Robot Behavioral Exploration and Multimodal Perception using POMDPs

Shiqi Zhang\textsuperscript{1,2}, Jivko Sinapov\textsuperscript{2}, Suhua Wei\textsuperscript{1}, and Peter Stone\textsuperscript{2}

\textsuperscript{1} Department of Electrical Engineering and Computer Science, Cleveland State University
\textsuperscript{2} Department of Computer Science, The University of Texas at Austin

1 Introduction

Service robots are increasingly present in everyday environments, such as homes, offices, airports and hospitals. A common task for such robots involves retrieving an object for a user. Consider the request, “Robot, please fetch me the red empty bottle”. A key problem for the robot consists of deciding whether a particular candidate object matches the properties in the query. For certain words (e.g., heavy, soft, etc.) visual classification of the object is insufficient as the robot would need to perform an action (e.g., lift the object) to determine whether it is empty or not. Furthermore, the robot would need to decide which actions (possibly out of many) to perform on an object, i.e., it would need to generate a behavioral policy for a given request.

Recent research in robotics has shown that robots can learn to classify objects using computer vision methods as well as non-visual perception coupled with actions performed on the objects (Högman, Björkman, and Kragic 2013; Sinapov et al. 2014; Thomason et al. 2016). For example, a robot can learn to determine whether a container is full based on the sounds produced when shaking the container (Sinapov and Stoytchev 2009); or learn whether an object is soft or hard based on the haptic sensations produced when pressing it (Chu et al. 2015). Nevertheless, there has been relatively little emphasis on enabling a robot to efficiently select actions at test time when it is tasked with classifying a new object. The few approaches for tackling action selection, e.g., (Rebguns, Ford, and Fasel 2011; Fishel and Loeb 2012), assume that only one target property needs to be identified (e.g., the object’s identity in the case of object recognition) and would not scale to requests such as the one presented earlier.

To address this limitation, we propose to generate behavioral exploration policies for a given request using the partially observable Markov decision process (POMDP) formalism. POMDP (Kaelbling, Littman, and Cassandra 1998) is a general framework that does not have the assumption of full observability over current state, so an agent needs to use its local, unreliable observations to estimate the underlying state and maintains a distribution over possible states. As a result, POMDPs have been used in object exploration in robotics. For instance, hierarchical POMDPs were used for suggesting visual operators for exploring multiple objects on a tabletop scenario (Sridharan, Wyatt, and Dearden 2010), and more recent work further used a robotic arm to move objects enabling better visual analysis (Pajarinen and Kyrki 2015). However, the sensing in these research is limited to robot vision and other modalities such as audio and haptics are not used.

Although multimodal perception and POMDP-based object exploration have been studied previously, to the best of our knowledge, there is no research that integrates both in robotics. In this work, given queries about object properties, we dynamically construct POMDPs using a data set collected from a real robot. Experiments on exploring new objects show that our POMDP-based object exploration strategy significantly reduces the overall cost of exploration actions without hurting accuracy, compared to a baseline strategy that uses a predefined sequence of actions.

2 POMDP-based Object Exploration

We construct a POMDP for guiding the robot’s exploration behavior based on the robot’s sensing and actuating capabilities. A simplified version of the observable aspect of our POMDP’s transition diagram is shown in Figure 1, where object properties, as a part of the world state, are not shown. A standard POMDP model is a 7-tuple $\langle S, A, T, R, \Omega, O, \gamma \rangle$, where $\gamma$ is the discount factor that represents how much immediate rewards are favored over more distant rewards. In our case, $\gamma = 0.99$, which means the robot has a relatively long horizon in planning.

- $\mathcal{S} : S^o \times S^h \cup \text{term}$ is the state set. It includes a Cartesian product of sets $S^o$ and $S^h$, and a terminal state term. $s^o \in S^o$ corresponds to one of the non-terminal states $(s^o_0, \cdots, s^o_5)$ in Figure 1. $s^h \in S^h$ is specified by all attributes of a given
overall cost 0.130 33.87 (8.78)
Two 0.245 0.860 29.85 (12.87) 37.10 (0.00)
Two 17.56 (30)

and reward functions of our POMDP are learned from an predefined action sequence includes all exploration actions.

Experiments have been conducted to evaluate our POMDP-based policy generation (Kurniawati, Hsu, and Lee 2009).

We use an approximate, point-based POMDP solver for previous experience of object exploration.

Markov decision processes (POMDPs) to help robots select actions for multimodal perception in object exploration tasks. Our approach can dynamically construct a POMDP model given an object description from a human user (e.g., “a blue heavy bottle”), compute a high-quality policy for this model, and use the policy to guide robot behaviors (such as “look” and “shake”) toward maximizing information gain. Experimental results show that our POMDP-based exploration approach enables the robot to identify object properties more accurately without introducing extra cost from exploration actions, compared to a baseline that suggests actions following a predefined action sequence.

In this paper, we investigate using partially observable Markov decision processes (POMDPs) to help robots select actions for multimodal perception in object exploration tasks. Our approach can dynamically construct a POMDP model given an object description from a human user (e.g., “a blue heavy bottle”), compute a high-quality policy for this model, and use the policy to guide robot behaviors (such as “look” and “shake”) toward maximizing information gain. Experimental results show that our POMDP-based exploration approach enables the robot to identify object properties more accurately without introducing extra cost from exploration actions, compared to a baseline that suggests actions following a predefined action sequence.

In the future, we plan to evaluate our approach using a larger, more recent data set we collected using a real robot (Thomason et al. 2016). Another direction is to better implement the question-asking action as a POMDP-based dialog system (Zhang and Stone 2015), and potentially use a single POMDP for both multimodal and language-based perception. Finally, we plan to implement and evaluate this approach on a real mobile robot platform.
References


Thomason, J.; Sinapov, J.; Svetlik, M.; Stone, P.; and Mooney, R. J. 2016. Learning multi-modal grounded linguistic semantics by playing I Spy. In *Proceedings of the Twenty-Fifth International joint conference on Artificial Intelligence (IJCAI)*.

Zhang, S., and Stone, P. 2015. CORPP: Commonsense reasoning and probabilistic planning, as applied to dialog with a mobile robot. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*. 