Learning to Manipulate in Open Environments

Khoshrav P. Doctor
kdoctor@cs.umass.edu

Michael W. Lanighan
mlanighan@traclabs.com

Dirk Ruiken
dirk.ruiken@honda-ri.de

Roderic A. Grupen
grupen@cs.umass.edu

Primates (including humans) have evolved subtle kinematic adaptations that enhance the capacity of their hands to perceive the environment as well as to manipulate it. This duality makes sense when we observe that natural environments are dynamic, uncertain, and partially observable. To successfully manipulate objects in this setting requires that uncertainties, specifically those that can cause catastrophic outcomes, are actively suppressed. We expect these fundamental challenges are not specific to biological systems and evolution, but are instead fundamental properties of the manipulation domain that arise due to the wide range of sensitivity that manual interactions can have to context. We propose that these observations suggest that the robotic analogs of manipulation behavior in the animal kingdom must actively control uncertainty in a manner that is object- and context-dependent and, thus, that balances the risk of acting against the cost of gathering additional information from the environment.

A natural mathematical framework for this study is the Partially Observable Markov Decision Process (POMDP). This paper reviews an experimental platform for evaluating our hypotheses using a class of model-based, belief-space planning techniques that generate approximate solutions for POMDPs. Our presentation reviews a vertically integrated architecture grounded in a "landscape of attractors" for representing mobile manipulation tasks and providing structure that makes planning in this domain tractable. Actions in this framework are parametric closed-loop controllers that either reference observed stimuli from the environment or sample goals from models of previous experiential data. The former establish concrete "facts" about the run-time environment and the latter exploit probabilistic models of robot-world interactions. In this framework, we define a visuotactile state called the generalized *aspect* and model a context dependent probabilistic transition function between such states called the generalized *Aspect Transition Graph (ATG)* that summarizes observed control affordances. An *Active Belief Planner (ABP)* rolls-out possible future states and uses information theoretic measures to identify actions that condense belief into the partition of the state space that conform with the task specification. These ABPs can also be arranged hierarchically to reason over different levels of complexity and leverage common structure at higher levels of planning.

After a brief introduction to this experimental architecture, we focus on extensions of previous work to learn the required models using an intrinsically-motivated structure learning technique that is designed to learn task-independent ATGs from direct exploratory interactions with objects in the mobile manipulation domain. We use no prior knowledge about these objects to build these ATGs. We believe this learning process should adapt to novel objects/environments and must also consider the uncertainty arising from its interactions while building these graphs.

Although learning probabilistic transition functions (models) can be accomplished using frequency-based empirical techniques, a large amount of direct exploration is required. Instead, we propose using Intrinsically-Motivated Structure Learning (IMSL) to efficiently generate the models. The resulting models are task-independent and significantly more sample efficient to obtain in practice. This system actively resolves the difference between observed transitions and expectations derived from a distribution over the generative models (ATGs). In particular, we use the information gain resulting from experiments *in situ* as a reward to focus learning on states and actions that are poorly understood or that are sensitive to the run-time context.

Unlike much of the literature in belief-space planning, our results show that this architecture yields active planning techniques that justify the computational expense of using belief-space planning techniques relative to a strawman based exclusively on random exploratory actions. Moreover, this framework supports "active belief" systems that can avoid committing to brittle or catastrophic policies in the context of realistic application domains.

The presentation will include a summary of experiments conducted on the uBot mobile manipulator that uses this approach to generate approximate, model-based solutions to POMDPS, including: recognition, pick-and-place and simple assemble tasks.