

## Theory

1. What is different about reinforcement learning and supervised learning? What is the same?
2. What is the difference between Q-Learning and Policy Gradients?

## Discussion

3. Do you think human level AGI can be created from our current DL algorithms or improvements to them? Do you think they can become self aware?

Note there is no correct answer here, this is just something to consider and discuss get involved in the discussion and listen to your tutor to make sure your thinking is on the right track.

## Reinforcement Learning

The programs are the two algorithms introduced in the lecture training on a Python AI gym problem where they need to learn to balance a pole on a moving block by choosing from the moves left and right at each time step. Each algorithm has one line to fill in marked by #FILL IN. PG works extremely well on this problem and on a CPU if you leave it running overnight you should come back to great results in the morning. Q-learning doesn't run well as is, I have tweaked the hyperparameters in the past to get it to run reasonably well (the tweaking is left as a task), but never as well as PG, so that is a challenge to anyone keen.

**WARNING:** All the programming problems take a bit of time to run to get good results (probably around an hour or two minimum on a basic laptop), but you can see that they are working within around 10-15 minutes in the lab.

**To really see optimal results I suggest leaving them to run overnight.**

The RL examples are quite interesting to see the progression of learning. Both reinforcement learning examples produce a graph of episode length. So, if you are interested, you can run the programs overnight.