KNOW YOUR ACTIONS: DECISION MAKING WITH ACTION INTERDEPENDENCE

Anonymous authors
Paper under double-blind review

ABSTRACT

In the recent development of deep reinforcement learning approaches, we have seen some successful attempts of dealing with the large action space, e.g., permutation or combination of primitive-actions(multi-discrete action space; Berner et al. (2019), Guss et al. (2019), recommender systems; Chen et al. (2019), Park et al. (2020)).maybe I can also talk about that conventinoally ppl use the point-wise prediction to deal with it? One of the promising approaches in this direction is framing the listwise action space as a sequential decision-making or directly learning a listwise value estimation. However, these approaches learn a policy and value function only from a fixed candidate set of primitive-actions, not by considering all the available primitive-actions at the time. Hence, this can result in poor performance when the characteristic of the available primitive-actions is different from the ones during training so that the relationship of primitive-actions in the candidate set is no longer true in another candidate set. To this end, we propose a generalisable policy architecture to a non-stationary and large set of candidate primitive-actions, called *Graph Cascading Deep Q-network(GCDQN)*. For a generalisable agent to candidate set, we learn two different inter-dependencies of primitive-actions in the candidate set that include (1) Intra-slate: How interaction among primitive-actions in a selected slate affects the task. (2) Complement: How availability of other primitive-actions affects an primitive-actions's relevance Then, we use the learned representations for decision-making, e.g., physical reasoning and recommendations. Specifically, we use graph neural networks(GNNs) to encode relations between primitive-actions in the candidate set and the agent utilises this information in decision-making. We evaluate our project on tasks in recommender systems(public simulator/dataset (e.g., RecSimIe et al. (2019a), MovieLensHarper & Konstan (2015)) and control tasks(CREATEJain et al. (2020)). Based on the results of those simulated experiments, we apply the agent to the real world task(LINE Wallet Recommender Systems Park et al. (2020)).

1 INTRODUCTION

Imagine you want to hang a frame on the wall. Your approach may involve placing a screw on the wall and then using a drill to screw it in. However, if you do not have a drill at home, you would not use the screw. Instead, you would apply alternative approaches such as using a hook and adhesive-strips, or a nail and a hammer. In general, we solve sequential decision-making tasks by choosing actions that interact with each other to achieve a desired outcome in the environment. Therefore, such interaction among actions makes it important for every action decision we make to account for other actions available to us.

In this work, we propose the problem of modeling action interdependence in reinforcement learning with a varying action space. While typical RL algorithms assume a fixed action space, prior work has explored varying action space in the context of generalization to unseen actions (Jain et al., 2020), adding new actions in lifelong learning (Chandak et al., 2020), and dealing with new items for recommendations in recommender systems Chen et al. (2019).In RecSys we assume to have the item embedding so that although I couldn't find any RL recsys work focussing on item cold start problem, ppl just say that by putting through the new item's embedding an agent can provide recommendation.. However, the policy architectures proposed by these works do not account for