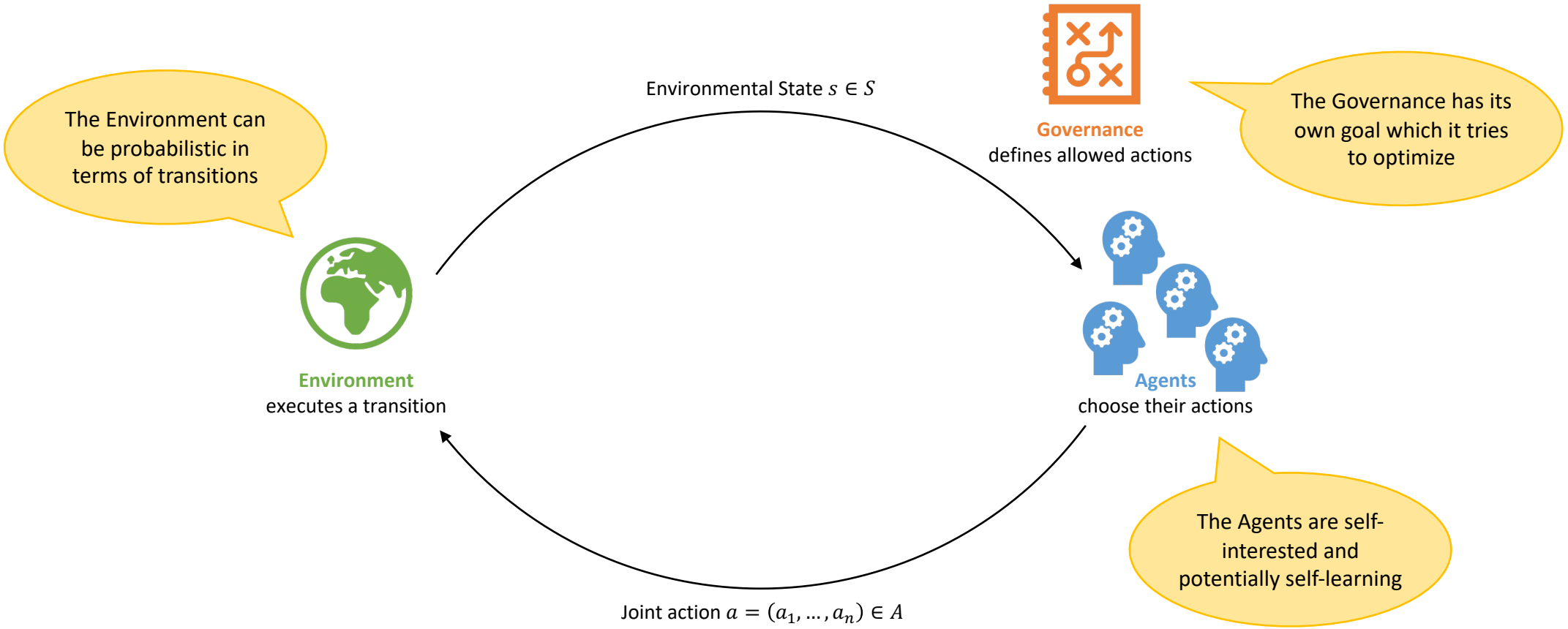
Governing Black-Box Agents in Competitive Multi-Agent Systems

Michael Pernpeintner, Christian Bartelt, Heiner Stuckenschmidt

EUMAS 2021

Governing a Competitive Multi-Agent System in a Nutshell



Motivation for Considering Agents as a Black Box

Problem	Agents in competitive MAS generally act strategically; therefore, actions cannot simply be equated with genuine intentions and goals
Challenge	Govern the system in order to achieve the system-level goal, but without destroying agent autonomy
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

Manually Tracking Agent Behavior and Deriving Reactions

EUMAS SUBMISSION

Approach/Algorithm

- 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

Main Theorem

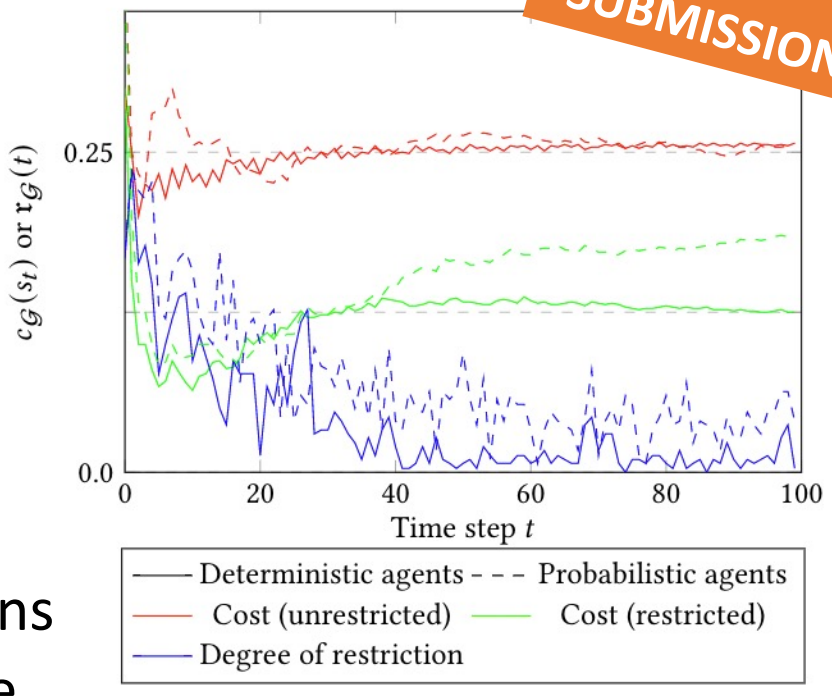
- This algorithm produces a restriction with an expected cost for the next step which is below a pre-defined value
- Moreover, it does not unnecessarily restrict the agents, i.e., the restriction is pareto-minimal

