CSCI6505 2021W P1: Update

Jasper Dupuis
Mani Teja Varma Kucherlapati

February 27, 2021

1 Status update

Topic: Iterated social dilemmas with imperfect communication using reinforcement learning models.

Proposal: `https://github.com/jasperdupuis/IPD-comms/blob/main/documents/csci_6505_proposal_dupuis_kucherlapati.pdf`

Git: `https://github.com/jasperdupuis/IPD-comms`

The next steps from the proposal stage were identified as:

1. Read and become familiar with the references included in the proposal,
2. Develop familiarity with Axelrod library to run the iterated prisoner's dilemma tournament,
3. In parallel, develop and use RL training models in the tournament.

We've accomplished these three points. The code for our implementations is available on the git at the top of this page. We first fixed the default q-learner in the Axelrod library, which had a weird state table set up. We then implemented a simple DQN learner, which had no memory from game to game. Then a DQN learner with inter-game memory was implemented.

The structure of the tournament and players provided by the Axelrod library suggests a way ahead to implement communication between players. We will try to derive a new class from the base Axelrod.Player class, which will have communication functionality added. The derived class will also containerize other players from the Axelrod library or our own RL models as appropriate, up to three such instances. There are three possible "game players" per "communicating player," labelled as Base Axelrod Agent, Trust, and Conviction in the below image.

Early implementation will treat Trust and Conviction as hyperparameters, and then we can develop something more sophisticated.

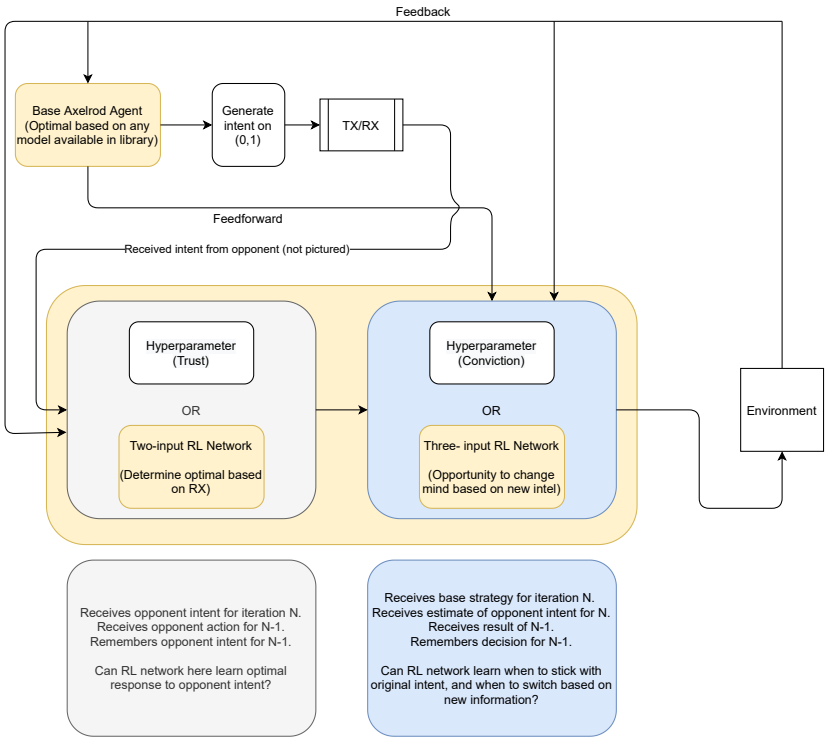

Figure 1: Detail diagram for implementation of communicating player, showing the three contained learners/agents/hyperparameters.