UNIVERSAL MECHANISM DESIGN FOR AI POLICY-MAKING

Anonymous authors
Paper under double-blind review

ABSTRACT

Artificial Intelligence (AI) holds promise as a technology that can be used to improve government and economic policy-making. This paper introduces the **Universal Mechanism Design Game**, a general framework for the use of AI for automated policy-making. The framework extends mechanism design to capture a fully general economic environment, including voting on policy objectives, and gives a direction for the systematic analysis of government and economic policy through AI simulation. We highlight key open problems for future research that we hope can enable achieving social-welfare objectives in AI-based policy, thereby promoting more ethical and responsible decision making.

1 INTRODUCTION

Macroeconomic policy formulation is a domain fraught with complexity, where traditional economic models provide limited foresight into the outcomes of policy decisions. Policy-makers must not only understand the immediate implications of individual policies but also their aggregate and long-term effects. In addition, human policy-makers may not be incentive-aligned with the general public, and instead may prioritize lobbyist interests or reelection (de Figueiredo & Richter, 2014). In this light, AI-based approaches to policy design, that can simulate economies and take-up different objectives, hold potential for improved policy understanding and formulation (Zheng et al., 2020; Koster et al., 2022).

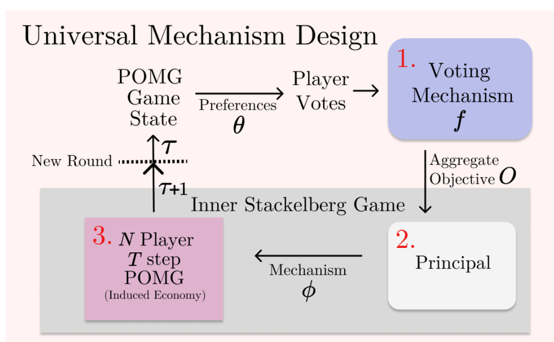

Figure 1: *The proposed game.* The process begins with voting, where human or AI players report preferences on social welfare objectives to a voting mechanism (1). This defines an objective for the Principal, who designs a mechanism (2) that parameterizes an N-player Partially Observable Markov Game (POMG). The players are the same as the voters. This POMG unfolds over several timesteps T (3). Following the POMG, game state information is extracted to initiate a new round of voting, with the last POMG state used as the first game state of the new round. This whole process is repeated for τ timesteps.

In our game, we suggest to address the concern of a misaligned policy-maker with "Voting on Values," (Hanson, 2013) coupled with a Principal who seeks to achieve the suggested policy goals. We capture the complexity of a general economic environment by modeling the economy as a Partially Observable Markov Game (POMG), similar in motivation to Agent-Based Computational Economics (Tesfatsion, 2023). Ideally, a theoretical framework tailored for AI-led policymaking in complex economic systems should balance modelling **real-world complexity** with **computational tractability**, while bringing **theoretical analyzability**. In this paper, we propose a new game which makes progress towards these desiderata and a concrete example of our game and accompanying code. We also discuss open problems within this game, taking forward a dialogue on AI's application to macroeconomic policy design. By introducing this game, we aspire to help leverage AI to assist policymakers in enhancing economic resilience and governance effectiveness.

2 Formal Definition of Universal Mechanism Design Game

We define the Universal Mechanism Design Game formally as repeatedly finding a Stackelberg Equilibrium in a Markov Game (Gerstgrasser & Parkes, 2023; Brero et al., 2022), iterated over several rounds of voting.