Results and Limitations

EUMAS SUBMISSION

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
- The effect gets weaker for larger sets of agents

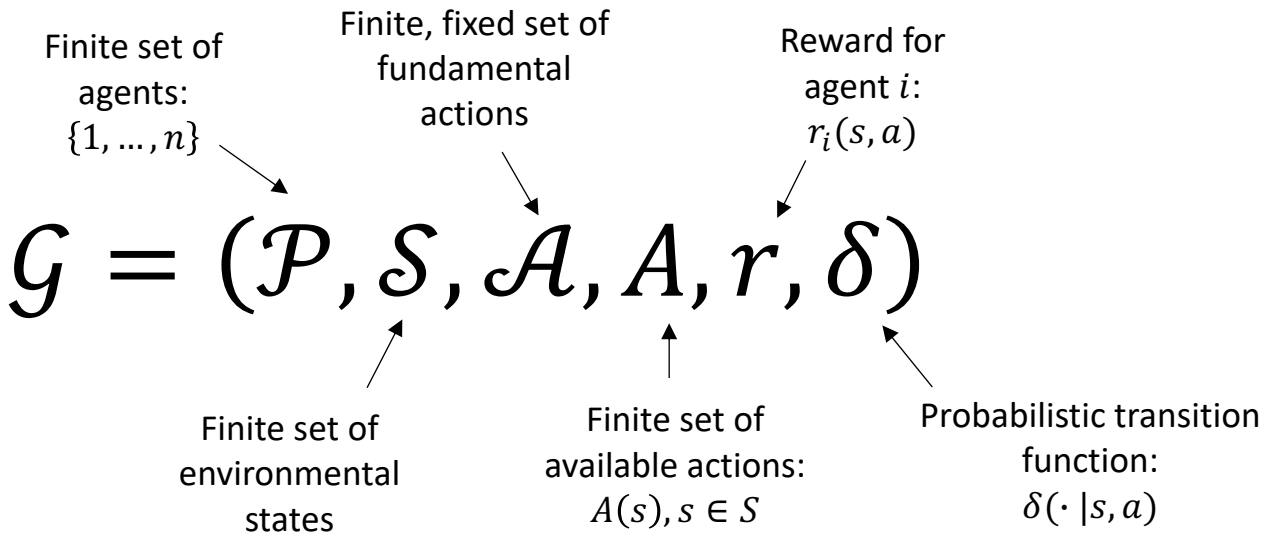


Limitations

- Scalability is poor for naïve tracking of observed actions
- Individual treatment of agents, therefore only feasible for static agent set (no generalization possible)
- Full storage of observation history and complete re-evaluation of knowledge in each step

