Agent-based Modeling for Policy-Making in Inequity Contexts

Alba Aguilera
Artificial Intelligence Research Institute (IIIA-CSIC), Barcelona, Spain
aaguilera@iiia.csic.es

Abstract

Recent advancements in AI have allowed for more nuanced agent-based models (ABMs). These simulations offer a non-invasive way to evaluate policies in diverse and complex social contexts. Our project focuses on developing ABMs to examine the impact of legal and social norms on inequity, particularly in scenarios where systemic discrimination affects human well-being. Our research is grounded on the Capability Approach, which provides a comprehensive framework to assess inequity in terms of real opportunities, underpinning the United Nations Sustainable Development Goals (SDGs). By defining a representative (i) agent population, (ii) agent decision-making, and (iii) agent environment, this work aims to enhance the realism of ABMs and provide valuable insights for policy-making. To achieve this, we are in the process of developing (i) a novel population synthesis method to generate agent profiles that include motivators of behavior, (ii) a decision-making model based on Markov decision processes (MDPs) that integrates values and needs in short-term and long-term rewards, and (iii) an ABM simulation to assess the impact of norms on inequity in terms of the real opportunities of individuals, among other indicators.

1 Research Problem and Motivation

In recent years, research in Artificial Intelligence (AI) has led to giant leaps forward in the field of agent-based simulations. These simulations have huge potential to be used as tools for policy-making, as their architecture allows to represent both bottom-up and top-down effects of policies in a non-invasive way. They have been applied in diverse social contexts where political action can, and should, drive meaningful change, such as in pandemics or housing crises.

Our research is centered around scenarios where inequity arises due to the presence (or absence) of legal and social norms. For agent-based models to be informative in these scenarios, they must accurately represent reality. This requires defining representative (i) agent profiles, (ii) agent decision-making, and (iii) agent environment. Our project aims to incorporate these three elements consistently to guide the evaluation of alternative policy approaches that aim to mitigate inequity, working towards the UN SDGs.

For the first element (i), agent profiles, population synthesis techniques, such as Bayesian networks, are employed. The second one (ii), agent decision-making, is defined as being driven by needs and values. We start with the needs-based model [Dignum, 2021] and then improve this by relying on a Markov decision process (MDP) framework with more complex motivators of behavior. For the third (iii), agent environment, we distinguish between the physical environment (e.g. discrete grids), and the regulatory environment (e.g. legal policies modelled with institutional syntax [Frantz and Siddiki, 2024]). Both elements constrain the agents' behavior in some way and, thus, the simulation's outcome.

2 State of the Art and Contributions

This research builds upon prior work in social simulations for policy evaluation, particularly [Montes *et al.*, 2023]'s research proposal. Their work explores how an ABM could be used to assess the impact that aporophobia [Cortina, 2017], represented within the legal policies, has on poverty levels. It adapts the needs-based model from [Dignum, 2021] to the socioeconomic problem under consideration (poverty). Our research introduces several key contributions:

Proof-of-concept Model. In this proof-of-concept [Aguilera *et al.*, 2024], (i) agent profiles are sampled from macro data, (ii) agent behavior is driven by needs, and (iii) the physical and regulatory environments are modeled using a simplified grid structure and a set of norms represented through preconditions and postconditions. This study provides innovative insights into how agent-based models can incorporate normative frameworks to evaluate the consequences of discrimination on vulnerable populations. However, technical and conceptual refinements are needed for this model to be informative in the real world.

Motivational Population Synthesis. In [Aguilera *et al.*, 2025], we present a novel population synthesis method to create synthetic populations that account for cultural variation, using Bayesian networks. This approach integrates motivational attributes, such as human values and needs, with demographic ones, such as gender, age or employment status. In this way, we can use these motivators of behavior, aligned