Definition 1. *A* **Universal Mechanism Design Game** $\mathcal{S} = (\Phi, P, \phi_0, D, \delta, \Theta, \mathcal{O}, f)$ *is a one-leader-n-follower online Stackelberg-Markov Game.*

Here, $\Phi \subseteq \mathbb{R}^k$ is the economic designer action space, and $P : \Phi \mapsto \mathcal{M}^{\phi}$ is a policy implementation map that maps from an economic designer action $\phi \in \Phi$ to a parameterized POMG $\mathcal{M}^{\phi}$. $\phi_0 \in \Phi$ is some initial action, $D : \Phi \times \Phi \mapsto \mathbb{R}_{\geqslant 0}$ is a divergence measure on the leader action space, $\delta > 0$ is the divergence constraint, and $\Theta \subseteq \mathbb{R}^{(n+1)\times m}$ is the type space. $\mathcal{O} = \{\mathcal{O}_i\}_{i \in [n]}$ is some set of predefined social welfare functions, where each $\mathcal{O}$ maps $\Phi \times \Pi \mapsto \mathbb{R}$. Finally, Π here refers to the set of all possible strategy profiles in the parameterized POMG, and $f : \Theta \mapsto \mathcal{O}$ is a social choice function Arrow (2012) representing the voting mechanism. For an optional more detailed breakdown of our model, please refer to Appendix A.

3 Example: Apple Picking Game

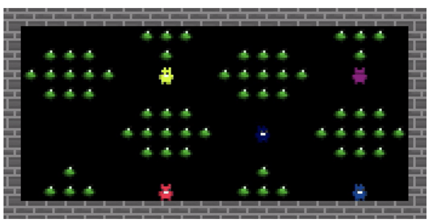

Figure 2: An example of universal mechanism design as an apple picking game, built with Melting Pot 2.0 (Agapiou et al., 2022).

In order to give a motivating example for how preference elicitation for the principal in Universal Mechanism Design can be used to align policy-maker incentives, we have created a Sequential Social Dilemma Game inspired by the *Harvest Game* proposed in Perolat et al. (2017). The aim in the Harvest game is to collect apples, with each apple yielding a reward. If all apples in an area are harvested, they never grow back. The dilemma arises when individual self-interest drives rapid harvesting, which could permanently deplete resources. Thus, agents must sacrifice personal benefit and cooperate for the collective well-being. One potential solution to this dilemma is through the use of a central government that taxes and redistributes apples. Thus, we have created a new game in which a principal designs tax rates on apple collection, and players vote on Utilitarian (productivity) vs. Egalitarian (equality) objectives for the principal, similar to Zheng et al. (2020). As players interact within this evolving environment, the principal faces the challenge of crafting policies that balance immediate economic incentives with sustainability goals. In order to achieve this, the principal must foster cooperation among players, guiding them towards the Pareto-Efficient equilibrium they have chosen. We release our code in the supplementary material for reproducibility.

4 Challenges, Open Problems and Conclusion

Based on the AI-led economic policy-making framework presented, the following key open problems of our framework are proposed for further exploration: **Preference aggregation and democratic representation** in voting mechanisms is a complex challenge that requires advanced algorithms to reflect collective preferences while respecting minority views, as well as having the simulated population be representative in the first place and their preferences correctly modeled. **Modeling human behavior** within the simulator is another key challenge, and points towards possibly incorporating bounded-rationality into MARL (Wen et al., 2019) or role-based modeling (Wang et al., 2020; 2021). To ensure responsible **AI governance and accountability**, responsible oversight mechanisms must be established, incorporating both AI and human collaboration. Furthermore, exploring socioeconomic interactions within these systems is critical, especially in understanding and deriving the conditions for **convergence to desired equilibria**, to provide a training objective for the Principal. For an optional more detailed discussion of these and other issues, please refer to Appendix B. In this paper, we present a theoretical framework for simulation that merges economic policy design with AI to potentially help better inform economic policy-making. It tackles issues such as preference aggregation and equilibria in complex economic systems. Significant challenges, including democratic representation and accountability in AI-driven systems, are highlighted. We hope to engage interdisciplinary expertise and foster collaborative innovation, and aspire to help create AI systems that not only enhance economic resilience and governance effectiveness but also uphold democratic ideals and ethical standards.

URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of the ICLR 2024 Tiny Papers Track.

