What's in an axiom?
On the Nature of Axioms in Social Choice Theory*

Ulle Endriss and Marie Christin Schmidtlein

Institute for Logic, Language and Computation
University of Amsterdam

Working Paper
Feedback welcome: `ulle.endriss@uva.nl`
Version: September 15, 2025

Abstract

In social choice theory we routinely use axioms to rigorously specify normatively appealing properties of mechanisms for collective decision making. However, we rarely reflect on the nature of those axioms themselves, beyond the technical needs arising in the context of one specific project at a time. Here we do so and ponder a number of fundamental questions regarding the definition and use of axioms in social choice theory. This includes the question of how to formally capture the meaning of an axiom, how to measure its logical strength, how to characterise the range of situations an axiom is talking about, and how to put order into the space of axioms by classifying them in terms of the dimensionality of the constraints they impose. We explore these questions in the context of voting rules but argue that the observations made have relevance also beyond this specific context.

1 Introduction

What's in a name? That which we call a rose
By any other name would smell as sweet.

— Shakespeare, *Romeo and Juliet*

The axiomatic method is a core staple of the systematic and principled study of collective decision making. In the context of social choice theory, an *axiom* is a

*Parts of this work have previously been presented at the *15th OSGAD Seminar on Ordered Structures in Games and Decision* held in Paris in November 2022, the *Informatik Kolloquium* at TU Munich in March 2024, the *17th Meeting of the Society for Social Choice and Welfare* held in Paris in July 2024, the *Microeconomic Theory Workshop* at the Market Design Center (UTMD) of the University of Tokyo in March 2025, and the *20th European Meeting on Game Theory (SING)* held in June 2025 in Maastricht. We are grateful for the helpful feedback received from numerous individuals, including in particular Antoinette Baujard, Franz Dietrich, Kenzo Imamura, Michihiro Kandori, Klaus Nehring, Eric Pacuit, Dominik Peters, Marc Pirlot, Marcus Pivato, Clemens Puppe, Colin Rowat, William Thomson, and Koji Yokote.

property of mechanisms for collective decision making that someone has argued for on normative grounds and that they have specified in precise mathematical terms. In the words of William Thomson, an axiom is "the mathematical expression of some intuition we have about how a [mechanism] should behave in certain situations" (Thomson, 2023, p. 76). And in those of Charles Plott, it is "a type of *minimal* expectation about system performance" for any mechanism we might consider using (Plott, 1976, p. 520).[1] Well-known examples for axioms include *Anonymity*, postulating that all individuals involved in the decision-making process should be treated equally; *Pareto Efficiency*, postulating that mechanisms should not return outcomes for which there are alternative outcomes that all stakeholders would prefer; and *Strategyproofness*, postulating that mechanisms should not incentivise participants to lie about their true preferences.

Scholars working in the field frequently use axioms to identify the qualities of and to differentiate between mechanisms. Many of the field's most celebrated results, starting with the seminal work of Kenneth Arrow (1963) and Amartya Sen (2017), take the form of either *characterisation results* or *impossibility results*. A characterisation result establishes a given mechanism (or family of mechanisms) as the only one that satisfies a certain combination of axioms of interest, while an impossibility result shows that for certain combinations of axioms of interest there does not exist—and cannot exist—any mechanism that would satisfy all of them.

Arguably, the axiomatic method is the single most important methodological tool used by economic theorists working on questions of social choice (Arrow et al., 2002). Yet, systematic reflections on the nature of axioms and the axiomatic method itself are rare—though notable exceptions include the works of Fishburn (1973), Plott (1976), Mongin (2003), and Thomson (2001, 2023).

This paper is such a reflection on the axiomatic method in social choice theory. Our chief interest is in the *nature of the axioms themselves*. While in a typical paper employing the axiomatic method one might find a philosophical discussion of the normative merits of different axioms, followed by a mathematical analysis of their logical consequences, here we want to think of axioms as formal objects that need to be defined, that can be compared to one another, and that might be classified in a number of different ways. We shall be guided by four questions: What options are available to us when we need to formally specify an axiom's *meaning*? How can we quantify the *strength* of an axiom and compare it to that of another axiom? What is the *scope* of an axiom, i.e., what are the situations it *talks about*? And finally, what is the *dimensionality* of the constraints imposed by an axiom, i.e., how many data points would we need before we might infer that the axiom has been violated?

In the aforementioned treatises on the axiomatic method, the focus usually is a different one. To begin with, we find advice on how to use the axiomatic method

[1] We stress that the use of the term 'axiom' in social choice theory is different from its use in mathematical logic and the foundations of mathematics, even though the former certainly has been inspired by the latter. To the mathematical logician an axiom is a property that is 'obviously true', while to a social choice theorist it is a property that is 'obviously desirable'. In both disciplines, we must argue for any axiom we wish to introduce *outside of the formal system* we are working in, usually employing arguments of a philosophical nature. Only once we have settled on a given set of axioms can we explore their logical consequences, now using mathematical arguments.

and what typical mistakes to avoid. For instance, Thomson (2001) reminds us that an axiomatic study should be driven by the properties of mechanisms—i.e., by the axioms—we deem important rather than by the goal to find a mathematically neat characterisation of a specific mechanism. Plott (1976) argues that the findings turned up by careful applications of the axiomatic method should be taken seriously, going as far as to say that "our philosophical positions must be altered accordingly" (p. 553) when impossibility theorems show that certain postulates regarding collective decision making are mutually inconsistent. On the other hand, Mongin (2003) cautions against the over-interpretation of technical results.

We also find proposals for classifying axioms. For instance, Thomson (2001) argues that most axioms proposed in the literature belong to a small number of natural classes, such as axioms identifying efficiency requirements, symmetry requirements, or monotonicity requirements. This kind of classification relates to the purpose of an axiom. Fishburn (1973) instead classifies axioms in terms of their form, e.g., in view of whether they are formulated using universal or existential quantifiers, and whether they prescribe behaviour for the mechanism for one situation at a time or rather impose constraints on how outcomes for different situations should relate to one another. Thomson (2023) also highlights this latter distinction as an important characteristic of an axiom. Relatedly, some authors, such as Richelson (1977) or Sertel and Van der Bellen (1979), have attempted to systematically chart the logical relationships between large numbers of axioms pertaining to a specific problem domain, which might also serve as a tool for classification.

There also is a somewhat more recent strand of literature concerned with the formal representation of axioms in a suitable logical language (Endriss, 2011), which necessarily involves at least some reflection on the nature of axioms. One motivation for investigating what kind of formal language (or logic) can express what kind of axiomatic principle is philosophical in nature. In this context, Pauly (2008) argues that the expressive power required for a formal language to encode a given axiom tells us something relevant about that axiom—and that we should generally prefer axioms that can be described in simple languages. The second motivation is pragmatic. A very fruitful research agenda has been to attempt to automate some of the process of obtaining axiomatic results using automated-reasoning tools, which always starts by encoding axioms in a formal language. The use of *satisfiability solvers* has been especially successful and has led to a growing literature resolving several open problems in the field (Geist and Peters, 2017; Endriss, 2023). The formal verification of known proofs using interactive *proof assistants* such as Isabelle (Nipkow, 2009) or Lean (Holliday et al., 2021) also bears significant promise.

Let us now look ahead to what is to follow in this paper. Any formal statements we are going to make will be formulated with respect to one concrete model of collective decision making, namely a widely used and fairly simple model of voting. We are going to introduce that model in Section 2. But we believe that many of the conceptual points we make also apply to a wide range of other models of collective decision making, and we occasionally are going to reference other such models.

We are going to discuss different approaches to specifying the meaning of axioms in Section 3. The common approach, of course, is to express definitions in a mix of plain English and set-theoretical notation, without any *a priori* limitations on the range of mathematical concepts being referred to. Such definitions, when well executed, are perfectly rigorous, but they are not *formal* in the sense of being written in a formal language with a clearly specified syntax and an unambiguously defined semantics. Depending on one's objectives, having definitions of axioms available that meet such higher standards of formality can be useful, and we are going to briefly discuss different approaches that have been taken in the literature to do so. Both the common approach and the logic-based approach to specifying axioms are *intensional*, focusing on the attributes of mechanisms that would satisfy the axiom being defined. We are going to contrast this with an *extensional* approach to fixing the meaning of an axiom, which amounts to enumerating the mechanisms that satisfy the axiom. While actually listing all such mechanisms explicitly will rarely be feasible in practice, as we are going to see, introducing notation to refer to the set of all such mechanisms in an abstract manner turns out to be very useful.

In Section 4, we use our notation for the extensional meaning of an axiom to define the *strength* of an axiom as the proportion of mechanisms that violate the axiom (out of all conceivable mechanisms). This notion of strength generalises the familiar idea that we think of one axiom a being at least as strong as another in case the latter is a logical consequence of the former. We then discuss two applications of this notion of strength. First, taking inspiration from Shapley's idea for how to quantify a player's contribution to the worth generated by the grand coalition in a transferable-utility game (Shapley, 1953), we propose a way of measuring the *contribution* of individual axioms to a characterisation or impossibility result of interest. Second, by considering the ratio between the combined strength of the set of axioms involved in the result and the sum of the strengths of the individual axioms, we suggest a way of measuring the degree of '*surprise*' one might want to attach to the result. For instance, three relatively weak axioms giving rise to an impossibility is more surprising than one involving three axioms that each on their own already rule out most conceivable mechanisms.

A typical axiom will *talk about* some 'situations' (which in the case of voting are profiles of preferences) and not have anything to say about others. When we are given the definition of an axiom in intensional form, be it in plain English or using a formal language, we might be able to extract which situations those are. But any given formulation of the axiom might obscure the fact that for some situations the constraints imposed by the axiom are vacuous, so even though the specific rendering of the axiom *mentions* the situation, it in fact has nothing to *say* about that situation. And if we are given the axiom's definition in extensional form, then we have no rendering in language we could refer to in the first place. In Section 5, we propose a definition of *scope*—the set of situations an axiom talks about—that is language-independent and works directly on the extensional semantics of the axiom.

It is important to distinguish the number of situations an axiom talks about in total from the number of situations it relates to one another through any one of the constraints it imposes. For example, the axiom of *Neutrality*, requiring us to treat all alternatives the same, talks about all profiles of preferences, because for any given profile the collective choice we settle on is constrained by the choices we would make for the profiles we can obtain by permuting the alternatives in the profile at hand. Yet, we can fully describe the effect of imposing Neutrality by looking at two profiles at a time. We say that Neutrality has *dimensionality* 2. We formally define and then discuss this notion of dimensionality in Section 6.

As mentioned earlier, Fishburn (1973) proposed a natural manner in which to classify axioms. At least parts of his classification—the differentiation between *intraprofile* and *interprofile axioms*—are widely used in the literature. While convincing at an intuitive level, Fishburn's account, however, is lacking a precise definition of the classes of axioms identified. Using the machinery developed here, in particular the notions of dimensionality and scope, in Section 7 we suggest a way of making Fishburn's classification mathematically precise. Finally, in Section 8 we conclude, raising a few remaining questions deserving of attention in future work.

Throughout, our interest will be in axioms as formal objects rather than in axioms as descriptions of normative principles. This is not to say that the latter is not important. On the contrary. Any meaningful axiomatic analysis must start and end with the normative content of the principles being analysed. The formal tools we develop here are intended to support and enrich this kind of analysis.

2 The Model

While many of the conceptual points we wish to make in this paper apply to a variety of formal models of collective decision making, for our technical exposition we are going to focus on one specific such model, namely the model of *irresolute social choice functions* with a *fixed electorate* and a *fixed agenda*. Examples for such social choice functions include many of the best-known voting rules, such as the Plurality rule, the Borda rule, and the Copeland rule (Zwicker, 2016).

Let X be a finite set of *alternatives*. We think of X as being fixed throughout and use $m = |X|$ to denote the number of alternatives. We think of *preferences* as strict linear orders on X, ranking them from best to worst, and we use $X!$ to denote the set of all $m!$ such strict linear orders.

Let N be a finite set of *voters*. We also think of this *electorate* N as being fixed throughout and use $n = |N|$ to denote its size. A *profile* of preferences is a function $R : N \to X!$, mapping each of the individuals in N to the preference order they report.[2] We use $top_i(R)$ to denote the top-ranked alternative in $R(i)$; and we use $x \succ_i^R y$ to express that voter i ranks alternative x above alternative y in profile R. We write $\textsc{Prof}$ as a shorthand for the set of all profiles:[3]

$$\textsc{Prof} \;=\; X!^N$$

For any given profile of preferences, we would like to be able to select a single alternative from X that best reflects the preferences reported. But accounting for the

[2] Equivalently, one can also think of profiles as n-vectors of preferences.

[3] For any two sets S_1 and S_2, we use the familiar notation of $S_2^{S_1}$ to denote the set of all functions from S_1 to S_2, so $X!^N$ is the set of all functions from voters to preferences.

fact that sometimes resolving ties between seemingly equally deserving alternatives will be difficult, we shall be content with being able to select a nonempty subset of X. We shall refer to such nonempty subsets of X as *outcomes* and introduce the following notation to refer to them:

$$\textsc{Out} \;=\; 2^X \setminus \{\varnothing\}$$

The mechanisms we are interested in are *voting rules* (or *social choice functions*) that map any given profile R to a valid outcome in $\textsc{Out}$:

$$F : \textsc{Prof} \to \textsc{Out}$$

So our voting rules are *irresolute*. But we are free to impose resoluteness, i.e., the need to always return just a single alternative, as an additional requirement.

Our notation is consciously chosen so as to hide most of the specifics of the model. Indeed, at the conceptual level, what matters is that we are interested in mechanisms F that map profiles of preferences to outcomes. How preferences are modelled is not crucial: besides strict linear orders, they could also be, for instance, weak orders or sets of approved alternatives. What constitutes valid outcomes also is not crucial: besides nonempty sets of alternatives, they could also be, for instance, single alternatives (to model resolute voting rules) or sets of a fixed cardinality (to model multiwinner voting rules). We also could expand the notion of profile and work with a variable-electorate model, where any finite subset of some given universe of potential voters might report preferences.[4]

Observe that, due to the fact that for our model of voting both $\textsc{Prof}$ and $\textsc{Out}$ are finite sets, the set of all voting rules is finite as well. This will be important for our discussion in Section 4, which deals with the strength of axioms, but it is not crucial for any of the other parts of the paper.

An *axiom* is a property of voting rules $F : \textsc{Prof} \to \textsc{Out}$. Taking a purely technical perspective, it could be absolutely any kind of property. Of course only certain properties are of interest in practice. These are the properties of voting rules that encode some kind of normative principle one might reasonably want to postulate, either in general or for a specific kind of scenario. What is and what is not a reasonable property to encode as an axiom is not a question we can answer using mathematical tools. Rather, this requires arguments that are grounded in philosophical or ethical considerations, or indeed in 'common sense'.

