Efficient Control Using Temporally Layered Architecture

Artificially intelligent agents trained by reinforcement learning algorithms often have a fast constant response times and action frequency. However, a fast response time comes at the cost of more FLOPs and energy consumption. We propose a novel temporally layered architecture that can adapt its response time based on the state of the environment. Our algorithm layers a fast controller on a slower controller allowing the agent to vary its response time between the slow and fast. The network is trained in a layer-wise manner using the TD3 actor-critic algorithm, so that the faster network is trained after the slower network. The actions picked by both the networks are added up. However, the fast actions are gated based on their effect on the final action. We evaluate our method in a real-time setting on the suite of OpenAI gym MuJoCo tasks and demonstrate that our method outperforms the constant frequency networks on most environments and obtaining a better return per action in every environment tested.