Logic and social choice theory.

Endriss, U.

Publication date
2011

Document Version
Final published version

Published in
Logic and Philosophy Today; vol. 2

To appear in A. Gupta and J. van Benthem (eds.), *Logic and Philosophy Today*, College Publications, 2011.

Logic and Social Choice Theory

Ulle Endriss

Institute for Logic, Language and Computation
University of Amsterdam

2011

Abstract

We give an introduction to social choice theory, the formal study of mechanisms for collective decision making, and highlight the role that logic has taken, and continues to take, in its development. The first part of the chapter is devoted to a succinct exposition of the axiomatic method in social choice theory and covers several of the classical theorems in the field. In the second part we then outline three areas of recent research activity: logics for social choice, social choice in combinatorial domains, and judgment aggregation.

Table of Contents

1 Introduction

When a group needs to make a decision, we are faced with the problem of aggregating the views of the individual members of that group into a single collective view that adequately reflects the "will of the people". How are we supposed to do this? This is a fundamental question of deep philosophical, economic, and political significance that, around the middle of 20th century, has given rise to the field of *social choice theory*. Logic has played an important role in the development of social choice theory from the very beginning.[1] In this chapter we give an introduction to social choice theory and we review a number of research trends that emphasise the place of logic in this field.

A typical (but not the only) problem studied in social choice theory is *preference aggregation*. Let us begin by considering an example.

Example 1 (Condorcet Paradox)**.** Suppose five individuals each declare their preferences by providing a ranking of the elements of a set of alternatives $\mathcal{X} = \{x, y, z\}$, as follows:

Individual 1:	$x \succ y \succ z$
Individual 2:	$x \succ y \succ z$
Individual 3:	$y \succ z \succ x$
Individual 4:	$z \succ y \succ x$
Individual 5:	$z \succ x \succ y$

What preference order would best represent the collective view of the group? An approach that suggests itself is to use the *majority rule:* rank x above y if and only if a majority of the individuals do, and similarly for all other pairs of alternatives. If we accept this rule, then we must rank x above y (as three out of five individuals do) and y above z (as, again, three out of five individuals do). This suggests that the collective preference order should be $x \succ y \succ z$. But this solution is in conflict with the fact that three out of five individuals rank z above x! This in an instance of the *Condorcet Paradox*, named after Marie Jean Antoine Nicolas de Caritat, the Marquis de Condorcet (1743–1794), the French philosopher, mathematician, and political scientist who first discussed the problem at length. Another reading of the paradox is that, when we use the majority rule to decide on the relative ranking of each pair of alternatives, then we end up with a cyclical preference relation: $x \succ y \succ z \succ x$. Yet another perspective is the following: Even if we are content with just finding a best alternative (rather than a full collective preference order), whichever alternative we decide to declare the winner, there will always be another alternative that is preferred by a strict majority of the individuals. □

[1] A curious anecdote in this regard is the fact that, during his final year as an undergraduate at City College of New York in 1940, Kenneth J. Arrow, the father of social choice theory and winner of the 1972 Nobel Prize in Economics, took a class with Alfred Tarski (Suppes, 2005).

The question now arises whether there are better methods of aggregation than the majority rule. Social choice theorists have approached this question using the so-called *axiomatic method*. This method amounts to formulating normatively desirable properties of aggregation rules as "axioms" in a mathematically rigorous manner, to then obtain precise characterisations of the aggregation rules that satisfy these properties. The best known example is the Impossibility Theorem from Kenneth J. Arrow's seminal work "*Social Choice and Individual Values*", originally published in 1951 (Arrow, 1963). Arrow argued that any acceptable method of aggregation should satisfy at least the following two axioms:

(1) If every individual ranks x above y, then so should society.

(2) It should be possible to determine the relative social ranking of x and y by considering only the relative ranking of x and y supplied by each of the individuals.

Observe that the majority rule, for example, does satisfy both of these requirements—but of course, as we have seen, it is not a "proper" aggregation method, as it may generate a preference order with cycles. Arrow then proved a truly astonishing result: If there are at least three alternatives, then the only kind of mechanism that will respect both of our axioms and that will return a collective preference that is a linear order for any combination of individual preferences is a *dictatorship*, i.e., a function that simply copies the preferences of a fixed individual and returns it as the collective preference! In other words, satisfying both axioms and the requirement of being nondictatorial is *impossible*.

Section 2 of this chapter is an introduction to the axiomatic method in social choice theory. As part of this introduction, we present a selection of the most important classical theorems in the field: Arrow's Theorem, Sen's Theorem on the Impossibility of a Paretian Liberal, May's characterisation of the majority rule for two alternatives, the Muller-Satterthwaite Theorem establishing the impossibility of devising an acceptable aggregation rule that is monotonic, and the Gibbard-Satterthwaite Theorem showing that any voting rule can be manipulated. We provide full proofs for all of these results.

The axiomatic method often makes reference to notions from logic, albeit only in an informal manner. For instance, the notion of axiom used here is inspired by, although different from, the use of the term in mathematical logic, and most of the results we will discuss establish the "logical inconsistency" of certain requirements. The fact that preferences are modelled as binary relations also provides a bridge to formal logic, and some further connections will be commented on in Section 2.

In recent years a number of other, and more formal, uses of logic have emerged. Some of these developments have been fuelled by the insight that the significance of social choice theory goes beyond the humanities and the social sciences and that topics such as preference aggregation also have applications in computer science, e.g., for collective decision making in multiagent systems or the aggregation of ranked search results in an Internet meta-search engine. This insight, together with the fact that formal methods traditionally associated with computer science have turned out to be helpful in the analysis of social choice problems, has lead to the emergence of a new research area called *computational social choice*.

In the second part of this chapter we will discuss a selection of recent developments in social choice theory and computational social choice in which formal logic plays a central role. Section 3 reviews approaches to *formalising* parts of social choice theory in a suitable logic, e.g., a modal logic or classical first-order logic. This line of research is useful not only in view of deepening our understanding of some of the concepts of social choice theory and clarifying the expressivity required from a language that will allow us to talk about these concepts, but it also bears the potential of leading to some very practical applications. Formalisation is a first step towards automation, and this kind of work may pave the way for the future development of tools that will allow us to automatically reason about mechanisms for collective decision making. Section 4 is an introduction to the area of *social choice in combinatorial domains*, i.e., the design and analysis of collective decision making mechanisms for problems where the set of alternatives has a combinatorial structure, as is the case, for instance, when a group has to decide which of a number of issues to accept and which to reject. Many social choice problems arising in practice have a combinatorial component and classical methods in social choice theory are often not adequate to deal with this computational challenge. Several of the approaches that have been proposed to tackle this difficult problem involve a language for the compact representation of preferences, and these languages are typically based on logic. Section 5 covers the basics of *judgment aggregation*, a relatively new topic in social choice theory dealing with the problem of aggregating the judgments of a group of individuals judges on a set of interrelated propositions. These propositions are modelled as formulas of propositional logic. We conclude in Section 6 with a few general remarks about the field and suggestions for further reading.

2 The Axiomatic Method in Social Choice Theory

In this section, we present the *axiomatic method*, as pioneered by Arrow (1963). It probably is the most important methodological tool in social choice theory (see, e.g., Arrow et al., 2002; Sen, 1986; Austen-Smith and Banks, 1999; Gaertner, 2006). We will prove several of the seminal theorems in the field. Throughout, let $\mathcal{N} = \{i_1, \ldots, i_n\}$ be a finite set of (at least two) *individuals* (or *voters*, or *agents*); and let $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$ be a (not necessarily finite) nonempty set of *alternatives* (or *candidates*, or *social states*).[2]

[2] In this exposition we do not distinguish between *feasible* and *infeasible* alternatives, as some of the literature does (particularly in the context of *social choice functions*, which we discuss in Section 2.2). We also do not model *variable electorates*.

Each voter in $\mathcal{N}$ is endowed with, and will be asked to express, a *preference* over the alternatives in $\mathcal{X}$. There are a number of options at our disposal when we want to model such preferences. In most of social choice theory, preferences are either linear or weak orders on $\mathcal{X}$. Recall that a linear order is a binary relation that is irreflexive, transitive, and complete, while a weak order is a binary relation that is reflexive, transitive, and complete. Throughout this section, we shall assume that preferences are linear orders, but all definitions given and all results proven can easily be adapted to the case of weak orders. By taking preferences to be linear orders, we accept certain basic principles. One of them is that we cannot compare preferences across individuals: our model cannot express whether individual 1 likes alternative x more than individual 2 likes alternative y.[3] There also is no notion of preference intensity: we cannot model whether individual 1's preference of x over y is more intense than her preference of y over z. Yet another such principle is that we take it that each individual has the cognitive capacity to rank any two alternatives. We stress that other models of preference, such as utility functions (Roemer, 1996) or partial orders (Pini et al., 2008), are also interesting and relevant, but here we shall restrict ourselves to linear orders.

Let $\mathcal{L}(\mathcal{X})$ denote the set of all linear orders on $\mathcal{X}$. A *profile* $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{L}(\mathcal{X})^{\mathcal{N}}$ is a vector of linear orders (i.e., preferences), where R_i is the linear order supplied by individual i. We write $N^{\boldsymbol{R}}_{x \succ y}$ to denote the set of individuals that rank alternative x above alternative y under profile $\boldsymbol{R}$. For instance, if $\boldsymbol{R}$ is the profile given in Example 1 above, then $N^{\boldsymbol{R}}_{z \succ x} = \{3, 4, 5\}$.

2.1 Social Welfare Functions: Arrow's Theorem

The first type of preference aggregation mechanism we consider are functions that map a profile of preference orders to a single (collective) preference order. Such a function is called a *social welfare function* (SWF). Formally, a SWF is a function $F : \mathcal{L}(\mathcal{X})^{\mathcal{N}} \to \mathcal{L}(\mathcal{X})$.

Let us now give a precise account of the two axioms mentioned in the introduction above, which Arrow (1963) argued to be basic requirements for any acceptable SWF. The first is a fundamental principle in economic theory, due to the Italian economist Vilfredo Pareto (1848–1923), that states that if x is at least as good as y for all and strictly better for some members of a society, then x should be socially preferred to y. Given that we assume that preferences are strict, i.e., no individual will be indifferent between two distinct alternatives, this simplifies to asking that x should be socially preferred to y if everybody prefers x to y. We now formulate this principle as an axiom (i.e., a property) that may or may not be satisfied by a given SWF.

[3] Note that while it is likely that, if individual 1 ranks x first and individual 2 ranks y last, then individual 1 likes x more than individual 2 likes y, this is a heuristic inference that is outside of the formal model.

Pareto. A SWF F satisfies the *Pareto condition* if, whenever all individuals rank x above y, then so does society: $N^{\boldsymbol{R}}_{x \succ y} = \mathcal{N}$ implies $(x, y) \in F(\boldsymbol{R})$.

In addition to the Pareto condition, a widely accepted standard requirement, Arrow proposed an independence axiom that states that the relative social ranking of two alternatives should not change when an individual updates her preferences regarding a third alternative. That is, social choices should be independent of irrelevant alternatives.

Independence of irrelevant alternatives (IIA). A SWF F satisfies *IIA* if the relative social ranking of two alternatives only depends on their relative individual rankings: $N^{\boldsymbol{R}}_{x \succ y} = N^{\boldsymbol{R}'}_{x \succ y}$ implies $(x, y) \in F(\boldsymbol{R}) \Leftrightarrow (x, y) \in F(\boldsymbol{R}')$.

That is, if the set of individuals ranking x above y does not change when we move from profile $\boldsymbol{R}$ to profile $\boldsymbol{R}'$, then the social preference order obtained under $\boldsymbol{R}$ should make the same judgment regarding the relative ranking of x and y as the social preference order obtained under $\boldsymbol{R}'$. Arrow argued that the Pareto condition and IIA are basic democratic principles that any acceptable SWF must satisfy. Note, however, that the Pareto condition and IIA *alone* do *not* guarantee a democratic procedure. Specifically, any *dictatorship*, i.e., any SWF that for a fixed "dictator" $i \in \mathcal{N}$ will map any profile $\boldsymbol{R}$ to the dictator's reported ranking R_i, satisfies both properties. Arrow's deeply surprising result shows that dictatorships are in fact *the only* SWFs that satisfy both properties, at least when there are three or more alternatives.