Let us now review a few well-known examples for axioms used frequently throughout the literature on social choice theory. We are not going to attempt to retrace their origins here, which in any case are somewhat unclear for several of them, but these definitions, or alternative definitions of the same axioms, can be found in most

[4]While we shall stick with the terminology of referring to the inputs to mechanisms F as "profiles", in other areas of collective decision making the inputs a mechanism is defined on might include additional information. For instance, when modelling a claims problem (Thomson, 2003), we might want to reason about how a mechanism is to respond when a given agent's endowment changes but her preferences stay the same. And in the context of voting with a variable agenda, the input to a mechanism will consist of both a profile of preferences and a set of available alternatives.

textbooks on the topic as well as in the introductory chapter by Zwicker (2016). We stress that none of these axioms necessarily need to be accepted by everyone or in all conceivable circumstances. Rather, this should be understood as a catalogue of properties that will be relevant in many circumstances and that as system designers we should give serious consideration to.

Some axioms encode basic symmetry requirements. *Anonymity* requires that all voters be treated the same and *Neutrality* does the same for alternatives. Observe that for any permutation $\sigma : N \to N$ and any profile $R \in \textsc{Prof}$, the function $R \circ \sigma$ is the profile we obtain if we swap the preferences in R according to σ.

Axiom 1. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Anonymity** if $F(R) = F(R \circ \sigma)$ for any profile $R \in \textsc{Prof}$ and any permutation $\sigma : N \to N$.*

To define Neutrality we instead require a permutation $\sigma : X \to X$ defined on the alternatives. While for Anonymity, the outcome should not change when we permute the voters, for Neutrality, the outcome should change in line with the chosen permutation of the alternatives. To formulate the axiom, we extend σ to both profiles and outcomes in the natural manner.[5]

Axiom 2. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Neutrality** if $F(\sigma(R)) = \sigma(F(R))$ for any profile $R \in \textsc{Prof}$ and any permutation $\sigma : X \to X$.*

Anonymity asks for all voters to have equal power. A very weak form of this requirement would be to say that, at the very least, no single voter should have all the power. This is known as *Nondictatorship*.[6]

Axiom 3. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Nondictatorship** if there is no voter $i^\star \in N$ such that $F(R) = \{top_{i^\star}(R)\}$ for all profiles $R \in \textsc{Prof}$.*

Another group consists of axioms that determine the outcome, either fully or partly, for certain special cases of profiles. The simplest such axiom is *Unanimity*, saying that unanimously held preferences should be respected by the voting rule.

Axiom 4. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Unanimity** if $F(R) = \{x^\star\}$ whenever $top_i(R) = x^\star$ for all voters $i \in N$.*

A more demanding variant of this principle is *Pareto Efficiency*, saying that the outcome set must not include any dominated alternatives.

Axiom 5. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Pareto Efficiency** if $y \notin F(R)$ whenever there exists an $x \in X$ such that $x \succ_i^R y$ for all voters $i \in N$.*

---

[5] That is, for any profile $R \in \textsc{Prof}$, we have $\sigma(x) \succ_i^{\sigma(R)} \sigma(y)$ if and only if $x \succ_i^R y$; and for any set $S \in \textsc{Out}$, we have $\sigma(x) \in \sigma(S)$ if and only if $x \in S$.

[6] Axiom 3 is sometimes called *Weak Nondictatorship*, as this variant of the axiom only rules out any would-be dictator *fully* determining the outcome. *Strong Dictatorship* also rules out the possibility of there being a *nominator*, i.e., a voter who has the power to always place their top-ranked alternative into the outcome set, albeit possibly alongside other alternatives.

Observe that, indeed, Pareto Efficiency implies Unanimity: for any profile in which there is unanimous agreement on the top-ranked alternative, Pareto Efficiency will rule out all other alternatives; as the outcome must not be empty, the voting rule then must respect Unanimity.

The axiom of *Condorcet Consistency* postulates that an alternative $x^\star$ should be the sole election winner in any profile where it would beat any other alternative in a one-to-one majority contest. In such a profile, $x^\star$ is called the *Condorcet winner*.

Axiom 6. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Condorcet Consistency** if $F(R) = \{x^\star\}$ whenever $\#\{\, i \mid x^\star \succ_i^R y \,\} > \frac{n}{2}$ for all $y \in X \setminus \{x^\star\}$.*

A further group of axioms includes principles that specify how the outcome for a given profile should—or should not—be altered when certain small changes are made to that profile. This includes well-known axioms such as *Strategyproofness* and *Independence of Irrelevant Alternatives*, as well as a number of different monotonicity requirements. Here we shall formally define only one such monotonicity requirement, namely *Positive Responsiveness*. It expresses the idea that when an alternative $x^\star$ is winning in a given profile R, either alone or tied with other alternatives, then $x^\star$ should become the sole winner if one of the voters moves that alternative up in their own preference (and there are no other changes).[7]

Axiom 7. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Positive Responsiveness** if $F(R') = \{x^\star\}$ whenever $x^\star \in F(R)$ and R' is the result of moving $x^\star$ upwards in the preference of one of the voters in profile R.*

A final group of axioms covers requirements regarding basic characteristics of the outcomes returned by voting rules. The first one is *Resoluteness*, which states that the voting rule should return just a single alternative for every possible profile.

Axiom 8. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Resoluteness** if $|F(R)| = 1$ for any profile $R \in \textsc{Prof}$.*

Finally, the axiom of *Nonimposition* encodes the requirement that the outcome should not be imposed *a priori*—by postulating that, for every alternative $x^\star$, there must exist at least one way of voting for everyone so that $x^\star$ will be the sole winner.

Axiom 9. *A voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies **Nonimposition** if, for every alternative $x^\star \in X$, there exists a profile $R \in \textsc{Prof}$ such that $F(R) = \{x^\star\}$.*

In parts of the literature, starting with the seminal work of Arrow (1963), some more fundamental conditions relating to the definition of voting rules $F : \textsc{Prof} \to \textsc{Out}$ themselves are also considered 'axioms'. One of them is the *universal domain condition*, saying that F must be defined on all profiles in $\textsc{Prof}$. We take this to be implicit in the specification of the function $F : \textsc{Prof} \to \textsc{Out}$ and thus do not treat this condition (or indeed variants of it, requiring, say, that F is defined on all

[7] So for the new profile $R' \neq R$ there is one $i^\star \in N$ such that $\{\, y \mid x^\star \succ_{i^\star}^{R'} y \,\} \supseteq \{\, y \mid x^\star \succ_{i^\star}^{R} y \,\}$ but $x \succ_{i^\star}^{R'} y$ if and only if $x \succ_{i^\star}^{R} y$ for all $x, y \in X \setminus \{x^\star\}$, while $R'(i) = R(i)$ for all $i \in N \setminus \{i^\star\}$.

profiles that belong to a certain subdomain) as 'axioms'. We also do not want to conflate axioms with *collective rationality conditions*, which express that outcomes returned by a voting rule really must be elements of OUT. For our model, a natural collective rationality condition would be that $F(R)$ must not be the empty set.[8] We do not talk about such conditions here, because, again, we take them to be implicit in the definition of voting rules as functions of the type $F : \text{PROF} \to \text{OUT}$.

We conclude this section by introducing some additional terminology to talk about axioms. First, we call an axiom *trivial* (or *tautological*) if it is satisfied by *all* voting rules and we call it *unsatisfiable* if it is satisfied by *none*. Trivial and unsatisfiable axioms are of no interest in and of themselves—they should never feature in a characterisation or an impossibility result, for instance—but they might come up when an analyst manipulating a number of axioms makes a mistake. So it is useful to have the vocabulary in place to talk about such cases.

Second, the *conjunction* of two axioms A and A', denoted by $A \wedge A'$, is the requirement on a voting rule to respect both A and A'.[9] For example, one might want to introduce a general *Symmetry* axiom as the conjunction of *Anonymity* and *Neutrality*. We use calligraphic letters, as in $\mathcal{A}$, to denote *sets of axioms*. A voting rule satisfies $\mathcal{A}$ exactly if it satisfies all of the axioms in $\mathcal{A}$. For any given axiom set $\mathcal{A}$, we often treat $\mathcal{A}$ and the conjunction of the axioms in $\mathcal{A}$ interchangeably.

3 The Meaning of Axioms

In the previous section we saw several examples for definitions of axioms, each of them a carefully crafted attempt to find an adequate compromise between readability and rigour. This certainly is the right mode of communication to adopt in a paper such as this. But if we wanted to provide a truly unambiguous specification of an axiom, one that cannot be misunderstood by even the most pedantic of mathematicians and one that could be parsed correctly by a machine, then our definitions would fall short—as would any of those one can find in even the most exemplary of instances of mathematical writing in economic theory literature.

In this section, we want to review possible approaches to formally specifying an axiom in social choice theory. We are going to argue that thinking about an axiom in terms of its *extensional* semantics can offer a helpful additional perspective, on top of the familiar *intensional* approach to defining the meaning of axioms.

The differentiation between intensional and extensional definitions of an object or a concept has a long history in philosophical logic and the philosophy of language. Intensional meaning is what Gottlob Frege (1892) calls *Sinn*, while extensional meaning is what he refers to as *Bedeutung* (usually translated as *sense* and *reference* in English). While the former is a characterisation of a concept in terms of its essential

[8]For other models of collective decision making, such as the model of Arrovian *social welfare functions* mapping profiles in $X!^N$ to linear orders in $X!$ (Arrow, 1963), one can formulate more interesting collective rationality conditions, such as the requirement that the outcome returned should be transitive—which of course is guaranteed if indeed that outcome is an element of $X!$.

[9]In Section 7, we are also going to discuss *disjunctions* of axioms.

attributes, the latter determines meaning on the grounds of what entities in the world the concept refers to. In the case of axioms, the intensional meaning thus is given by the idea that the axiom in question expresses, i.e., the normative principle that it is supposed to capture. This meaning can be expressed by describing the axiom in a suitable (formal) language. The extensional meaning of an axiom, on the other hand, is the set of entities it characterises, i.e., the set of all those voting rules that satisfy the normative principle encoded by the axiom.

Taking a formal approach to providing an intensional definition for an axiom means encoding that axiom in a logical language with formally defined syntax and semantics. There are a wide range of logics to pick from. As this is not the focus of this paper and as a good part of this literature is reviewed elsewhere (Endriss, 2011), we are only going to briefly sketch a couple of examples here. Let us start with classical *first-order predicate logic*, probably the best-known logical system and the one most commonly used across a plethora of scholarly disciplines.

Example 1 (First-order logic)**.** First-order logic is the familiar system in which we can use not only the boolean connectives, i.e., 'not' ($\neg$), 'and' ($\wedge$), 'or' ($\vee$), and 'implies' ($\rightarrow$), but also both universal and existential quantifiers ($\forall$ and $\exists$). We refer to van Dalen (2013) for a textbook introduction to its syntax and semantics.

Grandi and Endriss (2013) explore to what extent we can express the concepts required to state some of the seminal impossibility theorems in social choice theory in this logic. To provide an impression how one might go about this, but omitting most technical details, here is how to define Pareto Efficiency in first-order logic:

$$\forall r.\forall x.\forall y.[\,\underline{R}(r) \wedge \underline{X}(x) \wedge \underline{X}(y)\,] \rightarrow [\,(\forall i.\underline{N}(i) \rightarrow \underline{r}(r,i,x,y)) \rightarrow \neg\underline{w}(r,y)\,]$$

Here all underlined letters are predicate symbols in the logical language. The above formula can be parsed as follows: for all r, x, and y, if r is a profile, x is an alternative, and y also is an alternative, then it must be the case that, if for all i such that i is a voter it is the case that in profile r voter i *r̲anks* x above y, then it must not be the case that in profile r alternative y *w̲ins*. $\triangle$

Other logics that that have been used to model axioms (and other concepts) in social choice theory include various *modal logics* (see, e.g., Troquard et al., 2011), *dependence logic* (Pacuit and Yang, 2016), and various *higher-order logics* (see, e.g., Nipkow, 2009; Holliday et al., 2021). Of special interest, however, is classical *propositional logic* (which some readers may know as *boolean logic*), the simplest (and least expressive) of any of the widely used logical systems.

Example 2 (Propositional logic)**.** In propositional logic our expressive power is limited to using the aforementioned boolean connectives (van Dalen, 2013). To encode axioms of interest, we introduce one propositional variable $p_{R,x}$ for every profile $R \in \textsc{Prof}$ and every alternative $x \in X$. The intended reading is that, in profile R, alternative x is amongst the winning alternatives. Somewhat surprisingly, we can still encode Pareto Efficiency in this logic:

$$\bigwedge_{R \in \textsc{Prof}} \; \bigwedge_{x \in X} \; \bigwedge_{y \in \textsc{Dom}^R(x)} \neg p_{R,y}$$

Here $\textsc{Dom}^R(x)$ is short for $\{ y \in X \mid x \succ_i^R y \text{ for all } i \in N \}$, the set of alternatives dominated by x in profile R. So this formula is satisfied whenever, for all profiles R,[10] for all alternatives x, and for all alternatives y dominated by x in R, it is not the case that y is amongst the winning alternatives.

In fact, we can encode every one of the axioms defined in the previous section—and indeed *every* conceivable property of voting rules. To see this, observe that a long conjunction of literals of the form $p_{R,x}$ and $\neg p_{R,x}$ can be used to fully specify the behaviour of a single voting rule, which means that a disjunction of such conjunctions can be used to fully specify any arbitrary axiom of interest. △

The significance of the observation that propositional logic is fully expressive with respect to our model of voting (as well as many other models of collective decision making) is that this has enabled researchers to use *satisfiability solvers*—highly optimised tools to determine whether a given formula is satisfiable—to automatically derive impossibility theorems.[11] Having said this, the fine print here is important. Recall that we formulated our model with respect to (arbitrary but) *fixed* sets N and X. So we can express axioms (and prove theorems) for any given choice of N and X, but we cannot express that, say, Pareto Efficiency should be respected independently of the size of the electorate. If we want to be able to express this, we need to switch to a more expressive logic, such as first-order logic.[12] Of course, the disadvantage of switching to a more expressive logic is that reasoning in that logic becomes harder (and full automation almost impossible).

Let us now move on to the idea of defining axioms in extensional rather than intensional terms. In fact, the idea is very simple: we define an axiom A as the set of all the voting rules that meet the conditions imposed by the axiom.

Definition 1. *The **interpretation** (or **extension**) of an axiom A, denoted by $\mathbb{I}(A)$, is the set of voting rules that satisfy A:*

$$\mathbb{I}(A) \quad = \quad \{ F : \textsc{Prof} \to \textsc{Out} \mid F \textit{ satisfies } A \}$$

This definition and use of notation go back to the work of Boixel and Endriss (2020). Recall that, for our model, the set of *all* voting rules—those that satisfy A and those that do not—is finite. This is helpful, but in fact not critical.

Having this notation available makes it possible, at least in principle, to unambiguously define the meaning of any conceivable axiom A by 'simply' enumerating the voting rules we think of as meeting the conditions prescribed by A. Of course, an intensional definition—be it rigorous but informal or entirely formal—is a much more manageable object to construct. In fact, we do not propose that those wishing to define a new axiom really start enumerating voting rules explicitly. Rather, we claim

[10] Here the—very long—conjunction over all profiles has the effect of a universal quantification.

