Regret analysis in finitely-armed and continuously-armed bandits

Abstract:

The multi-armed bandit problem can be traced back to Robbins (1952), but it is only very recently that we obtained an (almost) complete understanding of this simple model. In the first part of the talk I will present two strategies that contributed to this recent improved understanding: MOSS (Minimax Optimal Stochastic Strategy) which is the first minimax optimal strategy (up to a numerical constant), and Robust UCB which is the first strategy that can deal with heavy tailed distributions. The second part of the talk will focus will on the continuously-armed bandit, and for this problem I will present the HOO (Hierarchical Optimistic Optimization) algorithm.

Biography: Sebastien Bubeck is an assistant professor in the department of Operations Research and Financial Engineering at Princeton University. He joined Princeton after a postdoc at the Centre de Recerca Matematica in Barcelona, where he was working with Gabor Lugosi. He received his Ph.D. in mathematics from the University of Lille 1, advised by Remi Munos, after undergraduate studies at the Ecole Normale Superieure de Cachan. His research focuses on the mathematics of machine learning, with emphasis on problems related to multi-armed bandits. His work was recognized by several awards, such as the COLT 2009 best student paper award, and the Jacques Neveu prize 2010 for the best French Ph.D. in Probability/Statistics.