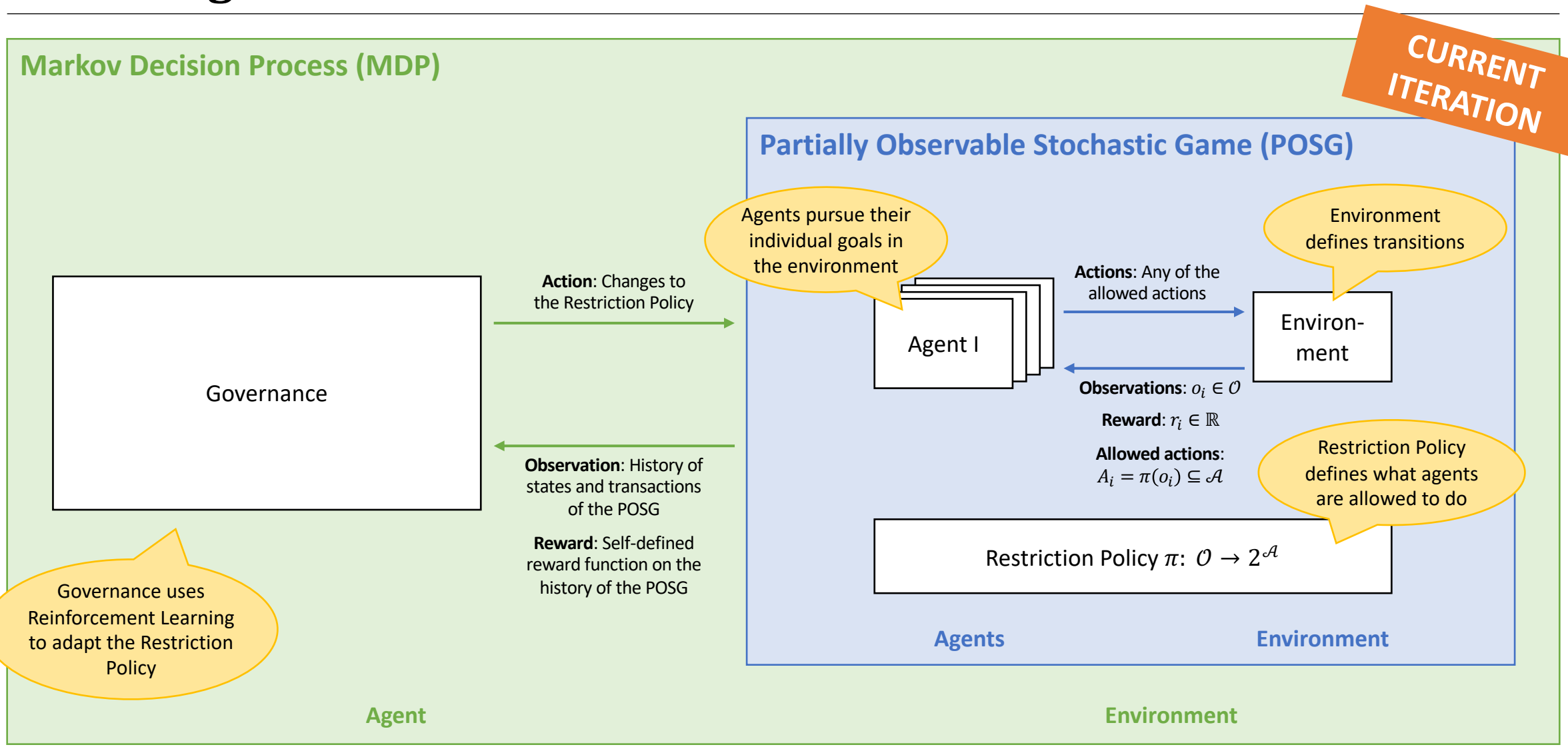
Markov Decision Processes and Stochastic Games

CURRENT ITERATION



- If there is only one agent, the system is a Markov Decision Process (MDP); otherwise, it is called a Stochastic Game
- In both concepts, observations can be a function of the full environmental state, in which case we call the system *partially observable*
- A Governance can act on this system by defining allowed (= available) actions

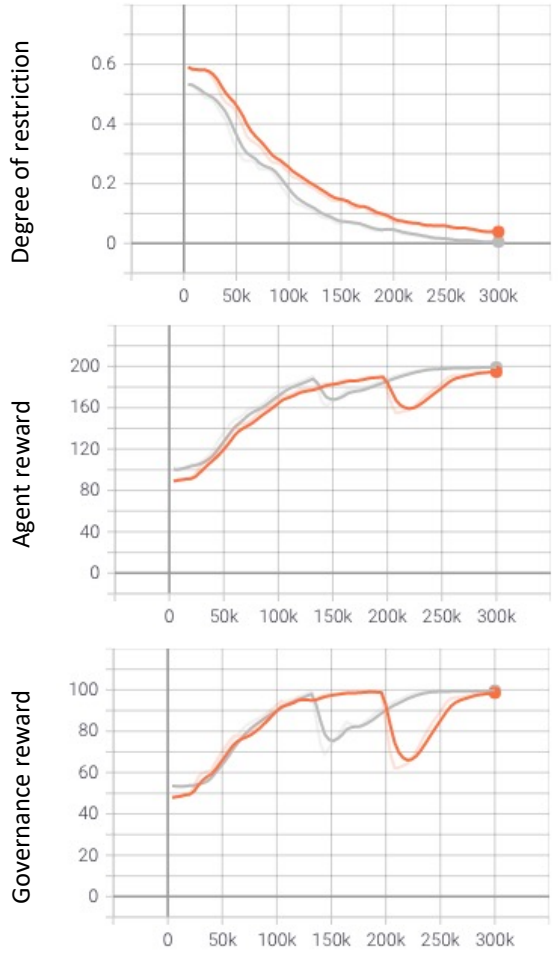
Modeling the Governance as an MDP over a Stochastic Game



Evaluation Setup and First Results

- KPIs: Agent reward, Governance reward, degree of restriction, stability of the system (in particular, stability of the restriction function)
- Only toy examples like the Coordination Game have been tested so far (see graphs to the right)
- Basic functionality can be observed already: Degree of restriction goes down, both agent rewards and Governance reward go up

Training result for the Coordination Game



I’m looking forward to your questions, comments and suggestions!

You can also reach out to me via e-mail (pernpeintner@es.uni-mannheim.de)