Behavioral Robotics Assignment№ 1

Sami SELLAMI, Innopolis University March 12, 2020

Exercice 1

Familiarization with the commands and environments variables in Gym as well as inspecting the source code of some of these environments.

Exercice 2

The task is to implement a neural network controller for a Gym problem, for this we implemented an evolutionary strategy depicted in the problem statement, we tested the code for $\lambda = 10$ and $\sigma = 0.02$ and for a number of evaluation episode equal to 200 each lasting 500 steps.

The algorithm (cart pole.py) shows satisfactory results regarding the stabilization , the fitness converge to the max value of 500 for all the parameters vectors after some iterations which correspond to the balancing of the cart pole, the problem is solved every time we run the training process

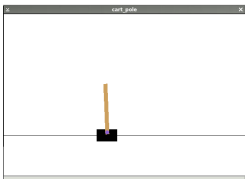

Figure 1: Representation of the Cart-Pole

Increasing the number of hidden units results in better performance, the algorithm converge faster to the max possible fitness, whereas increasing the variance of the perturbation vector σ results in poorer performance ie a higher convergence iteration number. We also tested the algorithm on another control problem which is the pendulum, the results show convergence of the fitness value for almost all the parameters at the end of the evaluation episodes.

1 Exercice 3

We run few replications of the experiment by using different seeds, $S = 9, 10, 11, 12, 15$ and 20 , and using the trained replications that was already present in the git folder, we plot the performance across generation (Fig 2)

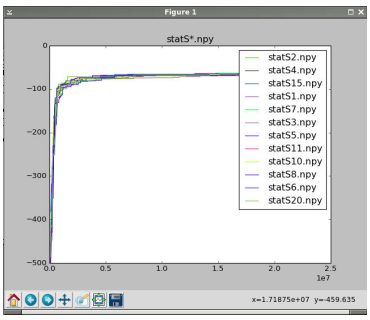

Figure 2: Algorithm performance across generation

We can see that the stats corresponding to some of the seeds are approximately the same and converge the value -90 which means the algorithm convrege in 90 steps approximately, and that for some others it equals to zero ($S = 11$ for example) which means that robot didn't manage to perform the swing-up. Moreover the standard deviation of the performance among multiple runs is equal to : -60.69 ± 18.37. Observing the behaviour of the evolved robots, we see that the Acrobot manage to swing up both links for some seeds number and doesn't for others, for example for $S = 11$ the fitness is -500 which means that training process was unsuccessful, and for $S = 1$ the fitness equals -69 which correspond to the swing up in 70 iteration steps hence the resolution of the problem. we can conclude that one need to carefully choose the initialization of the seed set number to successfully solve the problem.