ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Evaluation 3

1 QUESTIONS

1. What are the two main problematic issues with the on-line Q-learning algorithm used with function approximators? How can replay buffers address one of these two issues? Propose at least two sampling strategies from replay buffers and explain their pros/cons.
2. Explain the overestimation problem in Q-learning. Propose an approach to avoid it.
3. Is the Q-learning algorithm well suited to handle continuous actions? Explain why. What would be the simplest modification that you would add to the Q-learning algorithm to handle continuous actions? What are its limitations?
4. What are Normalized Advantage Functions? How would you modify these functions to extend their representational power?
5. Describe the DDPG algorithm, and propose at least two variants.
6. Describe at least two extensions to Q-Learning used in *Rainbow* and their pros/cons.
7. In which conditions ϵ-greedy policies show clear limitations? Describe how *Rainbow* addresses them.
8. Which components of *Rainbow*, according to the authors, were the most important? Explain their impact on results.