Social Mechanism Design: A Low-Level Introduction

Extended Abstract

Ben Abramowitz
Tulane University
New Orleans, LA, USA
babramow@tulane.edu

Nicholas Mattei
Tulane University
New Orleans, LA, USA
nsmattei@tulane.edu

ABSTRACT

When it comes to collective decisions, we have to deal with the fact that agents have preferences over both decision outcomes and *how* decisions are made. If we create rules for aggregating preferences over rules, and rules for preferences over rules for preferences over rules, and so on, it would appear that we run into infinite regress with preferences and rules at successively higher "levels." The starting point of our analysis is the claim that such regress should not be a problem in practice, as any such preferences will necessarily be bounded in complexity and structured coherently in accordance with some (possibly latent) normative principles. Our core contributions are (1) the identification of simple, intuitive preference structures at low levels that can be generalized to form the building blocks of preferences at higher levels, and (2) the development of algorithms for maximizing the number of agents with such low-level preferences who will "accept" a decision. We analyze algorithms for acceptance maximization in two different domains: asymmetric dichotomous choice and constitutional amendment. In both settings we study the worst-case performance of the appropriate algorithms, and reveal circumstances under which universal acceptance is possible. In particular, we show that constitutional amendment procedures proposed recently by Abramowitz et al. [2] can achieve universal acceptance.

KEYWORDS

Social Choice; Mechanism Design; Social Mechanism Design; Preference Aggregation; Voting

ACM Reference Format:

Ben Abramowitz and Nicholas Mattei. 2023. Social Mechanism Design: A Low-Level Introduction: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

In the literature on collective decision-making, a *mechanism* is, "a specification of how economic decisions are determined as a function of the information that is known by the individuals (agents) in the economy" [10]. A *mechanism designer* typically starts with an objective and a set of constraints, makes assumptions about the epistemic states and behaviors of the relevant agents, and designs a mechanism to elicit and aggregate information from the agents to produce a singular *outcome*, or state of the world, that achieves or approximates the chosen objective [6, 8]. But how are the objective and constraints determined? And who's to say whether the mechanism or decision outcome are any good? The *raison d'être* of what we call *Social Mechanism Design* (SMD) is to incorporate agents' views about mechanisms and collective choices into the identification, design, selection, and implementation of mechanisms. In SMD, the choices often relegated to mechanism designers are recognized as social choices and, where appropriate, are made by the agents themselves.

In the introduction to *The Calculus of Consent*, Buchanan and Tullock state, "The selection of a decision-making rule is itself a group choice, and it is not possible to discuss positively the basic choice-making of a social group except under carefully specified assumptions about rules. We confront a problem of infinite regression here." In Chapter 2, they state the implication explicitly, "...in discussing decision rules, we get into the familiar infinite regress if we adopt particular rules for adopting rules. To avoid this, we turn to the unanimity rule..." [4]. We find such fear of infinite regress to be unfounded. Rather, if one considers these "higher level" preferences over rules, and rules for choosing rules, etc., we claim that such preferences will necessarily be bounded in complexity and structured coherently. In other words, these preferences will not be arbitrary. For example, it is implausible that someone would insist that all voting rules should be Condorcet-consistent while the rule for choosing a Condorcet-consistent rule should necessarily be a scoring rule, and yet the rule for choosing a scoring rule must be by Plurality vote. While impossibility theorems abound and have been shown to extend to all such preference "levels" [5, 12], we believe that the inexorable structure of such preferences will circumvent many general impossibility theorems, much as single-peakedness evades Arrow's Theorem and Condorcet cycles [3, 7]. To this end, we begin investigating some plausible, intuitive structural properties at low levels. See Abramowitz and Mattei [1] for our working paper.

1.1 The Basic Idea

We propose a general framework in which a collective choice or *decision* (V, R, y) consists of a profile of information (V) provided by the agents, a rule (R) that is a function for aggregating such information, and a decision outcome $(y = R(V))$ that comes from applying the rule to the profile. The profile V is assumed to reflect agent preferences over outcomes. In our setting, agents also have preferences over decisions. In each instance exactly one decision gets made. In our low-level introduction, each agent either accepts a decision or rejects it.

