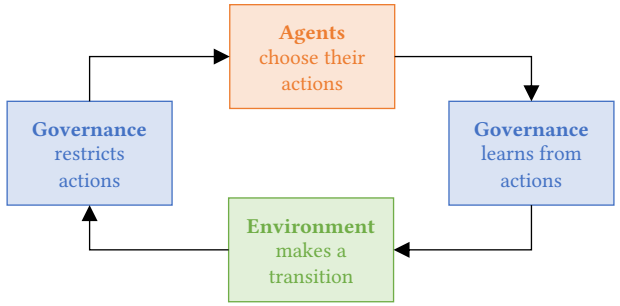
# Toward a Self-Learning Governance Loop for Competitive Multi-Attribute MAS

## Motivation

- Challenge: Govern competitive Multi-Agent Systems without destroying agent autonomy
- Problem: Agents in competitive MAS generally act strategically; therefore, actions cannot simply be equated with genuine intentions and goals
- Idea: Design a Governance component which learns to purposefully restrict agent actions in order to reach a system goal. Only use observations of actions and transitions, but do not try and derive agent goals




## Model

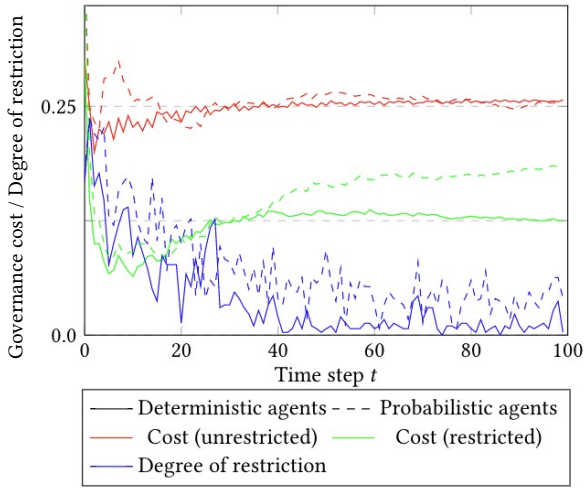
- Finite set of agents, individual action set for each agent
- Environment is a Finite-State Machine with deterministic transitions, determined by the joint action of all agents
- Action sets can be individually restricted by the Governance at each time step
- Cost function of the Governance guides the restrictions

## Approach

- Track observations as a state/action counter per agent
- Derive probabilities for the next actions from observations and compute the expected cost
- Successively remove actions with high cost and high probability until the expected cost falls below a threshold

## Results

- Governance can substantially decrease its cost by restricting agent actions
- A learning behavior can be observed for the degree of restriction which decreases over time
- Scalability is poor for naïve tracking of observed actions




## Drawbacks

- Individual treatment of agents, therefore only feasible for static agent set (no generalization possible)
- Full storage of observation history
- Complete re-evaluation of knowledge in each step

## Progress since AAMAS submission

- Model adaptation:
  - MAS is a Partially Observable Stochastic Game
  - Governance sees the MAS as the environment of a non-stationary MDP and learns from it
- This addresses both the scalability and openness problems from the original model