[11] This opportunity was first noted by Tang and Lin (2009). For expository discussions of this approach, we refer the reader to Geist and Peters (2017) and Endriss (2023).

[12] First-order logic also has its limitations. For instance, we cannot express that a given set is of *finite* (but otherwise arbitrary) size (Grandi and Endriss, 2013), which is an important detail, e.g., for Arrow's Theorem. If we need to express this feature, we can move up to higher-order logics.

and shall demonstrate throughout this paper that knowing that—*in principle*—it is possible to give an unambiguous definition and having available clear and simple notation to refer to that unambiguous definition is very useful when working with the axiomatic method in social choice theory.

Observe that $\mathbb{I}(A \wedge A') = \mathbb{I}(A) \cap \mathbb{I}(A')$ for any two axioms A and A'. We extend our notation from axioms to axiom sets in the natural manner. The interpretation of an axiom set $\mathcal{A}$ is the intersection of the interpretations of the axioms in $\mathcal{A}$:

$$\mathbb{I}(\mathcal{A}) \quad = \quad \bigcap_{A \in \mathcal{A}} \mathbb{I}(A)$$

Our notation is useful for expressing relevant relationships between axioms. For instance, $\mathbb{I}(A) = \mathbb{I}(A')$ means that A and A' are logically *equivalent*, in the sense of imposing the same conditions on voting rules.[13] $\mathbb{I}(A) \subseteq \mathbb{I}(A')$ means that A *implies* A', while $\mathbb{I}(A \wedge A' \wedge A'') = \varnothing$ means that it is *impossible* to construct a voting rule that would satisfy all three of those axioms. Finally, $\mathbb{I}(A \wedge A' \wedge A'') = \{F\}$ means that those three axioms fully *characterise* the voting rule F.

One reading of an equation such as $\mathbb{I}(A \wedge A' \wedge A'') = \{F\}$ is that axioms and the voting rules they characterise are interchangeable ways of describing the same mathematical objects. Interestingly, we can find an expression of this idea also in Philippe Mongin's critical essay on the use of the axiomatic method in economics:

> "From a formal point of view, the distinction between 'axioms' (which specify conditions) and 'solutions' (which specify [voting] rules) is *arbitrary*. In social choice theory [...] those conditions bear directly on the function F representing the rule. They therefore are of the exact same type[14] as the 'solution'." — Mongin (2003, p. 132, our translation)

But Mongin also points out that having established the mathematical equivalence between $A \wedge A' \wedge A''$ and F does not make these two objects interchangeable in all respects. A characterisation of F in terms of $A \wedge A' \wedge A''$ is interesting precisely because we can attach normative principles to those axioms that are not—at least not in an obvious or immediate way—incorporated by F as well.

On a related note, in certain 'small' instances of our model, such when there are just two voters and two alternatives (or when there is just a single voter, with any number of alternatives), the extensions of Pareto Efficiency and Condorcet Consistency coincide. Yet, their intensional meanings—the reasons for possibly wanting to accept these axioms—differ significantly.

To take yet another example,[15] from an extensional point of view, we should reject (the combination of) the axioms featuring in an impossibility result such as Arrow's Theorem (1963) or the Gibbard-Satterthwaite Theorem (Gibbard, 1973;

[13]In particular, if we have given intensional definitions for A and A' by encoding them using the same logic (say, classical first-order logic), then $\mathbb{I}(A) = \mathbb{I}(A')$ means that the two formulas must entail one another according to the formal semantics of that logic.

[14]Here we use the word '*type*' to translate Mongin's expression '*nature syntaxique*'.

[15]This example was suggested to us by Franz Dietrich.

Satterthwaite, 1975), because they reduce to the empty set, while from an intensional perspective these axioms are worth contemplating.

These dilemmas are not unique to the semantics of axioms but commonplace in the study of the philosophy of language. Maybe the most famous example is Gottlob Frege's puzzle regarding the '*Morning Star*' and the '*Evening Star*'. Intensionally, these two terms refer to different concepts, while extensionally they both denote the same celestial object (namely the one we nowadays tend to refer to as '*Venus*').

It also is worth pointing out that, sometimes, there will be more than one intensional definition of a given normative principle that suggests itself. Take the case of Anonymity. Recall how in Section 2 we defined Anonymity in terms of permutations $\sigma : N \to N$ defined on the set of voters (see Axiom 1). So the intuitive idea of wanting to treat all voters the same here has been operationalised by requiring that permuting the voters should not change the outcome. But how do we know that this is 'the right' way of operationalising this idea? We instead could have imposed this requirement for a special family of permutations $\sigma : N \to N$ only, namely those that swap around exactly two voters. The extension of both of these variants of the axiom is the same—but seeing this is not entirely trivial; it technically involves a proof by induction over the number of pairwise swaps needed to reconstruct a given arbitrary permutation. So, when moving from the informal expression of an idea to a formal intensional definition, we might encounter similar issues as when moving from a precise intensional definition to an extensional definition. In both cases, we may end up grouping together certain axioms as being 'the same', when at the other level of description we would rather keep them apart.

In conclusion, the intensional approach to defining axioms and the extensional approach to defining axioms provide complementary views that, together, allow for a deeper analysis than restricting oneself to just one approach.

4 The Strength of Axioms

In economic theory, when trying to formally capture a normative principle to be encoded in the shape of an axiom, we generally look for axioms that are especially weak. There are two reasons for this mantra of logical weakness. First, weakening a first formulation of an axiom will necessarily make that axiom less controversial. Anyone who accepts the stronger formulation of the axiom must also accept the weaker ones—lest they wish to quarrel with the rules of logic themselves. Second, the weaker the axioms involved in the formulation of a specific result, be it an impossibility result or a characterisation result, the more surprising—and thus interesting—that result will be judged to be. William Thomson puts it like this:

> "[I]f you encounter an impossibility theorem, you'll need to explore how serious it is, and one way to do that is to substitute weaker versions of the axioms [...] If the impossibility persists [...] [y]ou'd have a much deeper understanding of the situation." — Thomson (2023, p. 117)

It therefore is not surprising that, throughout the literature, one can find frequent claims regarding the logical weakness of the axioms being proposed.

But what does it actually mean to *weaken* an axiom? This raises a more general question: What is the *strength* of an axiom? In this section, we want to give a possible answer to this latter question (and thereby also to the former), and we want to illustrate some possible applications of such a definition. One will be to quantify the *contribution* of individual axioms to a result of interest. Another will be an attempt to quantify how *surprising* a given result should be taken to be.

For certain pairs of axioms A and A', it is entirely unambiguous what we mean by saying that one is stronger (or weaker) than the other:

> *If every voting rule that satisfies axiom A also satisfies axiom A', but the converse is not true, then we say that A is stronger than A'.*

Observe that we can neatly express this using our set-theoretical notation for the extensional semantics of an axiom: A is stronger than A' in case $\mathbb{I}(A) \subsetneq \mathbb{I}(A')$.

But when there is no such logical entailment relation between our axioms of interest, then this approach does not allow us to compare their strengths. Yet, certainly at an intuitive level, we frequently make appeals to the perceived strength of different axioms. For example, when considering the axioms involved in the Gibbard-Satterthwaite Theorem (usually formulated in a model for resolute voting rules), it seems very natural to refer to two of them, Nonimposition and Nondictatorship, as being very weak (and thus entirely uncontroversial), while the third, Strategyproofness, looks much stronger (and thus, conceivably, could be the subject of scholarly criticism). So, can we provide a definition of strength of an axiom that is absolute rather than relative (to that of a closely related axiom)?

Next, we propose such a definition of the strength of an axiom. The basic idea is that, the stronger an axiom, the fewer voting rules will satisfy that axiom. Recall that a voting rule is a function $F : \textsc{Prof} \to \textsc{Out}$. Note that $\textsc{Out}^{\textsc{Prof}}$ is the set of all functions from $\textsc{Prof}$ to $\textsc{Out}$, and thus the set of all voting rules.

Definition 2. *The* ***strength*** *of axiom A is the proportion of the voting rules—relative to the set of all voting rules—that are excluded by that axiom:*

$$\mathit{strength}(A) \;=\; \frac{|\textsc{Out}^{\textsc{Prof}} \setminus \mathbb{I}(A)|}{|\textsc{Out}^{\textsc{Prof}}|} \;=\; 1 - \frac{|\mathbb{I}(A)|}{|\textsc{Out}^{\textsc{Prof}}|}$$

We extend the definition of strength from individual axioms A to sets of axioms $\mathcal{A}$ by defining $\mathit{strength}(\mathcal{A})$ as the strength of the conjunction of the axioms in $\mathcal{A}$ (alternatively, we could simply replace A with $\mathcal{A}$ everywhere in the definition above).

Example 3 (Strength of Neutrality and Condorcet Consistency)**.** To exemplify our definition of axiom strength, let us calculate the strength of the axiom of Neutrality and let us then compare that strength to that of Condorcet Consistency. Intuition suggests that Condorcet Consistency is the stronger axiom: it seems to constrain our choice of voting rule much more than the seemingly mild requirement of respecting Neutrality. Indeed, essentially every common voting satisfies Neutrality, while asking for Condorcet Consistency means pushing for a fairly specific type of rule. But let us see whether this intuition also is borne out by our calculations.

Recall that n is size of the electorate and that m is the number of alternatives. There are $m!$ many preferences an individual might report, and thus $m!^n$ many profiles we might encounter. A voting rule needs to map each profile to one of the $2^m - 1$ possible outcomes (the nonempty subsets of X). So there are a total of $(2^m - 1)^{m!^n}$ many voting rules to consider.

To count how many of these voting rules satisfy Neutrality, let us fix one specific preference and call it the 'canonical' preference. For a neutral voting rule, it is sufficient to specify the outcome returned for those profiles in which the first voter reports this canonical preference; all other outcomes are then determined by the need to satisfy Neutrality. So there are $(2^m - 1)^{m!^{n-1}}$ neutral voting rules, meaning that the strength of Neutrality (N) is given by the following formula:

$$\begin{aligned} \textit{strength}(\mathrm{N}) &= 1 - \tfrac{(2^m-1)^{m!^{n-1}}}{(2^m-1)^{m!^n}} \\ &= 1 - \tfrac{1}{(2^m-1)^{(1-\frac{1}{m!})\cdot m!^n}} \end{aligned}$$

For instance, for $n = 3$ and $m = 2$, we get a strength of 0.9375. For larger choices of parameters, this number grows further (and it does so fast). This reflects the fact that even a single axiom such as Neutrality will exclude the vast majority of all conceivable voting rules.

The axiom of Condorcet Consistency fully determines the outcome for those profiles for which there is a Condorcet winner and it has no impact at all for any of the other profiles. Let $p_{n,m} \in [0,1]$ be the proportion of profiles with a Condorcet winner amongst all profiles for n voters and m alternatives. The value of $p_{n,m}$ has been computed (or at least approximated) for several choices of the parameters n and m. For instance, we know that $p_{3,3} = 0.9\bar{4}$ and that $p_{n,3} > 0.91$ for any odd number n (Gehrlein and Fishburn, 1976).

Given $p_{n,m}$, we can computer the strength of Condorcet Consistency (CC) as follows:

$$\begin{aligned} \textit{strength}(\mathrm{CC}) &= 1 - \tfrac{(2^m-1)^{(1-p_{n,m})\cdot m!^n}}{(2^m-1)^{m!^n}} \\ &= 1 - \tfrac{1}{(2^m-1)^{p_{n,m}\cdot m!^n}} \end{aligned}$$

Comparing the formulas for the strengths of Neutrality and Condorcet Consistency, we find that Condorcet Consistency is the stronger axiom if and only if $p_{n,m} > 1 - \frac{1}{m!}$. For $m = 3$ this is indeed the case for any odd value of n (as $0.91 > 0.8\bar{3}$). So this seems to confirm our intuitions about the relative strengths of these two axioms. But for $m = 4$, in fact the opposite is true, given that $1 - \frac{1}{4!} > 0.95$ and $p_{n,4} \leqslant 0.\bar{8}$ for any odd number n (Gehrlein and Fishburn, 1976).

So for larger numbers of alternatives, Neutrality in fact is the stronger axiom. Upon reflection, this is not entirely at odds with intuition either, given that the impact of Neutrality, which is defined in terms of permutations of the set of alternatives, might be expected to increase as the number of alternatives increases, while the likelihood of encountering a Condorcet paradox[16] is known to grow as the number of alterna-

[16] A *Condorcet paradox* is said to occur for a profile R if R does not have a Condorcet winner.

tives increases, meaning that the strength of the axiom of Condorcet Consistency reduces at the same time. So, as we increase m, there must be a turning point where the strength of Neutrality tops that of Condorcet Consistency. What still is surprising is that this turning point occurs between $m = 3$ and $m = 4$ already. $\triangle$

We note that for typical choices of n, m, and A, the value of $\textit{strength}(A)$ will be very close to 1, which conceivably could present practical difficulties when organising and comparing data points regarding the strengths of several different axioms. Unfortunately, this kind of difficulty seems unavoidable. Given the huge number of voting rules for any typical choice of n and m, we cannot but work either with very large numbers or with numbers that lie very close to one another.

Let us now take stock of some basic facts regarding our notion of axiom strength. According to our definition, an *unsatisfiable axiom*—or an axiom set that gives rise to an impossibility result—has strength 1, while a *trivial axiom* that is satisfied by all voting rules has strength 0. We say that a satisfiable axiom (or axiom set) has *maximal strength* in case it is satisfied by only a single voting rule. Any such axiom has strength $|\textsc{Out}^{\textsc{Prof}}|-1 / |\textsc{Out}^{\textsc{Prof}}|$. If axiom A implies axiom A', i.e., if $\mathbb{I}(A) \subseteq \mathbb{I}(A')$, then $\textit{strength}(A) \geqslant \textit{strength}(A')$, but the converse need not be true. The following fact, the proof of which is immediate given the relevant definitions, captures what we can say about the strength of a conjunction of two axioms of known strengths.

Fact 1. *For any two axioms A and A', the strength of their conjunction is bounded from above by the sum of their individual strengths:*

$$\textit{strength}(A \wedge A') \quad \leqslant \quad \textit{strength}(A) \ + \ \textit{strength}(A')$$

Now let us move on to a first application of the concept of axiom strength. Suppose $\mathcal{A}$ is the set of axioms involved in an impossibility result or a characterisation result. How can we quantify the contribution each of the individual axioms make to the overall result? The absolute strength $\textit{strength}(A)$ of each axiom $A \in \mathcal{A}$ gives a first indication, but it does not quite capture what we are interested in here, namely the strength of A *in the specific context* of $\mathcal{A}$. To capture this notion of an individual axiom's contribution to achieving a given result, let us think of the axioms in $\mathcal{A}$ as the players in a transferable-utility game,[17] where the worth achieved by any subset $\mathcal{A}$ is the combined strength of that subset. With this metaphor in place, it makes sense to measure an axiom's contribution to the result as the *Shapley value* of that axiom (Shapley, 1953), i.e., as the average marginal increase in strength across all possible ways of constructing $\mathcal{A}$ from the empty set by adding axioms one by one.

