\overline{\widehat{\mathbf{T}}}$ (resp. $\overline{\widehat{\mathbf{W}}}$).

Proof. For part 1, the only somewhat nontrivial part is to prove that the $\overline{D}$-generalization rule preserves truth over all CCRs in the relevant class.

For the $\overline{D}$-generalization rule, let K be any class of CCRs satisfying LD (or UD and SPA). Assume ($\star$): every CCR in K makes $\varphi \to ((t \sqsubseteq b) \to \widehat{D}_{x>y}(b))$ true, where $b \in \mathsf{Coal}$ does not occur in φ or t. Then we claim that every CCR in K makes $\varphi \to \overline{D}_{x>y}(t)$ true as well. To show that $f \in K$ makes $\varphi \to \overline{D}_{x>y}(t)$ true, we must show that *for every coalition assignment α and alternative assignment β*, f makes $\varphi \to \overline{D}_{x>y}(t)$ true relative to α, β. Suppose α and β are arbitrary coalition and alternative assignments, respectively. To show that f makes $\varphi \to \overline{D}_{x>y}(t)$ true relative to α, β, we assume f makes φ true relative to α, β and prove that f makes $\overline{D}_{x>y}(t)$ true relative to α, β. To show that f makes $\overline{D}_{x>y}(t)$ true relative to α, β, we must show that $\dot{\alpha}(t) \in \overline{D}_f(\beta(x), \beta(y))$. As observed in the proof of Theorem 7.3, to show that $\dot{\alpha}(t) \in \overline{D}_f(\beta(x), \beta(y))$, it suffices to show that for every $B \supseteq \dot{\alpha}(t)$, we have $B \in \widehat{D}_f(\beta(x), \beta(y))$. Given any $B \supseteq \dot{\alpha}(t)$, let α_B be a coalition assignment that differs from α at most in that $\alpha_B(b) = B$. Since b does not occur in t, we have $\dot{\alpha}(t) = \dot{\alpha}_B(t)$. Since b does not occur in φ, and φ is true relative to α, β, it follows that φ is true relative to α_B, β. Then by ($\star$), it follows that $(t \sqsubseteq b) \to \widehat{D}_{x>y}(b)$ is true relative to α_B, β. By the truth definition, this implies that if $\dot{\alpha}_B(t) \subseteq \dot{\alpha}_B(b)$, then $\dot{\alpha}_B(b) \in \widehat{D}(\beta(x), \beta(y))$. Then given the facts that $B \supseteq \dot{\alpha}(t) = \dot{\alpha}_B(t)$ and $\alpha_B(b) = B$, we have $\dot{\alpha}_B(t) \subseteq \dot{\alpha}_B(b)$ and hence $B = \dot{\alpha}_B(b) \in \widehat{D}(\beta(x), \beta(y))$, which completes the argument.

For part 2, we sketch the proof in the Appendix. □

As in Section 4, here too we have a decidability result showing that reasoning about decisive and almost decisive coalitions together can be carried out algorithmically.

Theorem 7.5. The set of theorems of $\overline{\widehat{\mathbf{T}}}$ (resp. $\overline{\widehat{\mathbf{W}}}$) is decidable.

Proof. The proof is exactly the same as that of Theorem 4.9. □

7.3 Collapse Lemma

Perhaps the most important fact about the relationship between almost decisiveness and decisiveness is that under the Pareto assumption, the two notions collapse.

Lemma 7.6. Assume $x \neq y$.

$$\vdash_{\overline{\widehat{\mathbf{T}}}} \widehat{D}(1) \to (\widehat{D}_{x>y}(a) \leftrightarrow \overline{D}_{x>y}(a)).$$

Proof. Given the easy theorem $\overline{D}_{x>y}(a) \to \widehat{D}_{x>y}(a)$, we have $\vdash \widehat{D}(1) \to (\overline{D}_{x>y}(a) \to \widehat{D}_{x>y}(a))$ by propositional logic. For the other direction, pick $b \in \mathsf{Coal}$ with $b \neq a$, $b \neq 1$.

(1) $(\widehat{D}_{y>z}(1) \wedge \widehat{D}_{z>y}(1)) \to ((\widehat{D}_{x>y}(a) \wedge (a \sqsubseteq b)) \to \widehat{D}_{x>y}(b))$ by Lemma 5.10, substituting $\widehat{D}$ for D

(2) $(\widehat{D}(1) \wedge \widehat{D}_{x>y}(a)) \to ((a \sqsubseteq b) \to \widehat{D}_{x>y}(b))$ from (1) by propositional logic

(3) $(\widehat{D}(1) \wedge \widehat{D}_{x>y}(a)) \to \overline{D}_{x>y}(a)$ from (2) by the $\overline{D}$-generalization rule, since b does not occur in $\widehat{D}(1) \wedge \widehat{D}_{x>y}(a)$ or a

(4) $\widehat{D}(1) \to (\widehat{D}_{x>y}(a) \to \overline{D}_{x>y}(a))$ from (3) by propositional logic. □

7.4 Inverse Decisiveness

The above method for combining reasoning about almost decisive coalitions and decisive coalitions in a single system can be used to add reasoning about inversely decisive coalition as well. The first step, as before, is to extend the representation theory.

Definition 7.7. Let f be a CCR for $\langle X, V\rangle$. We define functions $\widehat{I}_f : X^2 \to \wp(\wp(V))$ and $\overline{I}_f : X^2 \to \wp(\wp(V))$ as follows:

1. $\widehat{I}_f(x, y)$ is the set of all $A \subseteq V$ that are almost inversely decisive for x over y according to f;
2. $\overline{I}_f(x, y)$ is the set of all $A \subseteq V$ that are inversely decisive for x over y according to f.

Definition 7.8. A *dual decisiveness regime* for $\langle X, V\rangle$ is a quadruple $\langle \widehat{D}, \overline{D}, \widehat{I}, \overline{I}\rangle$ where each component is a function from X^2 to $\wp(\wp(V))$.

Definition 7.9. Let K be a class of CCRs for $\langle X, V\rangle$. A dual decisiveness regime $\langle \widehat{D}, \overline{D}, \widehat{I}, \overline{I}\rangle$ for $\langle X, V\rangle$ is *representable in* K if and only if there is an $f \in K$ such that $\widehat{D} = \widehat{D}_f$, $\overline{D} = \overline{D}_f$, $\widehat{I} = \widehat{I}_f$, and $\overline{I} = \overline{I}_f$.

All of the ideas required for the proof of the following theorem may be found in the proof of Theorem 7.3.

Theorem 7.10. Let X and V be nonempty sets with $|X| \geq 3$. A dual decisiveness regime $\langle \widehat{D}, \overline{D}, \widehat{I}, \overline{I}\rangle$ for $\langle X, V\rangle$ is representable in the class of CCRs for $\langle X, V\rangle$ satisfying LD, IIA, and TR (resp. FR) if and only if the following conditions hold:

1. $\langle \widehat{D}, \overline{D}\rangle$ satisfies the properties in Theorem 7.3 (for TR, resp. FR);
2. for all $x, y \in X$ and $A \subseteq V$: $A \in \widehat{I}(x, y)$ if and only if $A^c \in \widehat{D}(y, x)$;
3. for all distinct $x, y \in X$ and $A \subseteq V$: $A \in \overline{I}(x, y)$ if and only if for all $B \supseteq A$, we have $B \in \widehat{I}(x, y)$.

A useful consequence of properties 2 and 3 is the following:

4. for all distinct $x, y \in X$ and $A \subseteq V$: if for all $B \subseteq A^c$, $B \in \widehat{D}(y, x)$, then $A \in \overline{I}(x, y)$.

