Multi-agent reinforcement learning model of CPR

Topics in Intelligent Systems 2024/2025

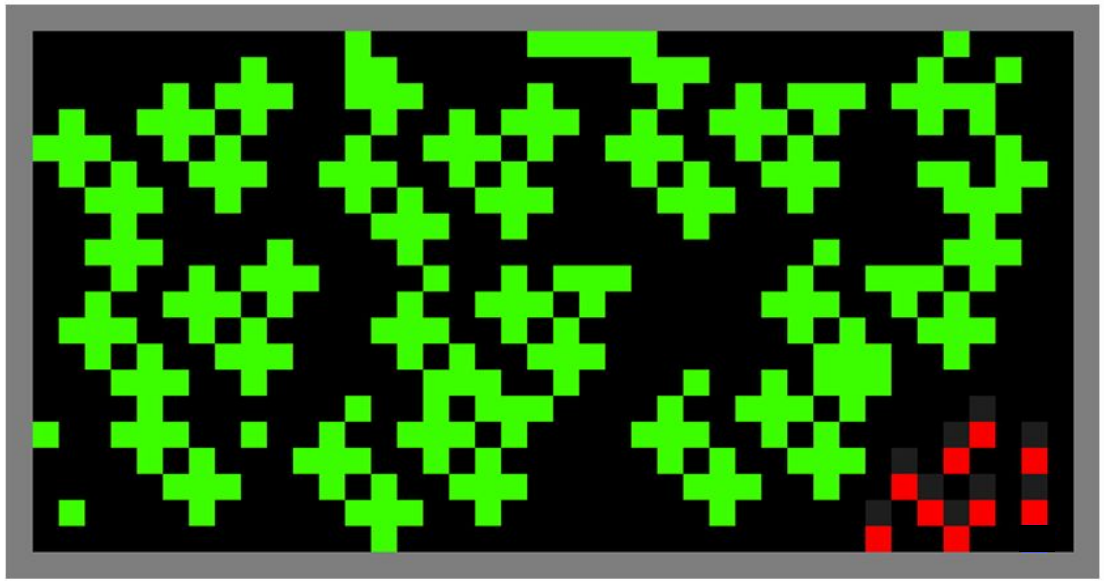
Group: Athos Freitas up202108792 | José Santos up202108729 | Luís Du up202105385

Common Pool Resources

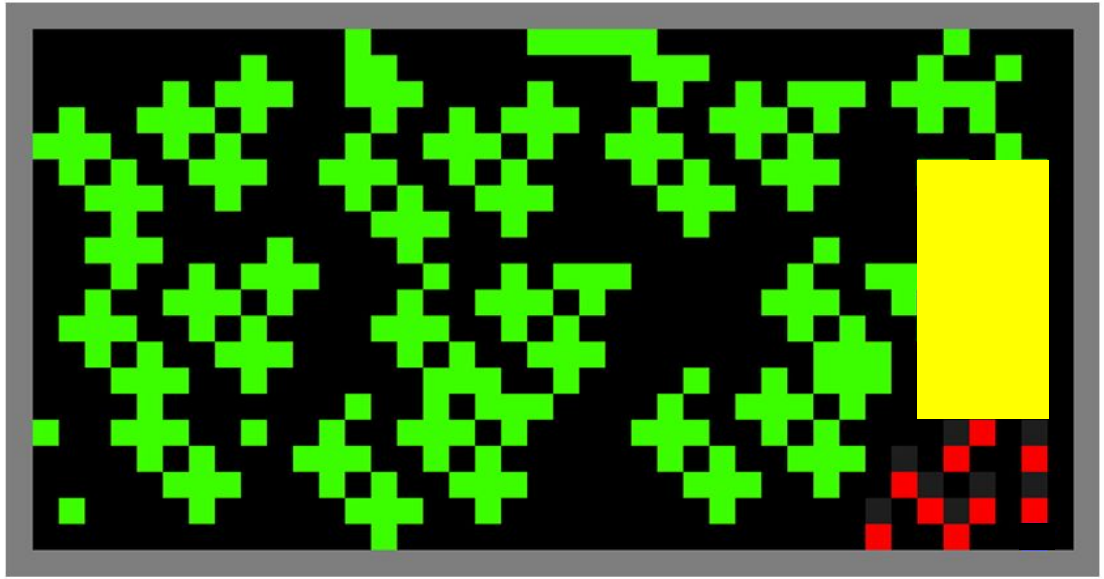
Common Pool Resources (CPR) are resources that benefit a group of people, but which provides diminished benefits to everyone if each individual pursues his or her own self-interest.

