

# Socially Intelligent Genetic Agents for the Emergence of Explicit Norms

Rishabh Agrawal<sup>1</sup>

Nirav Ajmeri<sup>2</sup>

Munindar P. Singh<sup>1</sup>

<sup>1</sup>North Carolina State University

<sup>2</sup>University of Bristol

July 2022

# Social Norms and Explanations

Contrasting applicable norms — *implicit* and *explicit*



Concert Hall, Orangerie, Schönbrunn Palace, Vienna



Messe Wien, Vienna

# Socially Intelligent Genetic Agent (SIGA)

A SIGA adapts to social context and supports meeting social expectations

- Ethical: Balances the needs of
  - Primary stakeholder, who directly interacts with the SIGA
  - Other stakeholders, who are affected by the SIGA's actions

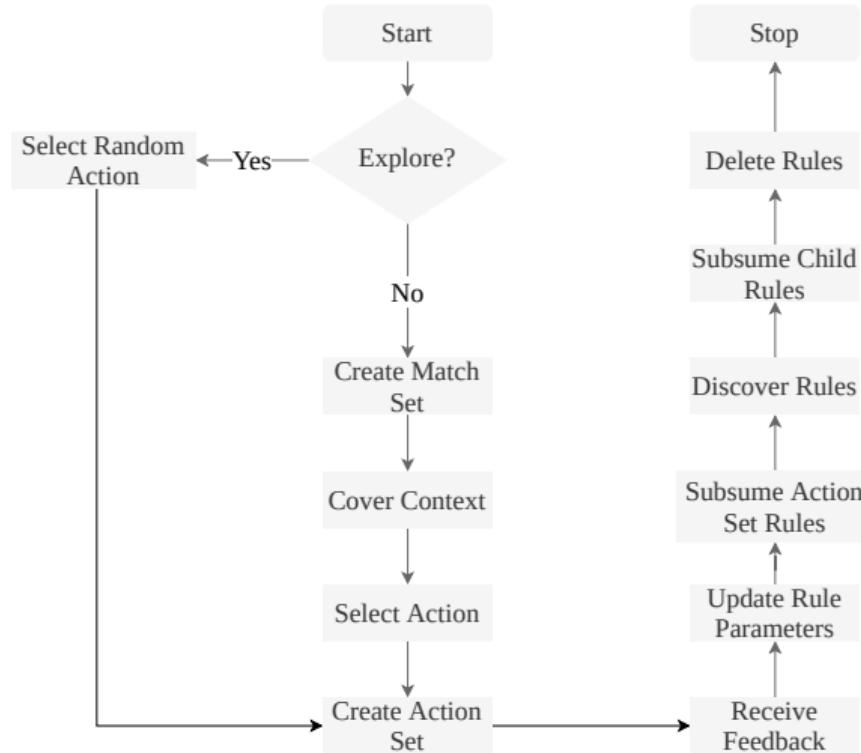
## XSIGA

Approach to engineering SIGAs that carry out enriched interactions

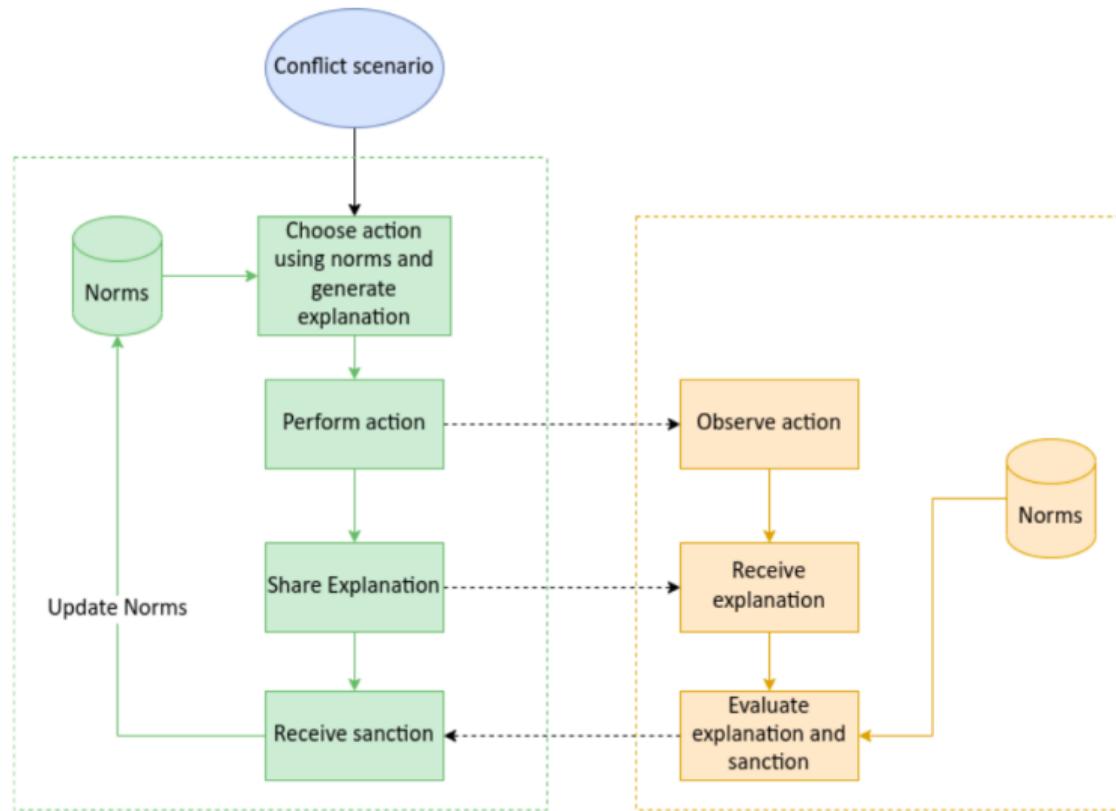
- where deviating SIGAs explain the deviations in form of explicit norms

# Realizing a SIGA

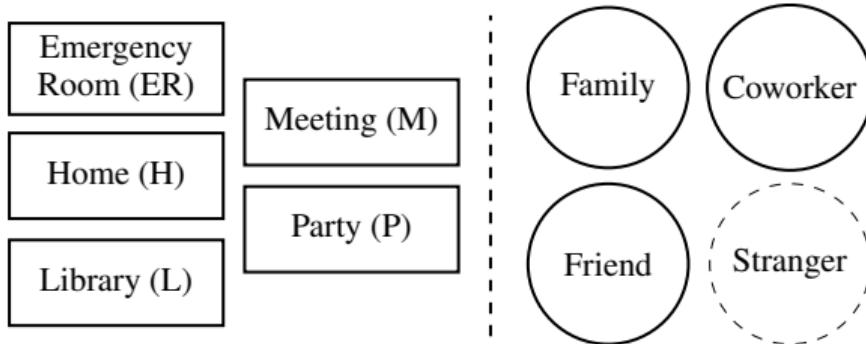
SIGAs use genetic algorithm and reinforcement learning



# Interaction and Learning in an XSIGA



# Evaluation: The Ringer Environment



Agent societies:

- Pragmatic
- Considerate
- Selfish
- Mixed

Learning strategies:

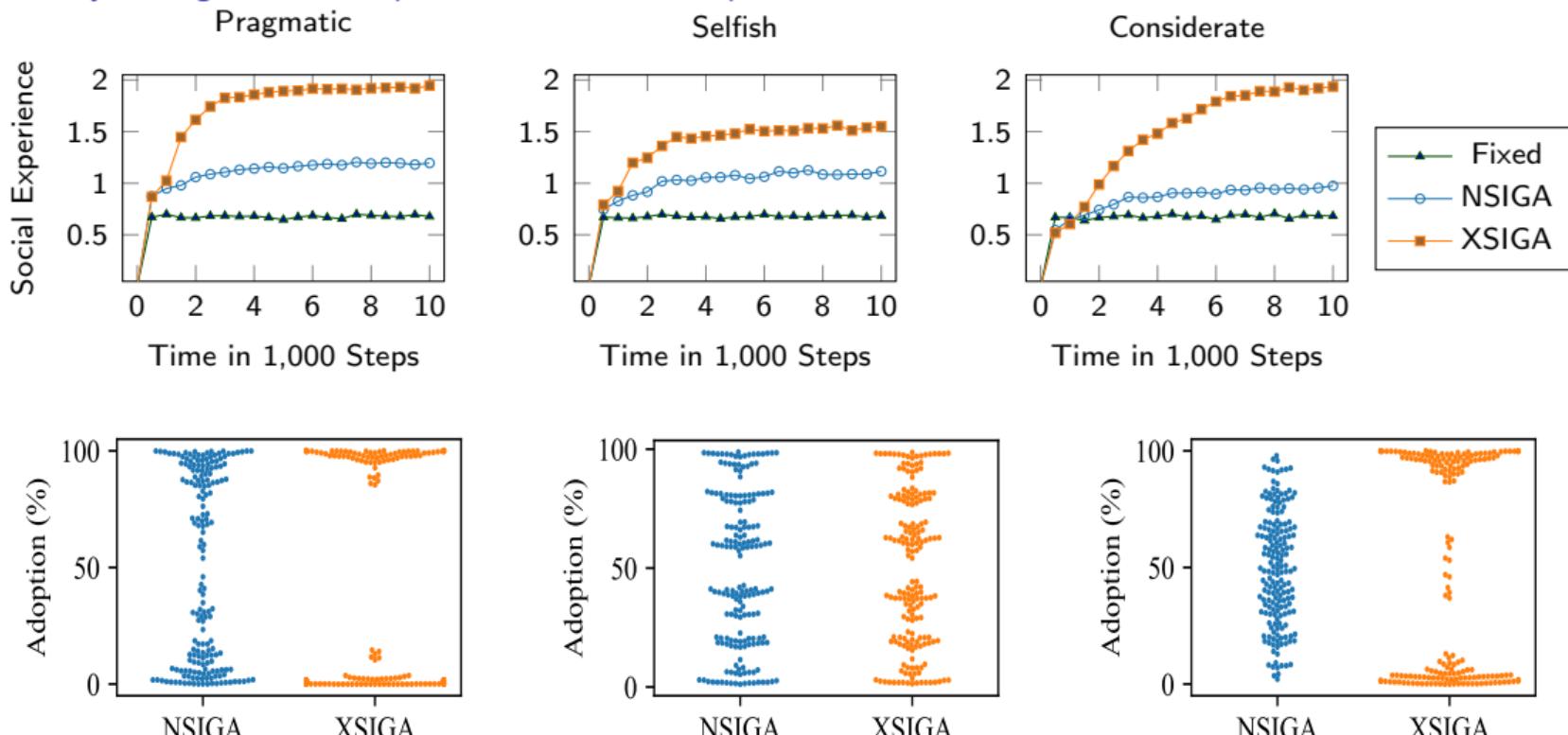
- Fixed
- NSIGA, learn by sanctions
- XSIGA, learn by sharing explanations

Metrics:

- Social experience
- Cohesion
- Adoption

# Results: Social Experience and Norm Adoption

XSIGAs yield higher social experience and norm adoption than baselines



# Conclusions and Directions

- SIGAs enable emergence of explicit norms in a decentralized system
- XSIGA yield to higher norm adoption
- Explanations in form of explicit norms lead to better social experience and cohesion

Directions:

- Abstractions for generating privacy preserving explanations
- Hierarchical ontology of terms for learning complex norm boundaries
- Incorporating values and preferences in explicit norm emergence

# Thank You

Q&A at **Stand 53 Row 2**

Contact at [nirav.ajmeri@bristol.ac.uk](mailto:nirav.ajmeri@bristol.ac.uk)

<https://niravajmeri.github.io>



## Acknowledgements:

- US National Science Foundation (grant IIS-2116751)
- University of Bristol