Graduate Research Plan Statement

Title: Deep reinforcement learning for robust plan modification and execution in robotic manipulation settings.

Research Objective: To explore the ways in which deep reinforcement learning algorithms can be used to fine-tune traditional robotic control paradigms for dexterous manipulation, without discarding the predictability and safety of these traditional paradigms.

Introduction: One major barrier for robots to be useful for manipulation tasks in household environments is the general lack of structure found in such environments. No two homes are alike, and household objects and furniture vary considerably from home to home. Humans are exceptionally good at handling these profoundly unstructured environments, and have no issue performing useful manipulation in nearly any environment. While working on robotic learning at the Samsung AI Center in New York, I observed how challenging these environments can be for traditional robotic planners. Traditional paradigms such as RRT and sampling-based planning work well when equipped with perfect knowledge of the environment and objects in it, but are quite brittle to novel situations and uncertainty within the environment [1].

Recently, deep reinforcement learning algorithms have shown promise in learning robust, sophisticated behaviors that can adapt to uncertainty and novelty in a range of environments [2]. However, these algorithms are notoriously challenging to apply to real-world robotics. Even using the sample-efficient learning algorithms our group worked on, these algorithms can often take tens or hundreds of thousands of trials to learn simple behaviors, which can be prohibitively expensive in robot time. Furthermore, it is difficult to make guarantees about learned behaviors in all situations, leading to potential issues of safety on real robots [3]. While there has been some compelling work on learning to correct physics models for robotic tasks [4], there has been little work that uses existing robotic planning and control methods in tandem with deep reinforcement learning to learn robust control policies quickly and safely.

Research Plan: In my previous research contributions, I explored the challenges of learning for dexterous manipulation [5] and investigated how the learning dynamics of deep reinforcement learning for control tasks can be improved [6]. As a Ph.D. student I propose to synthesize my experience in both domains, and examine how deep reinforcement learning can be used to augment existing classical robotic control techniques in order to increase their robustness to real-world dexterous manipulation tasks with less structure. At a high level, I will consider how the actions output by traditional planners can be improved by learning from interaction. I plan to execute this exploration in three stages:

Stage I: *Theoretical exploration of plan enhancement and simulation-driven evaluation*. The first stage of this investigation will involve the creation of a mathematical framework for unifying traditional planning methods and deep reinforcement learning. Consider an existing plan $\mathbf{P}$ consisting of $\mathbf{n}$ execution steps $\{\mathbf{p_1, p_2, \ldots, p_n}\}$. I will seek to design a deep reinforcement learning paradigm to learn a policy $\boldsymbol{\pi_\theta(s_i, p_i)}$ that incorporates both the existing planned action $\mathbf{p_i}$ and a current observation $\mathbf{s_i}$ to output an altered action $\mathbf{p_i}$'. The policy parameters $\boldsymbol{\theta}$ will be conditioned on a reward signal provided by the environment. I will then explore the learning dynamics and effectiveness of this formulation on various manipulation-based environments in simulation.

<u>Stage II:</u> *Application of learning framework to real-world robotic manipulation tasks*. In my previous work on learning to grasp from vision [5], I realized how challenging it could be to transfer algorithms from simulated robots to physical ones. Once the proposed learning method has been refined and validated on simulated manipulation tasks, I will proceed to design and construct a physical robotic manipulation task consisting of a functioning robotic manipulator and real sensors (RGB-D vision, proprioception) and a set of challenging household objects. I will then choose an appropriate classical planning algorithm with which to generate initial plans to be used in the above learning method, and evaluate the real-world characteristics of the learning framework, with an eye towards robustness to uncertainty and learning speed.
<u>Stage III:</u> *Expand the representational capabilities of the method to support plan modification*. Many of the errors I have encountered in working with traditional planners involved issuing plans with missing or fundamentally incorrect steps. After validation on a real robotic manipulation system, I will investigate formulations of this method that are capable of learning to add, remove, or replace steps to the original plan given by traditional planning algorithms. This will allow the system to be robust to inadequate initial plans, and to discover novel strategies not considered by traditional planners.

Intellectual Merit: If successful, this project will be an important contribution to bridging the gap between traditional control paradigms and the real manipulation environments humans operate in. Many reinforcement learning paradigms lack the ability to incorporate reasonable human intuition and prior knowledge during learning, and the proposed framework would be an important step towards leveraging human intuition and physical models in robotic learning. My background in both robotic learning and deep reinforcement learning is crucial to the success of this work, because both robotics and reinforcement learning independently have many nuanced components that can be challenging to get right. This project involves cross-domain expertise, so I am seeking to conduct this research at an institution which has strong robotics and machine learning faculty.

Broader Impacts: Robust robotic manipulation that can operate under diverse household conditions could help fill many of the much-needed labor gaps in homes. As demographics shift, the need for physical care for the aging will increase drastically, and will likely go unmet. Robotic manipulation systems have the potential to help reduce this burden, and empower the aging and physically-limited to maintain their independence and remain in their own homes for longer. My research contributions will make an important social impact by making these systems both computationally tractable and receptive to existing human intuition and analytical planning.

References:
[1] Saxena, A., Wong, L. L., & Ng, A. Y. (2008, July). Learning grasp strategies with partial shape information. In AAAI (Vol. 3, No. 2, pp. 1491-1494).
[2] Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. The Journal of Machine Learning Research, 17(1), 1334-1373.
[3] Kos, J., & Song, D. (2017). Delving into adversarial attacks on deep policies. arXiv preprint arXiv:1705.06452.
[4] Zeng, A., Song, S., Lee, J., Rodriguez, A., & Funkhouser, T. (2019). TossingBot: Learning to Throw Arbitrary Objects with Residual Physics. arXiv preprint arXiv:1903.11239.
[5] Tosun, T., Mitchell, E., Eisner, B., Huh, J., Lee, B., Lee, D., ... & Lee, D. (2019). Pixels to Plans: Learning Non-Prehensile Manipulation by Imitating a Planner. arXiv preprint arXiv:1904.03260.
[6] Simmons-Edler, R., Eisner, B., Mitchell, E., Seung, S., & Lee, D. (2019). QXplore: Q-learning Exploration by Maximizing Temporal Difference Error. arXiv preprint arXiv:1906.08189.