Toward a More Neurally Plausible Neural Network Model of Latent Cause Inference

Qihong Lu\textsuperscript{1}, Tan Nguyen\textsuperscript{2}, Uri Hasson\textsuperscript{1}, Thomas Griffiths\textsuperscript{1}, Jeffrey Zacks\textsuperscript{2}, Samuel Gershman\textsuperscript{3}, Kenneth Norman\textsuperscript{1}

1. Princeton University; 2. Washington University in St. Louis; 3. Harvard University

Main points

We developed a neural network model that uses latent cause inference (LCI) to support context-dependent behavior. The model:

- extracts shared structure across LCs while avoiding catastrophic interference
- captures human data on curriculum effects on schema learning
- infers the underlying event structure when processing naturalistic videos of daily activities

Leveraging the shared structure across tasks

We compared LCNet, SEM (Figure 1), and a regular neural network on a functional learning task, where each function is the sum of a shared component and an idiosyncratic component (Figure 2A,B). We found that our model can:

- factor knowledge shared across tasks vs. task-specific knowledge (Figure 2C).
- overcome catastrophic interference (Figure 2D).
- encode knowledge shared across tasks to learn new tasks with less data (Figure 2E).

Capturing curriculum effects on schema learning

Episodic memory was implemented as a mapping from the input states to inferred latent causes (from the full LCI process; Figure 5). We found that LCNet can...

- recapitulate the human data (Figure 3) while saving 94% of full inferences, making the LCNet significantly more computationally feasible.

Inferring the event structure of naturalistic videos

We trained a recurrent LCNet on a pre-processed naturalistic video dataset – META (Bezdek et al., 2022a) to predict the upcoming frame of the video (Figure 6). We found that ...

- the way our model segments events captured the human event boundaries data (Figure 7 A,B,C) and the ground truth event structure (Figure 7D), even though the model was only trained to predict upcoming frame of the video; these results are similar to SEM2.0 (Bezdek et al., 2022b).

References & Links

1. Franklin et al., Psych Review 2020
2. Beukers et al., PsyArXiv 2023
3. Bezdek et al., Behav Res Methods 2022a
4. Bezdek et al., PsyArXiv 2022b

Acknowledgement: This work was supported by a Multi-University Research Initiative grant to KAN, JMZ, SJG and UH (ONR/DoD N00014-17-1-2961).

Contact: qlu@princeton.edu