REFERENCES

John P Agapiou, Alexander Sasha Vezhnevets, Edgar A Duéñez-Guzmán, Jayd Matyas, Yiran Mao, Peter Sunehag, Raphael Köster, Udari Madhushani, Kavya Kopparapu, Ramona Comanescu, et al. Melting pot 2.0. *arXiv preprint arXiv:2211.13746*, 2022.

Kenneth J. Arrow. *Social Choice and Individual Values*. Yale University Press, 2012. ISBN 978-0-300-17931-6. URL https://www.jstor.org/stable/j.ctt1nqb90.

Gianluca Brero, Alon Eden, Darshan Chakrabarti, Matthias Gerstgrasser, Vincent Li, and David C Parkes. Learning stackelberg equilibria and applications to economic design games. *arXiv preprint arXiv:2210.03852*, 2022.

John M. de Figueiredo and Brian Kelleher Richter. Advancing the Empirical Research on Lobbying. *Annual Review of Political Science*, 17(1):163–185, 2014. doi: 10.1146/annurev-polisci-100711-135308. URL https://doi.org/10.1146/annurev-polisci-100711-135308. _eprint: https://doi.org/10.1146/annurev-polisci-100711-135308.

Matthias Gerstgrasser and David C Parkes. Oracles & followers: Stackelberg equilibria in deep multi-agent reinforcement learning. In *International Conference on Machine Learning*, pp. 11213–11236. PMLR, 2023.

Robin Hanson. Shall we vote on values, but bet on beliefs? *Journal of Political Philosophy*, 21(2): 151–178, 2013.

Raphael Koster, Jan Balaguer, Andrea Tacchetti, Ari Weinstein, Tina Zhu, Oliver Hauser, Duncan Williams, Lucy Campbell-Gillingham, Phoebe Thacker, Matthew Botvinick, and Christopher Summerfield. Human-centered mechanism design with democratic ai, 2022.

Julien Perolat, Joel Z. Leibo, Vinicius Zambaldi, Charles Beattie, Karl Tuyls, and Thore Graepel. A multi-agent reinforcement learning model of common-pool resource appropriation, 2017.

Leigh Tesfatsion. Agent-based computational economics: Overview and brief history. In Ragupathy Venkatachalam (ed.), *Artificial Intelligence, Learning and Computation in Economics and Finance*, pp. 41–58. Springer International Publishing, 2023. ISBN 978-3-031-15294-8. doi: 10.1007/978-3-031-15294-8_4. URL https://doi.org/10.1007/978-3-031-15294-8_4.

Tonghan Wang, Heng Dong, Victor Lesser, and Chongjie Zhang. Roma: Multi-agent reinforcement learning with emergent roles. In *Proceedings of the 37th International Conference on Machine Learning*, 2020.

Tonghan Wang, Tarun Gupta, Anuj Mahajan, Bei Peng, Shimon Whiteson, and Chongjie Zhang. Rode: Learning roles to decompose multi-agent tasks. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2021.

Ying Wen, Yaodong Yang, Rui Luo, and Jun Wang. Modelling bounded rationality in multi-agent interactions by generalized recursive reasoning. *arXiv preprint arXiv:1901.09216*, 2019.

Stephan Zheng, Alexander Trott, Sunil Srinivasa, Nikhil Naik, Melvin Gruesbeck, David C. Parkes, and Richard Socher. The AI Economist: Improving Equality and Productivity with AI-Driven Tax Policies, April 2020. URL http://arxiv.org/abs/2004.13332. arXiv:2004.13332 [cs, econ, q-fin, stat].

APPENDIX

A DETAILS OF UNIVERSAL MECHANISM DESIGN GAME

Here we give a more detailed analysis and breakdown of our proposed game.

