Tunable Behaviours in Sequential Social Dilemmas using Multi-Objective Reinforcement Learning

David O'Callaghan **Patrick Mannion**

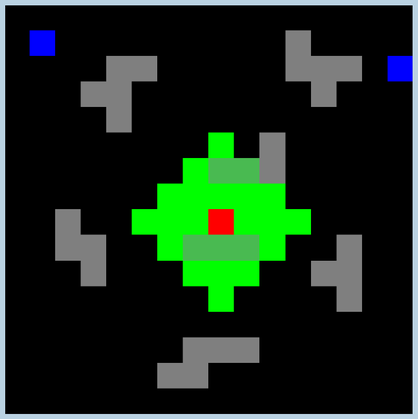
National University of Ireland Galway

Abstract

- Reinforcement learning agents are most commonly trained to learn some fixed behaviour.
- However, if a different form of agent behaviour is required, the agent typically needs to be retrained.
- In this study, we use a method to design agents whose behaviours can be tuned after training using techniques from multi-objective reinforcement learning.
- We empirically show that the tunable agents framework enables easy adaption between cooperative and competitive behaviours in sequential social dilemmas.

Wolfpack Experiment

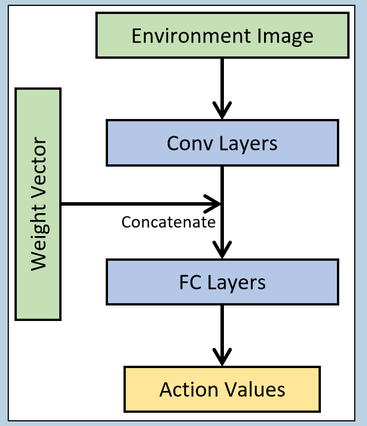
- 16x16 pixel-based gridworld.
- Multiple predator agents (blue) try to capture a prey (red).
- Team-capture if a predator is within the capture-radius (green area) when the other predator captures the prey. Lone-capture otherwise.
- Team-captures (cooperative) and lone-captures (competitive) are rewarded as separate objectives.



Full Paper: https://arxiv.org/abs/2101.11967
Code: https://github.com/docallaghan/tunable-agents
Videos: https://www.youtube.com/playlist?list=PLuIjfbXklqXbNY1tXl7gtEWj5J5pCH_Ub

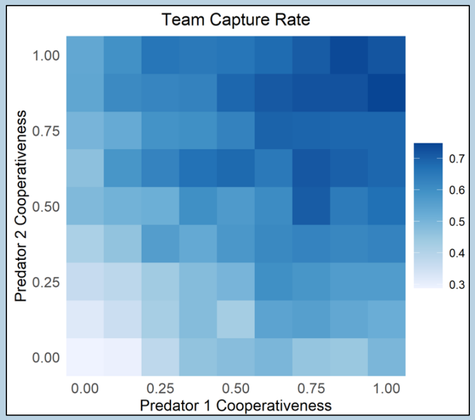
Tunable Agents

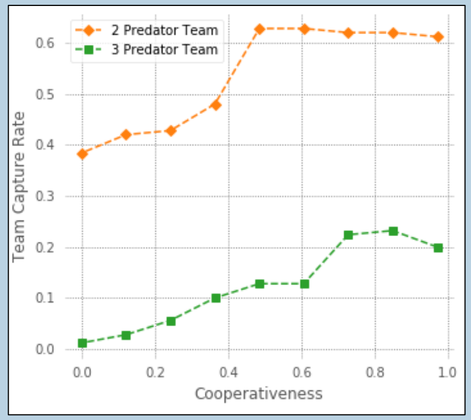
- The framework used for training tunable agents involves combining linear scalarization of rewards with the DQN algorithm.
- The neural network (right) is conditioned on both the environment state and weight vector.
- The weight vector is re-sampled repeatedly during training.



Results

- Analysed the impact of tuning the weight vectors after training.
- Found a strong correlation between predator agents' level of cooperativeness and their respective team-capture rates.





- Added a third pre-trained tunable predator into the Wolfpack environment.
- Results show that the neural network models generalise to tunable behaviours in unseen environment states.