- **CPR Stock –** total amount of the resource available at a given point in time. 
- **CPR Flow –** rate at which the stock is replenished and consumed over time. 

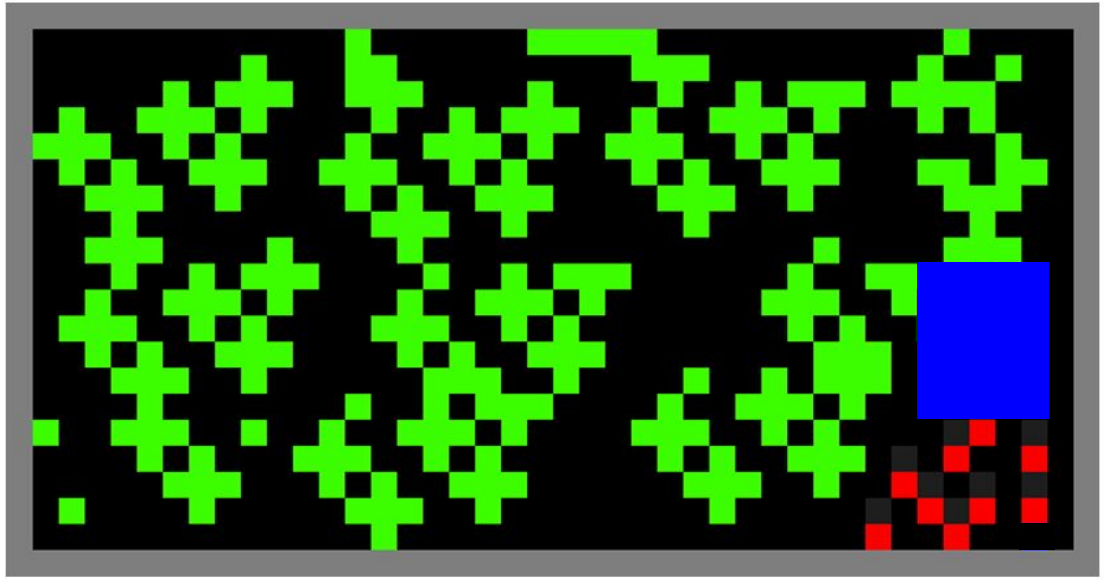
The Commons Game



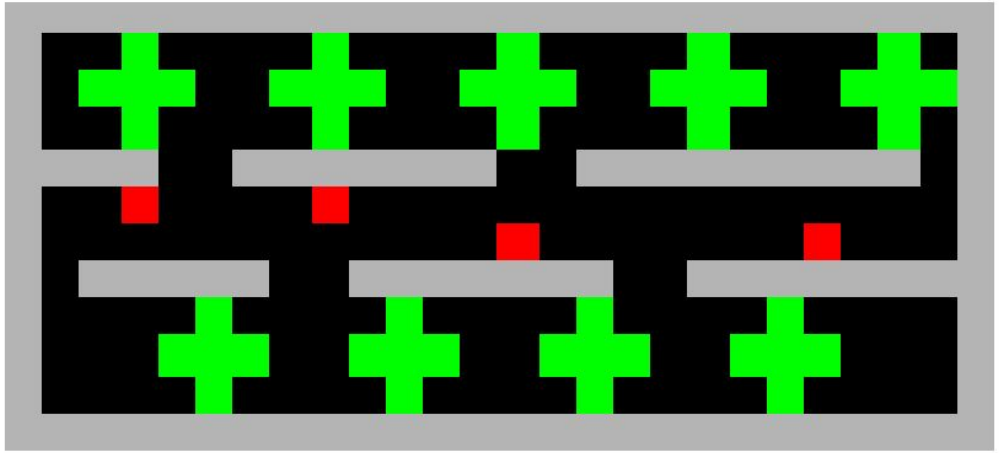
Tagging



Gifting



Walls



Main Dilemma

The interests of the **individual** lead toward harvesting **as rapidly as possible.** However, the interests of the **group** as a whole are advanced when i**ndividuals refrain from doing so**, especially in situations where many agents simultaneously harvest in the same local region.

Markov Game

The commons game is a **partially-observable general-sum Markov Game**. In each state of the game, agents take actions based on a partial observation of the state space and receive an individual reward.

Agents must learn through experience an appropriate behavior policy while interacting with one another.

Reinforcement Learning

