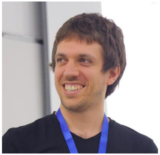
Robustness of Solutions in Game Theory:
Values and Strategies in Partially Observable, Perturbed, Stochastic, and Infinite Games



Raimundo Saona Urmeneta

Institute of Science and Technology Austria (ISTA)

July 2025

Supervisor

Krishnendu Chatterjee

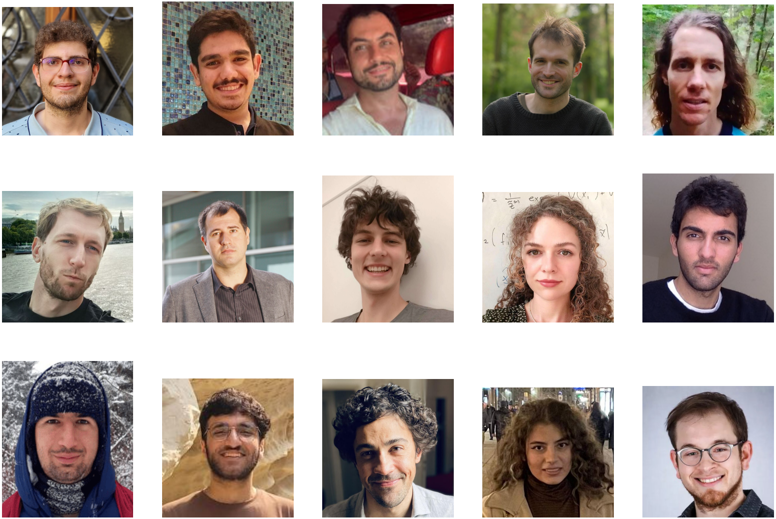
Rai, what have you done?

In the last 6 years,

- 5 journal publications
- 7 conference publications
- 5 submitted publications

- 5 research visits
- 16 talks as speaker
- 5 reviews in journals
- 11 reviews in conferences
- 7 students mentored

Coauthors



Research goal

Develop tools to analyze
dynamics involving **uncertainty**
providing **robust** conclusions

Guiding questions

1. **How to control a dynamic that can not be observed directly?**

2. How robust are game-theoretical solutions upon perturbations on the defining parameters?

3. How game-theoretic solutions inform continuous differential dynamics?

Studied models

1. **Partially Observable Markov Decision Processes**
2. Matrix games and Stochastic games
3. Random Zero-Sum Dynamic Games
   on Infinite Directed Graphs

Properties

- Partial Observation
- Perturbed Description
- Stochastic Transitions
- Infinite States

Partially Observable Markov Decision Processes

Partially Observable Markov Decision Processes

- Introduced in 1965 in optimal control.
- Applications in robot navigation, machine maintenance, artificial intelligence, automated planning, etc.
- Models planning under uncertainty.

Dynamic

At each step, the dynamic evolves as follows.

- Controller chooses an action.
- A state and a signal are drawn from a distribution that depends only on the current state and action.
- The signal is announced to the player.

Each state is assigned a reward.
The controller obtains the
liminf of the average reward of the states visited.

Previous results

- [Madani+ 2003]
  Undecidability of approximating
  the value of reachability objectives

$$\mathrm{val}(b) := \sup_{\sigma \in \Sigma} \mathbb{P}_b^\sigma \left(\exists n \geq 1 \quad S_n = \top \right) .$$

- [Rosenberg+ 2002], [Venel+ 2016]
  Existence of the uniform value

$$\mathrm{val}(b) := \sup_{\sigma \in \Sigma} \mathbb{E}_b^\sigma \left(\liminf_{n \to \infty} \frac{1}{n} \sum_{m \in [n]} r(S_m) \right) .$$

Our results

Define

$$\Sigma_0 := \{\sigma : \sigma \text{ uses finite memory } \} \ .$$

Then,

$$\mathrm{val}(b) = \sup_{\sigma \in \Sigma_0} \mathbb{E}_b^\sigma \left(\liminf_{n \to \infty} \frac{1}{n} \sum_{m \in [n]} r(S_m) \right) \ .$$

Consequences

For blind MDPs, there exists approximately optimal strategies that use finite-recall.

The decision version of the approximation problem is recursively enumerable.

The value as a mapping of the transition function is semi-lower continuous according to the relative distance.

