SYNOPSIS ON COOPERATION AND COMPETITION BASED MECHANISMS FOR MULTI-AGENT SELF PLAY AND TRANSFER LEARNING

SUBMITTED BY:

Shiv Kumar , 2016UIT2563

UNDER THE GUIDANCE OF:

Dr. Deepak K. Sharma

NETAJI SUBHAS UNIVERSITY OF TECHNOLOGY

Objective :

To research and analyse the different deep learning based reinforcement learning techniques in the domain of self play, and transfer learning and their applications.

Overview of the work :

Reinforcement learning (RL) is an area of machine learning concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learningand unsupervised learning.

It differs from supervised learning in that labelled input/output pairs need not be presented, and sub-optimal actions need not be explicitly corrected. Instead the focus is finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge).

Whereas **Deep learning** (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on artificial neural networks. Learning can be supervised, semi-supervised or unsupervised.

Techniques from these two domains independently are giving state of the art results in different problems that are known to be too complex to be solved using hard computing paradigms, and recently techniques from these two domains are being combined to produce even better results on tasks thought to be out of scope of artificial intelligence.

These techniques are being used to give machines the kind of thinking power which can be thought to be almost equivalent to that of humans where the machine is being observed to come up with its own strategies that in some cases have no correlation with those of human experiences on the same task and that is allowing humans not only a better automation performance on tasks but also giving humans better hindsight on different possible strategies that humans had not been able to devise on their own.

These allows machines to perform on unseen conditions and parameters in a way that supervised learning is not capable of achieving.

Inspired by the research done by **Google Deep Mind** in this field, this project is meant to research on how deep learning and reinforcement learning techniques can be employed to allow multi agents to cooperate and compete to better extrapolate individual and cumulative rewards using techniques like self play and transfer learning.

Self play is the technique where one instance of the environment plays against another instance of the same agent and then based on the reward signal chooses

the best actions upon which it builds based on the task, whether it is episodic or continuous.

Whereas **transfer learning** or **knowledge transfer** (called in reinforcement learning) is the overall technique where the agents apply the knowledge gained in one domain or task to a task in another domain. For example applying the knowledge of resistance gained in wrestling to combating the sense of wind resistance in football.

The project might even need a few concepts of **genetic algorithms like Neuro Evolution of Augmenting Topologies (NEAT)** which allows the agents to select the best networks out of the family and basing the next generation on their reproduction.

The testing of the algorithms will be done on multiple tasks and an analysis of the best combination of methods will be reported per task. These tasks will range from self curated to ones available with openai gym and the tools made available by NVIDIA for open use.

Hardware/Software used:

1. GPUs/ TPUs made available on clouds
2. 8 GB+ RAM
3. 256 GB+ Non-volatile storage
4. Openai Gym
5. Python
6. C++ (For optimisation if required)
7. Pytorch
8. Tensorflow 2.0
9. NVIDIA Cuda
10. NVIDIA Simulators

References :

1. https://deepmind.com/research/publications/general-reinforcement-learning-algorithm-masters-chess-shogi-and-go-through-self-play
2. https://www.ntu.edu.sg/home/sinnopan/publications/[AAAI17]Knowledge%20Transfer%20for%20Deep%20Reinforcement%20Learning%20with%20Hierarchical%20Experience%20Replay.pdf
3. https://arxiv.org/pdf/1701.07274.pdf
4. http://web.mst.edu/~gosavia/joc.pdf
5. https://dl.acm.org/citation.cfm?id=3234150
6. https://towardsdatascience.com/deep-deterministic-policy-gradients-explained-2d94655a9b7b
7. https://medium.com/@jonathan_hui/rl-proximal-policy-optimization-ppo-explained-77f014ec3f12