The algorithm used was **Q-learning** with function approximation, using **ε-greedy policy** for each agent.

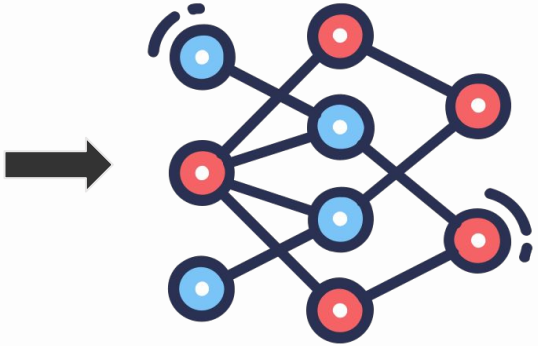
- **Actions**
  - step left ←, step right →, step up ↑, step down ↓
  - rotate left ↶, rotate right ↷, stand still
  - tag, **gift**
- **Reward**
  - Plus one whenever an agent collects an apple
  - **Minus one whenever gifting another agent**

Deep Q-Network (DQN)

DQN with replay buffer was used for the function approximation.

- The Main NN is used to predict Q values for each possible action in a given state
- The Target NN is used to get the target Q values and stabilize training by having soft updates that approximate it to the Main NN
- The replay buffer stores previous actions and their rewards to update the NNs in batches

Observable State (RGB matrix)



Q-values for each action

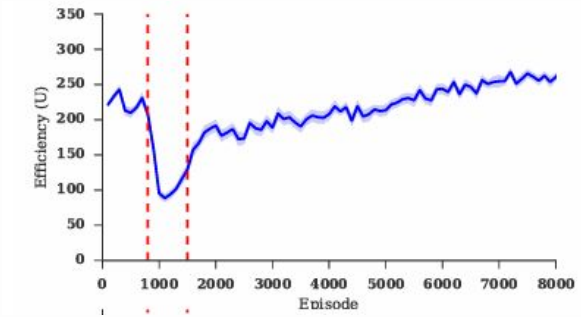
Social outcome metrics

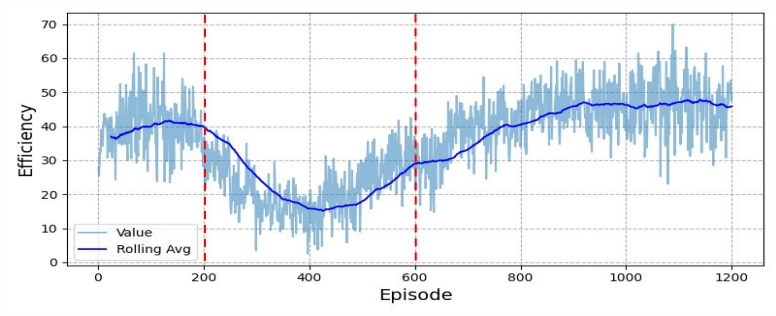
Five key social outcome metrics were used to summarize group behavior and facilitate its analysis:

