Optimism for decision making

Rémi Munos (INRIA Lille)

February 24, 2011

The "optimism in the face of uncertainty" principle has been recently successfully applied to complex numerical optimization and sequential decision making problems, e.g., the game of go. The idea is to use regret-minimization algorithms (the so-called bandits) to explore efficiently the set of possible strategies given noisy evaluations of their performance. In this talk I will discuss the Upper Confidence Bounds (UCB) approach when the number of possible strategies is large (possibly infinite), and present the idea of Hierarchical Optimistic Optimization (HOO) algorithms. Some related references are: