Behavioural and Cognitive-Tasks

Task1:

We tested different environments and observed their action spaces and we concluded that because the physics of all the environments is different that is why we got different action spaces.

Task2(a,b):

a) We run 20 episodes by initializing the random weights of the neural network consisting of four input neurons, 5 neurons in the hidden layer and 2 neurons in the output layer. We apply our neural network to the Cartpole-v0 and observe its behaviour because of the random initialization of waits the pendulum is unstable. We then observed the reward list which explained us the reason for instability of the pendulum.

b) This task is in continuation of the 2a part, in this part we initialized the neural weights at random and than after getting the reward list for 20 episodes, we updated the initial half parameters of the reward list with last half parameters as the reward list is sorted from bad to good rewards in an ascending order and than these updated rewards were used to train the weights of neural network and we trained it continuously until our Cartpole become stable. This is a really effective algorithm but the reward always depends upon the random initialization of weights.

Task3:

There were two methods using pre-trained models and the other one is to manually train the model. I have used the pre-trained x-pendulum robot. The training parameters I used were from the file bestgs2.npy.

The attached graph shows initially the performance was at zero but as the steps increases the performance is increased exponentially.