By now we assume it is clear how to extend the language, semantics, and logics to handle inversely decisive coalitions. A key rule of the extended logic is the following, matching property 4 above:

> $\widehat{D}$-to-$\overline{I}$-generalization: if $\varphi \to ((b \sqsubseteq -t) \to \widehat{D}_{y>x}(b))$ is a theorem and $b \in \mathsf{Coal}$ does not occur in φ or t, then $\varphi \to \overline{I}_{x>y}(t)$ is a theorem.

Using this rule, we can prove the collapse lemma for almost inverse decisiveness and inverse decisiveness.

Lemma 7.11. Assume $x \neq y$.

$$\vdash \widehat{I}(1) \to (\widehat{I}_{x>y}(a) \leftrightarrow \overline{I}_{x>y}(a)).$$

Proof. Given the easy theorem $\overline{I}_{x>y}(a) \to \widehat{I}_{x>y}(a)$, we have $\vdash \widehat{I}(1) \to (\overline{I}_{x>y}(a) \to \widehat{I}_{x>y}(a))$ by propositional logic. For the other direction, pick $b \in \mathsf{Coal}$ such that $b \neq a$ and $b \neq 0$.

(1) $(\widehat{D}_{z>y}(0) \wedge \widehat{D}_{y>z}(0)) \to ((\widehat{D}_{y>x}(-a) \wedge b \sqsubseteq -a) \to \widehat{D}_{y>x}(b))$ by Lemma 6.6, substituting $\widehat{D}$ for D

(2) $(\widehat{D}(0) \wedge \widehat{D}_{y>x}(-a)) \to ((b \sqsubseteq -a) \to \widehat{D}_{y>x}(b))$ from (1) by propositional logic

(3) $(\widehat{D}(0) \wedge \widehat{D}_{y>x}(-a)) \to \overline{I}_{x>y}(a)$ from (2) by the $\widehat{D}$-to-$\overline{I}$-generalization rule, since b does not occur in $\widehat{D}(0) \wedge \widehat{D}_{y>x}(-a)$ or $-a$

(4) $\widehat{D}(0) \to (\widehat{D}_{y>x}(-a) \to \overline{I}_{x>y}(a))$ from (3) by propositional logic

(5) $\widehat{I}(1) \to (\widehat{I}_{x>y}(a) \to \overline{I}_{x>y}(a))$ from (4) using $\widehat{I}_{x>y}(a) \leftrightarrow \widehat{D}_{y>x}(-a)$. □

In section 6.2, we proved a version of Wilson's theorem for almost decisive and almost inversely decisive sets. Using the collapse lemma above, we can now prove a version of Wilson's theorem for decisive and inversely decisive sets.[23]

Theorem 7.12 (Wilson's Theorem)**.**

1. $\vdash (WEA \wedge NE) \to (\overline{D}(1) \vee \overline{I}(1))$
2. $\vdash \overline{D}(1) \to (\overline{D}(a) \vee \overline{D}(-a))$
3. $\vdash \overline{D}(1) \to \big(\overline{D}(a \sqcap b) \leftrightarrow (\overline{D}(a) \wedge \overline{D}(b))\big)$
4. $\vdash \overline{I}(1) \to (\overline{I}(a) \vee \overline{I}(-a))$
5. $\vdash \overline{I}(1) \to \big(\overline{I}(a \sqcap b) \leftrightarrow (\overline{I}(a) \wedge \overline{I}(b))\big)$.

Proof. Immediate from Theorem 6.9 and Lemmas 7.6 and 7.11, using the easy theorems $\overline{D}_{x>y}(a) \to \widehat{D}_{x>y}(a)$ and $\overline{I}_{x>y}(a) \to \widehat{I}_{x>y}(a)$. □

8 From Rationality Postulates to Decisiveness Axioms

Having given ample illustration of our methodology of representation and derivation involving concepts of decisive coalitions, let us now take a "meta"-perspective on this methodology. How do we find the right axioms for reasoning about decisiveness?

It turns out there is an *algorithm* for doing so, at least given certain rationality postulates for social preference as inputs. For simplicity, we will focus on almost decisive coalitions, though the arguments can be adapted for decisive coalitions as well.

For example, suppose we are given a social rationality postulate R in the language of preference of one of the following forms, where for all $i \leq n$, $x_i \neq y_i$:[24]

(i) if for all $i \leq n$, $x_i P y_i$, then $x_{n+1} P y_{n+1}$;

(ii) if for all $i \leq n$, $x_i P y_i$, then $x_{n+1} P y_{n+1}$ or $x_{n+2} P y_{n+2}$.

Given R, we construct a list Σ_{R} of Boolean equations of the form

$$\pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2} = \varnothing,$$

where each $\pm_i$ denotes either the complement function or the identity function. We add such an equation to Σ_{R} if and only if for every profile $\mathbf{P} \in L(X)^V$ such that for all $i \leq n+2$, $A_i = \mathbf{P}(x_i, y_i)$, we have $\pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2} = \varnothing$. Clearly it can be checked in a finite time whether an equation belongs to Σ_{R}. Having thereby constructed our list Σ_{R} of equations, we translate the postulate R into a principle $\mathrm{R}^\star$ in the language of (almost) decisive sets as follows, depending on which case we were in above:

(i′) if for all $i \leq n$, $A_i \in D(x_i, y_i)$, and the equations in Σ_{R} hold, then $A_{n+1} \in D(x_{n+1}, y_{n+1})$;

(ii′) if for all $i \leq n$, $A_i \in D(x_i, y_i)$, and the equations in Σ_{R} hold, then $A_{n+1} \in D(x_{n+1}, y_{n+1})$ or $A_{n+2} \in D(x_{n+2}, y_{n+2})$.

[23] In Theorem 7.12, by WEA and NE, we do not mean exactly the formulas defined in Section 5.1, but rather the result of replacing D by $\widehat{D}$ (or $\overline{D}$) in those formulas.

[24] To make the following argument fully rigorous, we need to define the formal syntax of the language in which these postulates are given, as well as the formal syntax of the language into which we are translating them below. The former can be the language of predicate logic with a single binary predicate symbols P, and the latter can be the language of (almost) decisiveness in Section 4.

Given this translation, we will prove a generic representation theorem that implies our earlier Theorem 3.5, as well as others. Given rationality postulates $\mathrm{R}_1, \ldots, \mathrm{R}_m$ as above, let N be the highest number of alternative variables in any of the R_i's.

Theorem 8.1. Let X and V be nonempty sets with $|X| \geq N$. A function $D : X^2 \to \wp(\wp(V))$ is almost-decisively representable in the class of CCRs for $\langle X, V\rangle$ satisfying UD, IIA, and the rationality postulates $\mathrm{R}_1, \ldots, \mathrm{R}_m$ if and only if D satisfies properties 1 and 2 of Theorem 3.3 and $\mathrm{R}_1^\star, \ldots, \mathrm{R}_m^\star$.

Proof. Suppose D is represented by a CCR f satisfying UD, IIA, and rationality postulates $\mathrm{R}_1, \ldots, \mathrm{R}_m$. We must show that D satisfies $\mathrm{R}_1^\star, \ldots, \mathrm{R}_m^\star$. We will give the proof for a rationality postulate R of type (ii) above. The proof for type (i) is even easier.

