Monitoring and Intervening on Large Populations of Weakly Coupled Processes with Social Impact Applications

Andrew Perrault

The Ohio State University
perrault.17@osu.edu

Abstract

Many real-world sequential decision problems can be decomposed into processes with independent dynamics that are coupled via the action structure. We discuss recent work on such problems and future directions.

In recent years, there has been increased activity that leverages the tools of sequential decision-making to make the delivery of interventions more precise, impactful and fair. In many intervention delivery problems, the environment is made up of processes that are naturally modeled as having independent dynamics (e.g., patients, establishments, areas of land). The intervention designers have limited resources (e.g., phone calls, inspections, patrols) at each time step and must allocate them across processes to maximize their objective (e.g., number of positive behaviors, inspection-passing status, minimal poaching).

We consider the following broad category of models. Consider M Markov decision processes (MDPs) each with state space S_i. Let $\mathbf{A}$ be a joint action space consisting of an action for each process and may be subject to constraints (e.g., the number of patients who can be called in a given day). Reward is a function of the states and actions taken in each process and may be additive across processes or not (e.g., include a fairness component). Approached naively, these problems are often intractable due to their large size—e.g., they have exponentially many states in the number of processes. Despite this, leveraging the structure can yield efficient algorithms for optimal or near-optimal policies.

One such case where is the restless bandit (RMAB) (Whittle 1988) subclass, where actions are binary (active and passive), reward is additive, and there is a per-period action budget of k active actions. Under a technical condition known as indexability, an asymptotically optimal policy (as M becomes large) can be derived by combining policies that each only depend on the properties of a single arm.

Motivated by a medication adherence problem, we derive indexability conditions for a new class of restless bandits we call *collapsing bandits* (Mate et al. 2020). In this setting, community health workers (CHWs) have a limited budget of phone calls per day that they must allocate across a cohort of patients they are responsible for. Each phone call makes it more likely the patient will stay in the adherent state or reenter it if they have left. At the same time, each phone call reveals the current state of the patient called. Collapsing bandits address a common case where no information about a process' state is available until an action is taken, leading to chain-like belief dynamics.

In follow-up work, we generalize collapsing bandits in two ways. First, we broaden the class of objective functions that is permitted (Mate, Perrault, and Tambe 2021). In experiments, we found that some patients would receive no interventions at all (in contrast to current standard practice), due to low intervention effect for these processes. Risk-awareness in the objective provides a principled way to allocate interventions more broadly. Second, we generalize beyond RMABs to include multiple actions with different costs (Killian, Perrault, and Tambe 2021). In our motivation setting, this allows us to model CHW visits that require more time than calls, but have higher impact. A recent field study (Mate et al. 2022) shows empirical evidence that RMABs can be effective at preventing engagement drops.

Ongoing and future work addresses issues such as scheduling constraints on interventions, high dimensional observation spaces, and exploration.

References

Killian, J. A.; Perrault, A.; and Tambe, M. 2021. Beyond "To Act or Not to Act": Fast Lagrangian Approaches to General Multi-Action Restless Bandits. In *AAMAS 2021*.

Mate, A.; Killian, J. A.; Xu, H.; Perrault, A.; and Tambe, M. 2020. Collapsing Bandits and Their Application to Public Health Interventions. In *NeurIPS 2020*.

Mate, A.; Madaan, L.; Taneja, A.; Madhiwalla, N.; Verma, S.; Singh, G.; Hegde, A.; Varakantham, P.; and Tambe, M. 2022. Field study in deploying restless multi-armed bandits: Assisting non-profits in improving maternal and child health. In *AAAI 2022*.

Mate, A.; Perrault, A.; and Tambe, M. 2021. Risk-Aware Interventions in Public Health: Planning with Restless Multi-Armed Bandits. In *AAMAS 2021*.

Whittle, P. 1988. Restless bandits: Activity allocation in a changing world. *J. Appl. Probab.*, 25(A): 287–298.