Integrating Latent Cause Inference into Actor-Critic Algorithms for Continual Learning

Nora Xiao[1], Sashank Pisupati[2], Yael Niv[2,3]

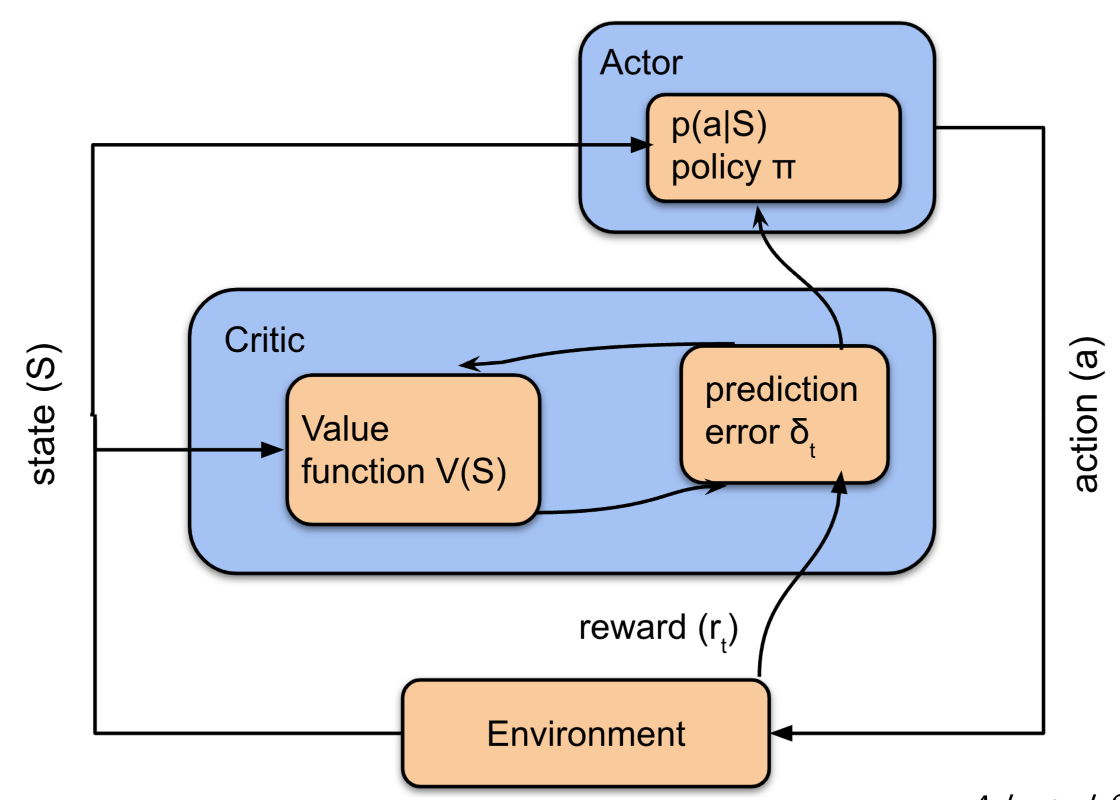
[1]University of North Texas • [2]Princeton Neuroscience Institute, Princeton University • [3]Department of Psychology, Princeton University

Background

- Reinforcement learning in the real world is often faced with changing environments that nevertheless have recurring, hidden structures.
- The challenge of "continual learning" in such situations involves adapting quickly to change while being able to retain and reuse old knowledge.
- One strategy for effective learning in such environments is for agents to discover useful state representations, by extracting latent causal structure and using it to control generalization and segregation across learning experiences.