Suppose $A_{n+1} \notin D(x_{n+1}, y_{n+1})$, so there is a profile $\mathbf{P}$ in which $A_{n+1} = \mathbf{P}(x_{n+1}, y_{n+1})$, $A_{n+1}^c = \mathbf{P}(y_{n+1}, x_{n+1})$, and *not* $x_{n+1} f(\mathbf{P}) y_{n+1}$. Further suppose that the equations in Σ_{R} hold for $A_1, \ldots, A_{n+2}$. First, we claim that there is a profile $\mathbf{P}' \in L(X)^V$ such that:

$$\text{for all } i \leq n+2,\ A_i = \mathbf{P}'(x_i, y_i) \text{ and hence } A_i^c = \mathbf{P}'(y_i, x_i).$$

The proof that there is such a profile follows the same partitioning strategy as in the proof of Theorem 3.3. If an equation $\pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2} = \varnothing$ belongs to Σ_{R}, then we do not need to specify a strict linear order for voters in $\pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2}$, because there are no such voters, by the assumption that the equations in Σ_{R} hold. On the other hand, if an equation $\pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2} = \varnothing$ does *not* belong to Σ_{R}, then by the definition of Σ_{R} there exists a profile $\mathbf{P}^* \in L(X)^V$ such that $\pm_1 \mathbf{P}^*(x_1, y_1) \cap \cdots \cap \pm_{n+2} \mathbf{P}^*(x_{n+2}, y_{n+2}) \neq \varnothing$. Let P^* be the strict linear order belonging to one of the voters in $\pm_1 \mathbf{P}^*(x_1, y_1) \cap \cdots \cap \pm_{n+2} \mathbf{P}^*(x_{n+2}, y_{n+2})$. Then for each $j \in \pm_1 A_1 \cap \cdots \cap \pm_{n+2} A_{n+2}$, define P_j' to be P^*. In this way, we assign to each voter a strict linear order on X, depending on the cell of the partition to which the voter belongs, that results in a profile $\mathbf{P}' \in L(X)^V$ satisfying the bulleted requirement above. Then since $\mathbf{P}_{|\{x_{n+1}, y_{n+1}\}} = \mathbf{P}'_{|\{x_{n+1}, y_{n+1}\}}$ by construction, from *not* $x_{n+1} f(\mathbf{P}) y_{n+1}$ we have *not* $x_{n+1} f(\mathbf{P}') y_{n+1}$ by IIA, and for all $i \leq n$, from $A_i \in D(x_i, y_i)$ we have $x_i f(\mathbf{P}') y_i$. Therefore, $x_{n+2} f(\mathbf{P}') y_{n+2}$ by the rationality postulate R. Now we claim that $A_{n+2} \in D(x_{n+2}, y_{n+2})$. Consider any profile $\mathbf{P}''$ in which $A_{n+2} = \mathbf{P}''(x_{n+1}, y_{n+2})$ and $A_{n+2}^c = \mathbf{P}''(y_{n+2}, x_{n+2})$. Then $\mathbf{P}'_{|\{x_{n+2}, y_{n+2}\}} = \mathbf{P}''_{|\{x_{n+2}, y_{n+2}\}}$, so by IIA, $x_{n+2} f(\mathbf{P}') y_{n+2}$ implies $x_{n+2} f(\mathbf{P}'') y_{n+2}$. Thus, $A_{n+2} \in D(x_{n+2}, y_{n+2})$. This completes the proof of $\mathrm{R}^\star$.

Conversely, by Lemma 3.4, it suffices to establish representability in the class of CCRs satisfying LD, IIA, and $\mathrm{R}_1, \ldots, \mathrm{R}_m$. For $\mathbf{P} \in L(X)^V$, we define f exactly as in the proof of Theorem 3.3: $x f(\mathbf{P}) y$ if and only if $x \neq y$ and $\mathbf{P}(x, y) \in D(x, y)$. We claim that $f(\mathbf{P})$ satisfies $\mathrm{R}_1, \ldots, \mathrm{R}_m$. Again consider an R of type (ii) above, as the proof for type (i) is similar. We must show that if for all $i \leq n$, $x_i f(\mathbf{P}) y_i$, then $x_{n+1} f(\mathbf{P}) y_{n+1}$ or $x_{n+2} f(\mathbf{P}) y_{n+2}$. From $x_i f(\mathbf{P}) y_i$, we have $\mathbf{P}(x_i, y_i) \in D(x_i, y_i)$. In addition, by the construction of Σ_{R}, all of the equations in Σ_{R} hold when we set $A_i = \mathbf{P}(x_i, y_i)$. Thus, by $\mathrm{R}^\star$ we have $\mathbf{P}(x_{n+1}, y_{n+1}) \in D(x_{n+1}, y_{n+1})$ or $\mathbf{P}(x_{n+2}, y_{n+2}) \in D(x_{n+2}, y_{n+2})$, which implies (given $x_i \neq y_i$) that $x_{n+1} f(\mathbf{P}) y_{n+1}$ or $x_{n+2} f(\mathbf{P}) y_{n+2}$. $\square$

Let us now show how Theorem 8.1 can be applied to concrete rationality postulates other than those we have seen so far. For example, Blair and Pollak [1979] and Blau

[1979] strengthened Arrow's theorem by showing that the assumption that the social strict preference relation P is negatively transitive can be replaced by the weaker assumption that P is a *semi-order*, which means that it satisfies both of the following conditions:

semi-transitivity: if xPy and yPz, then xPw or wPz.

interval order property: if xPy and zPw, then xPw or zPy.

Observe that semi-transitivity together with asymmetry implies transitivity, but not negative transitivity. From semi-transitivity, we have that if xPy and yPz, then xPz or zPz, which with asymmetry implies transitivity. Sen [1979, p. 543, fn. 1] observed that just replacing negative transitivity with semi-transitivity is enough to derive dictatorship from Arrow's other conditions—the interval order property is not needed.

The strengthening of Arrow's theorem with semi-transitivity can be analyzed using Theorem 8.1. First, let us give a name to the weaker rationality condition:

semi-transitive rationality (SR): for any $\mathbf{P} \in \mathrm{dom}(f)$, $f(\mathbf{P})$ is semi-transitive.

By calculating the set Σ_{R} of Boolean equations enforced by the semi-transitivity postulate, and then re-writing them as equivalent subset statements, we obtain the following representation theorem from Theorem 8.1.

Theorem 8.2. Let X and V be nonempty sets with $|X| \geq 4$. A function $D : X^2 \to \wp(\wp(V))$ is almost-decisively representable in the class of CCRs for $\langle X, V\rangle$ satisfying UD, IIA, and SR if and only if D satisfies properties 1 and 2 of Theorem 3.3, and for all $A, B, C \subseteq V$ and distinct $x, y, z, w \in X$:

- if $A \in D(x, y)$, $B \in D(y, z)$, $A \cap B \subseteq C_1 \cup C_2$, and $C_1 \cap C_2 \subseteq A \cup B$, then $C_1 \in D(x, w)$ or $C_2 \in D(w, z)$.

Let $\widehat{\mathbf{S}}$ (resp. $\overline{\mathbf{S}}$) be the decisiveness logic that replaces the negative transitivity axiom of $\widehat{\mathbf{W}}$ (resp. $\overline{\mathbf{W}}$) with the semi-transitivity axiom formalizing the condition in Theorem 8.2 (see (2) in the proof of Lemma 8.3 below). Since the negative transitivity axiom was only needed in the proof that the filter of (almost) decisive sets is complete (Lemma 5.14), we need only show how to modify that proof to use the semi-transitivity axiom instead.

Lemma 8.3. $\vdash_{\mathbf{S}} EA \to (D(a) \vee D(-a))$.

