COMP596 – Brain-inspired AI

Assignment 2

Winter 2021

Due: March 29th before 11:59 PM.

In this assignment, you will read a computational neuroscience article that describes a model of a rat solving the Morris water-maze task using temporal difference learning.

The purpose of this assignment is three-fold:

1) To help you see how temporal difference learning can explain some aspects of animal behaviour.

2) To get you familiar with a simple example of temporal difference learning in a neural network.

3) To have you think about the limitations of temporal difference learning.

You will be basing your model on a paper by Foster, Morris and Dayan (2000). Your ultimate goal from a coding perspective is to replicate their model. It should be doable for you, but remember, use office hours and the discussion board if you're getting stuck!

Submit your assignment on MyCourses. You must submit both your code (i.e. the modified Jupyter notebook) and also a **pdf** document with your answers to the questions below. You will be marked on both your code and the document, don't skip either!!!!

The assignment is out of 70 marks.

Part 0 – Getting started (no marks, though if you don't do this you won't get any marks!)

- As with Assignment 1, you will need to have Python installed, as well as numpy, a plotting library (e.g. matplotlib), and Jupyter notebook. One way to kill multiple birds with one stone here is to install Anaconda or Miniconda: https://docs.conda.io/en/latest/
- Download the Foster, Morris & Dayan (2000) paper, and read it!
- Make sure you can open the Jupyter notebook provided (`watermaze_template.ipynb`), which gives you the skeleton for your code. To open it you use Jupyter, e.g., on a Linux system if you have Jupyter installed you can type this into the command prompt:
  - `$ jupyter notebook watermaze_template.ipynb`
- Make sure you understand how to use the water-maze simulation components. There is a module provided that you can use to simulate the behaviour of a rat in the water-maze. Run the demo provided a few times to get a sense of how it works.

- You must put your code into this notebook, and submit it with a new name. The new name of the notebook must adhere to this format:
  - `watermaze_[firstname]_[lastname]_[mcgillid].ipynb`
  - Where the items in brackets must be replaced with your specific info.

Part 1 – Understanding the Foster, Morris and Dayan paper (30 marks total)

Read the paper, then answer the questions below, and put your answers in the pdf document that you will submit along with your code.

Read the Foster et al. (2000) paper, and answer the questions below.

P1Q1 (5 marks)

(a) In your own words, describe the global consistency problem. (2 marks) *Give your answer in 50 words or less.*

(b) What type of navigation learning model suffers from the global consistency problem, and why is it a problem for them? (3 marks.) *Give your answer in 100 words or less.*

P1Q2 (5 marks)

(a) In your own words, describe the distal reward problem. (2 marks) *Give your answer in 50 words or less.*

(b) What type of navigation learning model suffers from the distal reward problem, and why is it a problem for them? (3 marks.) *Give your answer in 100 words or less.*

P1Q3 (10 marks)

(a) What is the role of the actor in the Foster et al. (2000) model? Specifically, what is the actor learning to encode? (2 marks) *Give your answer in 50 words or less.*

(b) What is the role of the critic in the Foster et al. (2000) model? Specifically, what is the critic learning to encode? (2 marks) *Give your answer in 50 words or less.*

(c) What is the error signal used in TD learning? Give the equation for it, and explain all the terms/variables contained in it. (2 marks) *Give your answer in 50 words or less.*

(d) What is the role of the place cells in the Foster et al. (2000) model? Specifically, what do they encode? (2 marks) *Give your answer in 50 words or less.*

(e) What sorts of information can place cells not encode that a navigation system might want? (2 marks) *Give your answer in 50 words or less.*

P1Q4 (10 marks)

(a) Does TD learning with an actor-critic system solve the distal reward problem? Explain why or why not. (5 marks) *Give your answer in 150 words or less.*

(b) According to Foster et al. (2000), what information must be available to a rat if it is going to solve the global consistency problem with TD learning? Do you think that it is realistic that a rat would possess this information? (5 marks) *Give your answer in 150 words or less.*

Part 2 – Implementing the pure TD algorithm (20 marks total)

Here, you have to implement the pure actor-critic model described in the first part of the paper. You must replicate the basic design features as described by equations 1-10 and Figure 2 of the paper. Remember, a key part of this is that you must represent the position of the animal in the maze using a set of place cells. Base this off of the description in the paper. Additionally, you must generate plots demonstrating the performance of the model as the time it takes the simulated rat to find a single platform (as shown in Figure 1a and 4a of the paper up until day 7). You will be marked based on the following scheme:

P2Q1 (10 marks)

Can the TA run your code and have it generate plots (5 marks)? Does it make sense to them and is well commented such that they can understand what is going on by looking at it for 5 minutes (5 marks)?

P2Q2 (5 marks)

With a single platform location, does the model successfully reduce the latency over trials? Get your code to generate the plots, and also include them in your pdf file. (*Note: you should have a flag in your code that let's the TA set it to a single-platform or multi-platform case*.)

P2Q3 (5 marks)

What happens to the performance of the model when the flag is toggled to a multi-platform case (provide plots of the escape latency)? With reference to your answers in Part 1, explain why this happened?

Part 3 – Implementing the combined coordinate and TD algorithm (20 marks total)

Here, you have to implement the combines model described in the second part of the paper. You must replicate the basic design features as described by equations 1-12 and Figure 5 of the paper.[1] Additionally, as in Part 2, you must generate plots demonstrating the performance of the model as the time it takes the simulated rat to find the platform (as shown in Figure 8 of the paper). You will be marked based on the following scheme:

P3Q1 (10 marks)

Can the TA run your combined model code and have it generate plots (5 marks)? Does it make sense to them and is well commented such that they can understand what is going on by looking at it for 5 minutes (5 marks)?

---

1 ***Note:*** there is a small error in the Foster et al. paper in equations 11 and 12. Specifically, X(p_t+1) and Y(p_t+1) should be X(p_t) and Y(p_t) and vice-versa.

P3Q2 (5 marks)
With multiple platform location, does the model successfully reduce the latency over trials? Get your code to generate the plots, and also include them in your pdf file. (*Note: you should have a flag in your code that let's the TA set it to a single-platform or multi-platform case*.)

P3Q3 (5 marks)
What is your opinion of the coordinate model? Does it seem like a good strategy for an AI? Why or why not? Justify your answer. *Give your answer in 300 words or less.*