To formally state the definition, we require some additional notation. Observe that the set of all possible sequences of axioms in $\mathcal{A}$ is isomorphic to the set of all linear orders on $\mathcal{A}$. In view of this fact, we are going to slightly overload notation and use $\mathcal{A}!$ to denote the set of all such sequences. Then, for any sequence $\rhd \in \mathcal{A}!$, we write $A \rhd A'$ to express that A precedes A' in $\rhd$.

[17]Readers not familiar with the theory of transferable-utility games will find a helpful introduction in the textbook by Peters (2008).

Definition 3. *The **degree of contribution** (or the **Shapley value**) of axiom $A^\star$ within the set of axioms $\mathcal{A}$ is defined as follows:*

$$dcon(A^\star, \mathcal{A}) \;=\; \frac{1}{|\mathcal{A}!|} \cdot \sum_{\rhd \in \mathcal{A}!} strength(\{A \mid A \rhd A^\star\} \cup \{A^\star\}) \\ - \; strength(\{A \mid A \rhd A^\star\})$$

Thus, to compute the Shapley value of $A^\star$, we cycle through all possible orderings $\rhd$ of the axioms in $\mathcal{A}$. For each of them, we compute the marginal increase in strength of $A^\star$ as the difference between the combined strength of the axioms preceding $A^\star$ in $\rhd$ together with $A^\star$ and the combined strength of the axioms preceding $A^\star$ in $\rhd$ without $A^\star$. We add up all these marginal increases in strength, one for each sequence, and in the end divide by the number of sequences considered (which is $|\mathcal{A}!|$). When clear from context, we omit $\mathcal{A}$ from our notation.

Appendix A consists of a case study regarding the degree of contribution of each of the axioms involved in a simple impossibility result stating that there are no resolute voting rules for two alternatives and two voters that satisfy both Anonymity and Neutrality. This is a specific instance of a more general result due to Moulin (1983).

As a second application of our notion of axiom strength we want to suggest a possible route of quantifying the degree of 'surprise' one might want to associate with a given result of interest. Intuitively speaking, when comparing, say, two different impossibility results, we might find the one that involves the weaker axioms more surprising. In both cases, the combined strength of the axioms involved must be exactly 1, as in both cases the voting rules that are being excluded are the exact same ones—namely all of them. So, for each of the two results, the sum of the strengths of the individual axioms involved must be at least 1, but it typically will be more than that. This additional strength contributed by the axioms is, in some sense, redundant. If there is a lot of this redundancy, so if several of the axioms involved are fairly strong on their own, then the fact that together they achieve an impossibility seems less surprising or interesting.

Definition 4. *Let $\mathcal{A}$ be a set of two or more axioms, at least one of which is nontrivial. The **degree of complementarity** of $\mathcal{A}$ is defined as follows:*

$$dcom(\mathcal{A}) \;=\; \frac{strength(\mathcal{A})}{\sum_{A \in \mathcal{A}} strength(A)}$$

The assumption of there being at least one nontrivial axiom in the set ensures that we do not divide by 0.

Observe that the maximal value $dcom(\mathcal{A})$ can take for any set $\mathcal{A}$ is 1. This happens when no single voting rule that is excluded by one axiom in $\mathcal{A}$ is also excluded by one of the other axioms in the set. That is, the degree of complementarity of $\mathcal{A}$ is equal to 1 precisely when the axioms in $\mathcal{A}$ are perfectly complementary—when there are no redundancies whatsoever. At the other end of the spectrum, while $dcom(\mathcal{A})$ cannot reach or drop below 0, it can get arbitrarily close to 0, namely when $\mathcal{A}$ includes multiple axioms that all impose the exact same constraint on voting rules (i.e., multiple axioms with the same interpretation).

If $\mathcal{A}$ is the set of axioms involved in a result of interest, such as an impossibility theorem, we may interpret $dcom(\mathcal{A})$ as the degree of surprisingness of the result.[18] A high degree of complementarity means that the individual axioms involved are relatively weak in view of the result achieved by their conjunction.

Appendix B discusses a small case study regarding the complementarity of the axioms involved in two well-known results for the special case of scenarios with two voters and two alternatives. The first is the aforementioned impossibility of finding voting rules that satisfy Anonymity, Neutrality, and Resoluteness. The second is May's Theorem on the characterisation of the simple majority rule in terms of Anonymity, Neutrality, and Positive Responsiveness (May, 1952).

We had previously mentioned that in the literature we often find statements about how weakening axioms just enough without affecting the validity of a given result will make that result more interesting and surprising. Our notion of surprisingness-as-complementarity can capture this effect, as the following simple result shows.

Proposition 2. *Let $\mathcal{A}_1$ be a set of (at least two) nontrivial axioms giving rise to some result of interest, such as an impossibility result (meaning $\mathbb{I}(\mathcal{A}_1) = \varnothing$) or the characterisation of some voting rule F (meaning $\mathbb{I}(\mathcal{A}_1) = \{F\}$). Let $\mathcal{A}_2 = (\mathcal{A}_1 \setminus \{A_1\}) \cup \{A_2\}$ be the result of replacing axiom A_1 in $\mathcal{A}_1$ by another axiom A_2. If A_2 is a weakening of A_1 in the narrow sense (meaning $\mathbb{I}(A_1) \subsetneq \mathbb{I}(A_2)$) and if our result of interest is not affected by this change (meaning $\mathbb{I}(\mathcal{A}_1) = \mathbb{I}(\mathcal{A}_2)$), then the new result is more surprising than the original one in the sense of having a higher degree of complementarity: $dcom(\mathcal{A}_1) < dcom(\mathcal{A}_2)$.*

Proof. Let $\mathcal{A} = \mathcal{A}_1 \setminus \{A_1\} = \mathcal{A}_2 \setminus \{A_2\}$ be the set of axioms shared by $\mathcal{A}_1$ and $\mathcal{A}_2$. We now rewrite the degree of complementarity for each of our two axiom sets:

$$\begin{aligned} dcom(\mathcal{A}_1) &= \frac{strength(\mathcal{A}_1)}{strength(A_1) + \sum_{A \in \mathcal{A}} strength(A)} \\ dcom(\mathcal{A}_2) &= \frac{strength(\mathcal{A}_2)}{strength(A_2) + \sum_{A \in \mathcal{A}} strength(A)} \end{aligned}$$

The claim now follows from the observation that, by Definition 2, our assumption $\mathbb{I}(\mathcal{A}_1) = \mathbb{I}(\mathcal{A}_2)$ entails $strength(\mathcal{A}_1) = strength(\mathcal{A}_2)$, while our assumption $\mathbb{I}(A_1) \subsetneq \mathbb{I}(A_2)$ entails $strength(A_1) > strength(A_2)$. □

Proposition 2 remains true if we generalise from the narrow sense of axiom weakening to weakening with respect to axiom strength in the sense of Definition 2. We have chosen to state the more narrow variant of this basic insight here to emphasise the link between our proposed notion of 'surprise' and the broadly accepted and widely used sense of axiom weakening.

[18]If $\mathcal{A}$ is a singleton, this interpretation of the degree of complementarity as a degree of surprisingness breaks down, which is why this case is excluded in the formulation of Definition 4.

5 The Scope of Axioms

When discussing the impact a given axiom has in a variety of different situations of interest, an economic theorist will often refer to the concept of that axiom *talking about* a given profile, or equivalently, the concept of a given profile *being in the scope* of that axiom. When the axiom in question is spelt out in plain (yet precise) English or when it is defined in a formal language, then it is intuitively clear what we mean by this: the set of profile(s) being talked about will show up explicitly in such a definition. But does that mean that when we switch from one definition of the axiom in one language to an equivalent definition in a different language, that the set of profiles it talks about might change? That would seem unsatisfactory.

Our objective for this section is to provide a general definition of the concept of scope of an axiom that is independent of the (formal) language used to express that axiom. In other words, we are looking for a definition of scope that can be stated in terms of the extensional semantics of the axiom of interest. But first, to illustrate the idea of scope, let us consider a simple example.

Example 4 (Scope of Pareto Efficiency)**.** The axiom of Pareto Efficiency asks us to not select an alternative in case that alternative is dominated by another alternative in the profile under consideration. Intuitively, this axioms *talks about* certain profiles, while it has nothing to say about others. Take the following two profiles:

Voter 1 :	$a \succ b \succ c$	Voter 1 :	$a \succ b \succ c$
Voter 2 :	$b \succ c \succ a$	Voter 2 :	$b \succ c \succ a$
Voter 3 :	$b \succ a \succ c$	Voter 3 :	$c \succ a \succ b$

In the first profile, alternative c is dominated (by alternative b, given that all three individuals rank b above c). So the axiom of Pareto Efficiency talks about that profile: to satisfy the axiom, we must make sure that our voting rule does not select c in this particular profile. But the axiom does not talk about the second profile (where none of the three alternatives is dominated): under no circumstances does the question of whether a given voting rule satisfies the axiom depend on the outcome returned by that rule when applied to the second profile. So the second profile is not in the scope of the axiom of Pareto Efficiency. △

Whether or not a given axiom talks about a given profile will often be clear from the (intensional) definition of the axiom. Especially when a formal definition of the axiom is given, we should expect the profiles it talks about to show up explicitly in the scope of a quantifier—unlike those profiles it does not talk about. For the definition of Pareto Efficiency in Section 2 this is indeed the case.

But what if the formulation of an axiom might include redundancies? What if it *seems* to talk about a certain profile but *actually* has nothing to say about it—in the sense of not imposing any constraints on what outcome should be returned for that profile? The next example illustrates a subtlety of this kind.

Example 5 (Scope of Anonymity)**.** Recall the axiom of Anonymity, as defined in Section 2. What profiles does this axiom talk about? Intuition clearly suggests that it talks about *all* profiles. While it does not fix the outcome for any one profile, it seemingly imposes constraints on all profiles—by requiring that pairs of profiles in which the exact same preferences are being reported must be treated the same.

But in fact Anonymity does not talk about any of the *unanimous profiles*—the profiles where all voters report the same preference.[19] Indeed, the axiom does not constrain our voting rule on such profiles in any way; we are free to choose an outcome for a given unanimous profile without acquiring any obligations regarding the outcomes for other profiles. So the scope of Anonymity in fact is just the set of all those profiles that are not unanimous.[20] △

Now that we are aware of the potential pitfalls of relying too much on the particular manner in which the definition of an axiom is presented to us when attempting to determine its scope, let us formulate a definition of scope—i.e., of 'talking about'—that operates directly on the extensional semantics of the axiom of interest.

For two voting rules F and F', we write $F =_{-R} F'$ in case the two voting rules agree on all profiles, except possibly R.

Definition 5. *Let A be a satisfiable axiom. We say that axiom A **talks about** profile $R \in \textsc{Prof}$ if and only if there exist a voting rule $F \in \mathbb{I}(A)$ satisfying A and a voting rule $F' \in \textsc{Out}^{\textsc{Prof}} \setminus \mathbb{I}(A)$ violating A such that $F =_{-R} F'$. The set of all profiles that A talks about is the **scope** of A.*

In other words, axiom A talks about profile R if, whenever we wish to determine whether a voting rule F satisfies A, we require access to information about $F(R)$. Here *information about* $F(R)$ need not be of the form "the outcome is this and that". It might instead be something like "the outcome is *not* this and that" or "the outcome is a singleton". In the case of Pareto Efficiency, for example, it will be of the form "the outcome does not include this and that alternative".

We use *scope*(A) to denote the scope of axiom A and we extend the notion of scope from axioms to axiom sets in the natural manner, i.e., *scope*$(\mathcal{A})$ is the scope of the conjunction of the axioms in $\mathcal{A}$. When *scope*$(A) = \textsc{Prof}$, then we say that A has *full scope.* This notion of full scope is closely related to what Thomson (2001, 2023) calls *full coverage* and what Fishburn (1973) calls an axiom being *passive*—although these authors restrict this kind of terminology to certain types of axioms only. We are going to return to this point in Section 7.

Note that Definition 5 only applies to satisfiable axioms, i.e., axioms satisfied by at least one voting rule. How to define an appropriate notion of scope for an *unsatisfiable axiom* is not at all obvious—but fortunately also of fairly limited interest in practice; we briefly discuss this question in Appendix C. Observe that, according to

[19]We are grateful to Dominik Peters for sharing this observation with us.

[20]Interestingly, in a model of voting with *variable electorates*, where the set of voters reporting a preference can change from profile to profile, the situation is different and the scope of Anonymity really is the full set of profiles (because now the outcomes for unanimous profiles are constrained by the outcomes for profiles where a different set of voters report the same preferences).

Definition 5, the scope of a *trivial axiom*, i.e., an axiom satisfied by all voting rules, is the empty set.[21] This perfectly matches intuition.

While Definition 5 is a precise definition of the notion of scope, it arguably does not provide us with clear instructions what exactly we need to check to identify the scope of a given axiom. The next result provides us with such a recipe. (But we stress that this still is not an efficient algorithm that can be executed in practice.)

To state this result, we require some further notation. First, for any function $F : \textsc{Prof} \to \textsc{Out}$ and any set $S \subseteq \textsc{Prof}$, we use $F\restriction_S$ to refer to the restriction of F to S. In other words, this is the function $F\restriction_S : S \to \textsc{Out}$ with $F\restriction_S(R) = F(R)$ for all $R \in S$. Second, the *union* of two functions $F : S \to \textsc{Out}$ and $F' : \textsc{Prof} \setminus S \to \textsc{Out}$ is the function $F \sqcup F'$ that behaves like F on elements of S and like F' on elements of $\textsc{Prof} \setminus S$.[22] Finally, the *product* of two families of functions $\mathcal{F} = \{ F \mid F : S \to \textsc{Out} \}$ and $\mathcal{F}' = \{ F' \mid F' : \textsc{Prof} \setminus S \to \textsc{Out} \}$ is the family $\mathcal{F} \otimes \mathcal{F}' = \{ F \sqcup F' \mid (F, F') \in \mathcal{F} \times \mathcal{F}' \}$, i.e., the set of all possible unions we can construct.

For our alternative definition of scope of an axiom A, for a given subdomain $S \subseteq \textsc{Prof}$, we look at all rules F that behave like a rule that satisfies A on profiles in S, without imposing constraints for any of the other profiles. If those rules F still satisfy A, then S must have been chosen sufficiently large. So the smallest such subdomain S, which we can obtain as the intersection of all sufficiently large subdomains, must be the scope of A. We now state this formally.