Proof. We use the same strategy as in the proof of Lemma 5.14, only now we use the semi-transitivity axiom instead of the negative transitivity axiom and EA instead of $WEA \wedge NE$:

(1) $EA \to (D_{u>u'}(e) \wedge D_{u'>v}(e))$ by Lemma 5.2

(2) $\big(D_{u>u'}(e) \wedge D_{u'>v}(e) \wedge (e \sqcap e \sqsubseteq a \sqcup -a) \wedge (a \sqcap -a \sqsubseteq e \sqcup e)\big) \to \big(D_{u>w}(a) \vee D_{w>v}(-a)\big)$, instance of semi-transitivity axiom

(3) $(e \sqcap e \sqsubseteq a \sqcup -a) \wedge (a \sqcap -a \sqsubseteq e \sqcup e)$, valid Boolean inequalities

(4) $(D_{u>u'}(e) \wedge D_{u'>v}(e)) \to (D_{u>w}(a) \vee D_{w>v}(-a))$ from (2) and (3) by propositional logic

(5) $(D_{u>u'}(e) \wedge D_{u'>v}(e)) \to (\neg D_{u>w}(a) \to D_{w>v}(-a))$ from (4) by propositional logic

(6) $EA \to (D_{w>v}(-a) \to D(-a))$ by Lemma 5.3

(7) $EA \to (\neg D_{u>w}(a) \to D(-a))$ from (1), (5), and (6) by propositional logic. □

9 Conclusion

The theme of this paper has been that Arrow's concept of a decisive coalition gives rise to a social choice theoretic language and logic all of its own. We have investigated key aspects of this language and logic, but there remains a variety of questions for further study. An immediate question is whether our approach can be extended to a setting without the axiom of IIA. Indeed, in work in progress we analyze, via a new representation theorem and matching logic, an impossibility result that does not depend on IIA: Sen's theorem on the impossibility of the Paretian liberal [Sen, 1970]. Below we will mention more open-ended questions that are inspired by logic but also seem of interest for social choice theory. We will give three examples of avenues for investigation within the framework of the current paper and then one example of an avenue for extending the framework.

First, we can ask about the complete logic for reasoning about decisive coalitions with respect to some specific CCR, including CCRs that do not satisfy the axioms of IIA and TR that have been our focus. For example, dropping TR, we may consider the Condorcet CCR $f : O(X)^V \to P(X)$ that sets $xf(\mathbf{P})y$ if and only if a majority of voters prefer x to y in the profile $\mathbf{P}$. Key axioms of the Condorcet logic are the Pareto principle $D(1)$, the axiom $(D_{x>y}(a) \wedge (a \sqsubseteq b)) \to D_{x>y}(b)$ for closure under supersets, and (for an odd number of voters) the axiom $D_{x>y}(a) \vee D_{x>y}(-a)$. The first two axioms, but not the third,[25] are also valid for the Borda count CCR, which violates IIA. This also suggests the possibility of starting with a set of axioms about decisive coalitions, instead of social choice theoretic axioms of a more traditional form as in Section 2, and then investigating all of those CCRs that validate the desired decisiveness axioms (and none stronger).

Second, logicians often study the structure of whole classes of related logics (see, e.g., Chagrov and Zakharyaschev 1997). In that spirit, there are questions about the structure of the whole class of decisiveness logics. For example, let *Arrow's logic* be the smallest decisiveness logic that contains the ultrafilter axioms:

$$D(1) \to \neg D(0) \qquad D(1) \to (D(a) \vee D(-a)) \qquad D(1) \to (D(a \sqcap b) \leftrightarrow (D(a) \wedge D(b))).$$

We showed that the logic $\widehat{\mathbf{W}}$ (resp. $\overline{\mathbf{W}}$) is as strong as Arrow's logic, as it derives the axioms of Arrow's logic as theorems. In fact, the logic $\widehat{\mathbf{W}}$ (resp. $\overline{\mathbf{W}}$) is *stronger*: it has theorems that Arrow's logic does not, namely the negative transitivity axiom. The reason we know that the negative transitivity axiom is not a theorem of Arrow's logic is that the semi-transitivity logic $\widehat{\mathbf{S}}$ (resp. $\overline{\mathbf{S}}$) also derives the axioms of Arrow's logic, but as semi-transitivity does not imply negative transitivity (relative to UD and IIA), the logic $\widehat{\mathbf{S}}$ (resp. $\overline{\mathbf{S}}$) is weaker than $\widehat{\mathbf{W}}$ (resp. $\overline{\mathbf{W}}$). Let us say that a decisiveness logic $\mathbf{L}$ is *minimally super-Arrovian* if and only if $\mathbf{L}$ is stronger than Arrow's logic and there is no decisiveness logic intermediate in strength between Arrow's logic and $\mathbf{L}$. What is the cardinality of the set of minimally super-Arrovian logics? In intuitive terms, how many different minimal sets of assumptions about decisive coalitions are there for proving Arrow's theorem?

Third, a central question about any logical framework concerns the *expressivity* of the

[25] Thanks to Mikayla Kelley for this way of distinguishing the logic of decisive coalitions for Condorcet and Borda.

chosen formal language, given its formal semantics. For example, the language of (almost) decisive coalitions can express certain assumptions, such as the Pareto principle with the formula $D(1)$, but it cannot express other assumptions, such as the very weak *non-null* assumption that for some distinct $x, y \in X$ and some profile $\mathbf{P}$, we have $xf(\mathbf{P})y$.[26] Which classes of CCRs are definable by a formula in the language of decisive coalitions? In other areas of logic, there are characterizations of the definable classes of structures in terms of operations on structures that preserve truth in the formal language. We might hope for a similar characterization of definable classes of CCRs that illuminates the expressive power and limits of the language of decisive coalitions. Of course a language with quantification over profiles, alternatives, and voters is more expressive than our language, but such increased expressivity may come with a price of unwieldiness. The attractive decidability property of our logics will be lost with too much increase in expressivity.

A fourth avenue for investigation concerns extending our framework to reasoning about decisive coalitions in the context of *functional* collective choice rules (FCCRs), as in Sen 1969, 1993 (also see Blair et al. 1976, Suzumura 1983), which assign to each profile $\mathbf{P} \in O(X)^V$ a choice function $C = f(\mathbf{P}) : \wp(X) \to \wp(X)$ such that $C(S) \subseteq S$. Here we can allow choice functions that are not rationalizable by a binary preference relation, in the sense of Arrow 1959, so we can drop even the transitivity axiom that leads, together with IIA, to the Oligarchy Theorem of Section 6.1. For a menu $S \subseteq X$ of feasible alternatives, $x, y \in S$, and $A \subseteq V$, we may follow Sen [1993] and say that A is (*rejection*) *decisive for x over y relative to S* if and only if for all $\mathbf{P} \in \mathrm{dom}(f)$, if $A \subseteq \mathbf{P}(x, y)$, then $y \notin f(\mathbf{P})(S)$. To reason about this kind of decisiveness, we can replace our formulas $\overline{D}_{x>y}(a)$ by formulas $\overline{D}^S_{x>y}(a)$, where S is now a menu term (which may be built up from menu labels in the same way that coalition terms are built up from coalition labels), and add formulas $S \sqsubseteq T$ where S and T are menu terms. For example, in this language we may write axioms such as $(\overline{D}^S_{x>y}(a) \wedge S \sqsubseteq T) \to \overline{D}^T_{x>y}(a)$, expressing that a coalition that is decisive for x over y relative to one menu is also decisive for x over y relative to any larger menu include the first, which is valid for some FCCRs (e.g., the Condorcet method, allowing $C(S) = \varnothing$ when there is no Condorcet winner) but not for others (e.g., local Borda count as in Kelly 1988, p. 74). The same questions about obtaining representation theorems and complete logics that we have answered for our language make perfect sense for this extended language as well.

Our aim in this paper has been to stay close to the style of reasoning done by social choice theorists, while formalizing it in the style of symbolic logic. We would like to think that such a project would be congenial to Arrow, given the role that logic played in his creation of the mathematical theory of social choice.

