
A subcortical basis for policy compression

Alexander Tesmer^{*1}, Rafael Polania¹, and Denis Burdakov¹

¹ETH Zurich – Switzerland

Abstract

The optimal action for an individual is often dependent on the current state of its environment. In reinforcement learning (RL), state-dependent action selection is governed by a policy, whereby increasingly complex policies can improve task performance. However, adopting a policy incurs a cognitive cost, and the mechanisms by which the brain may "compress" a policy are not fully understood. Here, we use a spatial multi-armed bandit task, wherein mice could make self-paced decisions to forage for food rewards. Using an actor-critic framework to model behavior, we show that mice engage with the task and form state-dependent representations of optimal action selection, allowing us to quantify policy complexity. We combined this approach with neuron-specific pharmacological manipulations and recordings to begin to unravel neural causes and correlates of policy complexity and compression.

^{*}Speaker