Theorem 1 (Arrow, 1951)**.** *Any SWF for three or more alternatives that satisfies the Pareto condition and IIA must be a dictatorship.*

Proof. Our proof broadly follows Sen (1986) and is based on the idea of "decisive coalitions". Consider any SWF F for three or more alternatives that satisfies the Pareto condition and IIA. Let us call a coalition $G \subseteq \mathcal{N}$ *decisive* on alternatives (x, y) if $G \subseteq N^{\boldsymbol{R}}_{x \succ y}$ entails $(x, y) \in F(\boldsymbol{R})$. When G is decisive on *all* pairs of alternatives, then we simply say that G is decisive. Note that the Pareto condition is satisfied if and only if the grand coalition $\mathcal{N}$ is decisive, while an individual i is a dictator if and only if the singleton $\{i\}$ is decisive. The main idea of the proof is to show that, whenever some coalition G (with $|G| \geqslant 2$) is decisive, then there exists a nonempty $G' \subset G$ that is decisive as well (this property is known as the *Contraction Lemma*). Given the finiteness of $\mathcal{N}$, this means that F satisfying the Pareto condition will, by induction, entail that F is dictatorial.

Before proving the Contraction Lemma, let us first establish a fundamental property of the notion of decisiveness. Suppose society will rank x above y whenever *exactly* the individuals in G do: $(x, y) \in F(\boldsymbol{R})$ whenever $N^{\boldsymbol{R}}_{x \succ y} = G$. We will show that this is a sufficient condition for G being decisive on any given pair (x', y'). We prove the case where x, y, x', y' are all distinct (the other cases are similar and left as an exercise to the reader). Consider a profile where everyone in G ranks $x' \succ x \succ y \succ y'$ and all other individuals make the judgments $x' \succ x$, $y \succ y'$, and $y \succ x$ (note that we leave their judgments on x' vs. y' undetermined). That is, exactly the individuals in G rank $x \succ y$, which implies $x \succ y$ for the social ranking. The Pareto condition implies $x' \succ x$ and $y \succ y'$ for the social ranking. Thus, by transitivity, we get $x' \succ y'$ for the social ranking. Finally, by IIA and the fact that we had left the relative ranking of x' and y' undetermined for individuals outside of G, any other profile $\boldsymbol{R}$ with $G \subseteq N^{\boldsymbol{R}}_{x' \succ y'}$ will also result in $(x', y') \in F(\boldsymbol{R})$. Hence, if exactly the individuals in G ranking x above y is a sufficient condition for society to do the same, then G is decisive (on *all* pairs of alternatives).

Now we turn to the proof of the Contraction Lemma. Let $G \subseteq \mathcal{N}$ with $|G| \geqslant 2$, i.e., there are nonempty coalitions G_1 and G_2 with $G = G_1 \cup G_2$ and $G_1 \cap G_2 = \emptyset$. We want to show that G being decisive entails either G_1 or G_2 being decisive as well. (We now shall make use of the fact that there are at least three alternatives.) Consider a profile where all individuals in G_1 rank $x \succ y \succ z$, all individuals in G_2 rank $y \succ z \succ x$, and all others rank $z \succ x \succ y$. As G is decisive, we have $y \succ z$ in the social ranking. We distinguish two cases:

(1) Society ranks x above z. Note that it is exactly the individuals in G_1 that rank x above z. Thus, by IIA, in any profile $\boldsymbol{R}$ where exactly the individuals in G_1 rank x above z, society will do the same. But, as we have seen earlier, this means that G_1 is decisive.

(2) Society ranks z above x, and thus y above x. As exactly the individuals in G_2 rank y above x, by the same kind of argument as above, G_2 must be decisive.

As indicated earlier, this concludes the proof of the theorem: repeated application of the Contraction Lemma reduces the Pareto condition to the existence of a dictator. □

Arrow's Theorem may be read either as a *characterisation* of dictatorships in terms of the axioms of Pareto and IIA, or as an *impossibility theorem:* it is impossible to devise a SWF for three or more alternatives that is Pareto efficient, independent, and nondictatorial.

Observe that we have made explicit use of both the assumption that there are at least three alternatives and the assumption that the set of individuals is finite (the latter was required for the inductive application of the Contraction Lemma). Indeed, if either assumption is dropped, then Arrow's Theorem ceases to hold: First, for two alternatives, the *majority rule*, which returns the ranking made by the majority of individuals (with ties broken in favour of, say, the first alternative), satisfies both the Pareto condition and IIA and clearly is not dictatorial. Second, for an infinite number of individuals, Fishburn (1970) has shown how to design a nondictatorial SWF that is Pareto efficient and independent. Whether or not the set of *alternatives* is infinite is uncritical.

Several alternative proofs for Arrow's Theorem may be found in the literature (see, e.g., Barberà, 1980; Geanakoplos, 2005; Gaertner, 2006). We want to briefly mention one such proof here, due to Kirman and Sondermann (1972), which reduces Arrow's Theorem to a well-known fact in the theory of ultrafilters. Given the importance of ultrafilters in model theory and set theory, this proof provides additional evidence for the close connections between logic and social choice theory. Recall that an *ultrafilter* $\mathcal{G}$ for a set $\mathcal{N}$ is a set of subsets of $\mathcal{N}$ satisfying the following conditions (Davey and Priestley, 2002):

- (i) The empty set is not included: $\emptyset \notin \mathcal{G}$.
- (ii) $\mathcal{G}$ is closed under intersection: if $G_1 \in \mathcal{G}$ and $G_2 \in \mathcal{G}$, then $G_1 \cap G_2 \in \mathcal{G}$.
- (iii) $\mathcal{G}$ is maximal: for all $G \subseteq \mathcal{N}$, either $G \in \mathcal{G}$ or $(\mathcal{N} \setminus G) \in \mathcal{G}$.

Let us now interpret $\mathcal{N}$ as a set of individuals and $\mathcal{G}$ as the set of decisive coalitions for a given SWF satisfying the Pareto condition and IIA. It turns out that $\mathcal{G}$ satisfies the three conditions above, i.e., it is an ultrafilter. Condition (i) clearly holds, because the empty coalition is not decisive. Conditions (ii) and (iii) can be shown to hold (under the assumption that $|\mathcal{X}| \geqslant 3$) using the same kind of technique we have used in the proof of Theorem 1 to show that G being "almost decisive" implies decisiveness of G.

Hence, any property of ultrafilters will extend to the set of decisive coalitions. Recall that an ultrafilter is called *principal* if it is the set of all subsets containing a fixed element d. That is, a principal filter directly corresponds to the set of decisive coalitions for a dictatorial SWF. Arrow's Theorem now follows from the fact that any finite ultrafilter must be principal (Davey and Priestley, 2002). For recent examples of applications of ultrafilters in social choice theory, we refer to the works of Daniëls and Pacuit (2009) and Herzberg and Eckert (2011).

2.2 Social Choice Functions: Sen, May, Muller-Satterthwaite

A *social choice function* SCF is a function $F : \mathcal{L}(\mathcal{X})^{\mathcal{N}} \to 2^{\mathcal{X}} \setminus \{\emptyset\}$ mapping profiles of linear orders on alternatives to nonempty sets of alternatives. Intuitively, for a given profile of declared preferences, F will choose the "best" alternatives. If F always returns a singleton, then F is called *resolute*. We can think of a SCF as a *voting rule*, mapping profiles of ballots cast by the voters to winning candidates.

The first result we shall review for the framework of SCFs is Sen's Theorem on the Impossibility of a Paretian Liberal (Sen, 1970b). Sen introduced a new type of axiom, *liberalism*, which requires that for each individual there should be at least one pair of alternatives for which she can determine the relative social ranking (i.e., she should be able to ensure that at least one of them does not win). This idea makes sense if we think of $\mathcal{X}$ as the set of social states. For example, if x and y describe identical states of the world, except that in x I will paint the walls of my bedroom in white while in y I will paint them in pink, then I alone should be free to decide on the social ranking of x and y.

Liberalism. A SCF F satisfies the axiom of *liberalism* if, for every individual $i \in \mathcal{N}$, there exist two distinct alternatives $x, y \in \mathcal{X}$ such that i is *two-way decisive* on x and y in the sense that whichever of the two i ranks lower cannot win: $i \in N^{\boldsymbol{R}}_{x \succ y}$ implies $y \notin F(\boldsymbol{R})$ and $i \in N^{\boldsymbol{R}}_{y \succ x}$ implies $x \notin F(\boldsymbol{R})$.

In fact, as we shall see, for the purposes of Sen's Theorem it suffices to assume that there are at least two individuals with this power. The second axiom required to state the theorem is again the Pareto condition, which takes the following form in the context of SCFs:

Pareto. A SCF F satisfies the *Pareto condition* if, whenever all individuals rank x above y, then y cannot win: $N^{\boldsymbol{R}}_{x \succ y} = \mathcal{N}$ implies $y \notin F(\boldsymbol{R})$.

It turns out that the two axioms are incompatible:

Theorem 2 (Sen, 1970)**.** *No SCF satisfies both liberalism and the Pareto condition.*

Proof. For the sake of contradiction, suppose there exists a SCF F satisfying both liberalism and the Pareto condition. Let i_1 and i_2 be two distinguished individuals, let x_1 and y_1 be the alternatives on which i_1 is two-way decisive, and let x_2 and y_2 be the alternatives on which i_2 is two-way decisive. We shall derive a contradiction for the case where x_1, y_1, x_2, y_2 are pairwise distinct (the remaining cases are similar and left as an exercise to the reader). Consider a profile with the following properties:

(1) Individual i_1 ranks x_1 above y_1.
(2) Individual i_2 ranks x_2 above y_2.
(3) All individuals rank y_1 above x_2 and also y_2 above x_1.
(4) All individuals rank x_1, x_2, y_1, y_2 above all other alternatives.

Due to liberalism, (1) rules out y_1 as a winner and (2) rules out y_2 as a winner. Due to the Pareto condition, (3) rules out x_1 and x_2 as winners and (4) rules out all other alternatives as winners. As a SCF must return a nonempty set of winners, we have thus derived a contradiction and are done. $\square$

Note that Sen's Theorem, unlike Arrow's, does not rely on the number of individuals being finite. It does, however, presuppose that there are at least two individuals (which, technically, is not the case for Arrow's Theorem). Sen's Theorem also does not, *a priori*, make any assumptions on the number of alternatives, although the requirement that liberalism should apply to at least two individuals quickly rules out the case of two or fewer alternatives.

We now turn to yet another type of axiom: *monotonicity*. Intuitively, a SCF is monotonic if any additional support for a winning alternative will benefit that alternative. Somewhat surprisingly, not all commonly used voting rules do satisfy this property.

Example 2 (Failure of monotonicity)**.** Under *plurality with runoff*, the voting rule used to elect the French president, the two candidates that are ranked first by the largest number of voters are presented to the electorate in a second round of voting. Suppose 17 voters vote according to the following preferences:

6 voters:	$x \succ z \succ y$
5 voters:	$y \succ x \succ z$
6 voters:	$z \succ y \succ x$

Then x and z will make it into the second round, where x will beat z with 11 to 6 votes (because the 5 voters from the middle group will now vote for x). Now suppose that two of the 6 voters from the last group (who support z and rank x last) change strategy and instead vote according to $x \succ z \succ y$ (i.e., they join the first group). Then x and y will make it into the second round (with 8 and 5 points, respectively), where y will beat x with 9 to 8 votes (because the 4 voters from the diminished last group will now vote for y). That is, even though the old winner x received additional support, she lost the new election. $\square$

We shall consider two axioms that instantiate the generic idea of monotonicity, each of which is at the heart of a further classical theorem. The first of these is May's Theorem (May, 1952). It uses the axiom of *positive responsiveness*, which requires that whenever an alternative $x^\star$ is amongst the winners and some individuals raise $x^\star$ in their linear orders *without affecting the relative rankings of any other pairs of alternatives*, then $x^\star$ should become the *sole* winner.

> **Positive responsiveness.** A SCF F satisfies *positive responsiveness* if $x^\star \in F(\boldsymbol{R})$ implies $\{x^\star\} = F(\boldsymbol{R}')$ for any alternative $x^\star$ and distinct profiles $\boldsymbol{R}$ and $\boldsymbol{R}'$ with $N^{\boldsymbol{R}}_{x^\star \succ y} \subseteq N^{\boldsymbol{R}'}_{x^\star \succ y}$ and $N^{\boldsymbol{R}}_{y \succ z} = N^{\boldsymbol{R}'}_{y \succ z}$ for all $y, z \in \mathcal{X} \setminus \{x^\star\}$.