[26] This is because there can be CCRs f and f' such that: f and f' have the same (almost) decisive coalitions, i.e., for all $x, y \in X$, $\widehat{D}_f(x, y) = \widehat{D}_{f'}(x, y)$ and $\overline{D}_f(x, y) = \overline{D}_{f'}(x, y)$; f does not satisfy non-null; and f' does satisfy non-null. For example, let f be the CCR such that for all profiles $\mathbf{P}$ and $x, y \in X$, *not* $xf(\mathbf{P})y$. Given $x, y \in X$, we will define a CCR $f_{x,y}$ that differs from f only on the profile $\mathbf{P}^\star$ such that for all $x, y \in X$ and $i \in V$, *not* xP_iy; for this $\mathbf{P}^\star$, let $zf_{x,y}(\mathbf{P}^\star)w$ if and only if either $z = x$ and $w \neq x$, or $z \neq y$ and $w = y$, so $f_{x,y}(\mathbf{P}^\star)$ is a strict weak order. Then for all $z, w \in X$, if $z \neq w$, then $\widehat{D}_f(z, w) = \widehat{D}_{f_{x,y}}(z, w) = \overline{D}_f(z, w) = \overline{D}_{f_{x,y}}(z, w) = \varnothing$, while if $z = w$, then $\widehat{D}_f(z, w) = \widehat{D}_{f_{x,y}}(z, w) = \wp(V)$ and $\overline{D}_f(z, w) = \overline{D}_{f_{x,y}}(z, w) = \wp(V) \setminus \{\varnothing\}$; f does not satisfy non-null; and $f_{x,y}$ does satisfy non-null.

A Appendix

In this Appendix, we sketch the proofs of the completeness theorems stated in Theorems 4.8.2, 4.8.4, and 7.4.2. A similar approach can be used to prove completeness theorems for logics that include inverse decisiveness as in Section 7.4.

For a logic $\mathbf{L}$, a set Σ of formulas is $\mathbf{L}$-*inconsistent* if and only if there are $\sigma_1, \dots, \sigma_n \in \Sigma$ such that $\vdash_{\mathbf{L}} (\sigma_1 \wedge \dots \wedge \sigma_n) \to \bot$, where $\bot$ is any formula of the form $\varphi \wedge \neg\varphi$. Otherwise Σ is $\mathbf{L}$-*consistent*. We say that a formula φ is $\mathbf{L}$-consistent if and only if $\{\varphi\}$ is $\mathbf{L}$-consistent. Finally, a set Σ of formulas is *maximally* $\mathbf{L}$-*consistent* if and only if for every set Δ of formulas, $\Sigma \subsetneq \Delta$ implies that Δ is $\mathbf{L}$-inconsistent.

To say that a logic $\mathbf{L}$ is complete with respect to a class K of CCRs is to say that for any formula ψ, if ψ is true of all CCRs from K, then ψ is a theorem of $\mathbf{L}$. Equivalently, if ψ is *not* a theorem of $\mathbf{L}$, then there is a CCR from K of which ψ is *not* true. Since by propositional logic, we have $\nvdash_{\mathbf{L}} \psi$ if and only if $\nvdash_{\mathbf{L}} \neg\psi \to \bot$, the formula ψ not being a theorem of $\mathbf{L}$ is equivalent to $\neg\psi$ being $\mathbf{L}$-consistent. Moreover, there being a CCR from K of which ψ is not true is equivalent to there being a CCR from K of which $\neg\psi$ is true. Thus, to prove that $\mathbf{L}$ is complete with respect to K, it suffices to show that for any formula φ, if φ is $\mathbf{L}$-consistent, then there is a CCR from K of which φ is true.

We will give a fairly detailed proof of Theorem 7.4.2, where $\mathbf{L}$ is $\overline{\widehat{\mathbf{T}}}$ (resp. $\overline{\widehat{\mathbf{W}}}$) and K is the class of CCRs satisfying LD, IIA, and TR (resp. FR). The proof where $\mathbf{L}$ is $\widehat{\mathbf{T}}$ or $\overline{\mathbf{T}}$ (resp. $\widehat{\mathbf{W}}$ or $\overline{\mathbf{W}}$) and K is the class of CCRs satisfying UD, IIA, and TR (resp. FR) is quite a bit simpler, as we will explain at the end.

In what follows by 'consistent' we mean '$\overline{\widehat{\mathbf{T}}}$-consistent' or '$\overline{\widehat{\mathbf{W}}}$-consistent', and by '$\vdash$' we mean '$\vdash_{\overline{\widehat{\mathbf{T}}}}$' or '$\vdash_{\overline{\widehat{\mathbf{W}}}}$', depending on which case the reader has in mind. The proof is the same except at one point that we will flag. Indeed, the proof would be essentially the same for still other logics based on other rationality postulates (recall Section 8).

Suppose φ is consistent. Let T_1 be the set of all terms generated by the coalition labels in $\mathsf{Coal}(\varphi) = \{a \in \mathsf{Coal} \mid a \text{ appears in } \varphi\}$ using $-$, $\sqcap$, and $\sqcup$ as in Definition 4.1. Define an equivalence relation E on T_1 by: sEt if and only if $s \equiv t$ is a valid equation of Boolean algebra. Let T_1/E be the quotient of T_1 by E. It follows that T_1/E is a Boolean algebra in which the complement of $[t]$ is $[-t]$, and the meet of $[s]$ and $[t]$ is $[s \sqcap t]$. Since T_1 is the set of terms generated by $\mathsf{Coal}(\varphi)$, it follows that the Boolean algebra T_1/E is generated by the set of equivalence classes of elements of $\mathsf{Coal}(\varphi)$. Since the set of such equivalence classes is finite, T_1/E is finite by the well-known fact that any finitely generated Boolean algebra is finite (see, e.g., Givant and Halmos 2009, p. 82). For each $x, y \in \mathsf{Alt}$ and $[t] \in T_1/E$, pick a coalition label $c_{x,y,[t]} \in \mathsf{Coal} \setminus \mathsf{Coal}(\varphi)$ such that if $\langle x, y, [t] \rangle \neq \langle x', y', [t'] \rangle$, then $c_{x,y,[t]} \neq c_{x',y',[t']}$. Since Alt and T_1/E are finite, only finitely many coalitions labels $c_{x,y,[t]}$ are needed. Crucially, $c_{x,y,[t]}$ does not appear in φ or t. Now define

$$\Theta = \{(t \sqsubseteq c_{x,y,[t]}) \wedge (\neg\overline{D}_{x>y}(t) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]})) \mid t \in T_1, x, y \in \mathsf{Alt}\}.$$

We claim that the set $\{\varphi\} \cup \Theta$ is consistent. If not, then there are $\theta_1, \dots, \theta_{n+1} \in \Theta$ $(n \geq 0)$ such that $\{\varphi, \theta_1, \dots, \theta_{n+1}\}$ is inconsistent, but no proper subset is inconsistent. Thus, $\vdash (\varphi \wedge \theta_1 \wedge \dots \wedge \theta_{n+1}) \to \bot$ and hence $\vdash (\varphi \wedge \theta_1 \wedge \dots \wedge \theta_n) \to \neg\theta_{n+1}$ by propositional logic.

