

Know Your Enemy

A look at using Variational Autoencoders to model opponents in the Iterated Prisoner's Dilemma

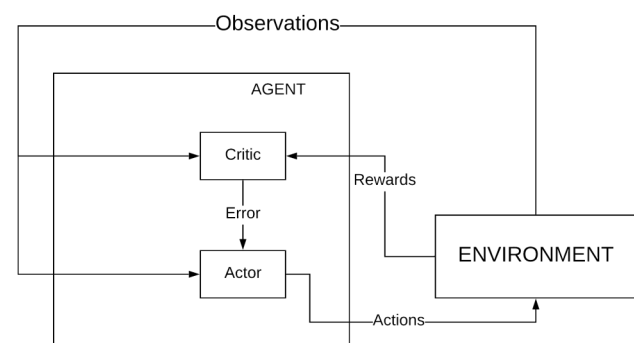
Research Project (CSE3000) by Eric van der Toorn

Supervised by Neil Yorke-Smith and Canmanie Ponnambalam

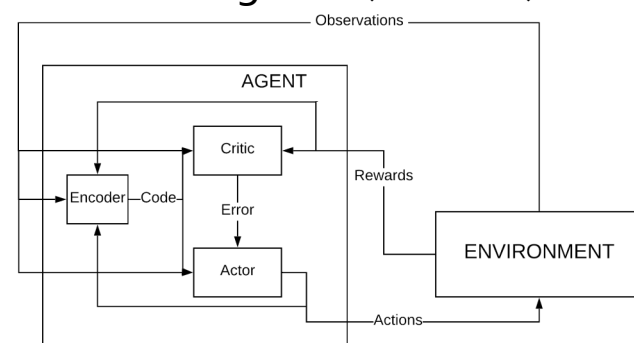
1. Problem

- Modelling opponents in Reinforcement Learning
- Do recent techniques work well in the Iterated Prisoner's Dilemma?
- Reproduced and expanded the recent Self Modelling Advantage Actor Critic work

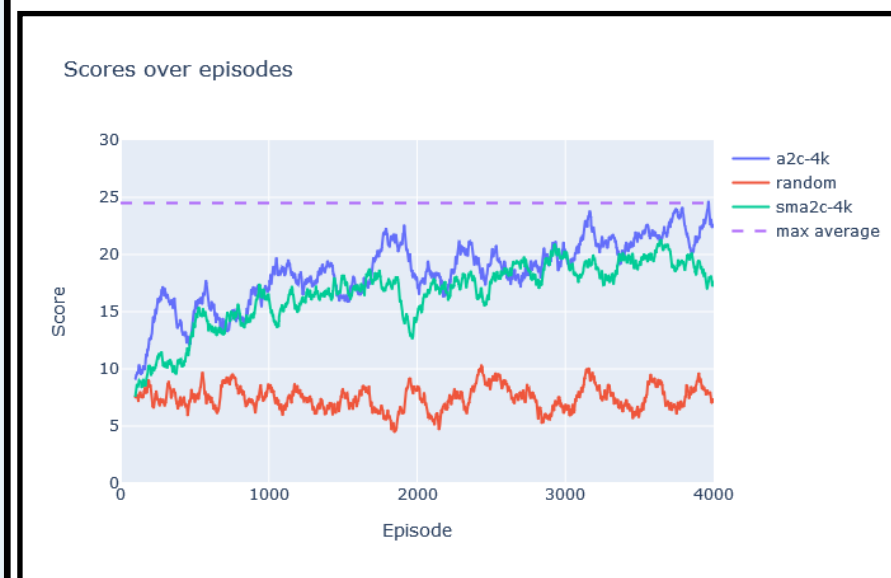
2. Advantage Actor Critic (A2C)



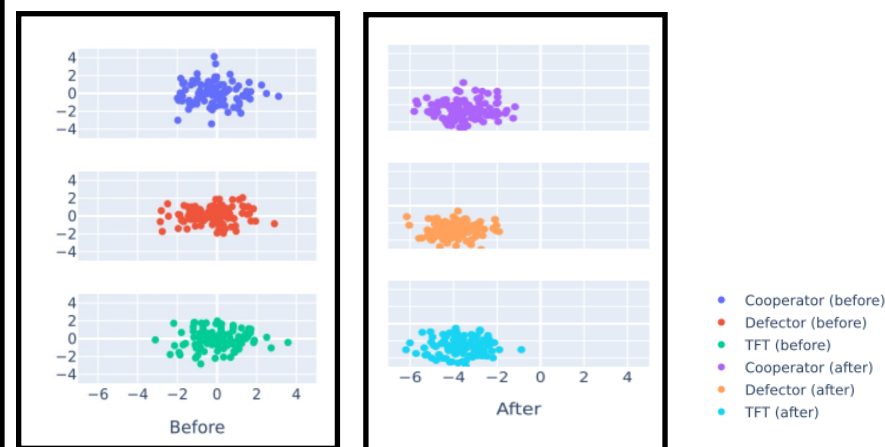
3. Self Modelling A2C (SMA2C)



4. Results



Rolling average (N=100) of 10 experiments—episodes are 25 turns of IPD



Encodings before training

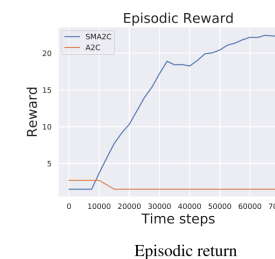
Encodings after training

5. Conclusion

- ⇒ Using an embedding did not improve results over baseline
 - ⇒ Both learn to exhibit Tit-For-Tat behavior
- ⇒ Embeddings only added to random noise
- ⇒ IPD perhaps too simplistic

6. Discussion

- ⇒ Goal was to test reproduction of SMA2C paper
- ⇒ Instead found baseline was better
- ⇒ Incredible dependence on hyperparameters



7. Future Work

- ⇒ Mutual Information Neural Estimation -> estimate effectiveness of embeddings
- ⇒ More complex examples like Sequential Social Dilemmas
- ⇒ Non-stationarity in environment