Learning Dirichlet Priors for Affordance Aware Planning

David Abel, Gabriel Barth-Maron
James MacGlashan, Stefanie Tellex
Brown University
Providence, RI, USA
{dabel, gabrielbm, jmacglashan, stefie10}@cs.brown.edu

Abstract

Planning algorithms for non-deterministic domains are often intractable in large state spaces due to the well-known "curse of dimensionality." In previous work, we introduced a novel, state- and reward- general approach to limiting the branching factor in large domains by reasoning about the domain in terms of *affordances* [1]. Our affordance formalism can be coupled with a variety of planning frameworks to create "affordance aware planning," allowing an agent to efficiently search the state-space by focusing on relevant action possibilities. This corresponds to highlighting useful actions on a state by state basis when solving an MDP.

The relevant actions that are returned by an affordance are a subset of the total available actions as defined by a Markov Decision Process (MDP). **JM: It might be better to only say that it's drawn from a probaility distribution, since in the case of determinisitic affordances currently tested, it's not a Dirichelt. Then, when you talk about learnign you can say a Dirichlet is used.** The probability that a set of actions is returned is given by a dirichlet-multinomial distribution, where the parameter N (the number of actions in the set) is a random variable that determines the size of the action set to be returned.

Previously, we provided planners with optimal affordances **JM: what an optimal affordance is will be unclear; maybe just say expert provided affordances?**, leading to massive speed ups in planning compared to the performance of the algorithms without affordances. To improve the transferability of affordance-aware planning, we propose learning affordances through a scaffolding process [**?**] to avoid hand crafting knowledge **JM: I'm not sure I'd say that the purpose of learning affordances is to improve tansferability, but rather to remove the expert from the equation**. We randomly generate a large number of simplified state spaces that are representative of more complicated tasks the agent must form a policy over. **JM: maybe instead:**

We randomly generate a large number of simplified environments for possible tasks the agent will need to solve that are representative of the more complicated environments in which the agent will ultimatly need to solve the tasks. Next, we form policies over each of the simplified state spaces and use these policies to create the parameters $\boldsymbol{\alpha}$ for the dirichlet-multinomial distribution. Additionally, we use a set of optimal trajectories taken from these policies to form a dirichlet distribution over N. **JM: I'm not sure it will be entirely clear what you're doing here. Something probably needs to be said about looking at the frequency of optimal actions for each possible affordance to model what the important actions are. I'm also not sure I understand why trajectories are analyzed in addition to the policy, since the former is a superset of the latter.**

We are still collecting experimental data on hard-coded vs learned affordances. We believe that the learned affordances will perform slightly worse than hand crafted affordances, and substantially better than planners without affordances.

References

[1] JJ Gibson. The concept of affordances. *Perceiving, acting, and knowing*, pages 67–82, 1977.