Observe that the requirement that $\boldsymbol{R}$ and $\boldsymbol{R}'$ be distinct ensures that there are at least one individual i and one alternative y such that i moves $x^\star$ from below to above y when switching from $\boldsymbol{R}$ to $\boldsymbol{R}'$. The axiom of *weak monotonicity*, which we only mention in passing here, is the axiom we obtain when we weaken $\{x^\star\} = F(\boldsymbol{R}')$ to $x^\star \in F(\boldsymbol{R}')$ in the above statement. As Example 2 has demonstrated, *plurality with runoff* violates both weak monotonicity and positive responsiveness.

Two simple further axioms feature in May's Theorem: *anonymity* and *neutrality*. Anonymity requires that F be symmetric with respect to individuals: if π is a permutation on $\mathcal{N}$, then $F(R_1, \ldots, R_n) = F(R_{\pi(1)}, \ldots, R_{\pi(n)})$ for any profile $\boldsymbol{R}$. Neutrality requires that F be symmetric with respect to alternatives: if π is a permutation on $\mathcal{X}$, then $\pi(F(\boldsymbol{R})) = F(\pi(\boldsymbol{R}))$ for any profile $\boldsymbol{R}$ (with π extended to sets of alternatives and profiles in the natural manner). May's Theorem provides a complete characterisation of the *simple majority rule* for social choice with *two* alternatives. The simple majority rule (more often referred to as the *plurality* *rule* when there are more than two alternatives) is the SCF that returns as winners those alternatives that have been ranked first by the largest number of voters. That is, in the case of two alternatives x and y, x wins under profile $\boldsymbol{R}$ if $|N^{\boldsymbol{R}}_{x \succ y}| > |N^{\boldsymbol{R}}_{y \succ x}|$, y wins if the opposite is true, and both win if the two figures are the same.

Theorem 3 (May, 1952)**.** *A SCF for two alternatives satisfies anonymity, neutrality and positive responsiveness if and only of it is the simple majority rule.*

Proof. The simple majority rule is easily seen to satisfy all three properties. For the opposite direction, suppose F is anonymous, neutral and positively responsive. Let $\mathcal{X} = \{x, y\}$. As there are only two alternatives, any profile $\boldsymbol{R}$ can be fully described in terms of $N^{\boldsymbol{R}}_{x \succ y}$ and $N^{\boldsymbol{R}}_{y \succ x}$. Due to anonymity, only the cardinalities of $N^{\boldsymbol{R}}_{x \succ y}$ and $N^{\boldsymbol{R}}_{y \succ x}$ are relevant to the determination of the outcome. Suppose the number of individuals is odd (the other case is similar and left as an exercise to the reader). Distinguish two cases:

(1) Suppose that for every profile $\boldsymbol{R}$ with $|N^{\boldsymbol{R}}_{x \succ y}| = |N^{\boldsymbol{R}}_{y \succ x}| + 1$, only x wins. Then, by positive responsiveness, only x will win whenever $|N^{\boldsymbol{R}}_{x \succ y}| > |N^{\boldsymbol{R}}_{y \succ x}|$, i.e., F is the simple majority rule.

(2) Suppose there exists a profile $\boldsymbol{R}$ with $|N^{\boldsymbol{R}}_{x \succ y}| = |N^{\boldsymbol{R}}_{y \succ x}| + 1$, but y wins (alone or together with x). Now suppose one individual that reported $x \succ y$ under $\boldsymbol{R}$ switches to $y \succ x$. Call the new profile $\boldsymbol{R}'$. By positive responsiveness, now *only* y wins. But this new situation, with $|N^{\boldsymbol{R}'}_{y \succ x}| = |N^{\boldsymbol{R}'}_{x \succ y}| + 1$, is symmetric to the earlier situation. Thus, by neutrality, x should win, i.e., we have arrived at a contradiction.

Hence, the only possibility is for F to coincide with the simple majority rule. □

Next on our list is the Muller-Satterthwaite Theorem (Muller and Satterthwaite, 1977), which shows how a stronger form of monotonicity can lead to an impossibility similar to Arrow's Theorem. This result applies to *resolute* SCFs. If F is resolute, we shall simply write $x = F(\boldsymbol{R})$ rather than $x \in F(\boldsymbol{R})$ to indicate that x is the winner under profile $\boldsymbol{R}$.

> **Strong monotonicity.** A resolute SCF F satisfies *strong monotonicity* if $x^\star = F(\boldsymbol{R})$ implies $x^\star = F(\boldsymbol{R}')$ for any alternative $x^\star$ and distinct profiles $\boldsymbol{R}$ and $\boldsymbol{R}'$ with $N^{\boldsymbol{R}}_{x^\star \succ y} \subseteq N^{\boldsymbol{R}'}_{x^\star \succ y}$ for all $y \in \mathcal{X} \setminus \{x^\star\}$.

Unlike for weak monotonicity (or positive responsiveness), here we do not require that the relative rankings of other pairs (y, z) need to be maintained when raising $x^\star$.

One further axiom we require is *surjectivity.* F is surjective if it does not rule out certain alternatives as a possible winners from the outset: for every $x \in \mathcal{X}$ there exists a profile $\boldsymbol{R}$ such that $F(\boldsymbol{R}) = x$. Finally, a resolute SCF F is *dictatorial* if there exists an individual $i \in \mathcal{N}$ such that the winner under F is always the top-ranked alternative of i.

Theorem 4 (Muller and Satterthwaite, 1977)**.** *Any resolute SCF for three or more alternatives that is surjective and strongly monotonic must be a dictatorship.*

Proof. We will show that any resolute SCF that is surjective and strongly monotonic must also satisfy the Pareto condition and an independence property similar to IIA, thereby reducing the claim to a variant of Arrow's Theorem for resolute SCFs.

First, let us show that strong monotonicity entails the following *independence* property: if $x \neq y$, $F(\boldsymbol{R}) = x$, and $N^{\boldsymbol{R}}_{x \succ y} = N^{\boldsymbol{R}'}_{x \succ y}$ then $F(\boldsymbol{R}') \neq y$. Suppose the three premises hold. Now construct a third profile, $\boldsymbol{R}''$, in which all individuals rank x and y in the top two positions, with $N^{\boldsymbol{R}''}_{x \succ y} = N^{\boldsymbol{R}}_{x \succ y}$. By strong monotonicity, $F(\boldsymbol{R}) = x$ implies $F(\boldsymbol{R}'') = x$. Again by strong monotonicity, $F(\boldsymbol{R}') = y$ would imply $F(\boldsymbol{R}'') = y$. Thus, we must have $F(\boldsymbol{R}') \neq y$.

Next, let us show that surjectivity and strong monotonicity together imply the *Pareto condition.* Take any two alternatives x and y. Due to surjectivity, there exists a profile under which x wins. Now move x above y in all individual ballots (if not above already). By strong monotonicity, x still wins (and y does not). Now, by independence, y does not win for *any* profile in which all individuals continue to rank x above y, i.e., we get Pareto efficiency.

We now prove that any resolute SCF for three or more alternatives that is independent and Pareto efficient must be a dictatorship. As for Theorem 1, we again use the "decisive coalition" technique. Call a coalition $G \subseteq \mathcal{N}$ *decisive* on (x, y) if $G \subseteq N^{\boldsymbol{R}}_{x \succ y}$ implies $y \neq F(\boldsymbol{R})$.

We start by proving that, if $G = N^{\boldsymbol{R}}_{x \succ y}$ implies $F(\boldsymbol{R}) \neq y$, then G is decisive on any given pair (x', y'). We prove the case where x, y, x', y' are all distinct (the other cases are similar and left as an exercise to the reader). Consider a profile $\boldsymbol{R}$ where everyone in G ranks $x' \succ x \succ y \succ y'$ and all other individuals rank $x' \succ x$, $y \succ y'$, and $y \succ x$. Furthermore, x, y, x', y' are ranked above all other alternatives by all individuals. Note that we do not specify the relative ranking of x' and y' by individuals outside of G. In this profile, x' must win: $G = N^{\boldsymbol{R}}_{x \succ y}$ excludes y as a possible winner and the Pareto condition rules out x (dominated by x') and y' (dominated by y). Now consider any profile $\boldsymbol{R}'$ with $G \subseteq N^{\boldsymbol{R}'}_{x' \succ y'}$. As we had left the relative ranking of x' and y' under $\boldsymbol{R}$ unspecified for individuals outside of G, w.l.o.g., we may assume $N^{\boldsymbol{R}}_{x' \succ y'} = N^{\boldsymbol{R}'}_{x' \succ y'}$. Thus, by independence, $F(\boldsymbol{R}') \neq y'$, i.e., G is decisive on (x', y').

We are now ready to prove a variant of the Contraction Lemma for SCFs. Let G be a coalition with $|G| \geqslant 2$ that is decisive (on all pairs), and let G_1 and G_2 be coalitions with $G = G_1 \cup G_2$ and $G_1 \cap G_2 = \emptyset$. We will show that either G_1 or G_2 must be decisive as well. Consider a profile where all individuals in G_1 rank $x \succ y \succ z$, all those in G_2 rank $y \succ z \succ x$, and all others rank $z \succ x \succ y$. Furthermore, all individuals rank all other alternatives below x, y, z. As G is decisive, z cannot win, which leaves two cases:

(1) The winner is x. Note that the individuals in G_1 were the only ones ranking x above z. By independence, z will lose (to x) in any profile where precisely the individuals in

G_1 rank x above z. By our earlier observation, this means that G_1 is decisive on all pairs of alternatives.

(2) The winner is y, i.e., x does not win. Considering that it is exactly the individuals in G_2 that rank y above x, by the same argument as above, it follows that G_2 is decisive.

Recall that the Pareto condition means that the grand coalition $\mathcal{N}$ is decisive. By induction, applying the Contraction Lemma at each step, we can reduce this fact to the existence of a singleton that must be decisive, i.e., there exists a dictator. □

Observe that if we drop the requirement of surjectivity, then we do obtain SCFs that are nondictatorial and strongly monotonic. For instance, any *constant* SCF that simply maps any input profile to a fixed winner $x^\star$ satisfies both conditions.

Theorem 4, which is what is nowadays usually referred to as the Muller-Satterthwaite Theorem, is in fact a corollary to the original result of Muller and Satterthwaite (1977) and the Gibbard-Satterthwaite Theorem, which we shall review next. In our exposition, we instead prove the Gibbard-Satterthwaite Theorem as a corollary to Theorem 4.

2.3 Strategic Manipulation: The Gibbard-Satterthwaite Theorem

So far there has been no need to distinguish between the preferences an individual *declares* when reporting to an aggregation mechanism (i.e., her ballot) and the *true* preferences of that individual. This distinction does become crucial when we want to reason about the incentives of individuals.

Example 3 (Manipulation)**.** Consider the following profile, broadly inspired by the situation in Florida during the United States presidential elections in the year 2000:

49% of the electorate:	Bush $\succ$ Gore $\succ$ Nader
20% of the electorate:	Gore $\succ$ Bush $\succ$ Nader
20% of the electorate:	Gore $\succ$ Nader $\succ$ Bush
11% of the electorate:	Nader $\succ$ Gore $\succ$ Bush

Suppose this represents the true preferences of the voters. Under the *plurality rule*, which awards one point to a candidate whenever he is ranked first by a voter, Bush will win this election (49% of the awarded points) ahead of Gore (40%), with Nader in last place (11%). Now suppose the Nader supporters (the last group) decide to misrepresent their true preferences on the ballot sheet and to instead rank Gore first. Then Gore (51%) will beat Bush (49%). As the Nader supporters prefer Gore over Bush, they have an incentive to engage in this kind of *strategic manipulation*. □

While the Nader voters would have been well advised to manipulate in this sense, it seems unfair to blame them. Rather, we may want to put the blame on the designers of the voting rule used, for allowing situations in which it is in the best interest of an individual to lie about her preferences. Are there voting rules that do not have this deficiency?

We encode this desideratum as another axiom. Informally speaking, a SCF is called *strategy-proof* if it never gives any individual an incentive to misrepresent her preferences.

> **Strategy-proofness.** A resolute SCF F is *strategy-proof* if for no individual $i \in \mathcal{N}$ there exist a profile $\boldsymbol{R}$ (including the "truthful preference" R_i of i) and a linear order R_i' (representing the "untruthful" ballot of i) such that $F(\boldsymbol{R}_{-i}, R_i')$ is ranked above $F(\boldsymbol{R})$ according to R_i.[4]

In the early 1970s, Gibbard (1973) and Satterthwaite (1975) confirmed a long-standing conjecture stating that there exists *no* acceptable resolute SCF that is strategy-proof.