Proposition 3. *The scope of a satisfiable axiom A is equal to the intersection of all sets $S \subseteq \textsc{Prof}$ that satisfy the following constraint:*

$$\mathbb{I}(A) \quad = \quad \{ F\restriction_S \mid F \in \mathbb{I}(A) \} \otimes \{ F' \mid F' : \textsc{Prof} \setminus S \to \textsc{Out} \}$$

Proof. Let $\textsc{Agr}(A, S) = \{ F\restriction_S \mid F \in \mathbb{I}(A) \} \otimes \{ F' \mid F' : \textsc{Prof} \setminus S \to \textsc{Out} \}$ for any given set $S \subseteq \textsc{Prof}$. This is the set of voting rules that agree with axiom A (at least) on all the profiles in S. It will be handy to take note of the fact that $S \subseteq S'$ implies $\textsc{Agr}(A, S) \supseteq \textsc{Agr}(A, S')$, and also of the fact that $\textsc{Agr}(A, \textsc{Prof}) = \mathbb{I}(A)$.

The claim we need to prove is that $F =_{-R} F'$ holds for some pair $(F, F') \in \mathbb{I}(A) \times (\textsc{Out}^{\textsc{Prof}} \setminus \mathbb{I}(A))$ if and only if $R \in S$ for all $S \subseteq \textsc{Prof}$ with $\textsc{Agr}(A, S) = \mathbb{I}(A)$. We proceed by proving one direction at a time.

($\Rightarrow$) We show the contrapositive. So assume there exists an $S \subseteq \textsc{Prof}$ with $\textsc{Agr}(A, S) = \mathbb{I}(A)$ and $R \notin S$. Now consider any $F \in \mathbb{I}(A)$ and thus also $F \in \textsc{Agr}(A, S)$. As $R \notin S$, we can freely change the behaviour of F on R without affecting the fact that it belongs to $\textsc{Agr}(A, S)$, and thus $\mathbb{I}(A)$. In other words, $F' \in \mathbb{I}(A)$ for any F' with $F =_{-R} F'$, so we are done.

($\Leftarrow$) We again show the contrapositive. So assume $F' \in \mathbb{I}(A)$ whenever $F =_{-R} F'$ and $F \in \mathbb{I}(A)$. We need to find a set $S \subseteq \textsc{Prof}$ with $\textsc{Agr}(A, S) = \mathbb{I}(A)$ and $R \notin S$.

[21] To see this, observe that when A is a trivial axiom, then there exists no voting rule $F' \in \textsc{Out}^{\textsc{Prof}} \setminus \mathbb{I}(A) = \varnothing$ to begin with, and thus certainly not one that meets the stated requirement.

[22] As we can think of a function as a set of input-output pairs, $F\restriction_S$ is well-defined even when $S = \varnothing$; in that case $F\restriction_S$ is simply the empty set of input-output pairs and $F\restriction_S \sqcup F'$ is simply F'.

But $S = \textsc{Prof} \setminus \{R\}$ fits the bill: $\textsc{Agr}(A, S) = \mathbb{I}(A)$ by assumption (which says that starting from a rule F within $\mathbb{I}(A)$ we do not move outside of $\mathbb{I}(A)$ by changing the rule's behaviour on R),[23] while $R \notin S$ holds by definition of S. □

What can we say about the scope of an axiom that is the conjunction of two other axioms? First, the following basic fact is easy to verify.

Fact 4. *Let A and A' be two axioms with the property that their conjunction $A \wedge A'$ is a satisfiable axiom. Then the following inclusion holds:*

$$\mathit{scope}(A \wedge A') \;\subseteq\; \mathit{scope}(A) \,\cup\, \mathit{scope}(A')$$

To see this, observe that we can decompose any check to see whether a voting rule F satisfies $A \wedge A'$ into two separate checks, one for A and one for A'. So the conjunction of the two axioms cannot possibly talk about profiles that neither one of the two original axioms talks about. (We must exclude the case of $A \wedge A'$ being unsatisfiable simply because $\mathit{scope}(A \wedge A')$ is not well-defined in that case.)

It is tempting to believe that Fact 4 can be strengthened to say that we must have $\mathit{scope}(A \wedge A') = \mathit{scope}(A) \cup \mathit{scope}(A')$. But this would be a fallacy, as the next example clearly illustrates.

Example 6 (Scope of Combined Axioms)**.** Consider the following two axioms:

- Axiom 1: *Respect the principle of Pareto Efficiency!*
- Axiom 2: *Either respect Pareto Efficiency or always return $\{a^\star\}$!*

Clearly, Axiom 1 (i.e., Pareto Efficiency) does not have full scope, while Axiom 2 does (because, in those cases where the voting rule under consideration violates Pareto Efficiency, we must go through all profiles to ensure only alternative $a^\star$ is being returned). But the conjunction of Axiom 1 and Axiom 2 is again equivalent to Pareto Efficiency, so its scope indeed is only a proper subset of the union of the scopes of Axiom 1 and Axiom 2. △

We conclude this section with a brief discussion on the relationship between axiom scope and axiom strength. Intuition suggests that, the stronger an axiom (or an axiom set), the larger the number of profiles it talks about. This intuition is certainly correct when we consider the extremes. We already mentioned earlier that trivial axioms, i.e., axioms without any logical strength, do not talk about any profiles. And satisfiable axiom sets of *maximal strength*, i.e., satisfiable axiom sets that are so strong as to only admit a single voting rule, must talk about every single profile in the domain. Otherwise such a set would not be able to fully specify that rule. We record these two insights below.

Fact 5. *Any satisfiable axiom or axiom set of maximal strength has full scope, while any trivial axiom or axiom set has empty scope.*

[23] Observe that this step is valid only in case we can be sure that $\mathbb{I}(A)$ is nonempty. So here we are making use of the assumption that A is a satisfiable axiom.

The trivial axioms furthermore are *the only* axioms that do not talk about any profiles. But the same is not true at the other end of the spectrum. There are many axioms (e.g., Neutrality or Resoluteness) that do not have maximal strength but that still talk about every possible profile.

Finally, the intuition, which some readers might share, that the strengthening of an axiom should always go hand in hand with a (possibly just weak) broadening of the axiom's scope turns out to be ill-founded, as the following example shows.

Example 7 (Strength *vs.* Scope)**.** We again focus on the case of two alternatives and two voters. Let R_1 be the profile in which both voters rank the first alternative at the top, and let R_2 the one in which both of them rank the second alternative at the top. Consider the following two axioms:

- Axiom 1: *If you respect unanimity in R_1, then do the same in R_2!*
- Axiom 2: *If you violate unanimity in R_1, then at least respect it in R_2!*

Furthermore, let Axiom 3 be the conjunction of Axiom 1 and Axiom 2. Observe that Axiom 3 is strictly stronger than Axiom 1. This must be so, because Axiom 1 and Axiom 2 are neither trivial nor identical to one another.

Now observe that we can simplify Axiom 3: As the antecedents of Axiom 1 and Axiom 2 are complementary to one another, what Axiom 3 really is saying is that we must *always* respect unanimity when we encounter profile R_2. So the scope of Axiom 3 is $\{R_2\}$, while the scope of Axiom 1 is its superset $\{R_1, R_2\}$. In other words, strengthening Axiom 1 to Axiom 3 results in a contraction rather than an expansion of scope. △

6 The Dimensionality of Axioms

In the previous section we tried to quantify for how many profiles a given axiom has some kind of impact. But this does not fully capture the structural complexity of an axiom. Indeed, we saw that all sorts of different axioms have full scope, meaning that they have some kind of impact on every possible profile. Yet, intuition suggests that some of these axioms impose more complex constraints on voting rules than others. Indeed, some axioms, such as Pareto Efficiency, only impose constraints on one profile at a time, while others impose constraints on tuples of profiles. For instance, Positive Responsiveness evidently impose constraints on pairs of profiles. In this section, we are going to develop the machinery to measure this *dimensionality* of the constraints imposed by different axioms.

We can use this new concept of the dimensionality of an axiom to construct a hierarchy of axioms, and—partly here but especially in Section 7—we are going to relate this hierarchy to earlier proposals in the literature to structure the space of all axioms. This includes, in particular, the separation of what Fishburn (1973) calls *intraprofile* and *interprofile* axioms, and what Thomson (2023) calls *punctual* and *relational* axioms. In combination with our concept on an axiom's scope, we will also be able to make precise some of the other proposals by Fishburn to classify axioms, notably the *active*, *passive*, *universal*, and *existential* axioms.

To get started, let us go through a small thought experiment, by contemplating the operation of '*applying*' an axiom A to a profile $R \in \textsc{Prof}$. The idea is to apply every voting rule F that satisfies A to R and to collect the resulting outcomes in a set:

$$A(R) \quad = \quad \{\, F(R) \mid F \in \mathbb{I}(A) \,\}$$

The set $A(R)$ captures the range of outcomes we might encounter for profile R if we commit to axiom A. Thus, if we observe an outcome that is *not* in $A(R)$ for profile R, then we can be sure that the voting rule that has been used to obtain that outcome must violate A. Producing an outcome not in $A(R)$ for R is a *sufficient* condition for the voting rule in use violating A, but it is not a *necessary* condition. For example, if A is the axiom of Neutrality, then we have $A(R) = \textsc{Out}$ for any given profile R, so we will never observe a case where an outcome is not in $A(R)$. Yet, it of course is very much possible for a voting rule to violate Neutrality.

For axioms such as Pareto Efficiency or Condorcet Consistency the situation is rather different. For axioms A of this kind, there being a profile R for which $F(R) \notin A(R)$ really is both a sufficient and a necessary condition for voting rule F violating A. So these axioms are of a different kind than Neutrality. They are what Fishburn (1973) calls *intraprofile* axioms and what Thomson (2023) calls *punctual* axioms.[24]

To formulate a condition for axiom satisfaction that works also for other kinds of axioms, let us generalise the operation of '*applying*' an axiom A from single profiles to k-tuples $(R_1, \ldots, R_k) \in \textsc{Prof}^k$ of profiles:

$$A(R_1, \ldots, R_k) \quad = \quad \{\, (F(R_1), \ldots, F(R_k)) \mid F \in \mathbb{I}(A) \,\}$$

Thus, $A(R_1, \ldots, R_k)$ is the range of k-tuples of outcomes we might encounter when we apply voting rules that satisfy A to the profiles in the k-tuple $(R_1, \ldots, R_k)$. Given a voting rule F, the requirement for any k-tuple $(F(R_1), \ldots, F(R_k))$ to belong to $A(R_1, \ldots, R_k)$ will never become less stringent as k increases, but there will be a k above which this stringency does not increase any further.

Definition 6. *The* ***dimensionality*** *of axiom A is the smallest integer $k \geqslant 0$ that satisfies the following constraint:*

$$\mathbb{I}(A) \quad = \bigcap_{(R_1,\ldots,R_k) \in \textsc{Prof}^k} \{\, F : \textsc{Prof} \to \textsc{Out} \mid (F(R_1), \ldots, F(R_k)) \in A(R_1, \ldots, R_k) \,\}$$

In other words, the dimensionality of A is the smallest integer $k \geqslant 0$ such that any given voting rule rule $F : \textsc{Prof} \to \textsc{Out}$ that violates A can be shown to do so by inspecting some k-tuple $(R_1, \ldots, R_k) \in \textsc{Prof}^k$ of profiles and the corresponding outcomes under F. We write $dim(A)$ for the dimensionality of axiom A.

[24]As Thomson (2023) focuses on resource allocation problems rather than voting problems, it would be more accurate to say that a punctual axiom, in the sense of Thomson, talks about one *situation* at a time, rather than one *profile* at a time. But in view of our chosen model of voting, such distinctions are beyond the focus of our discussion here.

Example 8 (Dimensionality of Anonymity)**.** The dimensionality of Anonymity is 2: for every pair of profiles (R, R') such that R' can be obtained from R through permutation, we must check that outcomes do not change.

But when inspecting our formulation of Anonymity in Section 2 (Axiom 1) one could be forgiven for momentarily believing that its dimensionality might be 1 or maybe even $n!$. Indeed, only a single profile, R, is being quantified over explicitly (suggesting dimensionality 1). Once one has understood that the quantification over permutations σ amounts to another quantification over profiles $R \circ \sigma$, the fact that there are $n! - 1$ permutations of R (besides the identity mapping $R \circ \sigma$ back to R), suggests that the axiom is referring to $n!$ profiles at a time. △

This illustrates our general point that even mathematically precise definitions of axioms can sometimes be hard to interpret correctly,[25] so having a formal apparatus at hand that allows the social choice theorist to be more precise if and when needed surely will be helpful.

The dimensionality of trivial axioms is 0, given that we do not need to inspect any profiles to check satisfaction,[26] and that of unsatisfiable axioms is also 0, for the same reason.[27] The dimensionality of any axiom A is at most $k = |\textsc{Prof}|$, as any rule that is consistent with A on all profiles in $\textsc{Prof}$ must satisfy A. We say that an axiom has *maximal dimensionality* in case it has dimensionality $k = |\textsc{Prof}|$.

Example 9 (Dimensionality of Nondictatorship)**.** Nondictatorship is such an axiom with dimensionality $k = |\textsc{Prof}|$. To see this, imagine a situation where you are inspecting a voting rule F and you are unsure whether it is the dictatorship of voter 1 or whether it is a rule that returns voter 1's top-ranked alternative in all but one profile—but you do not yet know which profile that is. Then, in the worst case, you need to check $F(R)$ for every single profile $R \in \textsc{Prof}$. △

Let us now try to get a better understanding of how the dimensionalities of different axioms relate to one another. When considering conjunctions of multiple axioms, the following basic fact provides an upper bound on dimensionality.

Fact 6. *For any two axioms A and A', the dimensionality of their conjunction is bounded from above by the dimensionalities of both of them:*

$$\mathit{dim}(A \wedge A') \quad \leqslant \quad \max\{\mathit{dim}(A), \mathit{dim}(A')\}$$

[25] There are also situations where determining the exact dimensionality of an axiom is objectively difficult, as illustrated by the following example from the domain of matching with contracts, communicated to us by Kenzo Imamura. An important axiom in this area is *Path Independence*, which has the appearance of an axiom of dimensionality 3. But a result by Aizerman and Malishevski (1981) shows that it is equivalent to the conjunction of two other axioms, *Substitutability* and *Irrelevance of Rejected Contracts*, that each have dimensionality 2, meaning that Path Independence in fact must have (at most) dimensionality 2 as well (see also the forthcoming Fact 6).

[26] To understand how Definition 6 correctly captures the case of $k = 0$, it is helpful to think of tuples of length 0 as some fixed symbol ϵ. We then get $A(\epsilon) = \{\epsilon\}$ and everything falls into place.

[27] To see that Definition 6 returns the correct dimensionality of 0 for any unsatisfiable axiom A, observe that in this case $A(R_1, \ldots, R_k) = \varnothing$ for any choice of k.

What can we say about the interplay between axiom strength and axiom dimensionality? Let us first look at the extremes. Recall that a satisfiable axiom, which might be the conjunction of several axioms, is said to have *maximal strength* in case it is satisfied by just a single voting rule. *Vice versa*, let us say that a nontrivial axiom has *minimal strength* in case it only rules out a singe voting rule.

