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

Qihong Lu[1], Tan Nguyen[2], Uri Hasson[1], Thomas Griffiths[1], Jeffrey Zacks[2], Samuel Gershman[3], Kenneth Norman[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

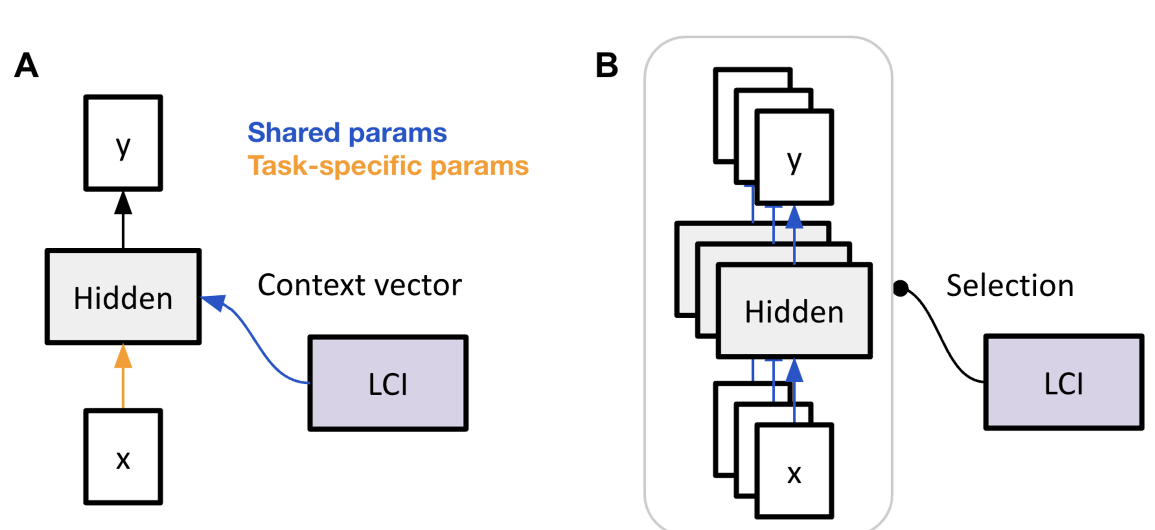

Figure 1:A) Latent Cause Network (**LCNet**) uses LCI to decide whether to use an existing context vector or a new context vector for the ongoing task. B) The Structured Event Memory model (**SEM**; Franklin et al., 2020) uses separate neural networks to represent different tasks and uses LCI to do network-selection.

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).