First we make a note regarding our type space. Since we define the first row of a specific type instantiation $\theta \in \Theta$ to be the type of the principal, Θ has $(n+1)$ rows. We thus refer to Θ_1 to be the type space of the principal and Θ_{-1} to be the type space of all participants. In addition, we remark that both the infinite-horizon and finite-horizon version of the Universal Mechanism Design Game can be considered. In constrast to standard Reinforcement Learning (RL), we do not need to introduce a discount factor for the infinite-horizon version, as our model's objective changes at each step. Thus, we consider the Principal only maximizing for the objective at the current timestep. In the finite horizon case, we add an additional time horizon $\mathcal{T}$ to our game. We now proceed to a detailed breakdown of our game.

The Universal Mechanism Design Game can be cleanly divided into a **Voting** mechanism and **Stackelberg** game, which is played with the Principal's objective determined by the Voting mechanism.

Definition 2. *The* **Voting** *mechanism is defined as* $\mathcal{S} = (\mathcal{O}, f, \Theta)$.

We use the standard axiomatic model (Arrow, 2012), where $\mathcal{O}$ is the set of alternatives, f is the social choice function, and Θ is the set of all preference profiles. Intuitively, a specific agent i's type θ_i for row i in $\theta \in \Theta$, can be thought of as some latent vector which represents the agent's values. This type contains all information necessary for recovering a partial ordering over alternatives, a more specific way of defining preferences. The goal of the Voting mechanism is then to define a objective for the Principal to optimize, given these types. To do so, we define the Voting Mechanism f and asks the players for a preference report $\theta_{-1} \in \Theta_{-1}$, which does not necessarily have to be truthful. It remains an open problem if some notion of approximate incentive compatibility can be achieved by the principal. The Voting Mechanism then computes the objective $\mathcal{O} = f(\theta_1, \theta_{-1})$ as a result of the vote. Here, we set $\theta_1 \in \Theta_1$ to be the preferences of the Principal. Importantly, the objective function includes the full θ, which allows expressing preferences of the principal if one wished to encoded a form of "moral objectivity", or other biases. We also make this modeling choice for generality, as it allows our model to express mechanisms such as auctions where the objective of the principal may be entirely selfish and not depend at all on the participant's types.

Social Welfare Examples. Examples of social welfare functions that could be included in the voting set are the Utilitarian objective $\mathcal{O}(\phi, \pi) = \sum_i J(\pi_i^\phi)$, where J is the expected discounted return $J = \sum_t (\gamma^\phi)^t r_i^\phi(s_t, a_{i,t}, a_{-i,t})$, π is the tuple of all agents $\pi = (\pi_i)_{i \in [n]}$, and π_i are singular agents that map $\Omega_i^\phi \to A_i^\phi$. Other possible choices include the Nash Welfare objective $\mathcal{O} = \left(\Pi_i J(\pi_i^\phi)\right)^{1/n}$ as well as the Egalitarian objective $\mathcal{O} = \min_i J(\pi_i^\phi)$. These are perhaps the most commonly discussed objectives, but bespoke or custom welfare functions could also be considered and added to the set of alternatives.

Definition 3. *The* **Stackelberg** *game* $I = (\Phi, P, D, \delta, \phi_0)$ *is a Stackelberg-Markov Game.*

The Stackelberg game is played subsequently after the Voting mechanism, and can be thought of as a single timestep of the full game. The economic designer (leader) will choose action $\phi \in \Phi$ which induces a parameterized **Induced Economy** $\mathcal{M}^\phi = (S, A^\phi, T^\phi, r^\phi, \Omega^\phi, O^\phi, \gamma^\phi, \mu_0^\phi)$ through the policy implementation map $P : \phi \mapsto \mathcal{M}^\phi$. Note that if agent preferences change over time, this can be modeled by adding agent types into the state space of the POMG. The transition function T would then be able to express changes in preferences over time. Thus, the objective of the leader in the Stackelberg game is then to design a POMG, given the objective $\mathcal{O}$ decided prior in the Voting mechanism: $\max_\phi \mathcal{O}(\phi, \pi; \theta)$ s.t. $D(\phi_0, \phi) \leqslant \delta, \mu_0^{P(\phi)} = \Delta(s_T)$. Again, π here is the tuple of all agents $\pi = (\pi_i)_{i \in [n]}$, and π_i individual agents that map $\Omega_i^\phi \to A_i^\phi$. Our notation $\mu_0^{P(\phi)}$ denotes the μ_0 of the tuple $P(\phi)$, and $\Delta(s_t)$ refers a Delta Dirac distribution centered on s_T. Therefore, the second constraint $\mu_0^{P(\phi)} = \Delta(s_T)$ forces the ϕ to choose a POMG that has the same initial state as the terminal state of the last round, so that continuity is kept between rounds.