Fact 7. *Any satisfiable axiom of maximal strength has dimensionality* $k = 1$*, while any nontrivial axiom of minimal strength has dimensionality* $k = |\textsc{Prof}|$*.*

The claim about maximally strong axioms follows from the fact that to verify that the axiom is being respected we simply need to check in each profile that the single rule being forced by the axiom indeed is being followed. The claim about minimally strong axioms follows from the fact that differentiating between the one rule being excluded by the axiom and a rule that agrees with it on all but one profile requires us to inspect the outcome for all profiles.

But in between these two extremes, as we increase strength, dimensionality might go up or down. The following example is an illustration of the latter possibility.

Example 10 (May's Theorem and Dimensionality)**.** May's Theorem says that for scenarios with two alternatives (and any number of alternatives), the axioms of Anonymity, Neutrality, and Positive Responsiveness together characterise a single voting rule, namely the simple majority rule, which returnss the alternative preferred by the majority of voters—or the set of both alternatives in case of a tie (May, 1952).

Each of these three axioms clearly has dimensionality 2. Thus, by Fact 6, the dimensionality of their conjunction can be at most 2. Now, in view of May's Theorem—stating that this conjunction characterises a single voting rule and thus is maximally strong (yet satisfiable)—and Fact 7, this conjunction of axioms actually only has dimensionality 1. So this is an example where the stronger axiom (the conjunction of the three axioms) has lower dimensionality than any of the weaker axioms. △

The next question that presents itself is what we might be able to say about the interplay between axiom scope and axiom dimensionality. Due to the fact that an axiom can impose constraints only on those profiles it talks about, we immediately obtain the following bound.

Fact 8. *The dimensionality of any axiom A is bounded from above by the size of its scope:*

$$dim(A) \quad \leqslant \quad |scope(A)|$$

But beyond this basic insight, we cannot make any general statements about the relationship between dimensionality and scope. For some axioms, $dim(A)$ and $|scope(A)|$ are the same (as for Nondictatorship), while for others they can come maximally apart (as for Resoluteness).

To conclude our discussion of the dimensionality of axioms, let us briefly ponder the question of what kind of dimensionality we should be looking for in an axiom. Broadly speaking, all else being equal, axioms of low dimensionality seem to be preferable and should be expected to be ‘better behaved’.[28] This is not just a philosophical matter but also a pragmatic one. The technique of automating the discovery of proofs for impossibility theorems in social choice theory using satisfiability solvers works better for some axioms than for others (Geist and Peters, 2017), and the dimensionality of the axioms involved appears to be a relevant factor. Indeed, the first step when using this approach is to encode the axioms involved in the conjectured theorem into propositional logic for the ‘base case’ of, say, two voters and three alternatives, to then attempt to show their unsatisfiability using the solver, and to eventually extract a human-readable proof from the proof trace produced. An axiom of high dimensionality will not only result in a long formula, but that long formula—in all likelihood—will also have to interact with many other formulas to generate the impossibility we are looking for. For instance, for the base case of the Gibbard-Satterthwaite Theorem, which involves two axioms of maximal dimensionality (Nondictatorship and Nonimposition), no short mechanical proof produced by a satisfiability solver is known (Endriss, 2023).

7 Fishburn’s Classification of Axioms

Peter Fishburn in his 1973 treatise on social choice theory proposed a hierarchical classification of axioms according to which each axiom is either *existential* or *universal*, each universal axiom is either an *interprofile* or an *intraprofile* axiom, and each intraprofile axiom is either *active* or *passive*.[29] This way of organising the space of axioms has been fairly influential, and especially the distinction between intraprofile and interprofile axioms is frequently made in the literature. The classification proposed is very natural; indeed, other authors, including Richelson (1977) and Thomson (2001, 2023), have made similar proposals. However, Fishburn describes his classes of axioms only at an intuitive level, without providing clear definitions. In this section, we are going to suggest a way of rationalising Fishburn’s classification by providing precise definitions. They will be formulated in terms of our notions of dimensionality and scope.

Let us begin with the top-level division of axioms into those that are existential and those that are universal in nature:

> “The existential [axioms] are based primarily on existential qualifiers (‘there exists . . . ’) although they may also use universal qualifiers (‘for

[28]One concrete example illustrating this view has been suggested to us by Koji Yokote. In the area of transferable-utility games, the original axiomatisation by Shapley (1953) of the value now bearing his name has at times been criticised for featuring an axiom, *Additivity*, that seems less attractive than the other axioms involved. It not only has a less convincing normative justification but, being of dimensionality 3, it also has higher dimensionality than the other axioms (which all have dimensionality 1). But alternative axiomatisations of the Shapley value, such as the one due to Young (1985) with core axiom *Strong Monotonicity* or the one due to Casajus and Yokote (2017) with core axiom *Weak Differential Marginality*, show that characterising the Shapley value in terms of axioms with dimensionality not exceeding 2 is possible as well.

[29]Fishburn also talks about “*structural conditions*”, such as there being at least three alternatives or preferences being modelled as linear orders. But these conditions do not qualify as axioms, so we do not include them in our discussion here.

> all …'). The universal [axioms] either do not use existential qualifiers in any way, or else they use such qualifiers in a secondary manner."
>
> — Fishburn (1973, p. 180)

Fishburn clearly acknowledges that these are not precise definitions (he says that his formulation "allows some question about the appropriate classification of a few [axioms]"). He therefore goes on to illustrate the intent of his definitions with examples for specific axioms that belong to either one of the two classes. For the existential axioms, the clearest examples are Nonimposition[30] and Nondictatorship. Both of these axioms postulate the *existence* of a certain kind of profile, be it one where some alternative of interest wins (for Nonimposition) or be it one where some voter of interest will not have their preference be respected in the outcome (for Nondictatorship). Most other commonly used axioms are universal. Typical examples including Condorcet Consistency and Anonymity. Both of them indeed involve a condition that needs to be met by *all* (relevant) profiles.[31]

The reason why it is difficult to give a precise definition of 'existential axiom' and 'universal axiom' is that, should we want such a definition to actually refer to quantifiers used in the definition of an axiom, we would need to settle on a specific formal language for expressing axioms, and we also would need to agree on some normal form for stating axioms in that language, to ensure that there is a unique way of stating any one axiom, allowing us to count and classify quantifiers in an unambiguous manner. This seems rather daunting a task, and it might be more fruitful to instead attempt to classify axioms on the basis of their extensional semantics.

The following observation is an attempt to relate Fishburn's top-level division of the space of axioms to unambiguously measurable features of axioms.[32]

Observation 9. *All standard axioms naturally classified as 'existential' axioms according to Fishburn have maximal dimensionality, while all those naturally classified as 'universal' axioms according to Fishburn have small constant dimensionality.*

Here, an axiom having *constant* dimensionality means that its dimensionality does not depend on n (the number of voterss) or m (the number of alternatives). Indeed, most universal axioms have a dimensionality of either 1 or 2. We only know of a single case of a commonly used axiom with dimensionality 3, namely the axiom of

[30] In fact, what Fishburn calls 'Nonimposition' is a weaker axiom than the one we have defined in Section 2 (which arguably is the most prevalent definition in use today). Fishburn's axiom is what we might want to call *Weak Nonimposition*, merely requiring the voting rule to not be constant, i.e., to admit at least two distinct outcomes.

[31] Thomson (2023) uses the term 'universal axiom' in an entirely different sense, namely to refer to axioms that "express universal ideas, ideas that are applicable to all domains of problems" (p. 78). For instance, both Anonymity and Pareto Efficiency are principles that are meaningful across a wide range of different models of, amongst others, voting and resource allocation.

[32] We state this and several subsequent findings as *Observations* rather than as *Propositions* or *Facts*, because they are—at least partly—of an empirical rather than a purely analytical nature. They are attempts to link incompletely or informally defined concepts from the literature to concepts formally defined in this paper.

Reinforcement, also known as *Consistency* (Young, 1974; Zwicker, 2016).[33] We are not aware of any cases of commonly used universal axioms in the theory of voting with a dimensionality higher than that.

As for the existential axioms, recall that having *maximal dimensionality* means that these axioms have dimensionality $k = |\textsc{Prof}|$. Besides the aforementioned Nonimposition and Nondictatorship, other examples for existential axioms include those requiring each voter i to be *essential* (or *not a dummy*) in the sense of there being a pair of profiles with distinct outcomes between which only i changed their reported preference. These also have maximal dimensionality. We are not aware of any commonly used axioms with a dimensionality that is neither constant nor maximal. But it is possible to *construct* such axioms, as we shall see next.

Example 11 (Restricted Nondictatorship)**.** Consider the following weak variant of the Nondictatorship axiom, which limits the power of one specific voter $i^\star$ regarding one specific alternative $x^\star$, by postulating that $x^\star$ should not be the sole winner in *all* profiles in which $i^\star$ places $x^\star$ at the top of their reported preference:

> *A voting rule* $F : \textsc{Prof} \to \textsc{Out}$ *satisfies the axiom of Restricted Nondictatorship for voter* $i^\star \in N$ *and alternative* $x^\star \in X$ *if there exists a profile* $R \in \textsc{Prof}$ *with* $top_{i^\star}(R) = \{x^\star\}$ *but* $F(R) \neq \{x^\star\}$*.*

This axiom is existential in nature and we expect that Fishburn would have classified it as such. Its dimensionality (and also its scope) is $1/m \cdot |\textsc{Prof}|$, because in 1 out of every m profiles will $i^\star$ rank $x^\star$ in the top position. So this is an example for an axiom where the dimensionality is neither constant nor maximal. △

Another (also somewhat artificial) example for an existential axiom that does not have maximal dimensionality would be the requirement to respect unanimously held preferences for at least one unanimous profile (but not necessarily all of them). This axiom has dimensionality $m!$, given that there are $m!$ unanimous profiles, all of which one would need to inspect. Interestingly, here the dimensionality, while not constant, does not depend on n.

The aforementioned difficulty of classifying axioms as either existential or universal on the basis of the quantifiers featuring in their description is nicely illustrated by the following example—which is not artificial. This may serve as an argument for relying on a classification in terms of axiom dimensionality instead.

Example 12 (Classifying Liberalism)**.** Amartya Sen proposed a simple axiom, *Liberalism*, that encodes the idea that when choosing between several alternatives (representing different social states), for every voter there should be a pair of social states such that she can block one of them from being adopted by society (Sen, 1970). The idea is that certain aspects of a social state will be some voter's 'private business' (e.g., between two social states that are identical except that in one of them you paint the walls of your bedroom in white and in the other in pink, you should be free to block either one of them). We can define the axiom as follows:

[33] Reinforcement is an axiom for models of voting with variable electorates. It states that, if two disjoint electorates elect overlapping sets of alternatives, then the intersection of those two sets should be returned when everyone in the union of the two electorates votes.

A voting rule F : $\textsc{Prof} \to \textsc{Out}$ satisfies the axiom of Liberalism if for every voter $i \in N$ there exist two alternatives $x, y \in X$ such that (i) $y \notin F(R)$ whenever $x \succ_i^R y$ and (ii) $x \notin F(R)$ whenever $y \succ_i^R x$.

The axiom, on its own, is satisfiable as long as the number of alternatives is large enough so as to be able to find alternatives to be controlled by each individual.[34] But what kind of axiom is it? Is it an existential axiom or is it a universal axiom?

We find both existential universal quantifiers in our formulation of the axiom. The quantification over voters is universal, while that over alternatives is existential. The quantification over profiles is left implicit, but it clearly is universal (the stated conditions must hold *for all* profiles R). Thus, the axiom has the following form:

$$\forall i. \exists x. \exists y. \forall R. \mathit{Condition}(i, x, y, R)$$

The leading universal quantification over voters seems least important. Indeed, we can think of Liberalism as a conjunction of smaller axioms, each postulating Liberalism with respect to one specific voter. The conjunction presumably should belong to the same class as each of the individual axioms, so it is sufficient to analyse Liberalism with respect to one fixed voter (with quantification $\exists x. \exists y. \forall R$).

One intuition might be that the quantification over profiles should be decisive when it comes to classifying the axiom (suggesting that Liberalism is a universal axiom). But another, conflicting, intuition would be that the leading quantifier should be decisive (suggesting that Liberalism is an existential axiom). So inspecting quantifiers alone does not lead to a clear-cut classification.

If instead we check the dimensionality of Liberalism, we find that it is an axiom of maximal dimensionality: there are situations where we need to inspect how a given voting rule performs on every single profile before we can say with certainty whether it does or does not satisfy Liberalism. Thus, in view of Observation 9, we propose to classify Liberalism as an existential axiom (but see Example 13 below). $\triangle$

Another family of axioms where it is unclear whether they should be classified as existential or universal are conditions that have the form of an implication (or, equivalently, a disjunction). Consider the following type of condition, which we would not want to claim to be a truly compelling normative desideratum, but which one could at least imagine some mechanism designer out there contemplating:

If a voting rule returns the same outcome on all unanimous profiles, then it also should return that same outcome on all other profiles.

That is, if we are to ignore voters' wishes in those clearest of cases, then we should be consistent and ignore their wishes in all cases. Observe that this axiom has maximal dimensionality. Yet, to some readers, it might have a universal flavour, which seems at odds with Observation 9. But, we would like to argue, it really is a *disjunction of two axioms*, one of which is an existential axiom (of dimensionality $m!$) and one

[34] With the interpretation of alternatives as 'social states' in mind, we would expect $|X|$ to be several orders of magnitude larger than $|N|$, so this is not a significant limitation.

of which is a universal axiom (of dimensionality 2).[35] The existential sub-axiom requires us to inspect all unanimous profiles and ensure that not all of them are mapped to the same outcome. The universal sub-axiom requires us to inspect all pairs of profiles and ensure that their outcomes coincide. The overall axiom is satisfied if at least one of the two sub-axioms is.

Here is another example for a problematic (but, again, artificial) axiom:

> *If there exists a profile in which one of the alternatives you consider returning as part of the outcome has not been top-ranked by any of the voters, then the full set of alternatives should be returned for all profiles.*

That is, if we are to ignore voters' wishes (as expressed through their top choices) in some cases, then we should do so systematically in all cases. This axiom also has maximal dimensionality. On the face of it, it is an implication between one existential and one universal statement. But things become clearer once we observe that it also can be rewritten as a disjunction of two universal axioms (both of which, remarkably, have dimensionality 1). The first sub-axiom asks us to only return alternatives that are top-ranked in at least one preference. The second sub-axiom asks us to always return the full set X. Thus, in terms of dimensionality, a disjunction of two universal axioms (with small constant dimensionality) can be indistinguishable from an existential axiom (with maximal dimensionality).

We conclude that Fishburn's division into existential and universal axioms does not adequately cover the space of all conceivable axioms—some 'compound axioms' such as the last two we discussed here, arguably, are neither. Thus, if we want to also cover such axioms, organising the space of all axioms in terms of their dimensionality seems the better approach. Having said this, for all commonly used axioms one can find in the literature—and indeed all truly natural axioms we are able to conjure ourselves—Fishburn's division is perfectly adequate and our Observation 9 allows us to make that division precise.[36]

Let us now move on to the next level in Fishburn's hierarchy:

> "The main division of universal [axioms] depends on whether more than one [profile] is actively involved in the statement of the [axiom]. Those with only one [profile] are called intraprofile [axioms]; the others are interprofile [axioms]." — Fishburn (1973, p. 181)

Examples for intraprofile axioms include Pareto Efficiency and Resoluteness, while examples for interprofile axioms include Positive Responsiveness and Neutrality. Thomson makes a very similar distinction:

[35] Recall from elementary logic that an implication between two universally quantified statements is equivalent to a disjunction between an existentially quantified statement and a universally quantified one. In general, rewriting implications as disjunctions tends to make it easier to grasp the precise impact of a statement.

