ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Evaluation 1

1 QUESTIONS

1. What is an intelligent agent? Give at least two applications of intelligent agents in real life.
2. Describe three generic difficulties with designing intelligent agents *with your own words*.
3. What is the definition of Reinforcement Learning? How does it relate to Machine Learning?
4. What is safe exploration ? What can make it so difficult?
5. Characterize the RL problem adopted in Assignment 1.
6. What is a policy? A stationary policy? An optimal policy? Should we consider all the possible policies to find an optimal one? Justify your answer.
7. Let us consider a policy which always select the action that maximizes the instantaneous reward signal. Can this policy be optimal? Justify your answer.
8. Define the recurrence equation used for computing J^{π} for a policy π and the Q-function. Can we stop the recurrence process after a few iterations? Provide bounds.
9. Define the state-action value function *with your own words*. Explain how to compute it. Define the value function *with your own words* and explain how to derive it from the state-action value function.
10. Let us consider the 3x3 grid world below:

18	-1	19
-1	-5	-5
-1	5	-5

An agent is located at one cell at each time. It can access the adjacent cells by making an appropriate move (up, left, down, right). When taking a move, the agent receives the reward from the cell it reaches.

- Describe the grid world as a RL problem.
- Compute Q_5 for each cell and each action with $\gamma = 0.95$ and $\gamma = 0.4$.
- Explain how μ_N^* policies are inferred from Q_N-functions. How does the value of γ modify μ_2^* and μ_5^*?
- Compute the sequence of actions u_0, u_1, u_2, u_3, u_4 that maximises $\sum_{t=0}^{4} \gamma^t r_t$ by exploiting the μ_N^* policies when the initial state is the bottom left cell of the grid world.
- Is this sequence the only one that maximises $\sum_{t=0}^{4} \gamma^t r_t$? Justify your answer.