
ROBERT KLEINBERG, Cornell

Recharging Bandits

Experimenting with different alternatives and learning from experience are quintessentially human behaviors that, until recently, were largely missing from the realm of computing. Online systems that self-improve by sequential experimentation are now becoming increasingly common. The theoretical foundation for studying such systems is furnished by the so-called multi-armed bandit problem, first formulated by statisticians in the middle of the 20th century.

Traditional multi-armed bandit models posit that the payoff distribution of each action (or "arm") is stationary over time, and hence that the goal of learning is to identify the arm with the highest expected payoff and choose that one forever after. However, in many applications the efficacy of an action depends on the amount of time that has elapsed since it was last performed. Examples arise in regulatory enforcement, online education, and music recommendations. In this talk we introduce a generalization of the multi-armed bandit problem that models such applications. In the course of analyzing algorithms for this problem, we will encounter some interesting questions about coloring the integers subject to constraints on the sizes of gaps between consecutive elements in a given color class.

This talk is based on joint work with Nicole Immorlica.