Example 1.1 (Friendly Dinner). Suppose three friends are deciding where to go for dinner. Friends v_1 and v_2 prefer restaurant A while v_3 prefers restaurant B. To which restaurant should they go?

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

The problem in Example 1.1 might seem simple. Majority rule dictates that the friends should go to restaurant A. If no further information about the alternatives, agents, or their preferences is known, then majority rule is arguably the only reasonable way to decide [9]. But suppose v_3 does not believe that majority rule is justified; perhaps because of the strength of their preference or in light of the history of past decisions the friends have made together. And suppose v_1 and/or v_2 agree that majority rule is not the right way to decide. Then what?

2 GENERAL MODEL

A set of agents N with $|N| = n$ must make a collective decision. The decision consists in selecting a single outcome from a set of feasible outcomes $\mathcal{Y} = \{y_1, y_2, \ldots\}$. The agents will provide information, which altogether constitutes a *profile* V, and the decision must be made by a rule R from a set of feasible rules $\mathcal{R}$.

A *decision* is a tuple (V, R) or (V, R, y) where $R(V) = y$. If $R(V) \neq y$, then (V, R, y) is not a decision. A decision (V, R, y) is *feasible* if $R \in \mathcal{R}$ and $y \in \mathcal{Y}$. The set of all possible decisions is denoted $\tilde{\mathcal{D}}$, while the set of all feasible decisions is $\mathcal{D} \subseteq \tilde{\mathcal{D}}$. Note that we do not assume that applying a feasible rule to a profile necessarily yields a feasible outcome. Even if $R \in \mathcal{R}$, the decision (V, R, y) may not be feasible if $y \notin \mathcal{Y}$, and similarly if $y \in \mathcal{Y}$ but $R \notin \mathcal{R}$.

For each agent $i \in N$, there is a set of decisions $D_i \subseteq \tilde{\mathcal{D}}$ that i will accept. We call D_i the satisfying set of i. An agent does not accept any decision outside their satisfying set, but nothing forbids their satisfying set from including infeasible decisions. We denote by $\mathcal{D}_N = (D_1, D_2, \ldots D_n)$ the collection of agents' satisfying sets.

2.1 Objective

An instance of our problem is denoted by $I = (V, \mathcal{R}, \mathcal{Y}, \mathcal{D}_N)$. Given an instance with profile V, set of feasible rules $\mathcal{R}$ and a set of feasible outcomes $\mathcal{Y}$, our goal is to compute a feasible decision that maximizes the *acceptance rate*, or the fraction of agents who accept the decision:

$$\arg\max_{d \in \mathcal{D}} |\{i \in N : d \in D_i\}|$$

In general, we want an algorithm M to compute an acceptance-maximizing feasible decision for all instances I within some class of relevant instances $\mathcal{I}$. We are principally concerned with the best achievable worst-case acceptance rate of any algorithm M over all instances in $\mathcal{I}$.

$$\alpha_{\mathcal{I}} = \max_{M} \min_{I \in \mathcal{I}} \frac{|\{i \in N : M(I) \in D_i\}|}{n}$$

3 ANALYSIS

We categorize types of agents based on the logical conditions on rules and outcomes that are necessary and sufficient for them to accept a decision and provide algorithms for making acceptance-maximizing decisions with such agents. We introduce a key property that we call implementation-indifference which captures the way agents view the counterfactuals about what rule was implemented. We also find that the relationship between profile V and satisfying sets $\mathcal{D}_N$ is key for compactly representing preferences. For each type of agent, with and without implementation indifference, we examine the worst-case acceptance rate for asymmetric dichotomous choices. Naturally, some agent types are more conducive to maximizing acceptance than others. Our analysis confirms the intuition that when agents are more accommodating, i.e., having larger sets R_i and Y_i or being implementation-indifferent, they are easier to satisfy. Finally, we examine constitutional amendment with implementation-indifferent agents and show that amendment procedures proposed by [2] can achieve universal acceptance.

4 CONCLUSIONS

Through an approach we term *Social Mechanism Design*, we seek to incorporate agents' views about collective choice into the identification, design, selection, and implementation of collective choice mechanisms or rules. It has been suggested that accounting for agents' preferences over rules leads to infinite regress. We do not believe this is a problem because agents' full preferences over decisions across all levels will be structured and bounded in complexity. We therefore offer an exploration of some basic, intuitive structures that agents' preferences over decisions may have, and demonstrate how such preferences can be aggregated to maximize the number of agents who will accept a decision.

