DENA FIROOZI, University of Toronto

Electronic Trading, Reinforcement Learning, and Mean-Field Games

We consider electronic markets with many participants trading a common asset that is impacted by all participants' actions and propose a batch reinforcement learning approach to estimate the price dynamics, the impact of trading, and the resulting equilibria. The price model and the impact resulting from trading is updated between epochs of active trading. During epochs of active trading, we use a mean-field game (MFG) framework to approximate the equilibria and obtain optimal trading strategies. To do so, we consider parameterized stochastic linear dynamics together with an entropy-regularized linear-quadratic reward function for each trader. Traders do not specify their trading action, but rather the distribution of their trading action. We obtain exploratory dynamics and rewards capturing repetitive learning under exploration – in the sense of relaxed controls – and show that (i) the optimal distribution of trading action for balancing exploration and exploitation is Gaussian, (ii) the exploitation is captured by the mean of the Gaussian distribution, which is influenced by the mean-field and coincides with the classical results for linear-quadratic MFG systems, (iii) the exploration is captured by the variance of the Gaussian distribution. In this talk, we present our ongoing work.