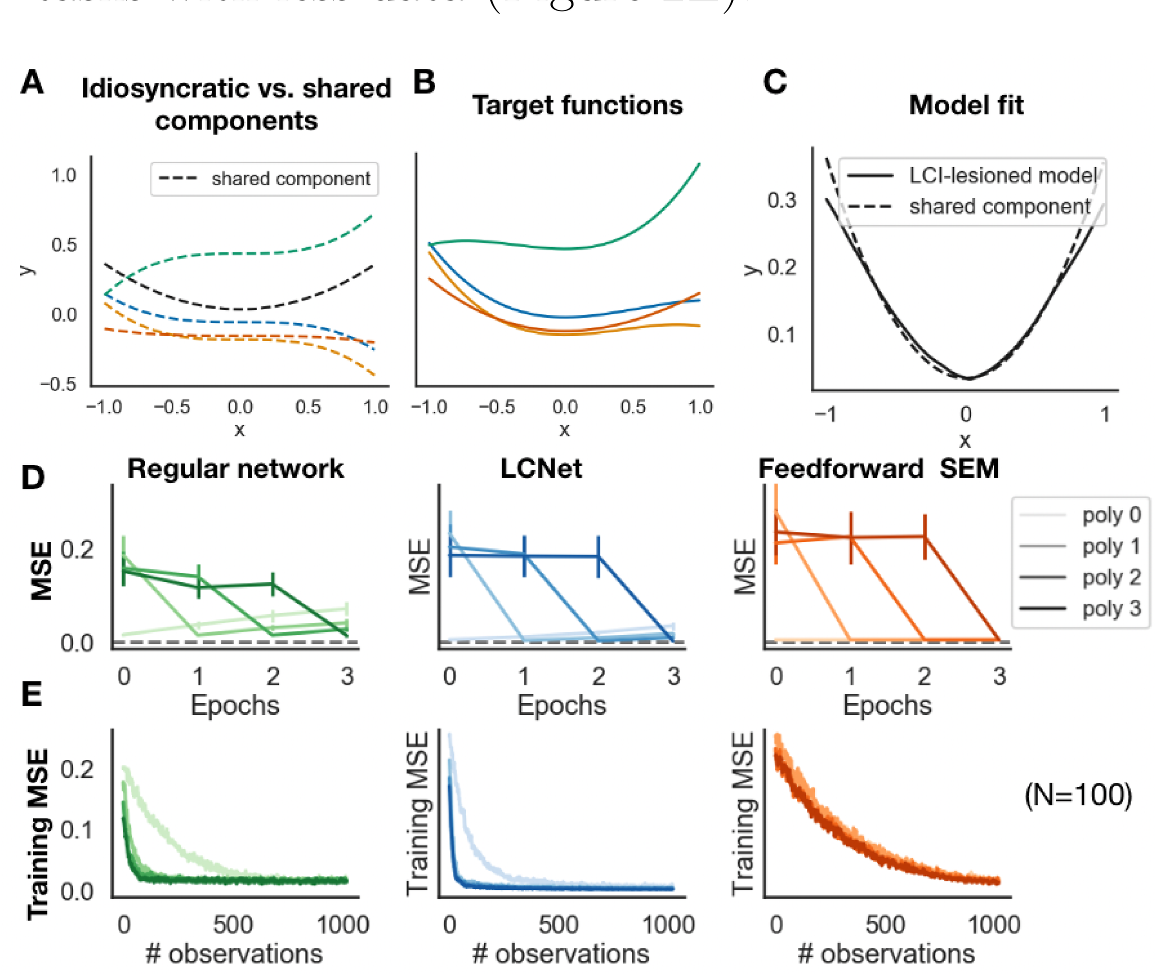

Figure 2:A, B) The target functions that the model had to learn. C) An LCI-lesioned LCNet reconstructs the shared component. D) Test MSE for all polynomials plotted separately over epochs in a blocked learning setting – the model was only trained on the i-th polynomial at epoch i. E) MSE for each polynomial plotted separately over the number of samples.