Theorem 5 (Gibbard and Satterthwaite, 1973/1975)**.** *Any resolute SCF for three or more alternatives that is surjective and strategy-proof must be a dictatorship.*

Proof. We will show that strategy-proofness implies strong monotonicity. The claim then follows from Theorem 4. Suppose F is *not* strongly monotonic. That is, there exist profiles $\boldsymbol{R}$ and $\boldsymbol{R}'$ as well as distinct alternatives x and x' such that $F(\boldsymbol{R}) = x$ and $F(\boldsymbol{R}') = x'$ even though $N^{\boldsymbol{R}}_{x \succ y} \subseteq N^{\boldsymbol{R}'}_{x \succ y}$ for all $y \in \mathcal{X} \setminus \{x\}$. Now suppose we move from $\boldsymbol{R}$ to $\boldsymbol{R}'$, with individuals changing their ballots one by one. There must be a *first* individual whose change affects the winner. Thus, w.l.o.g., we may assume that $\boldsymbol{R}$ and $\boldsymbol{R}'$ differ in exactly one ballot. Let the individual corresponding to that ballot be i. We can distinguish two cases:

(1) $i \in N^{\boldsymbol{R}'}_{x \succ x'}$. In this case, imagine i's true preferences are as in $\boldsymbol{R}'$. Then she can profit from instead voting as in $\boldsymbol{R}$ (causing x rather than x' to win). Thus, F is not strategy-proof.

(2) $i \notin N^{\boldsymbol{R}'}_{x \succ x'}$. As we have $N^{\boldsymbol{R}}_{x \succ x'} \subseteq N^{\boldsymbol{R}'}_{x \succ x'}$, this implies $i \notin N^{\boldsymbol{R}}_{x \succ x'}$ and thus $i \in N^{\boldsymbol{R}}_{x' \succ x}$. Now imagine i's true preferences are as in $\boldsymbol{R}$. Then she can profit from instead voting as in $\boldsymbol{R}'$, i.e., F is not strategy-proof.

Thus, lack of strong monotonicity implies lack of strategy-proofness in all cases. □

It is important to note that the Gibbard-Satterthwaite Theorem only applies to *resolute* SCFs. There are several variants of the theorem for irresolute rules, the best known of which is the Duggan-Schwartz Theorem (Duggan and Schwartz, 2000). On the other hand, there are also positive results that show that irresolute strategy-proof SCFs do exist, albeit SCFs

[4] Here $(\boldsymbol{R}_{-i}, R_i')$ denotes the profile we obtain when we replace R_i in $\boldsymbol{R}$ by R_i'.

that lack the discriminatory force of rules that would usually be considered acceptable in the context of voting, i.e., the set of winners will typically be large (Brandt and Brill, 2011).

There are a number of ways that have been suggested to *circumvent* the impossibility flagged by Theorem 5. The first approach is to *restrict the domain* (Gaertner, 2006). If some profiles can be assumed to never occur, then certain impossibilities will turn into possibilities. The best known example are Black's *single-peaked preferences* (Black, 1958). A preference profile is called single-peaked if there exists an ordering $\gg$ on $\mathcal{X}$ (e.g., reflecting the left-right spectrum of political parties) such that any individual prefers x to y if x is between (with respect to $\gg$) y and her most preferred alternative. If an electorate has single-peaked preferences, then there do exist attractive SCFs that are strategy-proof, e.g., Black's *median voter rule*, which asks each individual for her top choice and then elects the alternative proposed by the voter representing the median with respect to $\gg$.

Another approach, initiated by Bartholdi et al. (1989), has been to look for voting rules for which it is computationally intractable (i.e., NP-hard) to decide *how* to manipulate, even if an individual does have all the necessary information to do so. This indeed turned out to be possible for a select number of voting rules, even though more recent work suggests that worst-case notions such as NP-hardness fail to provide effective protection. Faliszewski et al. (2010) review the state of the art of using complexity as a barrier against manipulation. Given the close links between complexity theory and mathematical logic, this line of work provides yet another connection between logic and social choice theory.

A third approach consists in investigating variants of the standard formal framework of social choice theory, by altering the notion of preference, the notion of ballot, or both. For instance, the widely used rule of *approval voting* (Brams and Fishburn, 1978), in which you vote by nominating any number of "acceptable" candidates and the candidates receiving the most nominations win, does in fact not fit the framework in which the Gibbard-Satterthwaite Theorem is stated (because ballots are not linear orders). In the context of approval voting, and more generally, in the context of social choice mechanisms where the language used to express (declared) ballots does not coincide with the language used to model (true) preferences, positive results concerning incentives to vote *sincerely* are achievable (Endriss, 2007; Endriss et al., 2009). Finally, if we model preferences and ballots as *utility functions* rather than as binary relations, and thus increase their informational content, strategy-proofness will become feasible under certain assumptions. This is the point where social choice theory meets game-theoretical *mechanism design* (see, e.g., Nisan, 2007).

3 Logics for Social Choice Theory

Classical social choice theory, as presented in Section 2, is a mathematically *rigorous* but not a *formal* enterprise. For instance, the axioms proposed are not themselves expressed in a formal language, such as the language of classical first-order logic, endowed with a well-defined syntax and semantics. Indeed, in most work in social choice theory, references to "logic" are merely intended as references to the rigorous use of the axiomatic method.[5] However, in recent years we have witnessed more and more contributions that make use of logic in a more direct sense of that term. In particular, there have been several proposals for casting parts of the framework of social choice theory in a logical language. In this section, we review some of these proposals. Before we proceed, however, we should first ask ourselves *why* one would want to express problems of social choice in logic. One argument is surely that doing so will help us gain a deeper understanding of the domain we are formalising. This may be said in defence of any exercise in formalisation. But there are at least two further very good reasons, which we shall briefly elaborate on here.

The first reason is that formalisation is a necessary step towards automation. Just as logic has long been used in computer science to specify and automatically verify the properties of software and hardware systems, logic may also prove useful to formally specify and check the properties of procedures of social choice. This vision has been articulated in the *social software* programme of Parikh (2002).

The second reason is that, once we confine ourselves to expressing axioms in a formal language, we are able to compare the expressive power required to formulate different results in social choice theory. This point has first been made by Pauly (2008), who argued that besides its normative appropriateness and logical strength, a further important quality of an axiom is the richness of the language required to express it. For instance, it is natural to ask whether a given framework of social choice theory is definable in classical first-order logic, or whether there are certain inherently higher-order features that cannot be simplified. This perspective is related to the fact that some axioms are intuitively simpler (and thus less contestable) than others. For instance, an inter-profile axiom such as IIA (making reference to both the actual profile under consideration and a counterfactual other profile we are comparing to) is conceptually much more complex than the Pareto condition, which merely prescribes what to do if the profile under consideration meets certain requirements.

In the sequel, we review some of the proposals for using logic to model social choice that have been advanced in the literature. We concentrate on the problem of modelling the Arrovian framework of SWFs, which has received by far the most attention. In our review,

[5] Even the monograph of Murakami (1968), despite bearing the name "*Logic and Social Choice*", is essentially an exposition of and an investigation into the then, in 1968, still nascent axiomatic method. Murakami's work does however include one genuine application of formal logic: He considers voting with abstention for the case of two alternatives x and y, i.e., there are three possible inputs an individual may supply ($x \succ y$, $y \succ x$, abstention) and three corresponding outcomes (x wins, y wins, tie). This suggests an interpretation in three-valued logic and Murakami shows, by means of reference to a functional completeness result in three-valued logic, that any voting rule for this domain can be defined in terms of some nesting of the plurality rule for subelectorates of varying size, a "negation operator", and the set of constant rules.

we distinguish between work on logics that have been specifically designed for this purpose and work exploring possibilities for representing concepts from social choice theory within the confines of standard logical frameworks. Where applicable, we also comment on attempts at using the formalisation proposed as a basis for automating tasks in social choice theory.

3.1 Designing Logics for Modelling Social Choice

One approach is to design a logic specifically for our needs, either from scratch or by adapting an existing logical framework. This approach has the advantage that we can tailor our logic precisely to our requirements. A disadvantage of this approach is that we cannot directly draw on existing results for that logic, including algorithms for reasoning tasks such as model checking or satisfiability checking. For a new logic it will also be more difficult to interpret results stating that a given social choice problem is expressible in that logic, as it may be unclear how the expressive power of that logic relates to that of more established logics.

To date there have been relatively few examples for logics specifically designed to model social choice. Remarkably, most (if not all) contributions of this kind make use of the general framework of modal logic. To be sure, there has been substantial work on modal logics that can express concepts from *game theory*, both the notions of *strategy* and *preference* that are central to noncooperative game theory and the notion of *coalition* that is central to cooperative game theory. We will not review this literature on *logic and games* here, and instead only point to a few key references. An important early contribution has been Parikh's *game logic* (Parikh, 1985), which extends propositional dynamic logic and can express that in a given game a given player has a strategy to bring about a state in which a given formula φ is true. The *coalition logic* of Pauly (2002) is centred around the concept of a coalition of individuals having the power to move to a state in which a given formula φ holds. An early contribution to the study of games using *dynamic epistemic logic* is the work of Van Benthem (2001). For a review of and commentary on the field of logic and games we refer to Van der Hoek and Pauly (2006) as well as to Van Benthem et al. (2011).

While concepts such as the strategies of an individual or the power of a coalition, which are widely studied in that literature, are also important concepts for social choice theory, the works cited above do not focus on the explicit modelling of the *mechanism* used to implement a social choice. Instead, the focus has been on individuals (and sometimes coalitions of individuals) and how they act in a given environment, which is appropriate in game theory, but less so in social choice theory, where we want to take the perspective of the mechanism designer. Two works that have taken this perspective are those of Ågotnes et al. (2011) and Troquard et al. (2011). We shall review the former in some detail.

Ågotnes et al. (2011) define a modal logic for reasoning about SWFs.[6] The language of this logic is parametric in $\mathcal{N}$ (individuals) and $\mathcal{X}$ (alternatives), i.e., any formula in this logic will only make statements about SWFs for the particular number of individuals and alternatives chosen. The set of *states* is the Cartesian product of the set of possible profiles and the set of pairs of alternatives. A *model* is defined by such a set of states (i.e., by $\mathcal{N}$ and $\mathcal{X}$) and by a SWF F. There are three types of atomic propositions: p_i for $i \in \mathcal{N}$ is true in state $(\boldsymbol{R}, (x, y))$ if individual i ranks $x \succ y$ under profile $\boldsymbol{R}$; $q_{(x,y)}$ for $x, y \in \mathcal{X}$ is true in any state the second component of which is (x, y); and the special proposition σ is true in state $(\boldsymbol{R}, (x, y))$ if $(x, y) \in F(\boldsymbol{R})$, i.e., if the collective preference under $\boldsymbol{R}$ will rank $x \succ y$. There are two (universal) modal operators: $[\textsc{prof}]\varphi$ is true in a state $(\boldsymbol{R}, (x, y))$ if φ is true in state $(\boldsymbol{R}', (x, y))$ for every profile $\boldsymbol{R}'$; $[\textsc{pair}]\varphi$ is true in state $(\boldsymbol{R}, (x, y))$ if φ is true in state $(\boldsymbol{R}, (x', y'))$ for every pair of alternatives (x', y'). We are now able to express interesting properties of the SWF F (which determines the valuation of σ). For instance, F is *dictatorial* if and only if the following formula is true in every state of the model:

$$\textsc{dictatorial} \quad := \quad \bigvee_{i \in \mathcal{N}} [\textsc{prof}][\textsc{pair}](p_i \leftrightarrow \sigma)$$

The formula expresses that there exists an individual i (the dictator) such that, to whichever state $(\boldsymbol{R}, (x, y))$ we move in terms of the profile (by application of $[\textsc{prof}]$) and the pair of alternatives under consideration (by application of $[\textsc{pair}]$), it will be the case that the collective preference will rank $x \succ y$ (i.e., σ will be true) if and only if individual i ranks $x \succ y$ (i.e., p_i is true). Another example is the Pareto condition:

$$\textsc{pareto} \quad := \quad [\textsc{prof}][\textsc{pair}](p_1 \wedge \cdots \wedge p_n \rightarrow \sigma)$$

That is, in every state $(\boldsymbol{R}, (x, y))$ it must be the case that, whenever all individuals rank $x \succ y$ (i.e., all p_i are true), then also the collective preference will rank $x \succ y$ (i.e., σ is true). Ågotnes et al. (2011) show how to express Arrow's Theorem in this manner and give a sound and complete axiomatisation of their logic. It follows that the theorem is, in principle, derivable in the logic, even if in practice obtaining such a derivation may prove difficult.

