University of London

Assessment Coversheet

Complete this coversheet and read the instructions below carefully.

Candidate Number: KJ1932

Degree Title: BSc in Computer Science

Course/Module Title: Artificial Intelligence

Course/Module Code: CM3020

Enter the numbers, and sub-sections, of the questions in the order in which you have attempted them:

2 a b c d e f g
3 a b c d

Date: March 8, 2022

Instructions to Candidates

1. Complete this cover sheet and begin typing your answers on the page below, or submit the cover sheet with your handwritten answers (where handwritten answers are permitted or required as part of your timed online assessment).
2. Clearly state the question number, and any sub-sections, at the beginning of each answer and also note them in the space provided above.
3. For typed answers, use a plain font such as Arial or Calibri and font size 11 or larger.
4. Where permission has been given in advance, handwritten answers (including diagrams or mathematical formulae) must be done on light-colored paper using blue or black ink.
5. Reference your diagrams in your typed answers. Label diagrams.

The Examiners will attach great importance to legibility, accuracy, and clarity of expression.

Begin your answers on this page

Question 2

(a) MDPs are suitable for modeling stochastic sequential decision problems. These are problems that contain a partly stochastic or probabilistic and random element and are partly controlled by a decision-maker. It models problems where some process is in a state *s,* and the decision-maker can take action *a*, which leads to a new state *s'* stochastically, with an associated reward *r.*

(b)

i. I am choosing **Space Invaders**. This game gives the player control of a spaceship that can move left and right and shoot projectiles. The player fights against a wall of aliens, which shoot back. If the player shoots an alien, the alien dies and disappears. The player wins if they eliminate all aliens and survive the game.
ii. The control inputs are **moving right**, **moving left**, **staying in place**, and **shooting**.
iii. Points are gained by eliminating as many aliens as possible. For each eliminated alien, points are awarded. You gain bonus points if you survive the game and eliminate all aliens.
iv. An AI system would perceive the position of the player, the position of projectiles that both the player and the aliens are shooting, and the number and position of remaining aliens among other details.

(c) An MDP requires **states**, **actions,** and **rewards**. In our game, the states are the current position of the player and aliens, among other factors mentioned in the previous answer, the actions are the player's movements left and right and shooting, and the rewards are the points awarded by eliminating aliens and surviving the game.

(d) I believe it is suited for MDPs because Space Invaders has a stochastic element in the aliens that can randomly shoot at you. The game is sequential, in that actions taken lead from one state to another. The outcome of the action is not always pre-determined, as aliens can move, new aliens can appear, or shots eliminated with counter-shots.

(e) The Q function approximates the value function of a neural network. It infers the value of each possible action given an input state and accounting for future rewards. The inputs are the current observed state, the possible actions, and the weights of the neural network. The outputs are all the inferred values for all possible actions. The error function or loss function in deep Q networks is an equation that estimates the squared difference of errors between the expected value of the Q function (based on an old version of the network) and the actual value of the Q function (using the current version of the network). The inputs are the Q functions with the respective states, actions, and network weights (old and new), and a sample of the replay buffer. The output is the estimated loss, or performance, of our neural network.

(f) Q-learning is a way to find a suitable action policy for any given MDP. It does so by using the replay buffer of observed states, actions, and rewards to continually train a Q-function. It does this with Greedy Epsilon approximation, by which the Q function uses randomized weights at first, but gradually moves to use the trained weights from the reinforcement learning algorithm.

(g) What makes Q-learning difficult is that we do not have a state transition matrix to begin with. So, we do not know what actions lead from which state to which new state. Further, we do not have an action policy, so we don't know what to do given a certain state. This requires random exploration and progressive training of the neural network which makes Q-learning more difficult.

Question 4

(a) AI systems can assist in automated scientific discovery through their superhuman scientific reasoning powers. They can flawlessly remember a vast number of facts, execute flawless logical reasoning and near-optimal probabilistic reasoning, learn more rationally than humans, from vast amounts of data and extract information from millions of scientific papers. The advantages of automated discovery, therefore, include that hypotheses generation and experimentation can be executed thousands of times. This is especially useful as the amount of data produced far outpaces the amount of actual knowledge. Automated scientific discovery can aid in recording and sharing scientific results and solves the problem of the lack of scientists to tackle certain challenges. Automated scientific discovery produces unbiased results which are always accurate to within the scope of the problem, and can tackle the vast complexity of systems, such as biological systems, which are near-impossible for human scientists to map and process.

(b) Automated scientific discovery is limited in the sense that these systems have a limited knowledge model of the world. They rely on formalized knowledge representations that need to be fed in to understand the world they are operating in. This also means that automated systems cannot explore new areas of scientific discovery on their own unless they are told to do so. They are further not independent and rely on humans to design the computational model that drives them. In this regard, they are closer to "lab assistants" than actual scientists. They remove the chance for serendipitous discovery and thus potentially hinder innovation. Automated systems are also subject to any errors in the data they use and hence might lack the ability to critically assess errors in data or previous studies. This might lead to wildly different results in further hypothesis generation and validation. They might reach very different conclusions based on experiment results compared to humans. While they can operate at scale, their development and maintenance is also so cost-intensive that they are currently not cost-effective means of scientific discovery yet.

(c) I will describe an experiments planner as part of the robot scientist. The experiment planner needs to know about a set of testable hypotheses. Since not all hypotheses can be tested, potential candidate hypotheses need to be identified first and fed as the input to the planning agent. The planning agent needs to know what the available equipment and its functionality are, what materials are available, and what rules exist for planning experiments in the form of a knowledge base. The planning agent's knowledge base needs to contain information about what the controls for experimentation are. It then translates hypotheses to an experimental protocol, which is a list of experimental actions with specified parameters for the robot to execute it. Planning agents need an experimental protocol to operate.