Continual Deep Reinforcement Learning In Non Stationary Situations

**Description**

Standard RL is based on the assumption that the environment and tasks it has to solve remain stationary in its lifetime. However this is rarely the case in real world where the agent might undergo continuous changes and the problem needs us to think beyond standard MDP formalisms. We will thus try to formalise this continual / life-long learning setting and come up with RL algorithms that can tackle this problem in line with recent works\[3\]\[1\]\[2\].

Figure 1: A probabilistic graphical model for Dynamic MDP formulation from \[3\]

Figure 2: Examples for multi task RL tasks(Changing Targets, Half-Cheetah running under changing wind velocity and mobile robot moving under varying slopes).

**Tasks**

The tasks in this project will include:

- Literature Review: Getting familiar with Deep RL algorithms and in particular Multi-Task / Lifelong / Contextual RL.
- Experimental Setup: Choosing, understanding and setting up a multi task/non-stationary environment where the MDP changes either because of the changing goals / rewards or changing dynamics.
- Formalize the setting: A good start would be to formalize it under the hidden parameter MDP\[2\] and later on move to dynamic MDP\[3\] formalism.
- Algorithm and Evaluation: Come up with a model based RL algorithm to learn a policy under these formalisms with the non-stationary experimental setup.

**References**

