

Artificial Intelligence

Exercises week 10 - Human-aligned Intelligent Robotics - **Solutions**

COMP3411/9814

Question 1: Explainability

Consider an intelligent robot using reinforcement learning to find the exit of a maze autonomously. While navigating the maze, in a specific state, the robot faced three possible actions: turning to the left, turning to the right, or keeping going straight. The robot decided to turn to the left. At that moment, a human user required an explanation of why this action had been taken. Which of the following is a better explanation to be given by the robot?

- a. I chose turning to the left because maximizes future collected reward.
- b. I chose turning to the left because this action has the highest Q-value.
- c. I chose turning to the left because this action gave the highest probability to find the exit.
- d. I chose turning to the left because I performed an exploratory action.
- e. I chose turning to the left because was the next action following the optimal policy.

Answer: Alternative C.

Question 2: Interactive RL

Consider an intelligent robot using interactive reinforcement learning (IntRL) receiving advice from a human expert. As the given advice is not always correct, a persistent IntRL approach might use probabilistic policy reuse (PPR) in order to allow the agent to forget the previously received advice. Consider the robot in a particular state does not receive new advice but, in the same state, a previous advice was given. However, due to the PPR approach, the robot has forgotten what advice was received. How would be decide the next action?

- a. The robot uses the regular action selection policy.
- b. The robot takes the greedy action.
- c. The robot updates the agent advice model.
- d. The robot confirms the previous advice with the expert.
- e. The robot performs a random action

Answer: Alternative A

Question 3: Human-aligned decision making

A robot navigating a cluttered environment uses both contextual affordances and a memory-based explainability module (MXRL) to choose between actions. In a situation where two actions appear valid, grasp and push, but only one is contextually safe and explainable, which of the following best explains how the system resolves the decision and provides human-aligned reasoning?

- a. The MXRL module compares Q-values of all actions and selects the one with the least negative reward, regardless of safety constraints.
- b. The contextual affordance function filters out actions that are invalid in the current state, while MXRL selects the most probable successful action among the remaining ones and justifies it using Ps.

- c. The robot defers entirely to the human-in-the-loop to resolve ambiguities, using affordances only for object recognition.
- d. The MXRL module overrides contextual affordances if a high P_s value exists, enabling a safer exploration.
- e. The robot looks for the most suitable contextual affordance action-effect and pass it to the MXRL module to provide an explanation based on the highest Q-value.

Answer: Alternative B.