ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Evaluation 2

1 Questions

1. Define the recurrence equation used to compute $J^{\mu}(x)$. Provide a bound on $||J^{\mu}(x) - J^{\mu}_{N}(x)||_{\infty}$. Explain through words why this bound is important.
2. Describe the main high-level characteristics of the RL problem given in Assignment 1.
3. Define, for a given system dynamics f, reward function r and conditional disturbance probability distribution P_w, the corresponding components of the "equivalent" MDP. Describe an algorithm which computes these components from a given trajectory. How can you compute the sequence of Q_N-functions using the "equivalent" MDP?
4. Explain with your own words the following statement: *[...] if the estimated MDP structure lies in an 'ϵ-neighborhood' of the true structure, then, $J^{\hat{\mu}^*}$ is in a '$O(\epsilon)$-neighborhood' of J^{μ^*} where $\hat{\mu}^*(x) = \lim_{N\to\infty} \arg\max_{u\in U} \hat{Q}_N(x,u)$.*
5. For a given systems dynamics f, reward function r and conditional disturbance probability distribution P_w, is there always an MDP structure to which the algorithm defined in Question 3 will converge to when the length of the trajectory increases? Motivate your answer.
6. What is a contraction mapping? What is a fixed point of a mapping? What can be said about the set of fixed points of a contraction mapping?
7. Define the mapping H that corresponds to the recursive equation used for computing the Q_N-functions. Prove that it is a contraction mapping.
8. Write down the Q-Learning algorithm using the temporal difference. Under which conditions the sequence of $\hat{Q}$-functions computed by Q-learning eventually converges?
9. Give an example of learning ratio for which the Q-learning algorithm may in the general case converge towards the solution of the Bellman equation and another for which it may not converge.

10. Prove by using results related to contraction mappings that the Q-learning algorithm converges.
11. Show an example of a mapping between bounded real-valued functions which is not a contraction mapping. Justify your answer.