The multi armed bandit problem

Angelika Steger
ETH Zurich, Switzerland
steger@inf.ethz.ch

In this talk we cover two well studied problems in machine learning resp cognitive neuroscience: the multi armed bandit problem resp. reversal learning setups. The problem is the following: consider a machine with K arms, where each arm provides a random reward from an unknown probability distribution that potentially can change over time. The objective of the player is to maximize the sum of expected rewards by using an appropriate strategy. The crucial tradeoff that the player faces at each step is the tradeoff between exploitation" of the arm that she believes has the highest expected payoff and "exploration" to get more information on the underlying probability distributions of all arms. In this talk we provide new optimal algorithms for several versions of the problem.

Joint work with Maxime Larcher (ETH Zurich) and Robert Meier (ETH Zurich).