Research Proposal

A. Hypothesis

Interactive (deep) reinforcement learning aims to efficiently acquire new skills or align the learning agent's behavior to certain human preferences (i.e. learn a human reward function) by taking human advice/feedback (e.g. binary evaluative feedback, preference labels, or action advice). However, due to the high demand for training samples and human inputs, interactive (deep) RL hasn't been widely adopted in real-world scenarios. For instance, an autonomous agent usually needs hundreds of "good" or "bad" evaluative feedback labels to capture a simple objective of the user (e.g. a passenger wants the self-driving taxi to slow down when he/she is drinking coffee). This can result in an extremely poor user experience.

One way to make interactive RL more feasible is to resort to more informative human-agent interaction, such as **multi-modal human advice**. Lin's previous attempts include visual explanation augmented binary feedback [1] and gaze-informed demonstrations [2]. Although results indicate that the richer human guidance is a promising direction to pursue, Lin believes interactive RL can be further improved by addressing the following two fundamental problems:

a. AI system's inability to ground and accept **explicit knowledge-based** advice. Shaping agent behavior via plain numeric labels is extremely slow since numeric labels are not informative enough to convey human intents.

b. AI system's inability to leverage humans' symbolic abstractions in its own decision-making.

B. Methodology

Lin aims to tackle the aforementioned problems from a neuro-symbolic perspective:

a. Building a symbolic interface to construct **lingua franca** between humans and agents, and accept symbolic advice. The symbolic interface could be implemented with (few-shot) neural-symbolic concept learners; and the task of the interface is to interpret human advice given the scene and agent state, and ground the advice into the representation that can be used by the learning agent.

b. Neural-symbolic architectures as an inductive bias to leverage human symbolic abstractions.
Reasoning at the level of symbolic concepts and pixels simultaneously can be a key to efficient policy learning under human guidance.

C. Expected Outcome

Objectives to achieve: i) a more natural and informative way for the human to interact with the agent; ii) significantly lower human efforts in providing guidance; iii) more efficient policy learning (in terms of environment sample complexity). The ultimate goal is to build advisable RL systems that enable more natural and efficient human-agent interactions so that AI agents can seamlessly serve or collaborate with humans.

Reference

[1] **Guan, L.**, Verma, M.K., Guo, S., Zhang, R., & Kambhampati, S. (2020). Widening the Pipeline in Human-Guided Reinforcement Learning with Explanation and Context-Aware Data Augmentation. NeurIPS 2021.

[2] Zhang, R., Liu, Z., **Guan, L.**, Zhang, L., Hayhoe, M.M., & Ballard, D.H. (2020). Atari-HEAD: Atari Human Eye-Tracking and Demonstration Dataset. AAAI 2020.