

# Artificial Replay: How to get the most out of your data

*Thursday, June 20, 2024 4:30 PM (30 minutes)*

How best to incorporate historical data for initializing control policies is an important open question for using RL in practice: more data should help get better performance, but naively initializing policies using historical samples can suffer from spurious data and imbalanced data coverage, leading to computational and storage issues. To get around this, we will propose a simple meta-algorithm called Artificial Replay for incorporating historical data in control policies. We will first illustrate this for multi-armed bandits, showing how our approach uses a fraction of the historical data compared to a full warm-start, while achieving identical regret guarantees. Next we will extend this to a much more general class of problems we call Markov Decision Processes with Exogenous Inputs (or Exo-MDPS), where the uncertainty affecting the system can be represented as being exogenous to the system state. Here, we will show how our algorithms achieve data-efficiency by leveraging a key insight: using samples of the exogenous input we can infer counterfactual consequences, that then accelerate policy improvements. We will discuss how we can show formal regret guarantees for such systems using the compensated coupling, and also demonstrate its use in virtual machine allocation on real datasets from a large public cloud provider, where our approach outperforms domain-specific heuristics, as well as alternative state-of-the-art reinforcement learning algorithms.

**Primary author:** BANERJEE, Siddhartha (Cornell University)

**Co-authors:** LEE YU, Christina (Cornell University); SINCLAIR, Sean (MIT)

**Presenter:** BANERJEE, Siddhartha (Cornell University)

**Session Classification:** Parallel session: Learning and optimization