Review 1

PC member: Gabriele Leblanc-Spanu
Time: Nov 21, 18:51 UTC

Review

This paper presents a method for replacing continuous-time best-response dynamics in zero-sum games with a Poisson-sampled best-response dynamics. This method consists of computing the best response of the agents only at instants given by a homogeneous Poisson process, and letting the agents evolve deterministically in a certain direction given by the computed best response between these instants.

The approach is well explained, and the logic used to prove its convergence and convergence rates appears coherent but may require a more rigorous justification. In particular, the link between the three main steps of the method would benefit from additional detail, as the chaining of the reasoning is not clearly presented, and some variables need to be clearly defined and explained.

The implementation of the Reinforcement Learning (RL) formulation of this approach is very helpful for understanding and illustrating the Poisson-sampled method, but it could be better connected to the first part. Some parts of the RL presentation repeat content from the first part (e.g., the “Poisson-Sampled Updates” section). Additionally, the presentation order could be improved: the model architecture is introduced after the RL formulation, while it would be more natural to introduce it beforehand, as it could justify the choice of formulation.

Overall, the approach shows promise, although it would benefit from a discussion of its applicability to other problems.

Reviewer's confidence

2: Partly, I may be missing some concepts or elements of the state-of-the-art, but I got the main idea.

Usage of LLM

1: No, not at all.

Confidential remarks for the program committee

(None provided)

Review 2

PC member: Mouad Souissi
Time: Nov 21, 19:45 UTC

Review

This study investigates the convergence properties of Poisson-sampled Best-Response (BR) dynamics in continuous zero-sum games. The authors aim to bridge the gap between ideal continuous-time BR dynamics (which converge exponentially to equilibrium via a Lyapunov function) and a more practical piecewise-deterministic system where updates occur at random times (Poisson process). The theoretical goal is to establish a convergence rate for this sampled process using stochastic approximation techniques. Finally, the paper suggests an application of this algorithm within a Reinforcement Learning (RL) framework for a swarm deception environment involving drone navigation.

Strong Points

- **Solid Theoretical Foundation:** Builds effectively on established results, specifically the Hofbauer-Sorin findings on continuous BR dynamics and Gast-Gaujal theorems regarding differential inclusions.
- **Clear Theoretical Objective and Rigorous Analytical Approach:** The main aim is mathematically precise, and the methodology is strong, relying on established convergence results from stochastic approximation literature.
- **Integration with Applied RL:** Demonstrates immediate applicability by framing the dynamics within a complex swarm deception environment using a detailed RL formulation.

Weak Points: Form (Structure and Length)

- Abstract length exceeds 2 pages.
- Summary should be reserved for the end of the finished project.
- Typographical and grammatical errors (e.g., “This is too computationally”, “the agent pick frozen best-response”).
- Figure captions contain copy-paste errors.

Weak Points: Content

- Clarity of necessary hypotheses is lacking.
- Missing definition of dynamics function (h(u)).
- Connection between theoretical convergence bound and the RL formulation is missing.
- Discussion of simulation results is absent.

Suggestions for Improvement

- Explicitly define how the error bound interacts with the Lyapunov function to show “practical asymptotic stability.”
- Justify the zero-sum assumption in the application by mapping the drone scenario to the game-theoretic notation.
- Clarify the RL integration, explaining the “Poisson-RL” mechanism adapted to the addressed problem.

Writing Quality

The structural flow is logical (Background → Aim → Methods → Expected Results). The mathematical context is excellent, and the problem is clearly stated.

Reviewer's confidence

2: Partly, I may be missing some concepts or elements of the state-of-the-art, but I got the main idea.

Usage of LLM

1: No, not at all.

Confidential remarks for the program committee

(None provided)