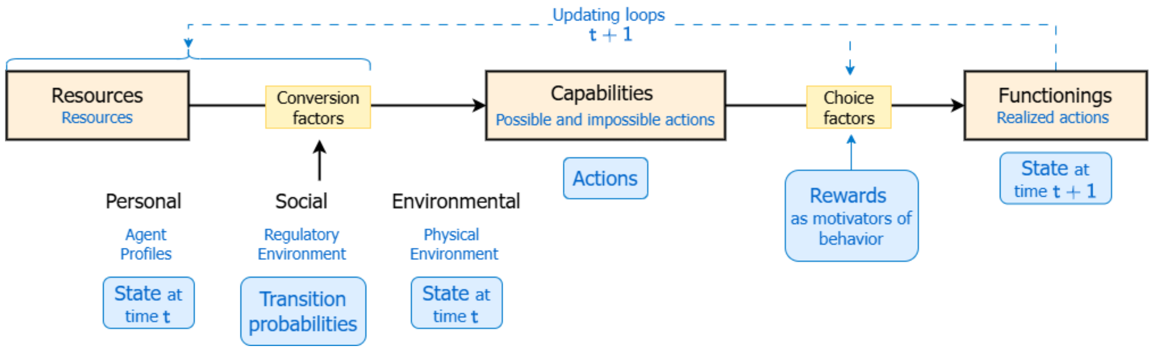

Figure 1: Overview of the CA [Robeyns, 2017] (in black and warm tones) with computational elements to operationalize it in the ABM domain within an ABM and MDP framework (in blue tones).

with the rest of the agent's profile, as input for the decision-making mechanisms in agent-based simulations.

CA-based Simulation Framework. Our latest research introduces a novel and enhanced agent-based simulation framework for policy-making. It is rooted in the Capability Approach (CA), a widely recognised framework in human development studies, used for analysing and assessing human well-being, development and social justice. Rather than looking only at resources or outcomes, it focuses on the real opportunities (capabilities) people have to lead the lives they value, such as being able to be adequately sheltered or being able to undertake employment. Policy interventions can then help restore these capabilities or even expand them.

In our framework, agent-based simulations are the tool to evaluate the impact of legal norms on people's capabilities, helping to identify the optimal ones in terms of mitigating inequity. We define and assess equity and inequity in terms of having (or lacking) central capabilities [Nussbaum, 2011], among other elements of value, in line with the CA. Beyond evaluation, we pay particular attention to the agents' decision-making. Our model defines behavior not only in terms of what people *value* or *need* to do but also in terms of what people *can* actually do given their circumstances. The CA emphasizes that our action space is constrained by resources and conversion factors, i.e. there are things we cannot do because we lack the means or the personal, social and environmental circumstances to do them. Capabilities provide this counterfactual information that is normally overlooked in decision-making architectures: they represent the set of an individual's impossible and possible actions. The choice of which possible action we realize in a particular instant of time, depends on the individual choice factors.

In Fig. 1, the key elements of the CA are mapped to the key elements of an ABM (agent profiles, decision-making, regulatory environment, etc.) and an MDP (states, actions, transition probabilities, rewards) to structure the simulation and define decision-making on a reinforcement learning basis. With this integration, we aim to enhance the modeling of human behavior and contribute to prior research into the computational representation of human values [Osman and d'Inverno, 2024]. The framework is currently being developed for a real-world case study on health inequity.

References

[Aguilera *et al.*, 2024] Alba Aguilera, Nieves Montes, Georgina Curto, Carles Sierra, and Nardine Osman. Can poverty be reduced by acting on discrimination? an agent-based model for policy making. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, pages 22–30, 2024.

[Aguilera *et al.*, 2025] Alba Aguilera, Miquel Albertí, Nardine Osman, and Georgina Curto. Value-enriched population synthesis: Integrating a motivational layer. In Nardine Osman and Luc Steels, editors, *Value Engineering in Artificial Intelligence - Second International Workshop, VALE 2024, Santiago de Compostela, Spain, October 19, 2024, Proceedings*, volume 15356 of *Lecture Notes in Computer Science*, pages 1–20. Springer, 2025.

[Cortina, 2017] Adela Cortina. *Aporofobia, el rechazo al pobre*. Paidós, Barcelona, Spain, 2017.

[Dignum, 2021] Frank Dignum. *Social simulation for a crisis*. Springer, 2021.

[Frantz and Siddiki, 2024] Christopher K. Frantz and Saba N. Siddiki. Institutional grammar 2.0 codebook, 2024. Available at https://arxiv.org/abs/2008.08937.

[Montes *et al.*, 2023] Nieves Montes, Georgina Curto, Nardine Osman, and Carles Sierra. An agent-based model for poverty and discrimination policy-making. *arXiv preprint arXiv:2303.13994*, 2023.

[Nussbaum, 2011] Martha C. Nussbaum. *Creating Capabilities: The Human Development Approach*. Harvard University Press, Cambridge, MA, USA, 2011.

[Osman and d'Inverno, 2024] Nardine Osman and Mark d'Inverno. A computational framework of human values. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2024, Auckland, New Zealand, May 6-10, 2024*, pages 1531–1539. International Foundation for Autonomous Agents and Multiagent Systems / ACM, 2024.

[Robeyns, 2017] Ingrid Robeyns. *Wellbeing, Freedom and Social Justice: The Capability Approach Re-Examined*. Open Book Publishers, Cambridge, UK, 2017.