Les mathématiques de l'apprentissage machine
« Atelier sur les défis modernes en théorie d'apprentissage »
23 au 26 avril 2018

Mathematics of Machine Learning
"Workshop on Modern Challenges of Learning Theory"
April 23-26, 2018

# Sub-linear reinforcement learning

Sham M. Kakade *

sham@cs.washington.edu

---

Suppose an agent is an unknown environment and seeks to maximize his/her long term future reward. We consider the basic question: does the agent need to learn an accurate model of the environment before he/she can start executing a near-optimal long term course of actions?

Specifically, this talk will consider the problem of provably optimal reinforcement learning for (episodic) finite horizon MDPs, i.e., how an agent learns to maximize his/her (long term) reward in an uncertain environment. The talk will present a novel algorithm, the Variance-reduced Upper Confidence Q-learning (vUCQ), which is the first algorithm which enjoys a regret bound that is both sub-linear in the model size and that achieves optimal minimax regret. The algorithm is sub-linear in that the time to achieve epsilon average regret is a number of samples that is far less than that required to learn any (non-trivial) estimate of the underlying model of the environment. The importance of sub-linear algorithms is largely the motivation for algorithms such as "Q-learning" and other "model-free" approaches.

vUCQ is a successive refinement method in which the algorithm reduces the variance in the "Q-value" estimates and couples this estimation scheme with an upper confidence based algorithm. Technically, this coupling of these techniques is what leads to the algorithm's strong guarantees, showing that "model-free" approaches can be optimal.

*Dept. of Computer Science & Engineering, University of Washington, Paul Allen Center, Seattle, WA 98195, USA