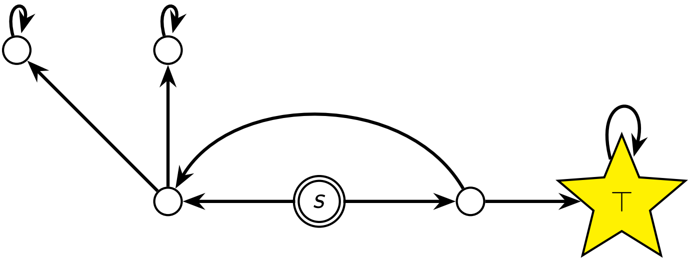
Intuition

Intuition for optimal strategies

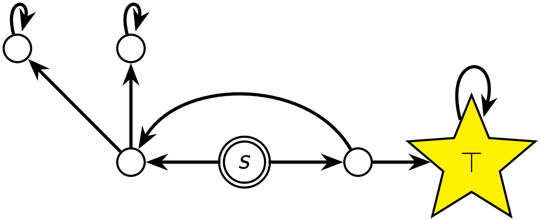
An optimal strategy does the following.

1. Guide the state to a good starting point.
2. Safely exploit your local environment.

Graphs



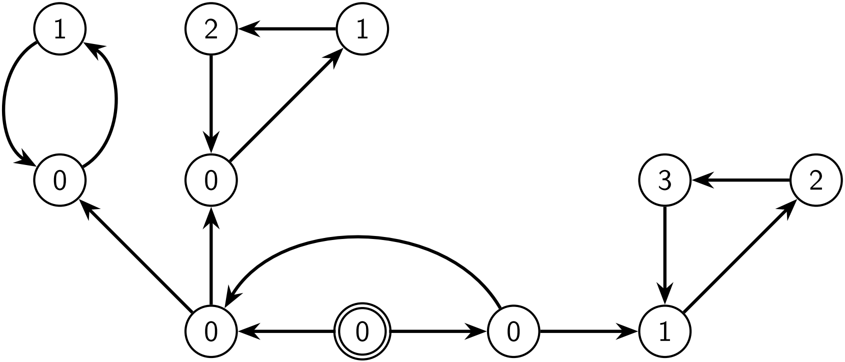
Graphs



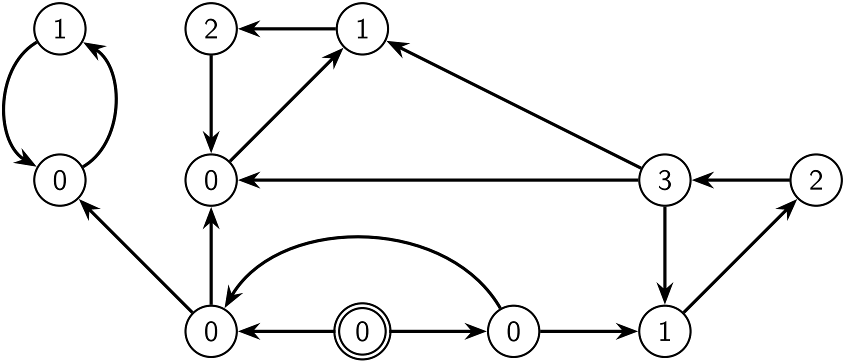
There is a strategy σ such that $\mathbb{P}^{\sigma}_{1[\top]}$-a.s.

$$\left(\frac{1}{n} \sum_{m \in [n]} 1[S_m = \top] \right) \xrightarrow[n \to \infty]{} 1 \,.$$

