Trial Parameters

* NN: 15,30,10
* Action: 0 joint movement in either joint not allowed
* Episodes: 200
* Iterations: 100
* reward = 200 \* (new\_x-old\_x)
* Penalty for no net joint movement (either joint at joint limit): yes
* Starting state: set
* Rollout state: set
* batch\_size =4
* C = 10
* agent = DQNAgent(epsilon\_decay=0.99995,actions\_params=(-pi/5, pi/5, pi/5))
* memory size: 1000
* t\_interval = 0.5
* learning rate = 0.0005