CERIAS

The Center for Education and Research in Information Assurance and Security

Rationality of Learning Algorithms in Repeated Games

Shivam Bajaj[1], Pranoy Das[1], Yevgeniy Vorobeychik[2], Vijay Gupta[1]

[1]Electrical and Computer Engineering, Purdue University [2]Computer Science and Engineering, Washington University in St. Louis

Introduction

- Game theory is a popular tool to model different types of interactions between multiple self-interested agents.
- Most learning algorithms designed for such agents in a game theoretic framework require every agent to adopt the same learning algorithm. Under this condition and for specific classes of games, these algorithms converge to an *equilibrium*, i.e., no agent has an incentive to deviate from the learning algorithm.
- An agent adopting a different learning algorithm may destabilize the system.
- A natural question is whether the agents have any incentive to adopt an alternative learning algorithm. If so, can we design algorithms in which an agent does not have any incentive to follow an alternative algorithm.

Problem Description

A game of rock-paper-Scissors

Agent 1 \ Agent 2	R	P	S
R	0,0	−1,1	1,−1
P	1,−1	0,0	−1,1
S	−1,1	1,−1	0,0

- Two player matrix games.
- Agents do not know the payoff matrix at start.
- Perfect monitoring assumption.
- A strategy profile (π_1^*, π_2^*) is a **Nash equilibrium** if

$$\mathcal{R}_1(\pi_1^*, \pi_2^*) \geq \mathcal{R}_1(\pi_1, \pi_2^*), \forall\, \pi_1 \neq \pi_1^*,$$
$$\mathcal{R}_1(\pi_1^*, \pi_2^*) \geq \mathcal{R}_1(\pi_1, \pi_2^*), \forall\, \pi_1 \neq \pi_1^*,$$

- The **value** of player i is

$$U_i(\mathcal{G}, \mathcal{A}_1, \mathcal{A}_2) = \lim_{T\to\infty} \inf E\left[\frac{1}{T}\sum_{t=0} \mathcal{R}_{1,t}\right]$$

- The **rationality ratio** of algorithm $\mathcal{A}$ is

$$s(\mathcal{A}', \mathcal{A}) \coloneqq \frac{U_1(\mathcal{A}', \mathcal{A})}{U_1(\mathcal{A}, \mathcal{A})}.$$

- For a constant $c \geq 1$, Algorithm $\mathcal{A}$ is ***c*-rational** if

$$\sup_{\mathcal{G}, \mathcal{A}'} s(\mathcal{A}', \mathcal{A}) \leq c.$$

- Algorithm $\mathcal{A}$ is **perfectly rational** if c=1.

Irrationality of Fictitious Play (FP)

Fictitious Play: *Best respond* to the empirical frequency of play of the other agent.

$$BR_i\left(\hat{a}_{-i}(t-1)\right) \coloneqq \arg\max_a \mathcal{R}_i\left(a, \hat{a}_{-i}(t-1)\right)$$

Theorem: Algorithm FP is not c-rational for any $c \geq 1$.

Rational Fictitious Play (R-GFP)

Exploration Phase: Agent i constructs E_i^t and selects action according to fictitious play. If agent $-i$ deviates, agent i enters punishment phase.

Examples of E_2^t for $t = 2$ and $t = 3$.

Agent 1 \ Agent 2		
	0	μ
	0	0

Agent 1 \ Agent 2		
	0	0
	μ	0

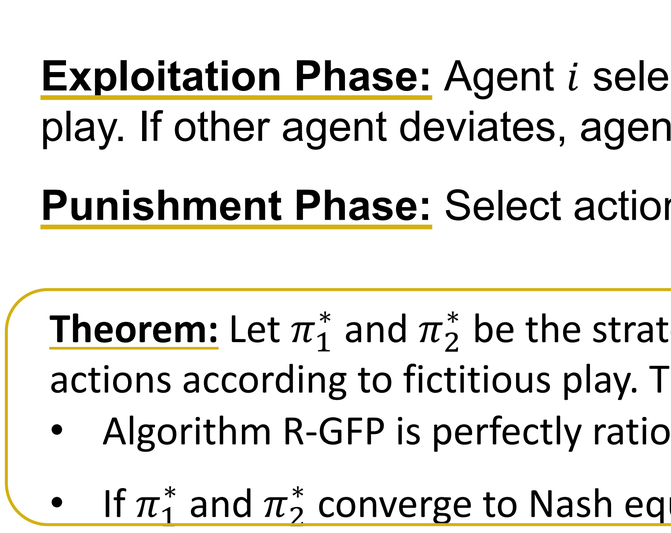
Exploitation Phase: Agent i selects action according to fictitious play. If other agent deviates, agent i enters punishment phase.

Punishment Phase: Select actions that yield lowest payoff to agent $-i$.

Theorem: Let π_1^* and π_2^* be the strategies of agent 1 and 2 if they selected actions according to fictitious play. Then,

- Algorithm R-GFP is perfectly rational if $\min_{a_2} \max_{a_1} r^1_{a_1,a_2} \leq U_1(\mathcal{A}, \mathcal{A})$
- If π_1^* and π_2^* converge to Nash equilibrium, then so does Algorithm R-GFP.

Numerical Results



Agent 1 \ Agent 2			
	(6,10)	(10,7)	(8,2)
	(5,1)	(15,8)	(7,9)

Future Work

Extension to multiple colluding agents, stochastic games, and imperfect monitoring scenarios.

PURDUE UNIVERSITY

Scan to more details of this work

CERIAS