Actor-Critic Model

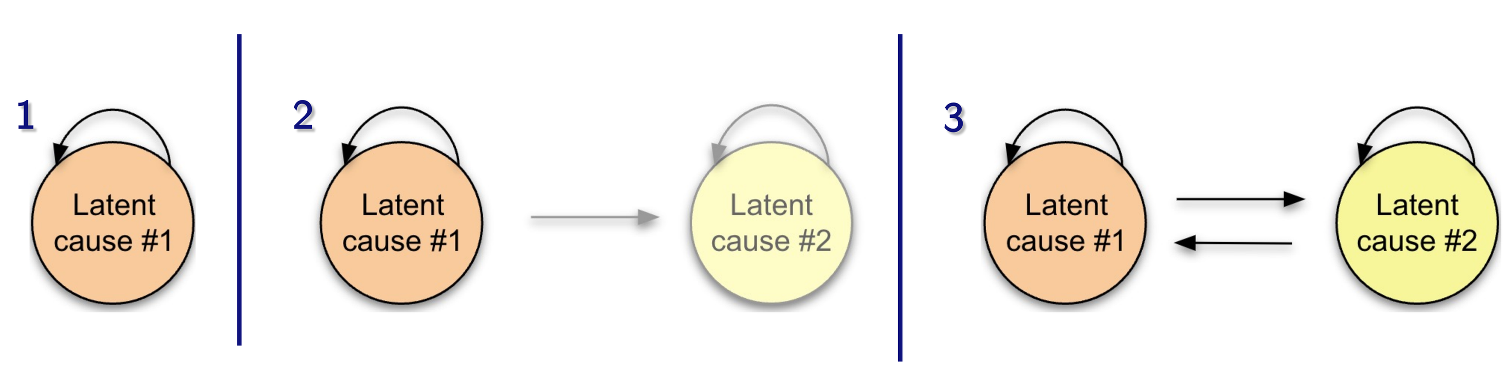
- Our methods involve an actor-critic algorithm connected to a latent cause inference module.



Adapted from Niv, 2008

Reinforcement Learning Agents

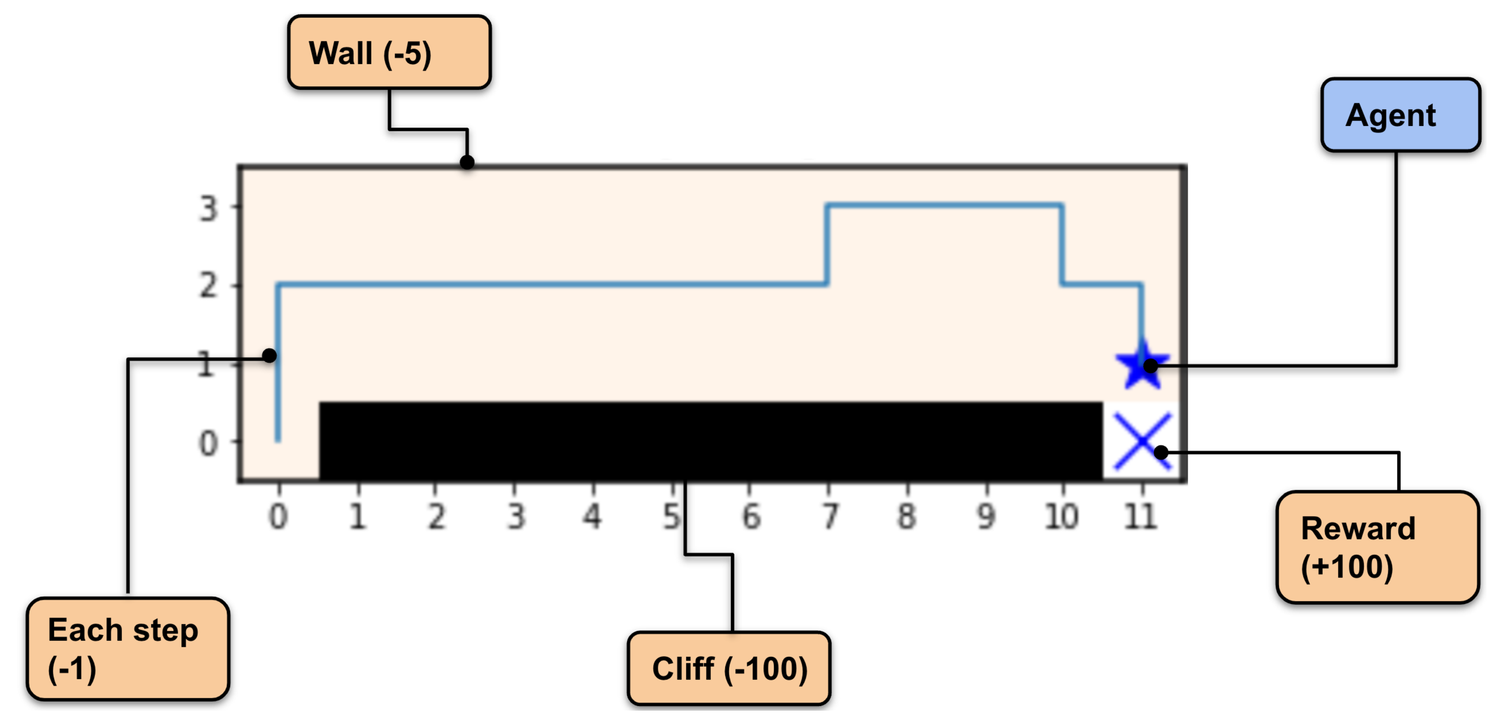
- We investigated the learning dynamics of three different reinforcement learning agents, with varying degrees of access to latent cause representations.



