ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Evaluation 3

1 QUESTIONS

1. What is supervised batch-mode learning? How is Fitted-Q-Iteration related to batch-mode learning?
2. Describe the Fitted-Q-Iteration (FQI) algorithm. Give examples of supervised learning algorithms that can be used inside FQI. Describe situations where the sequences of Q_N-functions computed with FQI (i) can diverge and (ii) may not converge to the true Q-function but still lead to high quality policies.
3. Give two possible practical stopping conditions for FQI and explain their drawbacks.
4. Can a sequence of Q_N-functions, computed with FQI and tree-based supervised learning, diverge to ∞? If not, provide bounds.
5. Explain how to compute the action maximising a Q-function estimated by a single regression tree for a continuous action space. Does this computation scale with ensemble of regression trees? Justify your answer.
6. Describe two classical results from dynamic programming theory exploited by FQI algorithm.
7. Describe the Bellman residual of a $\hat{Q}$-function. Explain the rationale of this metric *in plain words*. Give an example based on the paper when it is needed along with the score of the derived policy.