VINCENT LÉTOURNEAU, University of Ottawa
Complexity measures and regret bounds in reinforcement learning from classical statistical learning theory

I will present recent work that makes use of classical statistical learning theory, specifically complexity measures of supervised learning, to probe the complexity of reinforcement learning problems. We consider a family $\mathcal{M}$ of MDPs over given state and action spaces, and an agent that is sequentially confronted with tasks from $\mathcal{M}$. Although stated for this stepwise change in distributions, the insight we develop is informative for continually changing distributions as well. In order to study how structure of $\mathcal{M}$, viewed as a learning environment, impacts the learning efficiency of the agent, we formulate an RL analog of fat shattering dimension for MDP families and show that this implies a nontrivial lower bound on regret as long as insufficiently many steps have been taken. More precisely, for some constant c which depends on shattering d states, an inexperienced agent that has explored the learning environment for fewer than d steps will necessarily have regret above c on some MDP in the family.