The formula θ_{n+1} is $(t \sqsubseteq c_{x,y,[t]}) \wedge (\neg\overline{D}_{x>y}(t) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]}))$ for some $t \in T_1, x, y \in \mathsf{Alt}$. Thus, by propositional logic we have:

$$
\begin{aligned}
&\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to \neg((t \sqsubseteq c_{x,y,[t]}) \wedge (\neg\overline{D}_{x>y}(t) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]}))\\
\Rightarrow\ &\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to ((t \sqsubseteq c_{x,y,[t]}) \to \neg(\neg\overline{D}_{x>y}(t) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]}))\\
\Rightarrow\ &\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to ((t \sqsubseteq c_{x,y,[t]}) \to (\neg\overline{D}_{x>y}(t) \wedge \widehat{D}_{x>y}(c_{x,y,[t]})).
\end{aligned}
$$

Not only does $c_{x,y,[t]}$ not appear in φ or t, but also $c_{x,y,[t]}$ does not appear in $\theta_1, \ldots, \theta_n$. Toward a contradiction, suppose that $c_{x,y,[t]}$ does appear in some θ_i, so θ_i is the formula $(s \sqsubseteq c_{x,y,[t]}) \wedge (\neg\overline{D}_{x>y}(s) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]}))$ for some term $s \in [t]$. As an axiom of our logic, we have $s \equiv t \to (\theta_i[s/s] \leftrightarrow \theta_i[t/s])$, which is the same as $s \equiv t \to (\theta_i \leftrightarrow \theta_{n+1})$. Since $s \in [t]$, $s \equiv t$ is also an axiom of our logic, which with the previous step implies that $\theta_i \leftrightarrow \theta_{n+1}$ is a theorem of our logic. But then from the inconsistency of $\{\theta_1, \ldots, \theta_{n+1}\}$, its proper subset $\{\theta_1, \ldots, \theta_{n+1}\} \setminus \{\theta_i\}$ is also inconsistent, contradicting our assumption that no proper subset is inconsistent. Thus, we conclude that $c_{x,y,[t]}$ does not appear in $\theta_1, \ldots, \theta_n$.

On the one hand, since $\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to ((t \sqsubseteq c_{x,y,[t]}) \to \widehat{D}_{x>y}(c_{x,y,[t]}))$, and $c_{x,y,[t]}$ does not appear in $\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n$ or t, it follows by the $\overline{D}$-generalization rule (Section 7.2) that

$$\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to \overline{D}_{x>y}(t). \tag{1}$$

On the other hand, since

$$
\begin{aligned}
&\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to ((t \sqsubseteq c_{x,y,[t]}) \to \neg\overline{D}_{x>y}(t))\\
\Rightarrow\ &\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n \wedge \overline{D}_{x>y}(t)) \to \neg(t \sqsubseteq c_{x,y,[t]}),
\end{aligned}
\tag{2}
$$

and $c_{x,y,[t]}$ does not appear in $\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n \wedge \overline{D}_{x>y}(t)$, we can apply the substitution rule to replace $c_{x,y,[t]}$ in (2) by 1 to obtain $\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n \wedge \overline{D}_{x>y}(t)) \to \neg(t \sqsubseteq 1)$. But then since $t \sqsubseteq 1$ is a theorem of our logic, by propositional logic we obtain $\vdash \neg(\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n \wedge \overline{D}_{x>y}(t))$ and then $\vdash (\varphi \wedge \theta_1 \wedge \cdots \wedge \theta_n) \to \neg\overline{D}_{x>y}(t)$. It follows, given (1), that $\{\varphi, \theta_1, \ldots, \theta_n\}$ is inconsistent, contradicting the fact that no proper subset of $\{\varphi, \theta_1, \ldots, \theta_{n+1}\}$ is inconsistent. Thus, we conclude that $\{\varphi\} \cup \Theta$ is consistent.

We extend the consistent set $\{\varphi\} \cup \Theta$ to a maximally consistent set Γ using a standard construction. Fixing an enumeration $\varphi_0, \varphi_1, \ldots$ of the formulas of our language, we define:

$$\Gamma_0 = \{\varphi\} \cup \Theta \qquad \Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_n\} \text{ if this is consistent} \\ \Gamma_n \cup \{\neg\varphi_n\} \text{ otherwise} \end{cases} \qquad \Gamma = \bigcup_{n \in \mathbb{N}} \Gamma_n.$$

It is then a standard exercise to verify that Γ is a maximally consistent set.

For the next step of the proof, let T_2 be the set of all terms (including 0 and 1) generated by the coalition labels from the finite set $\mathsf{Coal}(\varphi) \cup \{c_{x,y,[t]} \mid x, y \in \mathsf{Alt}, [t] \in T_1/E\}$. Define an equivalence relation E_Γ on T_2 by: $sE_\Gamma t$ if and only if $s \equiv t \in \Gamma$. Then since Γ, as a maximally consistent set, contains all valid equations of Boolean algebra, it follows that the quotient T_2/E_Γ is a Boolean algebra in which the complement of $[t]$ is $[-t]$, and the meet of $[s]$ and $[t]$ is $[s \sqcap t]$. Let $\leq$ be the order of this Boolean algebra, so $[s] \leq [t]$ iff $[s] \sqcap [t] = [s]$.

By the same argument used above to show that T_1/E is finite, we have that T_2/E_Γ is finite.

Since T_2/E_Γ is a finite Boolean algebra (with at least two elements, since $\neg(0 \equiv 1)$ is an axiom), it is isomorphic to the powerset of a (nonempty) set V. Let f be the isomorphism sending elements of T_2/E_Γ to subsets of V. We take the set X of alternatives to be the set Alt of alternative labels. We then define $\widehat{D} : X^2 \to \wp(\wp(V))$ by:

$$f([t]) \in \widehat{D}(x, y) \text{ if and only if } \widehat{D}_{x>y}(t) \in \Gamma. \tag{3}$$

This is well defined, i.e., the choice of representative from $[t]$ does not matter, because if $\widehat{D}_{x>y}(s) \in \Gamma$ and $s \equiv t \in \Gamma$, then it is easy to see using the maximal consistency of Γ and the principles of our logic that $\widehat{D}_{x>y}(s) \leftrightarrow \widehat{D}_{x>y}(t) \in \Gamma$ and hence $\widehat{D}_{x>y}(s) \in \Gamma$ if and only if $\widehat{D}_{x>y}(t) \in \Gamma$. With this definition of D, the fact that Γ contains the axioms of $\overline{\widehat{\mathbf{T}}}$ implies that D satisfies the properties of Theorem 7.3 for TR (resp. the fact that Γ contains the axioms of $\overline{\widehat{\mathbf{W}}}$ implies that D satisfies the properties of Theorem 7.3 for FR).

Next we define $\overline{D} : X^2 \to \wp(\wp(V))$ by $f([t]) \in \overline{D}(x, y)$ if and only if $\overline{D}_{x>y}(t) \in \Gamma$. Again this is well-defined, for the same reasons as before. The fact that Γ contains the axiom $(\overline{D}_{x>y}(a) \wedge (a \sqsubseteq b)) \to \widehat{D}_{x>y}(b)$ implies that $\langle \widehat{D}, \overline{D} \rangle$ satisfies the left-to-right direction of property 6 of Theorem 7.3: if $A \in \overline{D}(x, y)$, then for all $B \supseteq A$, we have $B \in \widehat{D}(x, y)$. Proving the converse implication was the reason for our introduction of the set Θ of formulas. Suppose $f([t]) \notin \overline{D}(x, y)$, so $\overline{D}_{x>y}(t) \notin \Gamma$ and hence $\neg\overline{D}_{x>y}(t) \in \Gamma$ by the maximal consistency of Γ. Then given that $(t \sqsubseteq c_{x,y,[t]}) \wedge (\neg\overline{D}_{x>y}(t) \to \neg\widehat{D}_{x>y}(c_{x,y,[t]})) \in \Gamma$, from $\neg\overline{D}_{x>y}(t) \in \Gamma$ we have $\neg\widehat{D}_{x>y}(c_{x,y,[t]}) \in \Gamma$ by the maximal consistency of Γ, whence $\widehat{D}_{x>y}(c_{x,y,[t]}) \notin \Gamma$ by the consistency of Γ. Thus, $f([c_{x,y,[t]}]) \notin \widehat{D}(x, y)$. Since $t \sqsubseteq c_{x,y,[t]} \in \Gamma$, we also have $[t] \leq [c_{x,y,[t]}]$ and hence $f([t]) \subseteq f([c_{x,y,[t]}])$, which completes the proof of property 6.