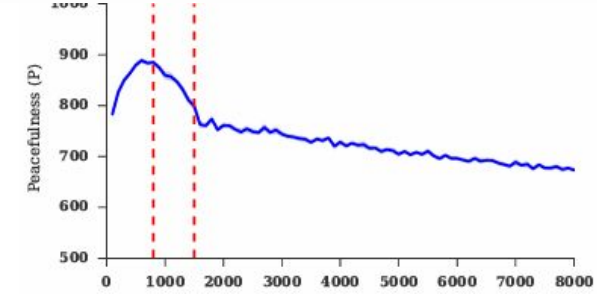
- **Utilitarian/Efficiency *(U)*** – the total sum of all rewards obtained by all agents.
- **Sustainability *(S)*** – the average time at which the rewards are collected.
- **Equality *(E)*** – the average amount of impurity in resource collection (GINI coefficient).
- **Peace *(P)*** – the average number of untagged agent steps.
- **Cooperability *(C)*** – the average number of gifting agent steps.

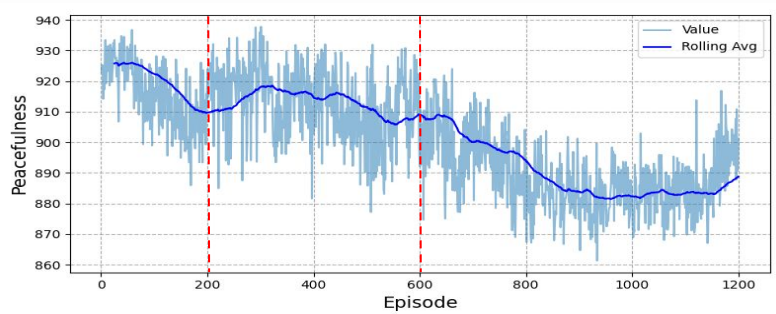
Result analysis

Map with Tagging





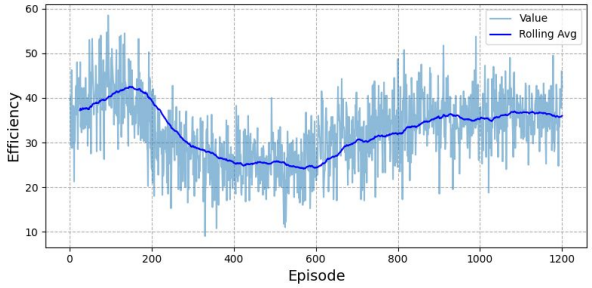


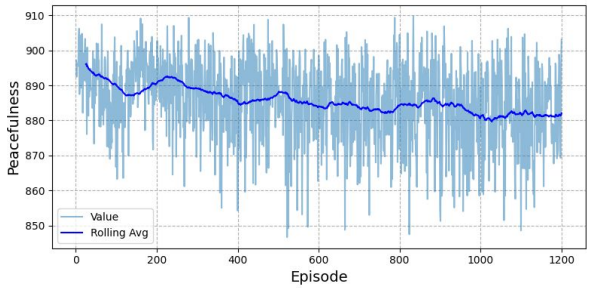


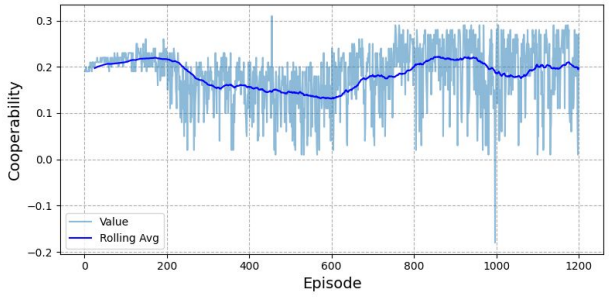
Map with Tagging & Gifting

Compared to map with tag only:

- **Efficiency** and **Sustainability** values are similar
- **Peacefulness** is slightly higher
- **Cooperability** decreases as **Efficiency** decreases



