FACULTY *of* SCIENCE *and* ENGINEERING

Department of Computer Science
and Information Systems

End-of-Semester Assessment Paper

Academic Year:	2021/22	Semester:	Spring
Module Title:	Deep Reinforcement Learning	Module Code:	CS6482
Duration of Exam:	1.5 Hours	Percent of Total Marks:	30
Lecturer(s):	J.J. Collins	Paper marked out of :	30

Instructions to Candidates:

- **Answer Question 1 and any two others.**
- Only the first three answers will be corrected unless other answers are explicitly pointed to by the student at the start of their script.
- Please do not use red ink

Q1 Answer ALL parts. Total marks awarded for this question: 10 marks.

a) How is experience captured, processed, stored, and sampled for training a Deep Q Networks (DQN) on Atari? Illustrate the discussion with coding fragments or pseudocode.

3 marks.

b) What is maximisation bias in the context of DQNs?
Provide an example to illustrate the discussion.
Describe two approaches that can be used to reduce maximisation bias

3 marks.

c) Describe the loss function used to train a DQN.
Illustrate the discussion with coding fragments or pseudocode.

4 marks.

Q2 Answer ALL parts. Total marks awarded for this question: 10 marks.

a) Explain why the number of parameters in GoogleLeNet using Inception modules is significantly less than AlexNet - 6 million as opposed to 60 million. The answer should focus exclusively on the Inception module. Illustrate the answer with a diagram and/or rough calculations.

3 marks.

b) What is a vanishing gradient?
Briefly describe the cause(s) of this phenomenon.
Describe technique(s) can be used to reduce the likelihood of vanishing gradients.
Include diagram(s) if discussing activation functions.

3 marks.

c) Describe the key concept(s) in ResNet. Include a discussion on the purpose of a kernel of size 1x1 with stride 2. Illustrate the discussion with a diagram.

4 marks.

Q3 Answer ALL parts. Total marks awarded for this question: 10 marks.

a) What is Reinforcement Learning?
What are the key issues that an RL agent must address?

3 marks.

b) Explain why it is stated that Temporal Difference (TD) methods bridge Dynamic Programming (DP) and Monte Carlo (MC) methods.

3 marks.

c) What is meant by the terms on-policy and off policy in the context of TD methods? The discussion should include the equations for an on policy and off policy update.

4 marks.

Q4 Answer ALL parts. Total marks awarded for this question: 10 marks.

a) Compare and contrast the Symbolic AI and Machine Learning (ML) paradigms, and illustrate the answer with examples.

3 marks.

b) Discuss four metric-based approaches used to quantify the performance of an ML algorithm, one example being precision and recall.

3 marks.

c) Describe Generative Adversarial Networks (GANs) **OR** Recurrent Neural Networks (RNNs). For the paradigm selected, discuss sample applications, training data, network structure, training, and constructs specific to the algorithm. For example, Long Short Term Memory (LSTM) in the case of RNNs.

4 marks.

Q5 Answer ALL parts. Total marks awarded for this question: 10 marks.

a) List the steps in the REINFORCE Algorithm by Williams 1992.

3 marks.

b) Compare and contrast Q Learning with Policy Gradient (PG) approaches. Briefly describe two issues that can arise in PG approaches.

3 marks.

c) The coding fragment is Figure Question-5 is an excerpt of an implementation of the REINFORCE method for the cartpole. Provide a detailed explanation for lines 7 and 8, and for the observation and action space in line 4.

4 marks.

```
1. if __name__ == "__main__":
2.     env = gym.make("CartPole-v0")
3.     writer = SummaryWriter(comment="-cartpole-reinforce")
4.     net = PGN(env.observation_space.shape[0], env.action_space.n)
5.     print(net)
6.     agent = ptan.agent.PolicyAgent(net, preprocessor=ptan.agent.float32_preprocessor,
       apply_softmax=True)
7.     exp_source = ptan.experience.ExperienceSourceFirstLast(env, agent, gamma=GAMMA)
8.     optimizer = optim.Adam(net.parameters(), lr=LEARNING_RATE)
```

Figure Question-5. Adapted from Maxim Lapan. Deep Reinforcement Learning Hands-On. 2018. Packt Publications.