[36] We note that the division suggested by Observation 9 is not exhaustive, as it only covers axioms of either maximal dimensionality or (small) constant dimensionality. One could define further classes in between, such as the class of axioms with a dimensionality that is a function of m but not of n, or the class of axioms with a dimensionality that is a function of both n and m but that is not maximal.

> "A *punctual axiom* applies separately to each problem in the domain under investigation, point by point. [...] On the other hand, a *relational axiom* prescribes how a rule responds to certain changes in a parameter of a problem, perhaps changes in several parameters at once, these changes being possibly linked." — Thomson (2023, p. 79)

So Fishburn's intraprofile axioms correspond to Thomson's punctual axioms, and Fishburn's interprofile axioms correspond to Thomson's relational axioms. Richelson (1977) also identifies the same two classes, calling them *ethical* and *aggregation* axioms, respectively. So these indeed seem to be classes of some universal significance and appeal. They are naturally captured by our notion of dimensionality as well.

Observation 10. *All standard axioms naturally classified as 'intraprofile' axioms according to Fishburn have dimensionality $k = 1$, while all those naturally classified as 'interprofile' axioms according to Fishburn have dimensionality $k > 1$.*

Example 13 (Classifying Liberalism, revisited)**.** Recall Amartya Sen's axiom of Liberalism, defined in Example 12, which we had classified as an existential axiom on the grounds that it has maximal dimensionality. But maybe our interpretation of the axiom was too weak? Another natural reading would be the following:[37]

> *Suppose X includes, for each voter $i \in N$, two alternatives x_i and y_i. Then a voting rule $F : \textsc{Prof} \to \textsc{Out}$ satisfies the axiom of Liberalism$^+$ if, for every voter $i \in N$, it is the case that (i) $y_i \notin F(R)$ whenever $x_i \succ_i^R y_i$ and (ii) $x_i \notin F(R)$ whenever $y_i \succ_i^R x_i$.*

So here we are fixing from the outset who should be given control over which pair of alternatives, while in our earlier formulation in Example 12 we were free to search for a fitting assignment when trying to satisfy the axiom. In other words, this new formulation is the normatively more demanding one. Note that every assignment of labels to alternatives (every '*rights system*') results in a new version of the axiom (and this includes versions where some alternatives receive more than one label).

Now, it is straightforward to verify that Liberalism$^+$ has dimensionality 1. Indeed, this axiom clearly should be classified as a universal intraprofile axiom.[38] △

Let us now turn to Fishburn's final division:

> "The intraprofile [axioms] further divide in a natural way into [axioms] which assume certain specific properties for the components of [the profile], and those that do not. We refer to the former as active intraprofile [axioms]; the latter are passive intraprofile [axioms] since they say nothing about the contents of [the profile]." — Fishburn (1973, p. 181)

[37] Which of the two readings Sen had in mind when formulating the axiom of Liberalism is not immediately clear. His impossibility theorem is true for both of them.

[38] Clemens Puppe shared with us an anecdote according to which Amartya Sen, already back around 1970 when he published his paper on the impossibility of a Paretian liberal (Sen, 1970), felt that an attractive feature of his result—especially when compared to Arrow's Theorem (1963)—was that his axioms (Liberalism and Pareto Efficiency) only talk about one profile at a time. In contrast, Arrow's Independence axiom is an interprofile axiom of dimensionality 2 and his Nondictatorship requirement is an existential axiom of maximal dimensionality.

Thomson (2001) says that such passive axioms have *full coverage*. In his words, an axiom has full coverage if it "applies to *every* problem in the domain" (p. 353).

For both active and passive axioms, all relevant conditions can be checked one profile at a time (as all such axioms have dimensionality 1). One way of reformulating Fishburn's definition of passive axioms would be to say that a passive axiom imposes a condition on outcomes that is the same for all profiles—the condition does not depend on the profile. From amongst the common axioms defined in Section 2, Resoluteness is the only such axiom.[39] It imposes the same condition on outcomes, regardless of the profile at hand. Clear examples for intraprofile axioms that are active are Pareto Efficiency[40] and Condorcet Consistency. Both of them impose constraints on outcomes only for very specific profiles.

We can capture Fishburn's division of the intraprofile axioms into active and passive axioms using our notion of scope.[41]

Observation 11. *All standard axioms with dimensionality* 1 *naturally classified as 'passive' axioms according to Fishburn have full scope, while all those naturally classified as 'active' axioms do not have full scope.*

This completes our rationalisation of Fishburn's classification of axioms.[42] We provide a schematic overview in Figure 1.

We conclude our discussion of Fishburn's classification with a brief example illustrating once more that this classification is intended for natural axioms rather than arbitrary properties of voting rules.

Example 14 (Classifying Surjectivity)**.** Recall that Nonimposition requires that, for every alternative $x \in X$, there is a profile for which the outcome is $\{x\}$. A more demanding variant of this axiom is *Surjectivity*, which requires that, for every nonempty set $S \subseteq X$, there is a profile for which the outcome is S. Clearly, Surjectivity is existential in nature.

Observe that Surjectivity is unsatisfiable when there are more nonempty subsets of X than there are profiles, which is the case when there are $m = 3$ alternatives and $n = 1$ voter, as then there are $2^3 - 1 = 7$ such sets but only $(3!)^1 = 6$ profiles. Recall from Section 6 that unsatisfiable axioms have dimensionality 0, in conflict

[39]We note that Fishburn (1973, p. 181) classifies Resoluteness (which he refers to as *Decisiveness*) as an active axiom. The reason is that he assumes a slightly different model of voting, one where abstention is possible, and his axiom requires the rule to not report a tie only in those cases where at least one voter did not abstain. So for this variant of the axiom we indeed need to inspect the profile before being able to decide whether the axiom has been respected for that profile.

[40]We note that Thomson (2023, p. 80) classifies Pareto Efficiency as an axiom with full coverage, i.e., as a passive axiom. This is because Thomson is focusing on resource allocation problems, where Pareto Efficiency indeed can potentially be violated in every situation rather than only situations that meet certain conditions on preferences.

[41]Fishburn also includes *collective rationality conditions* in the class of passive axioms, but as mentioned in Section 2 here we prefer drawing a clear terminological distinction between such conditions and axioms in their pure sense.

[42]Of course, further subdivisions would be conceivable, and Thomson (2023) indeed proposes such a subdivision, although several of the distinguishing features he discusses are tailored to resource allocation problems and seem less well suited to classifying axioms for voting rules.

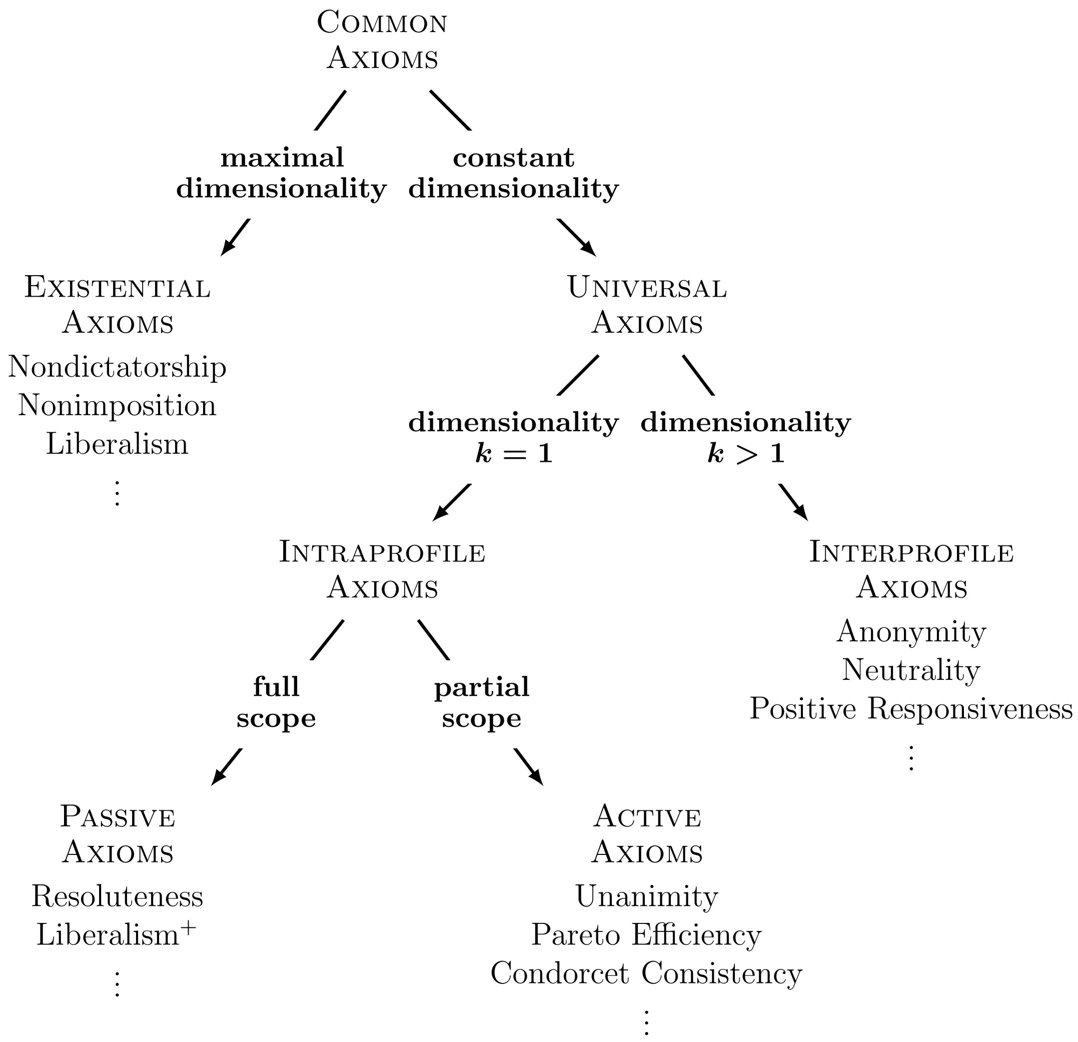

Figure 1: Rationalisation of Fishburn's Classification of Axioms.

with Observation 9 and our classification of Surjectivity as an existential axiom. (For other values of n and m, when Surjectivity is satisfiable, this anomaly vanishes.) $\triangle$

8 Conclusion and Open Problems

We have discussed different approaches to specifying the meaning of an axiom, as used in social choice theory, and we have developed a number of quantitative criteria for comparing and classifying such axioms: the strength of an axiom, the scope of an axiom, and the dimensionality of an axiom. We finally have illustrated an application of the latter two criteria by showing how they can be employed to rationalise a widely used but hitherto only informally defined classification scheme for axioms. We believe that this approach of treating axioms as formal objects with precisely measurable features offers a useful complement to the familiar approach of evaluating axioms in terms of their normative relevance and appeal.

We conclude with a brief discussion of promising directions for future work, stating a few specific open problems we believe are worth addressing.

We have developed our formal machinery for one specific model of voting only, namely the widely used model with a fixed set of voters and a fixed set of alternatives, where preferences take the form of rankings of the alternatives, and where the objective is to select a nonempty set of alternatives in response to the individual preferences reported. We believe that much of this machinery can (and should) be adapted to other models of collective decision making, be it in voting, matching under preferences, coalition formation, fair division, or judgment aggregation. In some cases this kind of adaptation might be a fairly simple technical exercise, while in others it likely will bring up new and interesting research challenges.

One specific challenge we foresee is that our definition of axiom *strength* (and the two definitions that build on it, namely those of *degree of contribution* of an axiom to a result and *degree of complementarity* of the axioms within a set) requires the set of mechanisms the axiom constrains to be finite. But in some other models of interest this set is not finite. An example is the standard model of voting with variable electorates (Young, 1974), where there are an infinite number of (finite) electorates that might report a preference, so where the set of profiles is infinite. An interesting open question thus is this: Is there a meaningful notion of axiom strength that would extend to scenarios with an infinite number of conceivable mechanisms?

Another set of research challenges relates to making the abstract measures we developed easier to use for analysts, by developing simple techniques for computing the strength, degree of contribution, degree of complementarity, scope, and dimensionality for a given axiom (set). This is particularly challenging for the measures related to axiom strength. In some cases, it may not be necessary to determine the exact strength of an axiom, but we might still wish to determine whether one axiom is stronger than another. Can we develop practical algorithms for such tasks?

The specific criteria for the analysis of axioms we developed likely are not the only such criteria worth considering. Specifically, we believe that it would be interesting to develop a theory of the *granularity* of axioms. Intuitively speaking, most common axioms naturally decompose into several *axiom instances*. For example, every way of permuting the alternatives corresponds to one instance of the axiom of Neutrality. To understand why a given mechanism fails a given axiom, it is most helpful if we are pointed to the specific axiom instance that is being violated. Indeed, axiom instances play an important role in the emerging literature on *explainability* in social choice (see, e.g., Boixel and Endriss, 2020). But, in fact, we are still lacking a precise definition of 'instance', because it is *a priori* unclear how fine-grained the decomposition of an axiom should be. The finest possible decomposition is always to rewrite an axiom as a conjunction of statements that each exclude only a single mechanism. But this rarely would be the right level of granularity. So, developing a formal theory of axiom granularity and axiom instances that matches our intuitions for how to interpret common axioms represents another important direction for future work (for initial ideas we refer to Schmidtlein, 2022, Section 2.2.3).

Finally, there is significant scope for developing formal criteria for classifying axioms in a manner that also takes some of the normative intent of an axiom into account. For example, it is natural to think of Anonymity and Neutrality as *symmetry conditions*, and of Positive Responsiveness and its many variants discussed in the literature as *monotonicity conditions* (see also Thomson, 2001). But is it possible to provide formal definitions of these and similar categories? For instance, a defining feature of a monotonicity condition seems to be that one agent moving closer to the collective decision of the moment should not affect that collective decision (or should only affect it in certain narrowly delimited ways). So making this precise would require us to fix a suitable notion of 'closeness'.

References

M. Aizerman and A. Malishevski. General theory of best variants choice: Some aspects. *IEEE Transactions on Automatic Control*, 26(5):1030–1040, 1981.

K. J. Arrow. *Social Choice and Individual Values*. John Wiley and Sons, 2nd edition, 1963. 1st edition published in 1951.

K. J. Arrow, A. K. Sen, and K. Suzumura, editors. *Handbook of Social Choice and Welfare*, volume 1. North-Holland, 2002.

A. Boixel and U. Endriss. Automated justification of collective decisions via constraint solving. In *Proceedings of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2020.