Since $\langle \widehat{D}, \overline{D} \rangle$ satisfies the properties of Theorem 7.3 for TR (resp. FR), it is representable by a CCR satisfying LD, IIA, and TR (resp. FR). Let α be the coalition labeling that sends each coalition label a to the image $f([a]) \subseteq V$ of the equivalence class of $[a]$ in T_2/E_Γ. Since the set of terms was defined inductively, one can prove by induction that the extended labeling $\dot{\alpha}$ from Definition 4.2 satisfies

$$\dot{\alpha}(t) = f([t]). \tag{4}$$

Let β be the alternative labeling that sends each alternative label x to itself, recalling from above that we take the set X of alternatives to be Alt.

We claim that for any formula ψ, we have:

$$f \models_{\alpha,\beta} \psi \text{ if and only if } \psi \in \Gamma. \tag{5}$$

The set of formulas was defined inductively, so we prove the claim by induction. For formulas of the form $t \equiv s$, we have the following equivalences:

$$\begin{aligned} f \models_{\alpha,\beta} t \equiv s \quad &\Leftrightarrow \quad \dot{\alpha}(t) = \dot{\alpha}(s) \\ &\Leftrightarrow \quad f([t]) = f([s]) \text{ by (4)} \\ &\Leftrightarrow \quad [t] = [s] \text{ since } f \text{ is injective} \\ &\Leftrightarrow \quad t \equiv s \in \Gamma \text{ by definition of the equivalence relation } E_\Gamma. \end{aligned}$$

For formulas of the form $\widehat{D}_{x>y}(t)$, we have:

$$\begin{aligned} f \models_{\alpha,\beta} \widehat{D}_{x>y}(t) \quad &\Leftrightarrow \quad \dot{\alpha}(t) \in \widehat{D}_f(\beta(x), \beta(y)) \\ &\Leftrightarrow \quad f([t]) \in \widehat{D}_f(\beta(x), \beta(y)) \text{ by (4)} \\ &\Leftrightarrow \quad f([t]) \in \widehat{D}_f(x, y) \text{ since } \beta \text{ is the identity map} \\ &\Leftrightarrow \quad f([t]) \in \widehat{D}(x, y) \text{ since } \widehat{D} \text{ is represented by } f \\ &\Leftrightarrow \quad \widehat{D}_{x>y}(t) \in \Gamma \text{ by (3).} \end{aligned}$$

The proof for formulas of the form $\overline{D}_{x>y}(t)$ is analogous. For formulas of the form $\neg\varphi$:

$$\begin{aligned} f \models_{\alpha,\beta} \neg\varphi \quad &\Leftrightarrow \quad f \not\models_{\alpha,\beta} \varphi \text{ by Definition 4.4} \\ &\Leftrightarrow \quad \varphi \notin \Gamma \text{ by the inductive hypothesis} \\ &\Leftrightarrow \quad \neg\varphi \in \Gamma \text{ since } \Gamma \text{ is maximally consistent.} \end{aligned}$$

The proofs for formulas of the form $\psi \wedge \chi$, $\psi \vee \chi$, etc., also use the maximal consistency of Γ, which ensures that $\psi \wedge \chi \in \Gamma$ if and only if $\psi, \chi \in \Gamma$; $\psi \vee \chi \in \Gamma$ if and only if $\psi \in \Gamma$ or $\chi \in \Gamma$; etc. Thus, (5) holds. Then since our initial consistent formula φ belongs to Γ, we have $f \models_{\alpha,\beta} \varphi$. This completes the proof that every $\widehat{\overline{\mathbf{T}}}$-consistent (resp. every $\widehat{\overline{\mathbf{W}}}$-consistent) formula is true of some CCR satisfying LD, IIA, and TR (resp. FR).

The proof of Theorem or 4.8.2 or 4.8.4 is even simpler. In this case, without the interaction between $\widehat{D}$ and $\overline{D}$ to worry about, we can skip the step involving the new coalition labels $c_{x,y,[t]}$ and the set Θ of new formulas. Simply extend φ to a maximally consistent set Γ and continue with the rest of the proof as above. At the point where the representation result from Theorem 7.3 (for decisiveness regimes) was used, use the representation result from Theorem 3.5 (for almost decisiveness) or Theorem 3.6 (for decisiveness) instead.

References

Samson Abramsky. Arrow's theorem by arrow theory. In Andrés Villaveces, Roman Kossak, Juha Kontinen, and Åsa Hirvonen, editors, *Logic Without Borders: Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics*, pages 15–30. De Gruyter, Berlin, 2015.

Thomas Agotnes, Weibe van der Hoek, and Michael Wooldridge. On the logic of preference and judgment aggregation. *Autonomous Agents and Multi-Agent Systems*, 22(1):4–30, 2009.

Kenneth J. Arrow. Rational choice functions and orderings. *Economica*, 26(102):121–127, 1959.

Kenneth J. Arrow. *Social Choice and Individual Values*. Yale University Press, New Haven, 3rd edition, 2012.

Kenneth J. Arrow. Origins of the impossibility theorem. In Eric Maskin and Amartya Sen, editors, *The Arrow Impossibility Theorem*, pages 143–148. Columbia University Press, New York, 2014.

Nicholas Baigent. Topological theories of social choice. In Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 2, pages 301–334, Amsterdam, 2010. North-Holland.

Ning Bao and Nicole Yunger Halpern. Quantum voting and violation of Arrow's impossibility theorem. *Physical Review A*, 95, 2017. doi 10.1103/PhysRevA.95.062306.

Salvador Barberá. Pivotal voters: A new proof of Arrow's theorem. *Economic Letters*, 6 (1):13–16, 1980.

Douglas H. Blair and Robert A. Pollak. Collective rationality and dictatorship: The scope of the Arrow theorem. *Journal of Economic Theory*, 21:186–194, 1979.

Douglas H. Blair, Georges Bordes, Jerry S. Kelly, and Kotaro Suzumura. Impossibility theorems without collective rationality. *Journal of Economic Theory*, 13:361–379, 1976.

Julian H. Blau. Semiorders and collective choice. *Journal of Economic Theory*, 21:195–206, 1979.

James Buchanan. Social choice, democracy and free markets. *Journal of Political Economy*, 62(2):114–123, 1954.

Donald Campbell. *Equity, Efficiency and Social Choice*. Clarendon Press Oxford, 1992.

Donald E. Campbell and Jerry S. Kelly. Impossibility theorems in the Arrovian framework. In Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1, pages 35–94. North-Holland, Amsterdam, 2002.

Alexander Chagrov and Michael Zakharyaschev. *Modal Logic*. Clarendon Press, Oxford, 1997.

Giovanni Ciná and Ulle Endriss. Proving classical theorems of social choice theory in modal logic. *Autonomous Agents and Multi-Agent Systems*, 30(5):963–989, 2016.

John Dryzek and Christian List. Social choice theory and deliberative democracy: A reconciliation. *British Journal of Political Science*, 33(1):1–28, 2003.

Ulle Endriss. Logic and social choice theory. In Amitabha Gupta and Johan van Benthem, editors, *Logic and Philosophy Today*, pages 333–377. College Publications, London, 2011.

Peter Fishburn. Arrow's impossibility theorem: Concise proof and infinitely many voters. *Journal of Economic Theory*, 2:103–106, 1970.