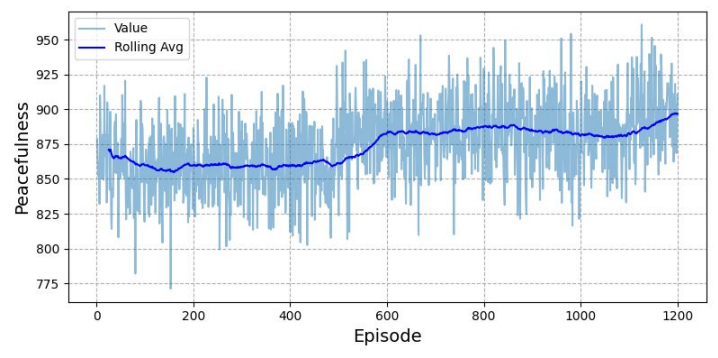


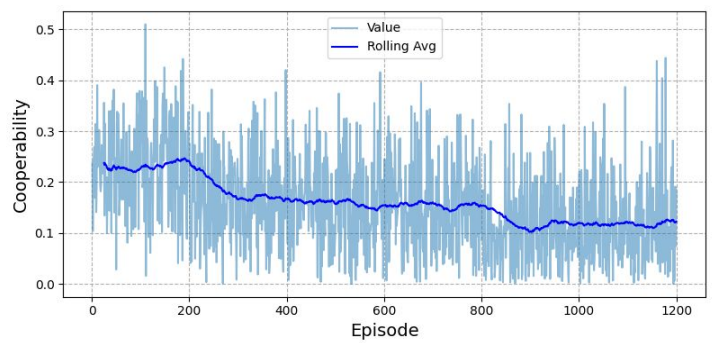


Map with Mixed Actions

Compared to map with tagging & gifting:

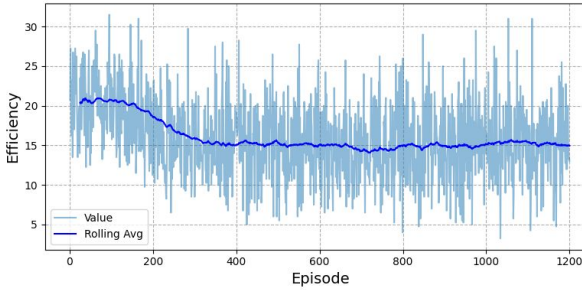
- **Peacefulness** is higher (fewer tagging agents)
- **Cooperability** is lower (fewer gifting agents)