A limitation of this logic is that we need to fix $\mathcal{N}$ before we can start writing down formulas. While Arrow's Theorem states that it is impossible to find a suitable SWF for *any finite* set $\mathcal{N}$, the variant of the theorem that is formalised here applies only to the particular set $\mathcal{N}$ chosen. That is, the theorem that has been formalised is actually weaker than the original theorem. In particular, the fact that Arrow's Theorem ceases to hold when we move to an infinite electorate cannot be modelled in this logic. (Similar remarks apply to the fact that we have to fix $\mathcal{X}$ before fixing the language of the logic, but this point is somewhat

[6]The following description of the logic introduced by Ågotnes et al. (2011) assumes some familiarity with basic modal logic (see, e.g., Blackburn et al., 2001).

less critical, as the formalisation of Arrow's Theorem given by Ågotnes et al. (2011) does in fact not involve any propositions of the form $q_{(x,y)}$.) As we shall see, modelling the Arrovian framework in a suitable logic without fixing the set of individuals in the language is a hard challenge in the field of logics for social choice theory, which to date has not been solved in a truly satisfactory manner. A further challenge is the fact that Arrow's Theorem requires that F must be defined on every possible combination of preferences (the so-called *universal domain* assumption). In our exposition in Section 2 this assumption has been implicit in the definition of a SWF, but any logic for modelling the Arrovian framework must account for it explicitly. Ågotnes et al. (2011) deal with this issue by including in their axiomatisation a large axiom that explicitly states for every possible profile that it must be part of every model. Again, to date no truly elegant and satisfactory approach to address this challenge has been put forward.

3.2 Embedding Social Choice Theory into Existing Logical Frameworks

Besides designing a new logic for social choice theory, a second approach is to attempt to embed the part of social choice theory we want to model into an existing and well-understood logic. This has the advantage that we can rely on known results and existing tools for that logic. We will review three such approaches, using first-order logic, propositional logic, and higher-order logics, respectively.

Let us begin by reviewing a proposal for embedding the Arrovian framework of SWFs into classical first-order logic (Grandi and Endriss, 2009). First-order logic is clearly a good language for talking about binary relations in general and linear orders in particular. What is less clear is whether it is a good language for modelling properties such as IIA. The formulation of IIA includes an implicit universal quantification over preference profiles, and those profiles themselves consist of several linear orders. This may suggest that simple quantification over plain objects, and thus first-order logic, is not enough to model this domain. As it turns out, these difficulties can be overcome. The central idea, due to Tang and Lin (2009), is to introduce the concept of a *situation*, a name referring to a profile that we can quantify over rather than quantifying over profiles directly.

Let N (for individuals), X (for alternatives), and S (for situations) be 1-place predicates. Furthermore, let p be a 4-place predicate (to model individual preferences) and let w be a 3-place predicate (to model the collective preference): $p(z, x, y, u)$ says that individual z ranks x above y under the profile associated with situation u, while $w(x, y, u)$ says that the collective preference order ranks x above y under the profile associated with situation u. In first-order logic, we can easily specify that all preferences should be linear orders. For instance, the following formula expresses that the individual preferences must be transitive:

$$\forall z.\forall x_1.\forall x_2.\forall x_3.\forall u.[N(z) \wedge X(x_1) \wedge X(x_2) \wedge X(x_3) \wedge S(u) \ \rightarrow \\ (p(z,x_1,x_2,u) \wedge p(z,x_2,x_3,u) \rightarrow p(z,x_1,x_3,u))]$$

That is, for any z that represents an individual, any x_1, x_2, x_3 representing alternatives, and any u representing a situation, if individual z ranks x_1 above x_2 in the profile associated with u, as well as x_2 above x_3, then she will also rank x_1 above x_3 in the same situation. Other simple formulas express that every object must belong to exactly one of the three available types (individual, alternative, situation) and that any two profiles associated with distinct situations differ in at least one preference judgment. The only difficulty consists in ensuring that there exists a situation for every possible preference profile (the *universal domain* assumption). This, unfortunately, requires a rather cumbersome formula that stipulates that for any individual z, alternatives x and y, and situation u with $p(z,x,y,u)$, there also must exist a situation v associated with the profile we obtain when we swap x and y in the preference order of z in the profile associated with u, but leave everything else the same.

Any model of these formulas corresponds to a SWF. We can now specify further formulas to express properties of SWFs. For instance, Arrow's IIA can be expressed as follows:

$$\forall u_1.\forall u_2.\forall x.\forall y.[S(u_1) \wedge S(u_2) \wedge X(x) \wedge X(y) \ \rightarrow \\ [\forall z.(N(z) \rightarrow (p(z,x,y,u_1) \leftrightarrow p(z,x,y,u_2))) \ \rightarrow \ (w(x,y,u_1) \leftrightarrow w(x,y,u_2))]]$$

That is, for any two situations u_1 and u_2 and for any two alternatives x and y, if each individual z makes the same judgment regarding x and y in u_1 and u_2, then also the collective preference order should agree on the judgment regarding x and y in u_1 and u_2. Arrow's Theorem now reduces to a statement saying that a particular set of first-order formulas does not have a finite model (Grandi and Endriss, 2009). It is *not* possible, using this approach, to reduce Arrow's Theorem to (the validity of) a first-order formula. The reason is that in first-order logic we cannot force models to be finite (and, as we have seen, Arrow's Theorem only holds when the set of individuals in finite). However, for any fixed finite set of individuals we can easily write a corresponding first-order formula modelling Arrow's Theorem. Under this restriction it should also, in principle, be possible to derive a proof of the theorem using an automated theorem proved for first-order logic, even if to date no such proof has been realised in practice.

In fact, for a fixed set of individuals and a fixed set of alternatives we also obtain a fixed set of profiles and we can "ground" above first-order representation, by replacing every universal quantification with a conjunction and every existential quantification with a disjunction, thereby obtaining a representation in propositional logic. Tang and Lin (2009) explain how to generate such a representation for the case of two individuals and three alternatives (the smallest nontrivial instance of Arrow's Theorem), consisting of 106354 clauses in propositional logic, using a simple computer program. They were able to verify that this set of clauses is indeed unsatisfiable by using a satisfiability solver (a state-of-the-art SAT-solver requires less than one second to verify this). This is a significant result and the first fully automated proof of an instance of Arrow's Theorem. Tang and Lin also prove two inductive lemmas (by hand), one showing that, if there exists a SWF meeting Arrow's conditions for n individuals and m alternatives, then there also exists such a SWF for $n + 1$ individuals and m alternatives, and the other showing that then there also exists such a SWF for n alternatives and $m + 1$ alternatives. Together with the "base case" proved using the SAT-solver this yields the full theorem, for any finite $\mathcal{N}$ and any finite $\mathcal{X}$. A possible criticism of this approach is that the (manual) proofs of the two lemmas are conceptually not much simpler than a manual proof of the full theorem. A very attractive feature of the approach, on the other hand, is that it can be used to quickly check whether a conjectured theorem does hold for a fixed small number of individuals and alternatives. That is, beyond verification of known results, this approach may also prove useful for the *discovery* of new theorems.[7]

A third approach has been followed by Nipkow (2009) and Wiedijk (2007), who have formally verified two of the proofs of Arrow's Theorem given by Geanakoplos (2005) using the interactive theorem provers Isabelle and Mizar, respectively, which are based on higher-order logics. The condition of the finiteness of $\mathcal{N}$ is easily expressible in such a language, which means that the theorem can be stated in its full generality. It is important to stress that these are formalisations and verifications of *proofs* rather than of a theorem. That is, the person carrying out the verification has to code an existing proof step by step; the interactive theorem prover can be used to verify these steps, but it does not derive a proof of the full theorem automatically.

In summary, all existing approaches to modelling the Arrovian framework do have some drawbacks. First, most of them require us to fix the set of individuals and alternatives up front, as their names form part of the language defined, and in any case the less expressive logics considered are not powerful enough to express the finiteness condition on the set of individuals. Second, expressing the *universal domain* assumption in an elegant manner is typically difficult. Third, a fully automated proof of Arrow's Theorem is still pending, even if several partial attempts have already succeeded.

[7]In an area of social choice theory different from preference aggregation, namely the area concerned with extending a preference order defined over some objects to a preference order defined over sets of such objects (Barberà et al., 2004), a refinement of the approach of Tang and Lin (2009) has recently led to the fully automated discovery of a number of genuinely new impossibility theorems (Geist and Endriss, 2011).

4 Social Choice in Combinatorial Domains

In practise, many collective decision making problems have a *combinatorial structure*. Examples include referenda, where voters are asked whether they wish to accept each of a number of proposals, or electing a committee of officials (rather than a single official), where we have to decide which of the candidates standing should receive a seat on the committee. If there are m proposals included in a referendum, each of which may be either accepted or rejected, then there are 2^m possible outcomes. If there are m seats to be filled on our committee and there are m' candidates standing, then there are $\binom{m'}{m}$ possible committees. In either case, the number of actual alternatives that our voters need to consider is (at least) exponential in m. These problems have been studied in political science for some time, but the insight that this is an inherently computational challenge that should be addressed with the tools of computer science is relatively recent (Lang, 2004). Today, social choice in combinatorial domains is one of the core topics studied in computational social choice (Chevaleyre et al., 2008). In this section, we review some of the work in the area.

For ease of exposition, we restrict attention to *binary* combinatorial domains, where each issue has to take one of two values: 1 ("yes") or 0 ("no"). Let $\mathcal{I}$ be a finite set of such *binary issues*. Each issue $k \in \mathcal{I}$ is associated with a *variable* X_k that may take either value 1 or 0. We will sometimes write x_k as a shorthand for $X_k = 1$ and $\bar{x}_k$ as a shorthand for $X_k = 0$. An assignment of values to all variables is a (combinatorial) *alternative*, i.e, $\mathcal{X}$ now has the form $\{0,1\}^{\mathcal{I}}$. By a slight abuse of notation, we will sometimes think of the X_k as propositional variables and use the corresponding propositional language to express properties of combinatorial alternatives. If we think of combinatorial alternatives as truth assignments, then such a formula may or may not be satisfied by a given alternative. For example, $(1,0,0)$ satisfies the formula $X_1 \wedge X_2 \rightarrow X_3$, while $(1,1,0)$ does not.

Example 4 (Paradox of Multiple Elections)**.** Suppose 13 voters are asked to vote on three binary issues. Our voters might be the members of a small city council, and the three issues might represent whether or not to fund a new museum, a new school, or a new metro line. Suppose we ask each voter for her most preferred outcome:

- 3 voters each support $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$.
- 1 voter each supports $(1,1,1)$, $(1,1,0)$, $(1,0,1)$, and $(0,1,1)$.

Now, if we decide *issue-by-issue* and use the simple majority rule for each issue, then the outcome will be $(0,0,0)$, with a 7 to 6 majority on each issue, even though this is the only combinatorial alternative that has not been voted for by a single voter. This is known as a *Paradox of Multiple Elections*. In this particular form, it is due to Brams et al. (1998). □

Why do we consider Example 4 a paradox? In fact, it might very well be that a voter's happiness is proportional to the number of issues on which her choice coincides with the collective choice, in which case $(0, 0, 0)$ is not an unreasonable outcome at all. Logic can help us to to give a precise definition of what we mean by "paradox" (Grandi and Endriss, 2011). Let us see how. Example 4 considered an aggregation procedure F mapping a profile of combinatorial alternatives, one for each individual, to a single combinatorial alternative. That is, in this context a *ballot* is simply a combinatorial alternative. Recall that we can use the propositional language defined over $\{X_k \mid k \in \mathcal{I}\}$ to describe properties of combinatorial alternatives. Now think of a formula in that language as an *integrity constraint.* For a given integrity constraint γ, we call a combinatorial alternative *rational* if it satisfies γ. Now we can define a *paradox* as a triple consisting of an aggregation procedure F, a profile of ballots $\boldsymbol{B}$, and an integrity constraint γ, where every individual ballot B_i in $\boldsymbol{B}$ does satisfy γ, but $F(\boldsymbol{B})$ does not. That is, even though all the individual ballots are rational, the outcome of the election is not. Under this definition, Example 4 qualifies as a paradox if we choose the integrity constraint $X_1 \vee X_2 \vee X_3$.

How can we avoid this kind of paradox? One approach to consider is to *directly vote on combinatorial alternatives.* If we apply the plurality rule to the data of Example 4, then $(1, 0, 0)$, $(0, 1, 0)$, and $(0, 0, 1)$ receive 3 points each and we need to pick a winner by means of a tie-breaking rule. Whichever way we break the tie, the paradox will be avoided, and in this highly symmetric example any of the three front-runners seems an equally deserving winner. In general, however, this is *not* a good approach, as it does rely excessively on the tie-breaking rule. For example, if we assume that every profile is equally likely to occur, then for 10 issues and 20 individuals, the probability that no combinatorial alternative receives more than a single vote is roughly 83% (because $2^{10} \times \cdots \times (2^{10} - 19)/(2^{10})^{20} \approx 0.83$).