1. Assumed that the same latent cause was active throughout training, leading it to incorrectly apply old knowledge even if the reward location changed, and overwrite previous knowledge through subsequent learning.
2. Only had access to changes in latent causes, leading it to adapt quickly to change-points but requiring it to re-learn every time, even if it had seen a latent cause before.
3. Had access to all past latent causes, enabling it to segregate learning when the task changed abruptly, but also "remember" and apply previous knowledge to tasks it had already seen.

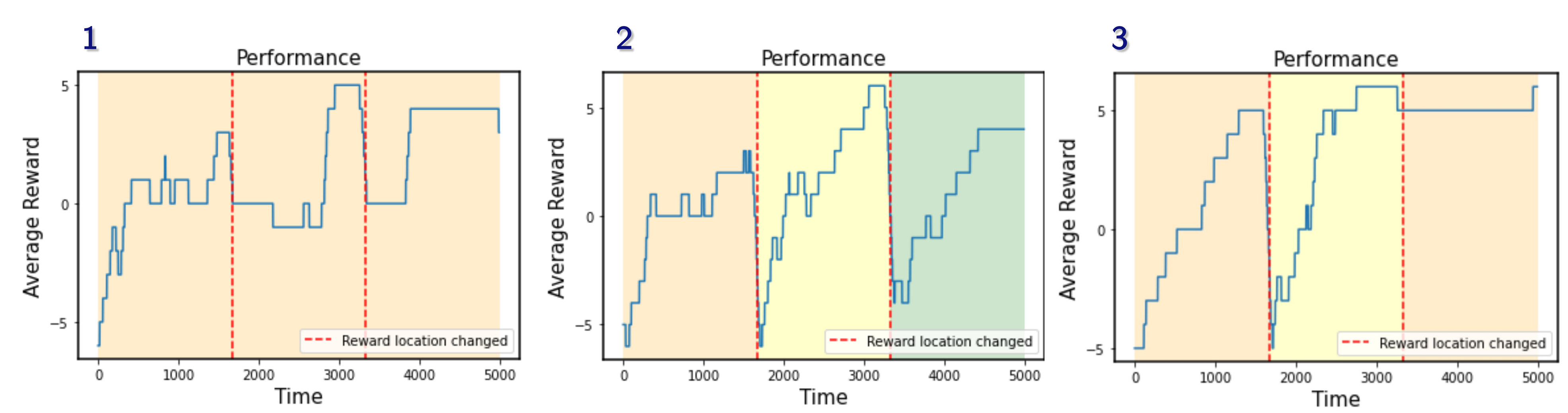
Grid-world Setting

- The agent's objective is to find the optimal path to a reward in a simulated grid-world setting.
- The reward location is determined by a hidden cause that could change and recur throughout training.



