RAPHAEL FONTENEAU, University of Liège (Belgium) / Inria Lille - Nord Europe (France)
Batch Mode Reinforcement Learning based on the Synthesis of Artificial Trajectories

Batch mode reinforcement learning (BMRL) is a field of research which focuses on the inference of high-performance control policies when the only information on the control problem is gathered in a set of trajectories. Such situations occur for instance in the case of clinical trials, for which data are collected in the form of batch time series of clinical indicators. When the (state, decision) spaces are large or continuous, most of the techniques proposed in the literature for solving BMRL problems combine value or policy iteration schemes from the Dynamic Programming (DP) theory with function approximators representing (state-action) value functions. While successful in many studies, the use of function approximators for solving BMRL problems has also drawbacks. In particular, the use of function approximator makes performance guarantees difficult to obtain, and does not systematically take advantage of optimal trajectories. In this talk, I will present a new line of research for solving BMRL problems based on the synthesis of "artificial trajectories" which opens avenues for designing new BMRL algorithms. In particular, it avoids the two above-mentioned drawbacks of the use of function approximator.