This problem could be overcome by using other voting rules than the plurality rule, namely rules that elicit more information from the voters. Most voting rules considered in the literature require each voter to provide a complete ranking of all alternatives (Brams and Fishburn, 2002). A well-known example is the *Borda rule*, proposed by the French engineer and political scientist Jean-Charles de Borda (1733–1799), under which an alternative receives $|\mathcal{X}| - 1$ points from every voter who ranks her first, $|\mathcal{X}| - 2$ points from every voter who ranks her second, and so forth. But for, say, 10 issues we would have to ask each voter to rank a total of $2^{10} = 1024$ combinatorial alternatives, which is not a realistic requirement.

Thus, if we want to avoid paradoxical outcomes, then we have to face serious challenges of a computational nature. A central point here is the fact that simply asking an individual to report her preferences can already become a challenging issue. Therefore, before reconsidering the full problem of social choice in combinatorial domains, we first introduce a number of approaches to modelling preferences in combinatorial domains.

4.1 Languages for Compact Preference Representation

A compact preference representation language is a formal language that can be used to express a given class of preference structures and that, usually, does require significantly less space to do so than an explicit representation. It is important to note that we cannot expect a magic solution. For example, suppose there are m binary issues and we merely want to represent which combinatorial alternatives an individual finds acceptable and which she does not find acceptable, i.e., we want to express a dichotomous preference structure on 2^m alternatives. There are 2^{2^m} such structures. Hence, even the most sophisticated representation language will have to use 2^m bits to be able to encode every possible structure. That is, we should look for representation languages that can represent those preferences that we are likely to encounter in a compact manner, but the same language may require more space on other preferences, or it may not be able to express all possible preferences. The study of compact preference representation languages is part of *knowledge representation and reasoning*, one of the major research areas in artificial intelligence (Goldsmith and Junker, 2008).

The most widely used compact preference representation language in computational social choice are *conditional preference networks*, or CP-nets for short (Boutilier et al., 2004). A CP-net consists of a directed graph, the nodes of which are the issues, and one so-called CP-table for each issue. The CP-table for issue $k \in \mathcal{I}$ specifies for each possible assignment of values to the variables associated with the parents of k a preference relation on the possible values for X_k. Each CP-table induces a partial order: a combinatorial alternative is preferred to another combinatorial alternative from which it differs only in one issue, if the pair matches one of the entries of the table. The preference relation induced by a CP-net is the transitive closure of the relation induced by its CP-tables.

Example 5 (CP-net)**.** The following is a CP-net for three binary issues, represented by the variables X, Y, and Z. Recall that we write x to say that X takes value 1 and $\bar{x}$ to say that X takes value 0, etc. Each variable is annotated with its CP-table:

$$\begin{tikzcd} X \arrow[r] \arrow[rr, bend left] & Y \arrow[r] & Z \\ \boxed{x \succ \bar{x}} & \boxed{\begin{array}{c} x : y \succ \bar{y} \\ \bar{x} : \bar{y} \succ y \end{array}} & \boxed{\begin{array}{c} xy : z \succ \bar{z} \\ x\bar{y} : z \succ \bar{z} \\ \bar{x}y : z \succ \bar{z} \\ \bar{x}\bar{y} : \bar{z} \succ z \end{array}} \end{tikzcd}$$

This CP-net generates the following partial order (with $\rightarrow$ representing $\succ$):

$$\begin{array}{ccccc} & \nearrow & x\bar{y}z & \searrow & \\ xyz & & & & x\bar{y}\bar{z} \rightarrow \bar{x}\bar{y}\bar{z} \rightarrow \bar{x}\bar{y}z \rightarrow \bar{x}yz \rightarrow \bar{x}y\bar{z} \\ & \searrow & xy\bar{z} & \nearrow & \end{array}$$

For instance, $xy\bar{z} \succ x\bar{y}\bar{z}$, which is another way of saying $(1,1,0) \succ (1,0,0)$, follows from the first entry in the CP-table for Y. Observe that we cannot rank $x\bar{y}z$ and $xy\bar{z}$. □

CP-nets can express a large class of partial orders, albeit not all of them. If the number of variables on which any single variable may depend is relatively small, which is a reasonable assumption to make in many problem domains, then a representation in terms of a CP-net will be relatively compact.

Another important family of preference representation languages is based on the idea that we can use the propositional language over variables introduced earlier to express *goals*. For example, $X_1 \vee X_2$ expresses the goal of accepting at least one of the first two issues. If we assign *weights* or *priorities* to these goals, then we can represent a wide range of preferences in this manner. For instance, a set of weighted goals defines a utility function by stipulating that the utility of a combinatorial alternative is the sum of the weights of the goals that are satisfied by that alternative. Similarly, a set of goals labelled with priority levels may be interpreted in a lexicographic manner by stipulating that one combinatorial alternative is preferred to another if there exists a priority level ℓ such that for each level of higher priority both alternatives satisfy the same number of goals and for level ℓ the first alternative satisfies more goals than the second. Forms of aggregation other than taking sums or using a lexicographic order have also been proposed and analysed (see, e.g., Lang, 2004; Uckelman and Endriss, 2010). This approach goes back to the work on *penalty logic* by Pinkas (1995); its relevance to preference representation in the context of social choice has first been recognised by Lafage and Lang (2000). For a discussion of the expressive power and relative succinctness of these languages, refer to the work of Coste-Marquis et al. (2004) for prioritised goals and Uckelman et al. (2009) for weighted goals. Note that classical propositional logic is not the only logic of interest for specifying goals. For instance, in a domain where variables specify whether or not a given individual will or will not receive a given resource, and where there are multiple identical copies of each resource available, weighed goals expressed in *linear logic* make for a useful preference representation language (Porello and Endriss, 2010).

4.2 Possible Approaches to Social Choice in Combinatorial Domains

The first approach to social choice in combinatorial domains we want to discuss here is known as *combinatorial vote* (Lang, 2004). The idea is to ask individuals to express their preferences in terms of a given compact preference representation language and to apply our voting rule of choice to these representations.

Example 6 (Borda Rule and Prioritised Goals)**.** Suppose three voters have to decide on two binary issues, associated with variables X and Y. We ask them to express their preferences as a set of prioritised goals (1 indicates high priority and 0 indicates normal priority):

Voter 1:	$\{X{:}1,\ Y{:}0\}$
Voter 2:	$\{X \vee \neg Y{:}0\}$
Voter 3:	$\{\neg X{:}0,\ Y{:}0\}$

Under a lexicographic interpretation, these goals induce the following weak orders:

Voter 1:	$xy \succ x\bar{y} \succ \bar{x}y \succ \bar{x}\bar{y}$
Voter 2:	$x\bar{y} \sim xy \sim \bar{x}\bar{y} \succ \bar{x}y$
Voter 3:	$\bar{x}y \succ \bar{x}\bar{y} \sim xy \succ x\bar{y}$

For voter 1, for instance, it is most important that $X = 1$, and getting $Y = 1$ has secondary importance, while voter 3 has two equally important goals. For voter 2, the three alternatives that satisfy at least one of $X = 1$ and $Y = 0$ are equally good and all better than the fourth alternative. Now, suppose we want to apply the Borda rule to elect a winning alternative. The standard Borda rule is defined for linear orders only, so we cannot use it here. There is however a natural generalisation: an alternative obtains as many points from a voter as there are other alternatives it dominates in the preference order reported by that voter. Alternative xy, for instance, obtains 3 points from voter 1, and 1 point each from voters 2 and 3. Let us summarise the points each combinatorial alternative will receive:

Alternative xy:	$3 + 1 + 1 = 5$
Alternative $x\bar{y}$:	$2 + 1 + 0 = 3$
Alternative $\bar{x}y$:	$1 + 0 + 3 = 4$
Alternative $\bar{x}\bar{y}$:	$0 + 1 + 1 = 2$

Thus, alternative xy wins the election. □

Example 6 is an illustration of the combinatorial vote approach—*except* for the fact that we have unravelled the preferences provided in terms of the compact representation language into an explicit representation before applying the voting rule. A full implementation of the ideal of combinatorial vote would require an algorithm for computing the Borda winner that can operate *directly* on preferences represented in terms of sets of prioritised goals. To date, no such algorithms are available. In fact, while the approach is promising, besides a host of work on compact preference representation languages that we can expect to eventually have a significant impact in the area of social in combinatorial domains, concrete results have so far been limited to complexity results pointing at some of the limitations of this approach. For instance, one of the basic results due to Lang (2004) is that even if each voter can specify only a single goal and even if we use the plurality rule (appropriately generalised so as to award a point to a combinatorial alternative for every voter whose goal it satisfies), deciding whether a given combinatorial alternative is an election winner is coNP-complete.

A second important approach to social choice in combinatorial domains is *sequential voting*. The basic idea is to vote on the issues in sequence and to publicly announce the decision on each issue as it is taken. For binary issues, in view of Theorem 3, the obvious rule to use for each "local election" is the simple majority rule. There is a growing literature devoted to sequential voting; here we only want to briefly mention two basic results. Both of these results involve notions closely related to the Condorcet Paradox, which we had introduced in Example 1. A *Condorcet winner* is an alternative that would win against any other alternative in a majority contest. Similarly, a *Condorcet loser* is an alternative that would lose against any other alternative in a majority contest. If a Condorcet winner does exist, then we would hope that it does get elected (recall that the Condorcet Paradox did demonstrate that a Condorcet winner need not exist though). If a Condorcet loser does exist, then we would hope that it does not get elected.[8]

One basic result in sequential voting, due to Lacy and Niou (2000), is that when all issues are binary and the simple majority rule is used in each local election (or, more generally, if each local voting rule has the property of never electing a local Condorcet loser), then we will never elect a combinatorial alternative that is a Condorcet loser. The proof is immediate: simply consider what happens in the final local election (at this stage, there will be two combinatorial alternatives left that might win; at most one of them can be the Condorcet loser; thus the assumption on the local rule will ensure that the Condorcet loser cannot win).

A second basic result, due to Lang and Xia (2009), applies to the case where voters express their preferences in terms of CP-nets. If the graphs underlying their CP-nets are all acyclic and if there exists an ordering of the issues that is compatible with all of them, then voting on the issues in that order using a local rule that will always elect a local Condorcet winner when it exists will ensure that whenever there exists a combinatorial alternative that is a global Condorcet winner, then that alternative will be elected.

In summary, if certain assumptions on the structure of preferences of the individuals are satisfied, then we can design good mechanisms for social choice in combinatorial domains. Solving significantly more general instances of the problem is, however, still out of reach. What makes this problem both interesting and challenging is the close interplay of computational and choice-theoretic concerns: on the one hand, we have to limit the amount of information handled by the mechanism so as to be able to manage the complexity of the problem, while, on the other, we have to limit the degree of uncertainty faced by the individuals to avoid paradoxical outcomes.

[8]Somewhat surprisingly, this basic property is violated by a number of standard voting rules, also outside of combinatorial domains. For example, the plurality rule will elect the Condorcet loser z when 2 voters report $x \succ y \succ z$, 2 voters report $y \succ x \succ z$, and 3 voters report $z \succ x \succ y$.

5 Judgment Aggregation

Not only does logic play an important role in the analysis of aggregation problems, but information expressed in terms of logic may itself be subject to aggregation. This kind of problem is studied in the field of *judgment aggregation* (List and Puppe, 2009). This section provides an introduction to judgment aggregation.