We have focused on the way agents balance their preferences between rules and outcomes, and how they treat counterfactuals related to the implementation of the rules. We have seen that the relationship between the agent preferences to which the rules are applied (V), and what decisions they will accept ($\mathcal{D}_N$), plays an important role. It is often reasonable to assume such a relationship when both sources of preferences are driven by the same normative principles, and this allows for more efficient elicitation. Future work may consider temporal problems, like whether satisfying sets are elicited before, after, or simultaneously with the profile. Alternatively, one might have agents express preferences over formal axioms or normative principles in natural language and need to infer their satisfying sets.

We have examined two settings, one in which the agents must make a single asymmetric binary choice, and the other in which agents must make decisions about amending a supermajority rule [1]. Many other settings are worth investigating. In particular, there are many open questions surrounding multi-winner voting. Concepts like justified representation, extended justified representation, and proportional justified representation, take the profile V can characterize what it means to represent the agents based on V [11]. In practice, the groups with which agents identify might not correspond exactly with V. Voters who identify as part of a minority group do not necessarily vote identically. In reality, voters decide for themselves whether they they have been adequately represented. Voting rules like Monroe and Chamberlain-Courant can be seen as minimizing a measure of misrepresentation, but this measure is assumed to be the same for all agents. Approaching multi-winner voting through the lens Social Mechanism Design suggests the immediate generalization to representing each voter based on what form of representation that voter cares about.

Lastly, we have treated rules as functions. In practice, agent preferences can reflect that the rules are algorithms or programs. Agents might care about the computational complexity of rules, whether rules are easy to understand, whether the rules preserve privacy, and other similar factors.

ACKNOWLEDGMENTS

Nicholas Mattei was supported by NSF Awards IIS-RI-2007955, IIS-III-2107505, and IIS-RI-2134857, as well as an IBM Faculty Award and a Google Research Scholar Award. Ben Abramowitz was supported by the NSF under Grant #2127309 to the Computing Research Association for the CIFellows Project.

REFERENCES

[1] Ben Abramowitz and Nicholas Mattei. 2022. Social Mechanism Design: A Low-Level Introduction. https://arxiv.org/abs/2211.08501

[2] Ben Abramowitz, Ehud Shapiro, and Nimrod Talmon. 2021. In the Beginning There Were n Agents: Founding and Amending a Constitution. In *International Conference on Algorithmic Decision Theory*. Springer, 119–131.

[3] Kenneth J Arrow. 2012. *Social choice and individual values*. Yale university press.

[4] James M Buchanan and Gordon Tullock. 1965. *The calculus of consent: Logical foundations of constitutional democracy*. Vol. 100. University of Michigan press.

[5] Nicolas Houy et al. 2004. *A note on the impossibility of a set of constitutions stable at different levels*. Technical Report. Université Panthéon-Sorbonne (Paris 1).

[6] Leonid Hurwicz. 1987. Inventing new institutions: the design perspective. *American Journal of Agricultural Economics* 69, 2 (1987), 395–402.

[7] Ken-ichi Inada. 1969. The simple majority decision rule. *Econometrica: Journal of the Econometric Society* (1969), 490–506.

[8] Eric S Maskin. 2008. Mechanism design: How to implement social goals. *American Economic Review* 98, 3 (2008), 567–76.

[9] Kenneth O May. 1952. A set of independent necessary and sufficient conditions for simple majority decision. *Econometrica: Journal of the Econometric Society* (1952), 680–684.

[10] Roger B Myerson. 1989. Mechanism design. In *Allocation, Information and Markets*. Springer, 191–206.

[11] Luis Sánchez-Fernández, Edith Elkind, Martin Lackner, Norberto Fernández, Jesús Fisteus, Pablo Basanta Val, and Piotr Skowron. 2017. Proportional justified representation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 31.

[12] Takahiro Suzuki and Masahide Horita. 2015. How to Order the Alternatives, Rules, and the Rules to Choose Rules: When the Endogenous Procedural Choice Regresses. In *International Conference on Group Decision and Negotiation*. Springer, 47–59.