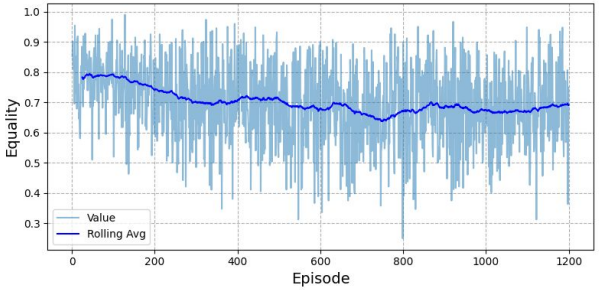


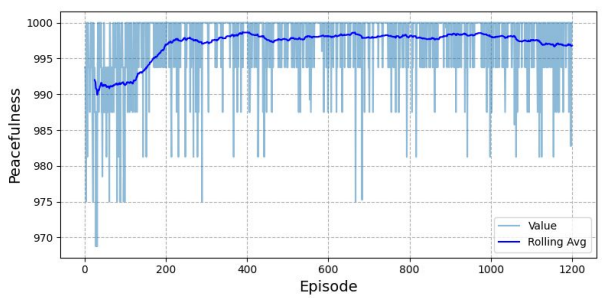


Walls with Tagging & Gifting

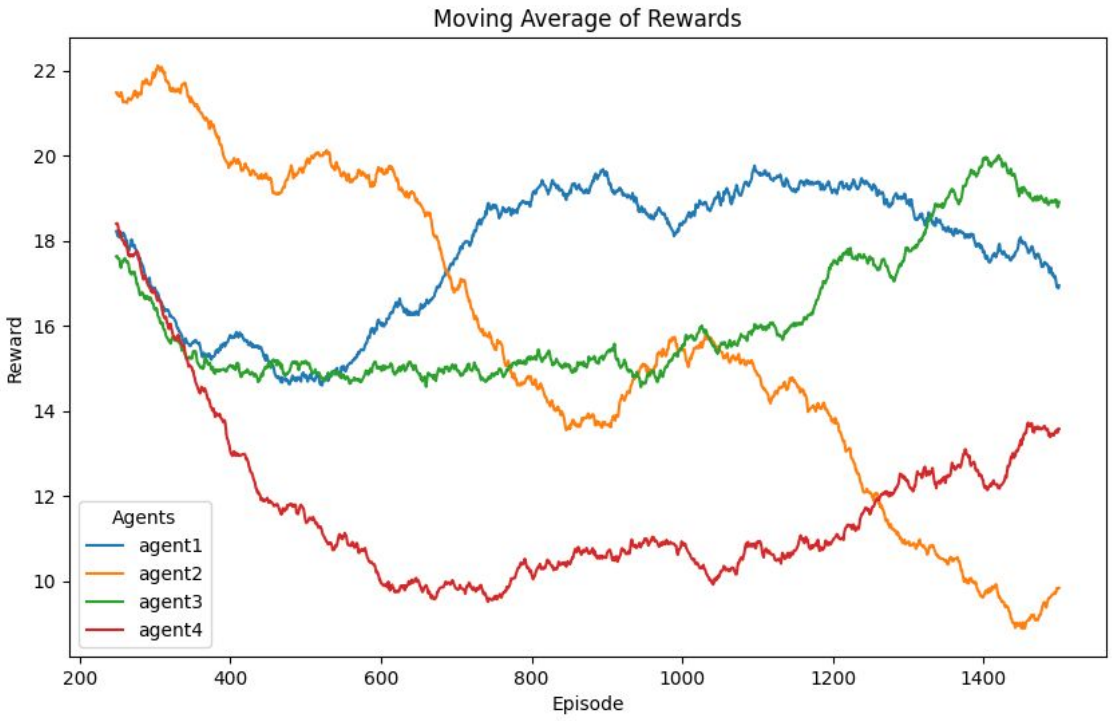
- **Efficiency** is, overall, lower
- **Equality** is constant
- **Peacefulness** is higher (the walls block tagging)
- **Cooperability** is lower (the walls block gifting)







Walls with Tagging & Gifting



Conclusions

- Even though our implementation was reduced in scale due to hardware limitations, we were still able to observe similar trends in the metrics defined in the paper
- The addition of gifting had a positive impact on the peacefulness metric
- Our agents were not as effective in the map with walls

This project allowed us to learn:

- How to create OpenAI Gym Reinforcement Learning environments
- How to implement Q-Learning using DQN, a deep RL algorithm
- How to apply this algorithm to multiple agents in the same environment
- How to implement cooperative and competitive environments

References

Our GitHub repository: https://github.com/athoscf/CPR-MAS

- Perolat, J., Leibo, J. Z., Zambaldi, V., Beattie, C., Tuyls, K., & Graepel, T. (2017). A multi-agent reinforcement learning model of common-pool resource appropriation. Advances in neural information processing systems, 30.
- Cuervo, S. (2020). *Commons Game, a multi-agent environment* [python]. https://github.com/tiagoCuervo/CommonsGame
- Haber, A. (2023). *Deep Q Networks (DQN) in Python From Scratch by Using OpenAI Gym and TensorFlow- Reinforcement Learning Tutorial.* https://aleksandarhaber.com/deep-q-networks-dqn-in-python-from-scratch-by-using-openai-gym-and-tensorflow-reinforcement-learning-tutorial/
- *OpenAI Gym Documentation* https://www.gymlibrary.dev/index.html