Example 7 (Doctrinal Paradox)**.** Suppose a court of three judges has to decide on a case in contract law. They are asked to judge whether the contract in question has been valid (p) and whether the contract in question has been breached (q). Legal doctrine dictates that the defendant be pronounced guilty if and only if both premises hold ($p \wedge q$). Judge 1 accepts both premises and the conclusion, while the other two judges each only accept one of the premises and thus are required to reject the conclusion:

	p	q	$p \wedge q$
Judge 1:	Yes	Yes	Yes
Judge 2:	Yes	No	No
Judge 3:	No	Yes	No

How should we aggregate this information to come to a collective judgment? If we believe the correct approach is to first decide on each of the premises by majority and to then infer the appropriate judgment for the conclusion ("premise-based procedure"), then we must accept p (as two out of three judges do), q (as, again, two out of three judges do), and $p \wedge q$ (because $p, q \models p \wedge q$). If, however, we believe that we should directly aggregate the individual judgments on the conclusion ("conclusion-based procedure"), then we must reject $p \wedge q$ (as a majority of the judges do). That is, two seemingly reasonable procedures return contrary results. This is known as the *Doctrinal Paradox* (Kornhauser and Sager, 1993). An alternative way of interpreting the paradox is the following: Each of the individual judges has specified a logically consistent set of formulas, namely $\{p, q, p \wedge q\}$, $\{p, \neg q, \neg(p \wedge q)\}$, and $\{\neg p, q, \neg(p \wedge q)\}$, while the collective judgment set we obtain when we decide on each formula by majority is inconsistent: $\{p, q, \neg(p \wedge q)\}$. $\square$

Kornhauser and Sager (1993) observed instances of the doctrinal paradox in their analysis of actual court cases and their work is a contribution to legal theory. Pettit (2001) discusses the philosophical implications of the paradox in view of the ideal of deliberative democracy. Here we focus on the technical side of the field instead.

Let us now define the formal framework of judgment aggregation (List and Pettit, 2002; Dietrich, 2006; List and Puppe, 2009). For any formula φ of propositional logic, let $\sim\!\varphi$ denote its *complement:* $\sim\!\varphi := \psi$ if $\varphi = \neg\psi$ and $\sim\!\varphi := \neg\varphi$ otherwise. An *agenda* Φ is a finite set of propositional formulas that is free of doubly-negated formulas and that is closed under complementation (i.e., $\sim\varphi \in \Phi$ whenever $\varphi \in \Phi$). A *judgment set* J for agenda Φ is a subset of Φ. J is called *complete* if $\varphi \in J$ or $\sim\varphi \in J$ for every formula $\varphi \in \Phi$; J is called *consistent* if $J \not\models \bot$. The set of all consistent and complete judgment sets for agenda Φ is denoted as $\mathcal{J}(\Phi)$. Now let $\mathcal{N} = \{i_1, \ldots, i_n\}$ be a finite set of (at least two) *individuals* (or *judges*, or *agents*) and suppose each individual $i \in \mathcal{N}$ provides a judgment set $J_i \in \mathcal{J}(\Phi)$. That is, we are given a *profile* $\boldsymbol{J} = (J_1, \ldots, J_n) \in \mathcal{J}(\Phi)^{\mathcal{N}}$ of judgment sets. Let $N_\varphi^{\boldsymbol{J}} := \{i \in \mathcal{N} \mid \varphi \in J_i\}$ denote the set of individuals that accept formula φ in profile $\boldsymbol{J}$. A (resolute) *judgment aggregation procedure* is a function $F : \mathcal{J}(\Phi)^{\mathcal{N}} \to 2^\Phi$ mapping any profile of complete and consistent judgment sets to a single collective judgment set. The latter need *not* be complete and consistent. For instance, as Example 7 has shown, if F is the majority rule, then the collective judgment set may fail to be consistent.

5.1 Axioms and Procedures

The axiomatic method extends to the framework of judgment aggregation. Several of the axioms that have been formulated for preference aggregation are naturally adapted to this framework. Let us briefly review some of them here:

- **Unanimity.** If all individuals accept a given formula, then so should society: if $\varphi \in J_1 \cap \cdots \cap J_n$ then $\varphi \in F(\boldsymbol{J})$.
- **Anonymity.** The aggregation procedure should be symmetric with respect to individuals: $F(J_1, \ldots, J_n) = F(J_{\pi(1)}, \ldots, J_{\pi(n)})$ for any permutation $\pi : \mathcal{N} \to \mathcal{N}$.
- **Neutrality.** If two formulas have the same pattern of individual acceptance in a profile, then both or neither should be accepted: if $N_\varphi^{\boldsymbol{J}} = N_\psi^{\boldsymbol{J}}$ then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$.
- **Independence.** If a formula has the same pattern of individual acceptance in two different profiles, then it should be accepted under both or neither of these two profiles: if $N_\varphi^{\boldsymbol{J}} = N_\varphi^{\boldsymbol{J}'}$ then $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \varphi \in F(\boldsymbol{J}')$.
- **Monotonicity.** If an accepted formula receives additional support, then it should still be accepted: if $\varphi \in J'_{i^\star} \setminus J_{i^\star}$ and $J_i = J'_i$ for all $i \neq i^\star$ then $\varphi \in F(\boldsymbol{J}) \Rightarrow \varphi \in F(\boldsymbol{J}')$.[9]

Let us now review a few concrete judgment aggregation procedures and see which axioms they satisfy. As the reader may easily verify, the *majority rule*, which accepts a formula φ if and only if a strict majority of the individuals do, satisfies *all* of the above axioms. But, as we have seen, the majority rule may return an inconsistent judgment set.

[9] An alternative (intra-profile) monotonicity axiom stipulates that when ψ is accepted by a strict superset of the individuals accepting φ, then ψ should be accepted by society whenever φ is (Endriss et al., 2010).

The *premise-based* and the *conclusion-based procedures*, which we have also mentioned already, first require us to declare which formulas in the agenda are to be treated as premises and which are to be treated as conclusions. This already points at one of their weaknesses, because there is no obvious criterion by which to make this distinction. The premise-based procedure first applies the majority rule to the premises and accepts a conclusion φ if φ is a logical consequence of the set of accepted premises. Suppose we declare all literals premises, and all other formulas conclusions. Then, if every atomic formula occurring anywhere within a complex formula in the agenda Φ does also occur as formula in Φ in its own right and if the number n of individuals is odd, then the premise-base procedure is easily seen to always return a complete and consistent judgment set. It also satisfies anonymity and all other axioms with respect to those formulas that are declared premises, but it might violate them with respect to the formulas that have been declared conclusions. The conclusion-based procedure, which accepts a conclusion whenever a strict majority of the individuals do and which does not pass judgment on any of the premises, is strictly speaking not a judgment aggregation procedure $F : \mathcal{J}(\Phi)^{\mathcal{N}} \to 2^{\Phi}$ meeting our formal definition.

Arguably the most attractive type of judgment aggregation procedure are *distance-based procedures* (Konieczny and Pino Pérez, 2002; Pigozzi, 2006; Miller and Osherson, 2009; Lang et al., 2011). The idea is the following. First, define a metric on judgment sets that, intuitively, specifies how distant two different judgment sets are. A common choice is the *Hamming distance* H, which is defined as $H(J, J') := \frac{1}{2} \cdot |(J \setminus J') \cup (J' \setminus J)|$, i.e., it counts on how many formulas J and J' differ. The distance-based procedure based on H then returns that complete and consistent judgment set that minimises the sum of the Hamming distances to the individual judgment sets. As there may be more than one such optimal judgment set, this is technically an *irresolute* procedure that may have to be combined with a tie-breaking rule. This procedure is unanimous and—depending on the tie-breaking rule chosen—can be made to be either anonymous or neutral, but it violates independence (and monotonicity). Observe that, whenever the collective judgment set produced by the majority rule is complete and consistent, then it will coincide with the outcome of the distance-based procedure (for any reasonable metric, including H). Miller and Osherson (2009) discuss several variants of the basic idea of distance-based judgment aggregation.

A *distance-based generalised dictatorship* is a procedure that works like a distance-based procedure, but where we choose from amongst the judgment sets *submitted by the individuals* (Grandi and Endriss, 2011). Despite its off-putting name, this is actually a very attractive procedure that has similar axiomatic properties as the distance-based procedure choosing from the set of *all* consistent and complete judgment sets, but that also has the additional advantage of having very low computational complexity.

5.2 An Impossibility Theorem and Further Directions of Research

As we have seen, one interpretation of the Doctrinal Paradox is the fact that the majority rule will not always return a judgment set that is complete and consistent. List and Pettit (2002) showed that this problem is not restricted to the majority rule. In fact, any judgment aggregation procedure that meets certain minimal requirements will sometimes fail to produce complete and consistent judgment sets.

Theorem 6 (List and Pettit, 2002). *No judgment aggregation procedure for an agenda Φ with $\{p, q, p \wedge q\} \subseteq \Phi$ that satisfies anonymity, neutrality, and independence will always return a collective judgment set that is complete and consistent.*

Proof. Observe that for any anonymous, neutral, and independent aggregation procedure F, collective acceptance of a formula will depend only on the number of individuals accepting it. In particular, $|N^{\boldsymbol{J}}_{\varphi}| = |N^{\boldsymbol{J}}_{\psi}|$ entails $\varphi \in F(\boldsymbol{J}) \Leftrightarrow \psi \in F(\boldsymbol{J})$. We distinguish two cases:

(1) Suppose the number of individuals n is *even*. Consider a profile $\boldsymbol{J}$ under which half of the individuals accept p and the other half accept $\neg p$, i.e., $|N^{\boldsymbol{J}}_{p}| = |N^{\boldsymbol{J}}_{\neg p}|$. Thus, $p \in F(\boldsymbol{J}) \Leftrightarrow \neg p \in F(\boldsymbol{J})$, meaning that the collective judgment set must accept either both of p and $\neg p$, or neither. But the former would violate consistency, while the latter would violate completeness.

(2) Suppose n is *odd* (and $n \geqslant 3$).[10] Consider a profile $\boldsymbol{J}$ under which $\frac{n-1}{2}$ individuals accept p and q, 1 individual accepts p and not q, 1 individual accepts q and not p, and the remaining $\frac{n-3}{2}$ individuals accept neither p nor q. Then $|N^{\boldsymbol{J}}_{p}| = |N^{\boldsymbol{J}}_{q}| = |N^{\boldsymbol{J}}_{\neg(p \wedge q)}|$. Hence, either all or none of p, q, and $\neg(p \wedge q)$ must be in $F(\boldsymbol{J})$. If the former is the case, then $F(\boldsymbol{J})$ is not consistent. If the latter is the case, then completeness would require that all of $\neg p$, $\neg q$, and $p \wedge q$ are in $F(\boldsymbol{J})$, which would again violate consistency.

Thus, for no number of individuals will we be able to devise an F satisfying all three axioms that always returns complete and consistent judgment sets. $\square$

Theorem 6 is the original impossibility theorem in the field of judgment aggregation. Since 2002 several other impossibilities have been established (List and Puppe, 2009). While List and Pettit's result applies to agendas with a specific structure, namely those that include at least two formulas and their conjunction, more recent results provide precise *characterisations* of the class of agendas for which it is possible to find an aggregation procedure that satisfies a given set of axioms and that will always return a collective judgment set that is complete and consistent. For instance, for the axioms of unanimity, neutrality, independence, and monotonicity the agendas Φ that are characterised in this manner are those for which any

[10] Recall that we assume that $n \geqslant 2$. Theorem 6 does not apply if there is just a single individual.

inconsistent subset of Φ does itself have an inconsistent subset of size at most 2 (Nehring and Puppe, 2007; List and Puppe, 2009). Recent work in computational social choice has furthermore studied the computational complexity of deciding whether we can guarantee consistent outcomes for a given agenda and a procedure meeting a given set of axioms. This problem of the *safety of the agenda* turns out to be intractable (and more precisely, Π_2^p-complete) for most interesting sets of axioms (Endriss et al., 2010).

The connections between the impossibilities arising in the context of preference aggregation and those arising in the context of judgment aggregation have been explored by a number of authors and the two frameworks have been shown to be closely related (Dietrich and List, 2007; Porello, 2010; Grossi, 2010; Grandi and Endriss, 2011), the most obvious (but by no means the only) link being the insight that we can think of preference statements such as "$x \succ y$" as judgments that may be true or false.

Finally, while originally associated with problems in legal reasoning and chiefly discussed in the philosophical literature, it is not hard to see that judgment aggregation can have a range of significant applications in other fields as well, e.g., in the Semantic Web, and more specifically the aggregation of knowledge distributed over a number of different ontologies (Porello and Endriss, 2011). These opportunities are yet to be explored in depth.

6 Conclusion and Further Reading

We have reviewed a number of classical and more recent contributions to the theory of social choice and argued that logic has played, and continues to play, and important role in its development as a scientific discipline. Naturally, given both this specific angle on the topic and the proverbial lack of space, our exposition had to remain incomplete. Social choice theory, and its cousin computational social choice, cover considerably more ground than has been possible to give justice to here.

It is not easy to predict what direction the field is going to take in the future, but the research questions raised in this chapter may be expected to play some role in its further development: What is the "right" logic to model social choice? How far can we push the automatisation of reasoning about problems in social choice theory? What is the best way of balancing complexity concerns and the need to limit uncertainty amongst decision makers when designing mechanisms for social choice in combinatorial domains? What can the methodology of judgment aggregation contribute to more general problems of information aggregation arising in applications such as ontology merging?

