ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Evaluation 5

1 QUESTIONS

1. Describe the *K*-armed bandit setting. Explain how to assess the performance of a given strategy in this setting.

2. How does a K-armed bandit problem differs from a classical reinforcement learning problem in an MDP?

3. Describe at least three approaches which address the exploration-exploitation problem in the *K*-armed bandit setting and discuss their drawbacks.

4. Describe the *Upper Confidence Bound* (*UCB*), and explains how it follows the *optimism in the face of uncertainty* principle.

5. When does the *UCB* algorithm should switch from exploration to exploitation? Provide and prove bounds.

6. Describe the *Upper Confidence Trees* (*UCT*) algorithm, and explain why, despite the fact that *UCT* is consistent with respect to *UCB*, it may demonstrate poor performance in practice.