Bridging Action and Perception: Towards Continual Reinforcement Learning

Khimya Khetarpal, Doina Precup

McGill University/Mila

My **research goal** is to develop Artificial Intelligence (AI) systems that can understand the data observed ("see"), learn to represent the world knowledge ("think"), and plan with acquired information ("act"), with the ability to adapt to changes, and become better learners over time.

Attend Before You Act

Human foveated attention images
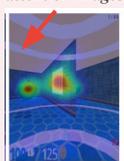
Training an agent to see in the same way

Distractor (Hue, Flickering, etc.)

Knowing where to look ("*see*") in an image allows RL agents to be more robust to distractors when *"acting"* in tasks. [1]

Temporal Abstraction with Interest Functions

Learn & leverage attention mechanism (interest functions) to discover *specialized* action representations ("*think*") to *adapt* faster ("*act*"). [2a,2b,4]
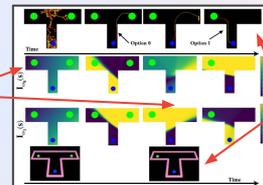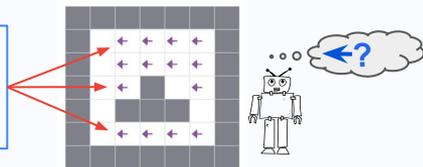
Task Changes: Goal is now located only on left end of the T-Maze

A Theory of Affordances in RL

Interactive behavior requires also *dynamically* tracking and updating *action possibilities* (aka ``affordances'').

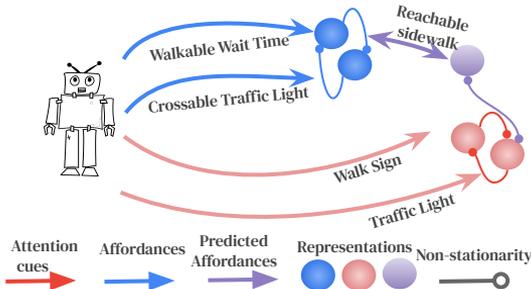Affordances are the subset of states and actions which complete the intent to go left.
$$\mathcal{AF}_\mathcal{I} \subseteq \mathcal{S} \times \mathcal{A}$$

$$\left\| V_M^* - V_M^{\pi^*_{\hat{M}_{\mathcal{AF}_\mathcal{I}}}} \right\|_\infty \leq \frac{2\text{Rmax}}{(1-\gamma)^2} \left(2\gamma\epsilon + \sqrt{\frac{1}{2n} \log \frac{2|\mathcal{AF}_\mathcal{I}||\Pi_\mathcal{I}|}{\delta}} \right)$$

Our main result shows that *planning* with affordances decreases the *computational complexity*! [3]

Towards Continual RL: A Unifying Framework

Reachable sidewalk

Walkable Wait Time

Crossable Traffic Light

Walk Sign

Traffic Light

Attention cues — Affordances — Predicted Affordances — Representations — Non-stationarity

References and Contact

[1] [Khetarpal & Precup, ICML 18– LLARLA] Attend Before you Act: Leveraging Human Visual Attention for Continual RL
[2] [Khetarpal et al. AAAI 19-SA, AAAI 20] Options of Interest: Temporal Abstractions with Interest Functions
[3] [Khetarpal et al. ICML 20] What can I do here? A Theory of Affordances in RL
[4] [Khetarpal, AAAII9-DC] Learning Generalized Temporal Abstractions across Both Action and Perception
[5] [Pezzulo et al. CogSci Trends] Navigating in the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition

Webpage: https://kkhetarpal.github.io/ **Email:** khimya.khetarpal@mail.mcgill.ca