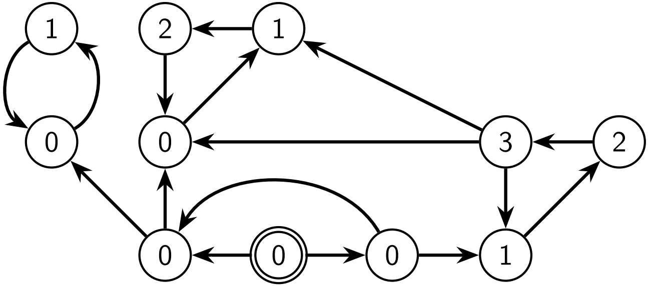
Graphs



Graphs



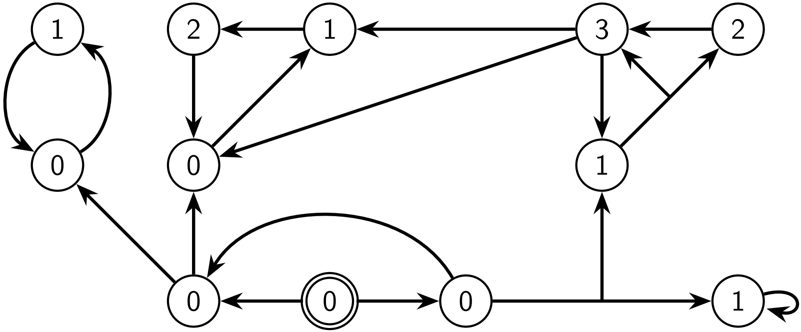
Graphs



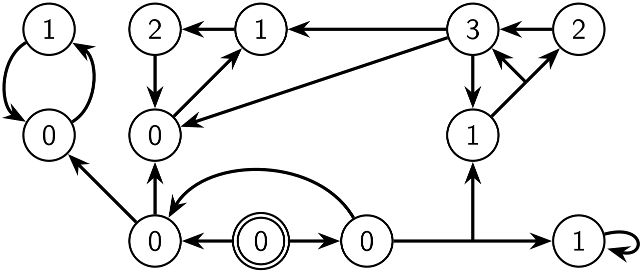
There is a strategy σ such that, for all states s in the cycle, $\mathbb{P}_s^\sigma$-a.s.

$$\left(\frac{1}{n} \sum_{m \in [n]} r(S_m) \right) \xrightarrow[n \to \infty]{} 2 \,.$$

Stochastic transitions



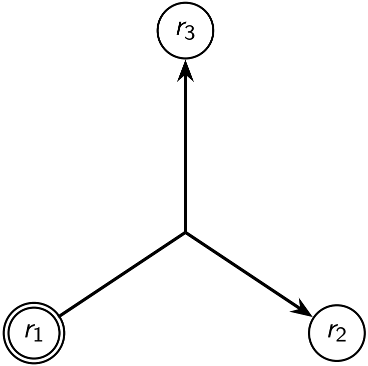
Stochastic transitions



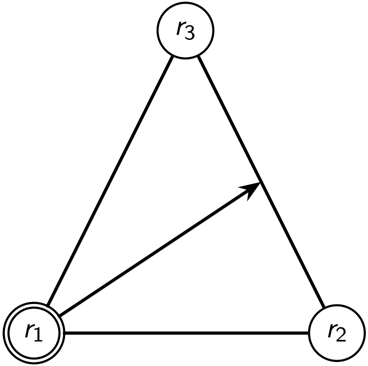
There is a strategy σ such that, for all states s in the "cycle", $\mathbb{P}_s^\sigma$-a.s.

$$\left(\frac{1}{n} \sum_{m \in [n]} r(S_m) \right) \xrightarrow[n \to \infty]{} \rho(\text{"cycle"})\,.$$

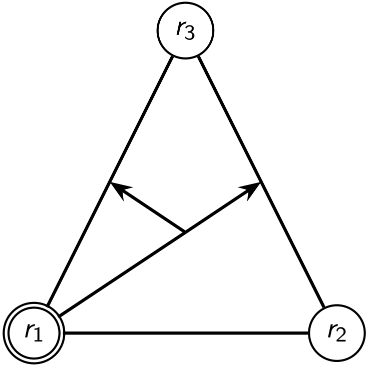
Partial observation: Beliefs



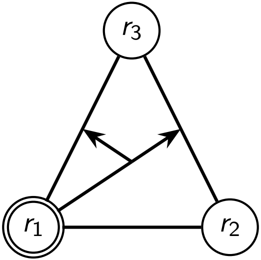
Partial observation: Beliefs



Partial observation: Random Beliefs



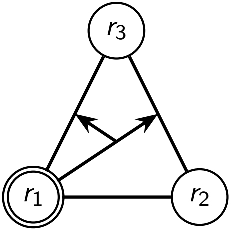
Our result



There is a strategy σ such that,
for all beliefs b in the "belief-cycle", $\mathbb{P}_b^\sigma$-a.s.

$$\left(\frac{1}{n} \sum_{m \in [n]} r(S_m) \right) \xrightarrow[n \to \infty]{} \rho(\text{"belief-cycle"}) \,.$$

Our result



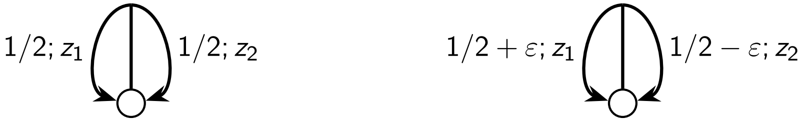
There is a strategy σ such that,
for all **states** s in the support of
a belief in the "belief-cycle", $\mathbb{P}_s^\sigma$-a.s.

$$\left(\frac{1}{n} \sum_{m \in [n]} r(S_m) \right) \xrightarrow[n \to \infty]{} \rho(s)\,.$$

Consequences for the value

The value as a mapping of the transition function is semi-lower continuous according to the relative distance.

The value as a mapping of the transition function can be discontinuous.



Other works on POMDPs

- Existence of approximately optimal strategies
  using only finite memory

- Existence of algorithms to solve
  a class of partially observable stochastic games

- Complexity of a classic objective
  under memory constraints

Thank you!