Definition 4. *The* **Induced Economy** *is a Partially Observable Markov Game* $\mathcal{M}^\phi = (S, A^\phi, T^\phi, r^\phi, \Omega^\phi, O^\phi, \gamma^\phi, \mu_0^\phi)$.

Finally, the **Induced Economy** is defined as the POMG produced as the output of the principal. Agents within the POMG interact with one another and attempt to maximize their utility according to their true preferences. The n economic participants (followers) will play strategically in the parameterized POMG $\mathcal{M}^\phi$. At each step t of the game, every follower i chooses an action $a_{i,t}$ from their action space A_i, the game state evolves according to the joint action $\boldsymbol{a}_t = (a_{1,t}, \cdots, a_{n,t})$ and the transition function T, and agents receive observations and reward according to O and r. An agent's behavior in the game is characterized by its policy $\pi_i : \Omega_i^\phi \to A_i^\phi$, which maps observations to actions. Each follower in the POMG $\mathcal{M}^\phi$ individually seeks to maximize its own (discounted) total return $\sum_t (\gamma^\phi)^t r_i^\phi(s_t, a_{i,t}, a_{-i,t})$.

B Detailed Discussion of Open Problems

Here we outline a more detailed discussion of the open problems discussed in the main paper, for further consideration and thought.

Preference Aggregation and Democratic Representation.

Aggregation algorithms within the Voting Mechanism: The development of sophisticated algorithms that can effectively aggregate disparate and potentially conflicting preferences of diverse agent populations is a significant challenge. These algorithms must ensure that the outcomes represent collective preferences without overwhelming the minority views.

Incorporating diverse decision-making models: The framework must be flexible enough to respect various cultural, ethical, and socio-economic decision-making paradigms that different groups of agents might exhibit. In addition, such agents should imitate humans well.

AI Governance and Accountability.

Transparent decision-making processes: AI systems involved in policy-making need to have their decision-making processes fully transparent. The creation of interpretable AI models that can provide explanations for suggested policies is essential for trust and accountability.

Legal and ethical frameworks for AI decisions: There is an urgent need to establish legal and ethical frameworks that delineate the responsibilities and liabilities associated with AI-driven decision-making. These frameworks should set guidelines for what constitutes fair and lawful AI behavior in an economic context.

Oversight and human-AI collaboration: Establishing effective oversight mechanisms that involve both AI and human collaboration is critical. The role of human experts in supervising and guiding AI decisions and their ability to intervene when AI-driven policies deviate from desired outcomes is still to be determined.

Convergence to Desired Equilibria.

Existence and characterization of equilibria or other forms of convergence: Fundamental work is required to characterize the conditions under which a stable equilibrium might exist in such complex socio-economic interactions. The uniqueness or multiplicity of equilibria and the conditions leading to each scenario need in-depth exploration. In addition, conventional game-theoretic equilibria may not be the right object of study in the first place in such settings, as empirically these economic systems may converge somewhere else.

Algorithmic stability and multi-agent coordination: Researching advanced multi-agent reinforcement learning algorithms that can demonstrably converge is another open problem. In addition, coordination among several agents, with varying objectives and possibly divergent strategies, also remains a large open problem.

Influence of dynamic changes on convergence points: The complex dynamics of economic systems call for a deep understanding of the sensitivity of equilibria to shocks and changes in the environment and agent behavior from variables that may have been unforeseen by the principal. Ensuring the robustness and stability of the principal to be able to recover to such shocks is also of importance.