Dealing with Non-Stationarity in Multi-Agent Deep Reinforcement Learning

Georgios Papoudakis , Filippos Christianos , Arrasy Rahman , Stefano V. Albrecht
School of Informatics, The University of Edinburgh
{g.papoudakis, f.christianos, arrasy.rahman, s.albrecht}@ed.ac.uk

Objective of the Paper

The objective of the paper is to explore different methods for handling Non-stationarity in multi-agent deep reinforcement learning (RL). The paper discusses approaches that learn models of opponents or use them to condition the agents' policy.

Non-stationarity in Multi-Agent Deep Reinforcement Learning

Non-stationarity refers to the continuously changing decision-making policies of agents in multi-agent environments.

In the context of multi-agent deep reinforcement learning, Non-stationarity poses a challenge because the transitions and rewards depend on the actions of all agents, and these actions keep changing during the training process.

Methods discuss by paper:

1. **Centralized Critic Techniques:** This approach uses an actor-critic algorithm with a centralized critic and decentralized actors.
2. **Decentralized Learning Techniques:** Self-play is an alternative decentralized approach to handle Non-stationarity. Agents train by playing against their current or previous versions, allowing them to learn policies that can generalize to any opponents.
3. **Meta-Learning:** The paper proposes a meta-learning approach that outperforms other adaptation methods in single-agent environments.
4. **Communication:** Communication methods allow training agents to exchange information about their observations, actions, and intentions to stabilize their training.
5. **Opponent Modeling:**One approach uses a shared network to predict the behavior of other agents, while the other approach uses a separate network to encode the opponent's behavior.

Research Questions in Multi-Agent Deep RL

1.Transfer Learning for Non-Stationarity:
How can agents effectively transfer learned knowledge?
How to leverage it for quicker adaptation to non-stationarity
2. Limited Access to Opponent Information:
How can agents Learn effective policies with limited or no access to opponent actions?

Learning Against Non-Stationary Agents with Opponent Modelling and Deep Reinforcement Learning

Richard Everett, Stephen Roberts
Department of Engineering Science
University of Oxford
{richard, sjrob}@robots.ox.ac.uk

Objective of the Paper

The objective of the paper is to address the challenge of learning in the presence of non-stationary agents in a decentralized learning problem.

The paper introduces the Switching Agent Model (SAM), which combines traditional deep reinforcement learning with opponent modeling and uncertainty estimation.

The paper demonstrates the success of SAM in a continuous-action environment and highlights the benefits of opponent modeling and uncertainty estimation in improving performance.