Udacity - Deep Reinforcement Learning Course

Project 2: Continuous Control

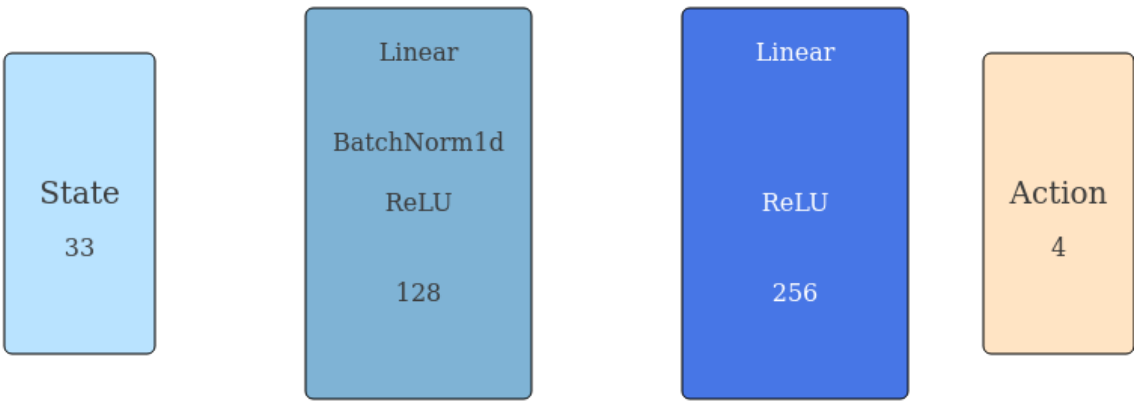
by Mauricio Ciprián Rodríguez (@mauriciociprian)

This document summarizes the relevant details of the proposed solution to Project 2 of the course, specifically the Multi-agent version (Version 2).

Actor-Critic Agent Implementation

The implementation of the agent was made starting from the code provided in the DDPG Udacity code, making several modifications: manage multiple agents, adapt the network architecture and include multiple learnings each some episodes. The code for this is in the **ddpg_agent.py** file.

The agent uses a Feed Forward neural network with the following architecture

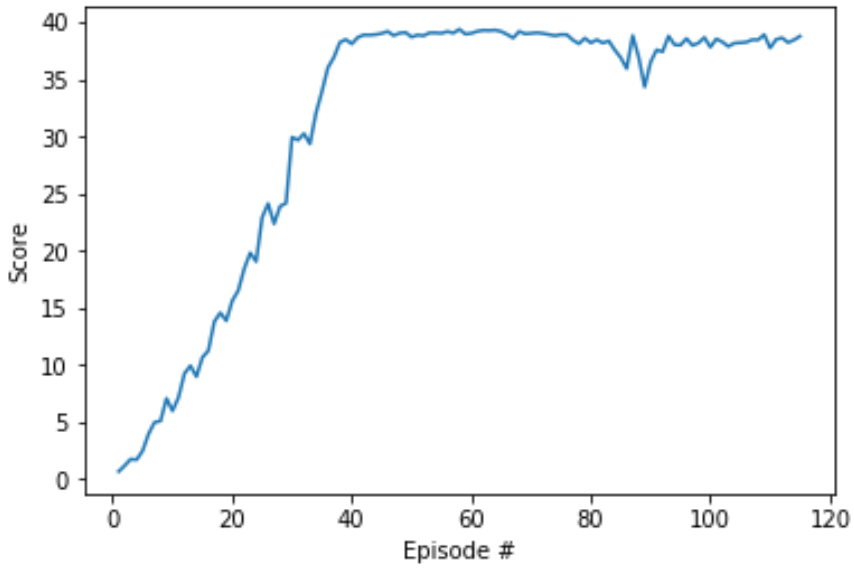


In addition, the agent uses the following hyperparameters:

- **Replay Buffer size:** 1,000,000
- **Batch Size:** 128
- **Gamma (Discount factor):** 0.95
- **Tau (Parameter update):** 0.001
- **Learning rate Actor:** 0.0001
- **Learning rate Critic:** 0.001
- **Weight Decay (Critic Network):** 0
- Learn **10 times** every **20 episodes**

Results

The graph below shows the evolution of the agent's reward during training



After 100 episodes, the agent reached the minimum score required to pass the project (Average Score = 30 in 100 Episodes). However, the agent trained to achieve an average score of 35. The models resulting from this training are stored in the **checkpoint_actor.pth** and **checkpoint_critic.pth** files.

Further Work

- A hyper parameters search (GridSearch) and other network architecture (Dropout, Layers sizes, etc) could improve the learning capacity of this DDPG implementation.
- Evaluate other Actor-Critic approaches (ie. A3C, PPO).