We conclude with a few pointers to the literature for readers who want to delve deeper into the subject. There are a number of excellent textbooks on the market that cover significant parts of classical social choice theory. The book by Gaertner (2006) is particularly broad in its coverage and highly accessible; the approach taken by Austen-Smith and Banks (1999) is somewhat more technical. Taylor (2005) focuses on strategic manipulation, while Moulin (1988) provides clear links to welfare economics and distributive justice. Like the present chapter, these books concentrate very much on the technical aspects of social choice theory rather than on the ethical or economic justifications for its assumptions or the philosophical and political ramifications of its results. The interested reader may want to consult the works of Roemer (1996) and Riker (1982) for pointers in these directions. Finally, the classics of Arrow (1963) and Sen (1970a) still make for instructive and fascinating reading today.

For the more recent developments described in Sections 3–5, no comprehensive exposition is available. However, both List (2011) and Grossi and Pigozzi (2011) provide good introductions to judgment aggregation; and a more detailed exposition of problems in social choice in combinatorial domains than we have given here may be found in the expository paper of Chevaleyre et al. (2008). For a wider discussion of questions in computational social choice, we refer to Chevaleyre et al. (2007).

Acknowledgements. I would like to acknowledge the numerous discussions with Umberto Grandi, Daniele Porello, Eric Pacuit and Jérôme Lang on logic and social choice theory, which greatly helped shape my views on the topic. I'm also very grateful for comments on an early draft of this chapter received from Johan van Benthem, Umberto Grandi and Daniele Porello.

References

T. Ågotnes, W. van der Hoek, and M. Wooldridge. On the logic of preference and judgment aggregation. *Autonomous Agents and Multiagent Systems*, 22(1):4–30, 2011.

K. J. Arrow. *Social Choice and Individual Values.* John Wiley and Sons, 2nd edition, 1963. First edition published in 1951.

K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare.* North-Holland, 2002.

D. Austen-Smith and J. S. Banks. *Positive Political Theory I: Collective Preference.* Universitu of Michigan Press, 1999.

S. Barberà. Pivotal voters: A new proof of Arrow's Theorem. *Economics Letters*, 6(1):13–16, 1980.

S. Barberà, W. Bossert, and P. Pattanaik. Ranking sets of objects. In S. Barberà, P. Hammond, and C. Seidl, editors, *Handbook of Utility Theory*, volume 2. Kluwer Academic Publishers, 2004.

J. J. Bartholdi, C. A. Tovey, and M. A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989.

J. van Benthem. Games in dynamic-epistemic logic. *Bulletin of Economic Research*, 53(4): 219–248, 2001.

J. van Benthem, E. Pacuit, and O. Roy. Toward a theory of play: A logical perspective on games and interaction. *Games*, 2(1):52–86, 2011.

D. Black. *The Theory of Committees and Elections*. Cambridge University Press, 1958.

P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge University Press, 2001.

C. Boutilier, R. I. Brafman, C. Domshlak, H. H. Hoos, and D. Pool. CP-nets: A tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.

S. J. Brams and P. C. Fishburn. Approval voting. *The American Political Science Review*, 72(3):831–847, 1978.

S. J. Brams and P. C. Fishburn. Voting procedures. In K. J. Arrow, A. K. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*. North-Holland, 2002.

S. J. Brams, M. D. Kilgour, and W. S. Zwicker. The paradox of multiple elections. *Social Choice and Welfare*, 15(2):211–236, 1998.

F. Brandt and M. Brill. Necessary and sufficient conditions for the strategyproofness of irresolute social choice functions. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2011)*, 2011.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. A short introduction to computational social choice. In *Proceedings of the 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM-2007)*. Springer-Verlag, 2007.

Y. Chevaleyre, U. Endriss, J. Lang, and N. Maudet. Preference handling in combinatorial domains: From AI to social choice. *AI Magazine*, 29(4):37–46, 2008.

S. Coste-Marquis, J. Lang, P. Liberatore, and P. Marquis. Expressive power and succinctness of propositional languages for preference representation. In *Proceedings of the 9th International Conference on Principles of Knowledge Representation and Reasoning (KR-2004)*, 2004.

T. R. Daniëls and E. Pacuit. A general approach to aggregations problems. *Journal of Logic and Computation*, 19(3):517–536, 2009.

B. A. Davey and H. A. Priestley. *Introduction to Lattices and Order*. Cambridge University Press, 2nd edition, 2002.

F. Dietrich. Judgment aggregation: (im)possibility theorems. *Journal of Economic Theory*, 126(1):286–298, 2006.

F. Dietrich and C. List. Arrow's Theorem in judgment aggregation. *Social Choice and Welfare*, 29(1):19–33, 2007.

J. Duggan and T. Schwartz. Strategic manipulation without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized. *Sociall Choice and Welfare*, 17(1):85–93, 2000.

U. Endriss. Vote manipulation in the presence of multiple sincere ballots. In *Proceedings of the 11th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2007)*, 2007.

U. Endriss, M. S. Pini, F. Rossi, and K. B. Venable. Preference aggregation over restricted ballot languages: Sincerity and strategy-proofness. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI-2009)*, 2009.

U. Endriss, U. Grandi, and D. Porello. Complexity of judgment aggregation: Safety of the agenda. In *Proceedings of the 9th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2010)*, 2010.

P. Faliszewski, E. Hemaspaandra, and L. Hemaspaandra. Using complexity to protect elections. *Communications of the ACM*, 53(11):74–82, 2010.

P. C. Fishburn. Arrow's Impossibility Theorem: Concise proof and infinite voters. *Journal of Economic Theory*, 2(1):103–106, 1970.

W. Gaertner. *A Primer in Social Choice Theory*. LSE Perspectives in Economic Analysis. Oxford University Press, 2006.

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

C. Geist and U. Endriss. Automated search for impossibility theorems in social choice theory: Ranking sets of objects. *Journal of Artificial Intelligence Research*, 40:143–174, 2011.

A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.

J. Goldsmith and U. Junker. Preference handling for artificial intelligence (editorial). *AI Magazine*, 29(4):9–12, 2008.

U. Grandi and U. Endriss. First-order logic formalisation of Arrow's Theorem. In *Proceedings of the 2nd International Workshop on Logic, Rationality and Interaction (LORI-2009)*. Springer-Verlag, 2009.

U. Grandi and U. Endriss. Binary aggregation with integrity constraints. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-2011)*, 2011.

D. Grossi. Correspondences in the theory of aggregation. In *Logic and the Foundations* *of Game and Decision Theory: Revised and Selected Papers from LOFT-2008*. Springer-Verlag, 2010.

D. Grossi and G. Pigozzi. Introduction to judgment aggregation. Lecture Notes, 23rd European Summer School in Logic, Language and Information (ESSLLI-2011), Ljubljana, 2011.

F. Herzberg and D. Eckert. Impossibility results for infinite-electorate abstract aggregation rules. *Journal of Philosophical Logic*, 2011. Forthcoming.

W. van der Hoek and M. Pauly. Modal logic for games and information. In P. Blackburn, J. van Benthem, and F. Wolter, editors, *Handbook of Modal Logic*. Elsevier Science, 2006.

A. P. Kirman and D. Sondermann. Arrow's Theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(3):267–277, 1972.

S. Konieczny and R. Pino Pérez. Merging information under constraints: A logical framework. *Journal of Logic and Computation*, 12(5):773–808, 2002.

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

D. Lacy and E. M. S. Niou. A problem with referendums. *Journal of Theoretical Politics*, 12(1):5–31, 2000.

C. Lafage and J. Lang. Logical representation of preferences for group decision making. In *Proceedings of the 7th International Conference on Principles of Knowledge Representation and Reasoning (KR-2000)*, 2000.

J. Lang. Logical preference representation and combinatorial vote. *Annals of Mathematics and Artificial Intelligence*, 42(1–3):37–71, 2004.

J. Lang and L. Xia. Sequential composition of voting rules in multi-issue domains. *Mathematical Social Sciences*, 57(3):304–324, 2009.

J. Lang, G. Pigozzi, M. Slavkovik, and L. van der Torre. Judgment aggregation rules based on minimization. In *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge (TARK-2011)*, 2011.

C. List. The theory of judgment aggregation: An introductory review. *Synthese*, 2011. Forthcoming.

C. List and P. Pettit. Aggregating sets of judgments: An impossibility result. *Economics and Philosophy*, 18(1):89–110, 2002.

C. List and C. Puppe. Judgment aggregation: A survey. In *Handbook of Rational and Social Choice*. Oxford University Press, 2009.

K. O. May. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20(4):680–684, 1952.

M. K. Miller and D. Osherson. Methods for distance-based judgment aggregation. *Social Choice and Welfare*, 32(4):575–601, 2009.

H. Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, 1988.

E. Muller and M. A. Satterthwaite. The equivalence of strong positive association and strategy-proofness. *Journal of Economic Theory*, 14(2):412–418, 1977.

Y. Murakami. *Logic and Social Choice*. Routledge, 1968.

K. Nehring and C. Puppe. The structure of strategy-proof social choice. Part I: General characterization and possibility results on median spaces. *Journal of Economic Theory*, 135(1):269–305, 2007.

T. Nipkow. Social choice theory in HOL: Arrow and Gibbard-Satterthwaite. *Journal of Automated Reasoning*, 43(3):289–304, 2009.

N. Nisan. Introduction to mechanism design (for computer scientists). In N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, editors, *Algorithmic Game Theory*. Cambridge University Press, 2007.

R. Parikh. The logic of games and its applications. In *Topics in the Theory of Computation*, volume 24 of *Annals of Discrete Mathematics*. North-Holland, 1985.

R. Parikh. Social software. *Synthese*, 132(3):187–211, 2002.

M. Pauly. A modal logic for coalitional power in games. *Journal of Logic and Computation*, 12(1):149–166, 2002.

M. Pauly. On the role of language in social choice theory. *Synthese*, 163(2):227–243, 2008.

P. Pettit. Deliberative democracy and the discursive dilemma. *Philosophical Issues*, 11(1): 268–299, 2001.

G. Pigozzi. Belief merging and the discursive dilemma: An argument-based account of paradoxes of judgment aggregation. *Synthese*, 152(2):285–298, 2006.

M. S. Pini, F. Rossi, K. B. Venable, and T. Walsh. Aggregating partially ordered preferences. *Journal of Logic and Computation*, 19(3):475–502, 2008.

G. Pinkas. Reasoning, nonmonotonicity and learning in connectionist networks that capture propositional knowledge. *Artificial Intelligence*, 77(2):203–247, 1995.

D. Porello. Ranking judgments in Arrow's setting. *Synthese*, 173(2):199–210, 2010.

D. Porello and U. Endriss. Modelling combinatorial auctions in linear logic. In *Proceedings* *of the 12th International Conference on the Principles of Knowledge Representation and Reasoning (KR-2010)*, 2010.

D. Porello and U. Endriss. Ontology merging as social choice. In *Proceedings of the 12th International Workshop on Computational Logic in Multiagent Systems (CLIMA-2011)*. Springer-Verlag, 2011.

W. H. Riker. *Liberalism against Populism: A Confrontation between the Theory of Democracy and the Theory of Social Choice*. Freeman, San Francisco, 1982.

J. E. Roemer. *Theories of Distributive Justice*. Harvard University Press, 1996.

M. A. Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

A. K. Sen. *Collective Choice and Social Welfare*. Holden-Day, San Francisco, 1970a.

A. K. Sen. The impossibility of a Paretian liberal. *The Journal of Political Economics*, 78 (1):152–157, 1970b.

A. K. Sen. Social choice theory. In K. J. Arrow and M. D. Intriligator, editors, *Handbook of Mathematical Economics*, volume 3. North-Holland, 1986.

P. Suppes. The pre-history of Kenneth Arrow's Social Choice and Individual Values. *Social Choice and Welfare*, 25(2–3):319–326, 2005.

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

A. D. Taylor. *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, 2005.

N. Troquard, W. van der Hoek, and M. Wooldridge. Reasoning about social choice functions. *Journal of Philosophical Logic*, 2011. Forthcoming.

J. Uckelman and U. Endriss. Compactly representing utility functions using weighted goals and the max aggregator. *Artificial Intelligence*, 174(15):1222–1246, 2010.

J. Uckelman, Y. Chevaleyre, U. Endriss, and J. Lang. Representing utility functions via weighted goals. *Mathematical Logic Quarterly*, 55(4):341–361, 2009.

F. Wiedijk. Arrow's Impossibility Theorem. *Formalized Mathematics*, 15(4):171–174, 2007.