Peter Fishburn and Ariel Rubinstein. Algebraic aggregation theory. *Journal of Economic Theory*, 38:63–77, 1986.

Marc Fleurbaey and Philippe Mongin. The news of the death of welfare economics is greatly exaggerated. *Social Choice and Welfare*, 25:381–418, 2005.

John Geanakoplos. Three brief proofs of Arrow's theorem. *Economic Theory*, 26(1):211–215, 2005.

Allan F. Gibbard. Intransitive social indifference and the Arrow dilemma. *Review of Economic Design*, 18:3–10, 2014.

Steven Givant and Paul Halmos. *Introduction to Boolean Algebras*. Springer, New York, 2009.

Umberto Grandi and Ulle Endriss. First-order logic formalisation of impossibility theorems in preference aggregation. *Journal of Philosophical Logic*, 42(4):595–618, 2013.

A. S. Guha. Neutrality, monotonicity, and the right of veto. *Econometrica*, 40(5):821–826, 1972.

Peter Hammond. Equity, Arrow's conditions, and Rawl's difference principle. *Econometrica*, 44:793–804, 1976.

Bengt Hansson. The existence of group preference functions. *Public Choice*, 28:89–98, 1976.

Frederik Herzberg and Daniel Eckert. The model-theoretic approach to aggregation: Impossibility results for finite and infinite electorates. *Mathematical Social Sciences*, 64:41–47, 2012.

Susan Hurley. Supervenience and the possibility of coherence. *Mind*, 94(276):501–525, 1985.

Ehud Kalai, Eitan Muller, and Mark Satterthwaite. Social welfare functions when preferences are convex and continuous: Impossibility results. *Public Choice*, 34:87–97, 1979.

Jerry S. Kelly. *Social Choice Theory: An Introduction*. Springer-Verlag, Berlin, 1988.

Alan P. Kirman and Dieter Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5:267–277, 1972.

Charles H. Kraft, John W. Pratt, and A. Seidenberg. Intuitive Probability on Finite Sets. *The Annals of Mathematical Statistics*, 30(2):408–419, 1959.

Thomas Kroedel and Franz Huber. Counterfactual dependence and Arrow. *Nous*, 47(3): 453–466, 2013.

Luc Lauwers and Luc Van Liedekerke. Ultraproducts and aggregation. *Journal of Mathematical Economics*, 24:217–237, 1995.

Michel Le Breton and John Weymark. Arrovian social choice theory on economic domains. In Kenneth J. Arrow, Amartya Sen, and Kotaro Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 2, pages 191–299. North-Holland, Amsterdam, 2010.

Christian List and Ben Polak. Introduction to judgment aggregation. *Journal of Economic Theory*, 145(2):441–466, 2010.

William MacAskill. Normative uncertainty as a voting problem. *Mind*, 124:967–1004, 2016.

Gerry Mackie. *Democracy Defended*. Cambridge University Press, Cambridge, 2003.

David Makinson. Combinatorial versus decision-theoretic components of impossibility theorems. *Theory and Decision*, 40:181–190, 1996.

Andreu Mas-Colell and Hugo Sonnenschein. General possibility theorems for group decisions. *The Review of Economic Studies*, 39(2):185–192, 1972.

Bernard Monjardet. Remarques sur une classe de procédures de votes et les théorèmes de possibilité. In *La Décision*, pages 117–184. Colloque du CNRS, Aix-en-Provence, 1967.

Bernard Monjardet. Une autre preuve du théorème d'Arrow. *R.A.I.R.O. Recherche Opérationnelle*, 12:291–296, 1978.

Bernard Monjardet. On the use of ultrafilters in social choice theory. In Prasanta K. Pattanaik and Maurice Salles, editors, *Social Choice and Welfare*, pages 73–78. North-Holland, Amsterdam, 1983.

Michael Morreau. Theory choice and social choice: Kuhn vindicated. *Mind*, 123(493): 239–262, 2015.

Y. Murakami. *Logic and Social Choice*. Dover Publications, New York, 1968.

Tobias Nipkow. Social choice theory in HOL. *Journal of Automated Reasoning*, 43(3): 289–304, 2009.

Samir Okasha. Theory choice and social choice: Kuhn versus Arrow. *Mind*, 120(477):83–115, 2011.

Eric Pacuit and Fan Yang. Dependence and independence in social choice: Arrow's theorem. In Samson Abramsky, Juha Kontinen, Jouko Väänänen, and Heribert Vollmer, editors, *Dependence Logic*, pages 235–260. Springer, 2016.

John Patty and Elizabeth Maggie Penn. *Social Choice and Legitimacy: The Possibilities of Impossibility*. Cambridge University Press, Cambridge, 2014.

Philip J. Reny. Arrow's theorem and the Gibbard-Satterthwaite theorem: a unified approach. *Economics Letters*, 70:99–105, 2001.

William H. Riker. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. William H. Freeman, San Francisco, 1982.

Ariel Rubinstein. The single profile analogues to multi profile theorems: Mathematical logic's approach. *International Economic Review*, 25(3):719–730, 1984.

Donald G. Saari. *Disposing Dictators, Demystifying Voting Paradoxes: Social Choice Analysis*. Cambridge University Press, Cambridge, 2008.

Maurice Salles. On Quine on Arrow. *Social Choice and Welfare*, 48:877–886, 2017.

Amartya Sen. Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies*, 36(2):381–393, 1969.

Amartya Sen. The impossibility of a Paretian liberal. *Journal of Political Economy*, 78(1): 152–157, 1970.

Amartya Sen. Personal utilities and public judgements: Or what's wrong with welfare economics. *The Economic Journal*, 89(335):537–558, 1979.

Amartya Sen. Internal consistency of choice. *Econometrica*, 61(3):495–521, 1993.

Amartya Sen. Arrow and the impossibility theorem. In Eric Maskin and Amartya Sen, editors, *The Arrow Impossibility Theorem*, pages 29–42. Columbia University Press, New York, 2014.

Amartya Sen. *Collective Choice and Social Welfare: An Expanded Edition*. Harvard University Press, Cambridge, Mass., 2017.

Saharon Shelah. On the Arrow property. *Advances in Applied Mathematics*, 34:217–251, 2005.

Jacob Stegenga. An impossibility theorem for amalgamating evidence. *Synthese*, 190(12): 2391–2411, 2013.

Kotaro Suzumura. *Rational choice, collective decisions, and social welfare*. Cambridge University Press, Cambridge, 1983.

Pingzhong Tang and Fangzhen Lin. Computer-aided proofs of Arrow's and other impossibility theorems. *Artificial Intelligence*, 173(11):1041–1053, 2009.

Alan Taylor and William Zwicker. A characterization of weighted voting. *Proceedings of the American Mathematical Society*, 115(4):1089–1094, 1992.

Alan D. Taylor and William S. Zwicker. *Simple Games: Desirability Relations, Trading, Pseudoweightings*. Princeton University Press, Princeton, 1999.

Nicolas Troquard, Wiebe van der Hoek, and Michael Wooldridge. Reasoning about social choice functions. *Journal of Philosophical Logic*, 40(4):473–498, 2011.

John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, New Haven, 2nd edition, 1947.

John A. Weymark. Arrow's theorem with social quasi-orderings. *Public Choice*, 42:235–246, 1984.

Freek Wiedijk. Arrow's impossibility theorem. *Formalized Mathematics*, 15(4):171–174, 2007.

Robert Wilson. Social choice theory without the Pareto principle. *Journal of Economic Theory*, 5:478–486, 1972.

Ning Neil Yu. A one-shot proof of Arrow's impossibility theorem. *Economic Theory*, 50(2): 523–525, 2012.