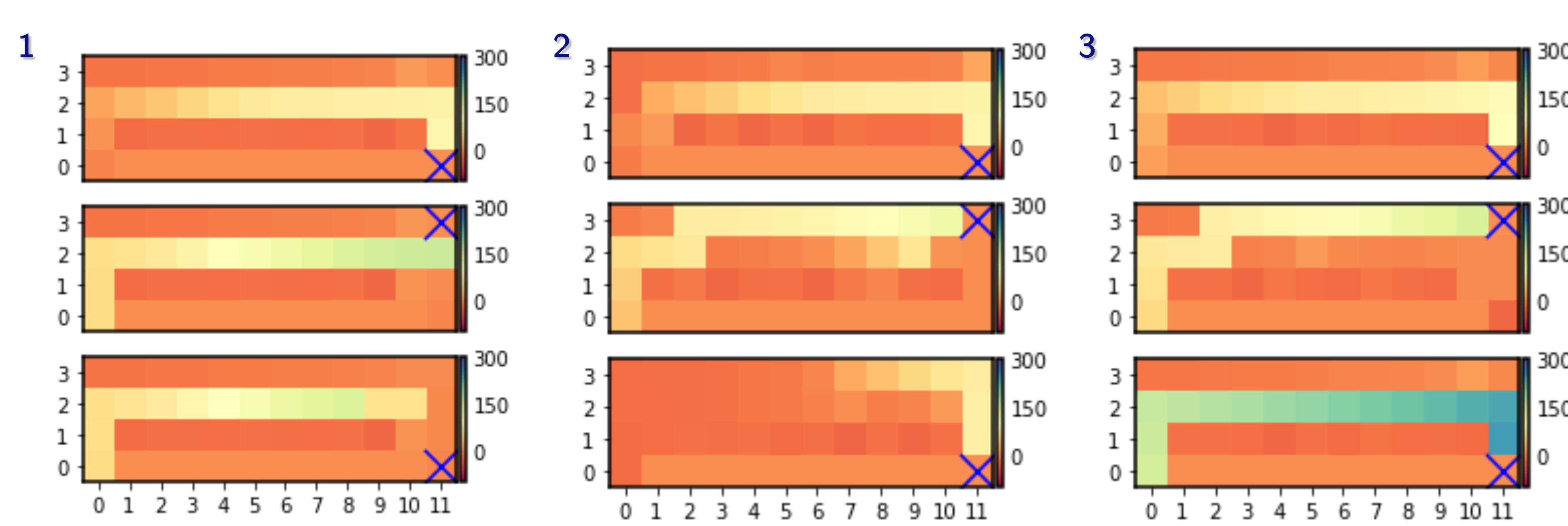
Performance Curves

- The average reward received during training across 5000 timesteps was plotted for each of the 3 agents.
- A difference in shading between each third of each plot represents a change in what the agent "thinks" the latent cause is.



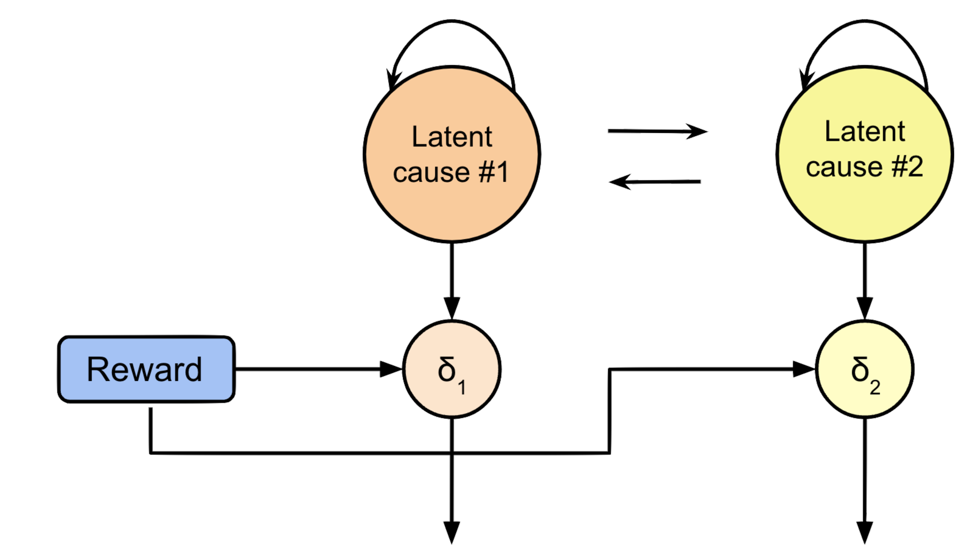
State Value Functions

- Heatmaps of the state value functions for each of the 3 reward locations for each agent.



Ongoing Work

- Tighter coupling between the latent cause and actor-critic modules.



RL in the Brain

- The brain discovers latent causes by interacting with its environment. (1)
- Failures of under-segregation or over-segregation of old and new learning in the brain may underlie psychiatric conditions such as PTSD and anxiety. (3)

Acknowledgments

Special thanks to

- The members of the Niv Lab at the Princeton Neuroscience Institute.
- The directors of the PNI SIP, Dr. Ed Clayton and Paryn Wallace, Ed.D.
- This work was supported by NSF REU award #1757554

References

1. Gershman, Samuel J., et al. "Discovering latent causes in reinforcement learning." Behavioral Sciences, vol. 2015, no. 5, 2015, pp. 43-50. Gershman Lab, https://gershmanlab.com/pubs/GershmanNormanNiv15.pdf.
2. Niv, Yael. "Reinforcement learning in the brain." Journal of Mathematical Psychology, vol. 53, no. 2009, 2008, pp. 139-154. CCNSS, http://www.ccnss.org/ccn_2013/materials/pdf/botvinick/NivMathPsych09.pdf.
3. Pisupati et al., RLDM 2022, SOBP 2021

Connect with me!