Talk Title: Fair Exploration for Online Learning

Abstract: Exploration is often necessary in online learning to maximize long-term reward, but it comes at the cost of short-term 'regret'. We study how this cost of exploration is shared across multiple groups. For example, in a clinical trial setting, patients who are assigned a sub-optimal treatment effectively incur the cost of exploration. When patients are associated with natural groups on the basis of, say, race or age, it is natural to ask whether the cost of exploration borne by any single group is 'fair'.

So motivated, we introduce the 'grouped' bandit model. We leverage the theory of axiomatic bargaining, and the Nash bargaining solution in particular, to formalize what might constitute a fair division of the cost of exploration across groups. On the one hand, we show that any regret-optimal policy strikingly results in the least fair outcome: such policies will perversely leverage the most 'disadvantaged' groups when they can. More constructively, we derive policies that are optimally fair and simultaneously enjoy a small 'price of fairness'. We illustrate the relative merits of our algorithmic framework with a case study on contextual bandits for warfarin dosing.

Bio: Jackie Baek is a PhD candidate in the Operations Research Center at MIT, advised by Vivek Farias. She received her undergraduate degree from the University of Waterloo, and she has interned at several tech companies including Dropbox, Snap, Bloomberg, and Grail. Her research interests are in the areas of machine learning, algorithmic fairness, and healthcare. Her work has been awarded finalists in the George Nicholson Student Paper Competition and the RMP Student Paper Award.