A. Casajus and K. Yokote. Weak differential marginality and the Shapley value. *Journal of Economic Theory*, 167:274–284, 2017.

U. Endriss. Logic and social choice theory. In A. Gupta and J. van Benthem, editors, *Logic and Philosophy Today*, volume 2, pages 333–377. College Publications, 2011.

U. Endriss. *Tutorial on Automated Reasoning for Social Choice Theory*. Zenodo, 2023. Available at `https://doi.org/10.5281/zenodo.10444889`.

P. C. Fishburn. *The Theory of Social Choice*. Princeton University Press, 1973.

G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und Philosophische Kritik*, 100(1):25–50, 1892.

W. V. Gehrlein and P. C. Fishburn. The probability of the paradox of voting: A computable solution. *Journal of Economic Theory*, 13(1):14–25, 1976.

C. Geist and D. Peters. Computer-aided methods for social choice theory. In U. Endriss, editor, *Trends in Computational Social Choice*, chapter 13, pages 249–267. AI Access, 2017.

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

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

W. H. Holliday, C. Norman, and E. Pacuit. Voting theory in the Lean theorem prover. In *Proceedings of the 8th International Workshop on Logic, Rationality and Interaction (LORI)*. Springer, 2021.

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

P. Mongin. L'axiomatisation et les théories économiques. *Revue Économique*, 54(1): 99–138, 2003.

H. Moulin. *The Strategy of Social Choice.* Advanced Textbooks in Economics. North-Holland, 1983.

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

E. Pacuit and F. Yang. Dependence and independence in social choice: Arrow's theorem. In S. Abramsky, J. Kontinen, J. Väänänen, and H. Vollmer, editors, *Dependence Logic*, pages 235–260. Birkhäuser, 2016.

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

H. Peters. *Game Theory: A Multi-Leveled Approach*. Springer, 2008.

C. R. Plott. Axiomatic social choice theory: An overview and interpretation. *American Journal of Political Science*, 20(3):511–596, 1976.

J. Richelson. Conditions on social choice functions. *Public Choice*, 31(1):79–110, 1977.

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.

M. C. Schmidtlein. Voting by axioms. Master's thesis, ILLC, University of Amsterdam, 2022.

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

A. K. Sen. *Collective Choice and Social Welfare: Expanded Edition*. Penguin, 2017. 1st edition published in 1970.

M. R. Sertel and A. Van der Bellen. Synopses in the theory of choice. *Econometrica*, 47(6):1367–1389, 1979.

L. S. Shapley. A value for n-person games. In H. W. Kuhn, , and A. W. Tucker, editors, *Contributions to the Theory of Games*, volume 28 of *Annals of Mathematics Studies*, pages 307–318. Princeton University Press, 1953.

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

W. Thomson. On the axiomatic method and its recent applications to game theory and resource allocation. *Social Choice and Welfare*, 18(2):327–386, 2001.

W. Thomson. Axiomatic and game-theoretic analysis of bankruptcy and taxation problems: A survey. *Mathematical Social Sciences*, 45(3):249–297, 2003.

W. Thomson. *The Axiomatics of Economic Design: An Introduction to Theory and Methods*. Springer, 2023.

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

D. van Dalen. *Logic and Structure*. Springer, 5th edition, 2013.

H. P. Young. An axiomatization of Borda's rule. *Journal of Economic Theory*, 9 (1):43–52, 1974.

H. P. Young. Monotonic solutions of cooperative games. *International Journal of Game Theory*, 14(2):65–72, 1985.

W. S. Zwicker. Introduction to the theory of voting. In F. Brandt, V. Conitzer, U. Endriss, J. Lang, and A. D. Procaccia, editors, *Handbook of Computational Social Choice*, chapter 2, pages 23–56. Cambridge University Press, 2016.

A Case Study: Shapley Value

To illustrate the concept of the Shapley value of an axiom within a given set of axioms, in this case study, we are going to calculate the degree of contribution (i.e., the Shapley value) for each of the axioms involved in a simple impossibility result.

As is well-known, for a fixed electorate of two voters and a set of two alternatives, there exists no voting rule that is resolute, anonymous, and neutral. Indeed, the symmetry requirements imposed by Anonymity and Neutrality entail that in case the two voters disagree we need to declare a tie, which however would be at odds with the requirement of Resoluteness. This is an instance of a more general result due to Moulin (1983). So let us compute the Shapley value of each of the three axioms for this simple setting.

In what follows, A denotes Anonymity, N denotes Neutrality, and R denotes Resoluteness. We first compute the strength for each of the 8 sets of axioms we can construct. There are 3 possible outcomes (the first alternative can win, the second can win, or there can be a tie). There are $2! = 2$ possible preferences a voter can report and thus $2^2 = 4$ possible profiles to consider. Hence, the overall number of voting rules is $3^4 = 81$. For each combination of axioms, we now count how many of those 81 rules satisfy those axioms. The results are shown in Table 1 and explained

Axioms	Strength	#Rules	Explanation
$\not{A}\not{N}\not{R}$	$^{0}/_{81}$	81	3 outcomes and 4 profiles, so $3^4 = 81$ rules
$A\not{N}\not{R}$	$^{54}/_{81}$	27	3 voting situations, so $3^3 = 27$ rules
$\not{A}N\not{R}$	$^{72}/_{81}$	9	2 relevant profiles, so $3^2 = 9$ rules
$\not{A}\not{N}R$	$^{65}/_{81}$	16	2 outcomes, so $2^4 = 16$ rules
$AN\not{R}$	$^{78}/_{81}$	3	only design choice is for unanimous profiles
$A\not{N}R$	$^{73}/_{81}$	8	2 outcomes and 3 voting situations, so $2^3 = 8$ rules
$\not{A}NR$	$^{77}/_{81}$	4	2 outcomes and 2 relevant profiles, so $2^2 = 4$ rules
ANR	$^{81}/_{81}$	0	impossibility result

Table 1: Strength of each of the eight subsets of the set of axioms consisting of Anonymity (A), Neutrality (N), and Resoluteness (R) for scenarios with 2 alternatives and fixed-electorates with 2 voters.

below. There are $3^3 = 27$ anonymous rules, because there are 3 'voting situations' (both voters prefer the first alternative, both prefer the second, or there is a split). There are $3^2 = 9$ neutral rules, as it is sufficient to count how many options there are in case the first voter prefers the first alternative. There are $2^4 = 16$ resolute rules, as for such rules there are only 2 possible outcomes. There are just 3 rules that are both anonymous and neutral: if the two voters disagree, such a rule must declare a tie, so we only need to count the possibilities for what to do in case the profile is unanimous (in which case the rule could agree with the voters, disagree with them, or declare a tie). Thee are 2^3 resolute rules that are anonymous (2 outcomes and 3 voting situations). There are $2^2 = 4$ resolute rules that are neutral (2 outcomes and keeping the first voter's preference fixed during counting). Finally, there are 0 rules that satisfy all three axioms (that's the statement of the impossibility result). Given the number of rules satisfying a given set of axioms, we obtain the corresponding strength by dividing by 81 and then subtracting the result from 1.

Considering the six possible sequences of axioms in the order [ANR], [ARN], [NAR], [NRA], [RAN], [RNA], we obtain the following Shapley values for the three axioms:

$$
\begin{aligned}
dcon(\mathrm{A}) &= \tfrac{1}{6}\cdot\left(\tfrac{54-0}{81}+\tfrac{54-0}{81}+\tfrac{78-72}{81}+\tfrac{81-77}{81}+\tfrac{73-65}{81}+\tfrac{81-77}{81}\right) &= \tfrac{130}{486}\\
dcon(\mathrm{N}) &= \tfrac{1}{6}\cdot\left(\tfrac{78-54}{81}+\tfrac{81-73}{81}+\tfrac{72-0}{81}+\tfrac{72-0}{81}+\tfrac{81-73}{81}+\tfrac{77-65}{81}\right) &= \tfrac{196}{486}\\
dcon(\mathrm{R}) &= \tfrac{1}{6}\cdot\left(\tfrac{81-78}{81}+\tfrac{73-54}{81}+\tfrac{81-78}{81}+\tfrac{77-72}{81}+\tfrac{65-0}{81}+\tfrac{65-0}{81}\right) &= \tfrac{160}{486}
\end{aligned}
$$

Observe that these three values add up to 1. This is a consequence of the fact that the Shapley value is known to be *efficient* (Shapley, 1953).

A possible interpretation of these values is that Neutrality has the strongest impact on the impossibility result and Anonymity has the least impact.

B Case Study: Axiom Complementarity

To illustrate the concept of complementarity of the axioms in an axiom set and the idea that results based on axiom sets with higher levels of complementarity might be classified as being more ‘surprising’, in this case study we calculate the degree of complementarity for the axiom sets involved in two well-known results, for the special case of two voters and two alternatives. The first result is the impossibility of satisfying Anonymity, Neutrality, and Resoluteness together (Moulin, 1983). The second result is May’s characterisation of the simple majority rule in terms of Anonymity, Neutrality, and Positive Responsiveness (May, 1952). So these two results differ in just one axiom.

Recall that in Appendix A we calculated the strengths of Anonymity (A), Neutrality (N), and Resoluteness (R), also for scenarios with two alternatives and two voters. The figures are shown in Table 1. As they are impossible to satisfy together, their combined strength must be 1 (rendered below as $^{81}/_{81}$). We now can compute the degree of complementarity for this set of three axioms as follows:

$$dcom(\{\mathrm{A},\mathrm{N},\mathrm{R}\}) \;=\; \frac{^{81}/_{81}}{^{54}/_{81} + {}^{72}/_{81} + {}^{65}/_{81}} \;=\; \frac{81}{191} \;\approx\; 0.424$$

How should we interpret such a number? One thing it shows is that, from a purely technical point of view, there would be significant room for weakening some or all of the three axioms without giving up on the impossibility found. Of course, whether there are such weakenings that are also *natural* is another question.

Now let us compare this to the degree of complementarity of the set of axioms involved in May’s Theorem, for the same parameters (that is, two voters and two alternatives). We still need to count how many of the 81 voting rules for this setting satisfy Positive Responsiveness (PR). Suppose the alternatives are called a and b. To indicate a voter’s preference it is sufficient to specify which of the two alternatives they prefer. There are 4 possible profiles, which for simplicity we denote as aa, ab, ba, and bb. There are 3 possible outcomes, namely $\{a\}$, $\{b\}$, and $\{a, b\}$. We distinguish five cases:

- Suppose the outcome for profile bb is $\{a\}$. Then PR fully determines the outcomes for all other profiles. So there is 1 rule for this case.
- Suppose the outcome for profile bb is $\{a, b\}$. Then PR fully determines the outcomes for all other profiles. So there is 1 rule for this case.
- Suppose the outcome for profile aa is $\{b\}$. Then PR fully determines the outcomes for all other profiles. So there is 1 rule for this case.
- Suppose the outcome for profile aa is $\{a, b\}$. Then PR fully determines the outcomes for all other profiles. So there is 1 rule for this case.
- Suppose the outcome for profile aa is $\{a\}$ and that for profile bb is $\{b\}$. Then our axioms do not impose any restrictions for which outcome to return for the other two profiles, ab and ba. So there are $3 \cdot 3 = 9$ rules for this case.

So there are $1 + 1 + 1 + 1 + 9 = 13$ voting rules overall. We are now calculate the degree of complementarity of May’s axioms, keeping in mind that the full set has strength $^{80}/_{81}$, as it admits exactly one out of the 81 possible voting rules.

$$dcom(\{\mathrm{A}, \mathrm{N}, \mathrm{PR}\}) \;=\; \frac{^{80}/_{81}}{^{54}/_{81} + {}^{72}/_{81} + {}^{68}/_{81}} \;=\; \frac{40}{97} \;\approx\; 0.412$$

Thus, we find that May's Theorem is moderately less surprising than Moulin's observation regarding the incompatibility of Anonymity and Neutrality for resolute voting rules—once again, for the specific scenario of two voters and two alternatives. Indeed, two out of three axioms are the same for the two results, while (*i*) Resoluteness is—both intuitively and formally—a weaker axiom than Positive Responsiveness and (*ii*) an impossibility is a stronger claim than a characterisation. So in the case of May's Theorem we are going from somewhat stronger assumptions to a somewhat weaker conclusion.

C Scope of Unsatisfiable Axioms

Our definition of scope only covers satisfiable axioms. This has been a deliberate choice. The question of what an appropriate definition of scope for an unsatisfiable axiom might be is debatable. Arguably, a case could be made for saying that (*i*) an unsatisfiable axiom talks about *no profiles*, that (*ii*) an unsatisfiable axiom talks about *all profiles*, and even that (*iii*) the concept of 'talking about' is *not well-defined* for unsatisfiable axioms. Here we review each of these three options in turn.

(*i*) **Empty scope:** We might want to think of an axiom's scope as an answer to the following question: "For which profiles do we need to check the outcomes returned by a voting rule to determine whether the rule satisfies the axiom?" In this case, we would have to define the scope of an unsatisfiable axiom as the empty set—meaning that such an axiom does not talk about any profiles—as we can say with confidence before having inspected any outcomes that whatever rule we are asked to check will not satisfy the axiom.

Observe that, if we extend Definition 5 in a mechanical way so as to cover also unsatisfiable axioms, they will be declared as having empty scope, because there exists no voting rule F that satisfies axiom A when A is unsatisfiable.

(*ii*) **Full scope:** But we might also want to think of an axioms's scope as an answer to this subtly different question: "For which profiles is the range of acceptable outcomes returned by a voting rule (potentially) affected by the need to satisfy the axiom?" Now we would have to define the scope of an unsatisfiable axiom as the set of all profiles, as such an axiom rules out any outcome for all profiles.

Observe that if we extend the alternative definition of scope provided by Proposition 3 in a mechanical manner to also cover unsatisfiable axioms, then such axioms will get declared as having full scope.

(*iii*) **Ill-defined scope:** The case for claiming that the concept of scope is ill-defined for unsatisfiable axioms is grounded in the fact that all unsatisfiable

> axioms are logically equivalent to one another—and thus arguably should all be talking about the same profiles. The argument goes as follows.
>
> Consider what might be the simplest unsatisfiable axiom we can construct, asking us to "select alternative a in profile R—and also *not* select a in R". When presented like this, the only natural choice for defining the axiom's scope seems to be to say that it talks about R and no other profiles. But we can reformulate the axiom, without changing its extensional semantics, to postulate that we should "select alternative a in profile R'—and also *not* select a in R'". Now the only natural choice would be to say that the axiom talks only about profile R', in direct contradiction to our earlier choice.

Of course, no axiom of practical interest will be unsatisfiable in its own right, so the relevance of the question of how to settle the status of such axioms might not seem relevant at first. But if we think of the conjunction of elementary axioms—such as the conjunction of the axioms involved in the Gibbard-Satterthwaite Theorem—as simply yet another axiom, then we sometimes will encounter unsatisfiable axioms and thus might also want to discuss their scope.

We note that the other concepts introduced in this paper—extensional meaning, strength, and dimensionality—are all well-defined for unsatisfiable axioms.