Behavioral Robotics Assignment№ 2

Sami SELLAMI, Innopolis University April 15, 2020

Exercice 4

The task is to evolve robots with the original and the modified reward function and then compare the behavior of the evolved robots, we performed the evolution using the hopper robot, the evolved robot with the reward function suitable for reinforcement learning strategy exhibits bad behavior whereas the robot evolved with the evolutionnary policy show better performances where the hopper manage to move forward.

The reason for this result comes from the fact that reward function for the evolutionnary strategy is simply the progress

$$reward =' progress'$$

which is the distance travelled during a time step $v \times \delta t$, whereas for reinforcement policy the reward is calculated like this

$$rewards =' \_alive + progress + electricity\_cost + joints\_at\_limit\_cost + feet\_collision\_cost'$$

The evolutionary strategy is by definition simpler and doesn't need back-propagation, moreover it has fewer hyper-parameters which is our case. Moreover, the reward of the evolutionary strategy is composed of the progress only because we don't want the population of the agents to be influenced by other factors than the progress which is the most important, so, unlike the reinforcement learning reward which has electricity cost, joints limits cost and others factors, the best agents of the evolutionary strategy that will survive during the policy evolution will reflect the biggest progress thus the biggest distance travelled and not the other terms related to electricity cost, joints at limits cost ...etc.

Exercice 5

The task is to implement a new Gym/Bullet environment from scratch, it consist of a simple wheeled robot than can be trained for the ability to balance. The code is given in the github folder and the algorithm manage to balance the robot after some episode steps.
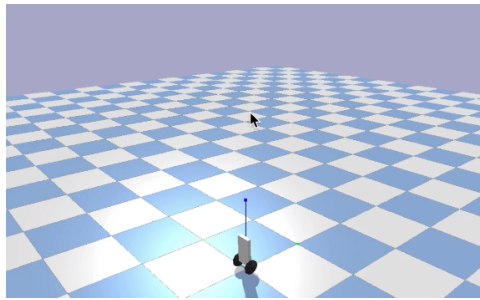
Figure 1: Balance bot robot

Exercice 6

The task is to run 10 replications of the experiment from evorobotpy/xdiscrim folder by using different seeds. we need to describe the strategies of the robots by grouping them in families and try to explain why the robot of each family behave in that manner. We see two kinds of behaviors, the first shows the robot reaching the target and oscillating around it , and in the second the robot stops when it reaches the goal.

The robots of the first family react in this manner because the robot action consist of turning right/left with continuous torque applied to the wheels, so when the robot reach the cylindrical object, it try to turn right increasing the reward and turn left when it goes too far, reaching a stable position corresponding to a local minimum. The second family reflect the case when the robot turn too fast and diverge from the local minimum indefinitely.

Next, we run other experiment using a feed-forward architecture, we see that the evolved robots behave differently from the LSTM architecture, indeed, the robot tends to cover the whole workspace in an attempt to find the cylinder and may find it rapidly or not depending on its position, whereas the LSTM architecture shows evolved robots which tends to have memory and remember the best strategy (path) taken to find the cylinder rapidly and efficiently especially when it has the same position as before. (Fig 2) depicts the architecture of some some neural network including feedforward and LSTM
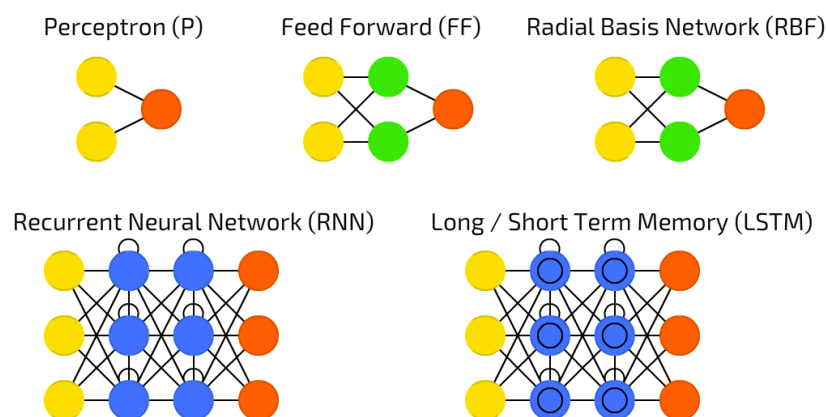

Figure 2: Neural Network architectures