Capturing curriculum effects on schema learning

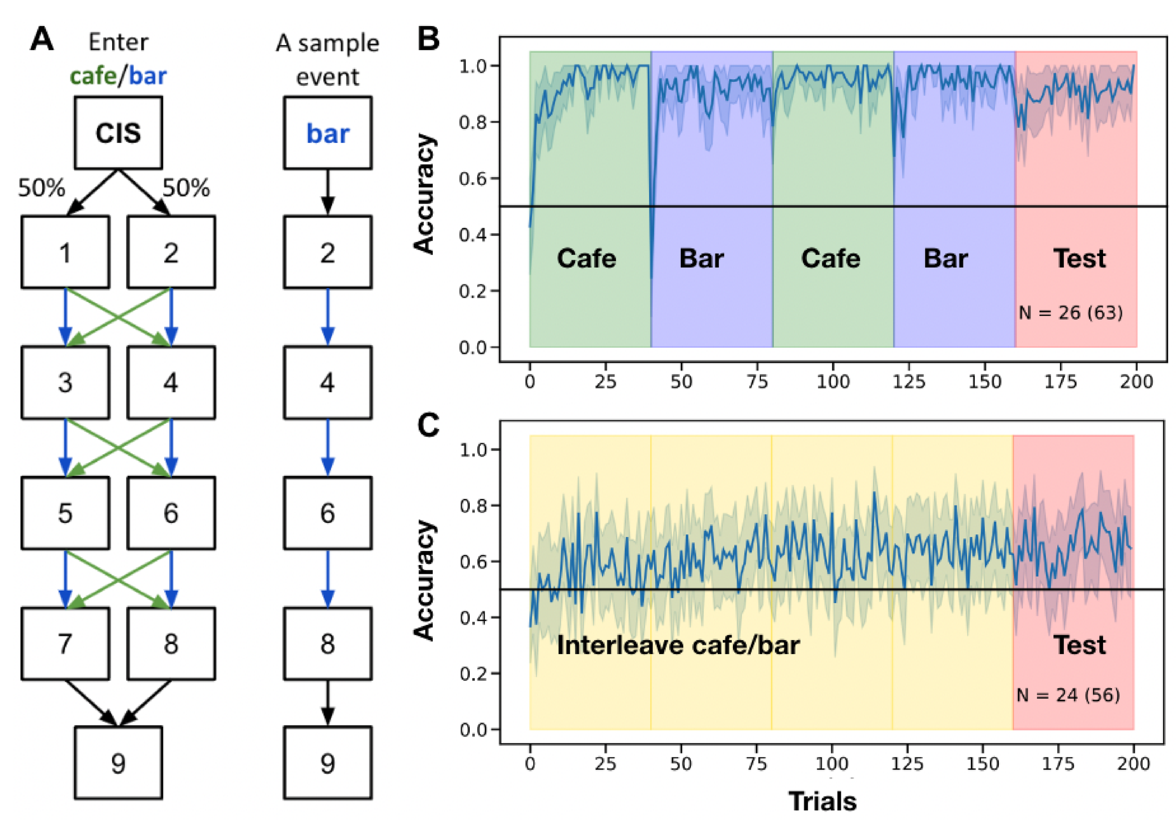

Figure 3:Human data; A) The state-transition graph used in a context-dependent sequence learning task by Beukers et al. (2023). B) Empirically, humans learned much better under the blocked curriculum than the interleaved curriculum (Beukers et al., 2023).

We found that LCNet can ...

- account for human data (shown in Figure 3): LCI was more accurate in the blocked curriculum than in the interleaved curriculum (Figure 4E).

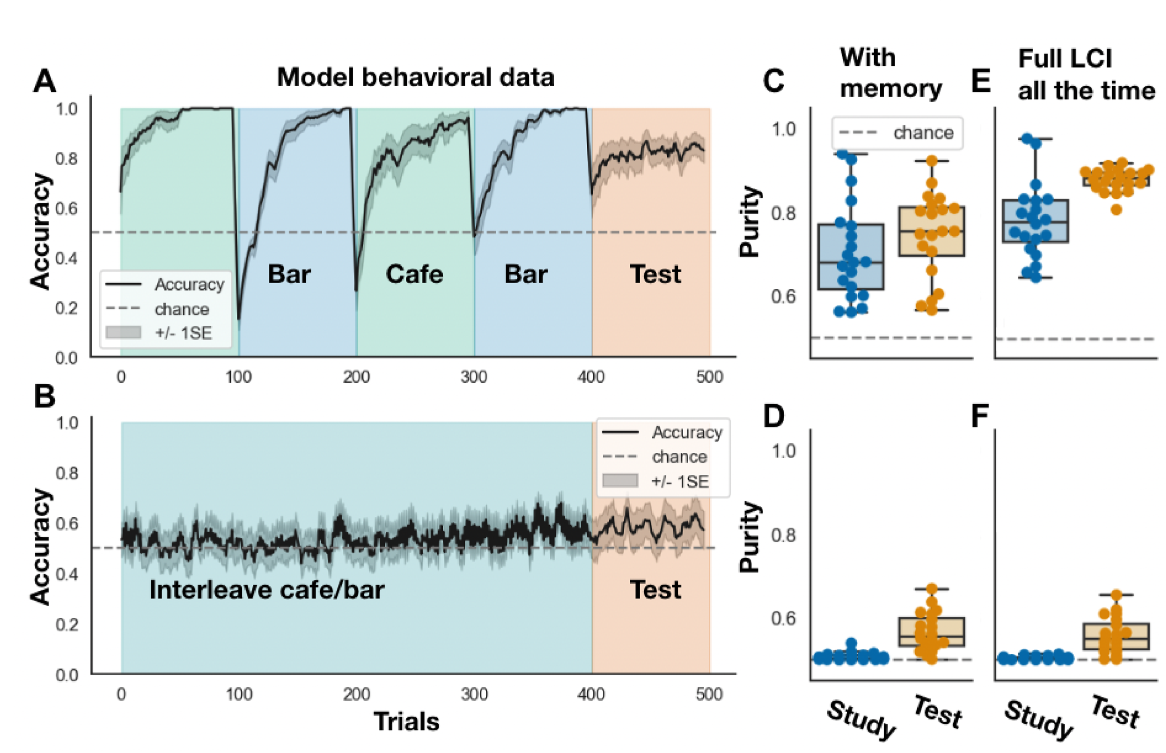

Figure 4:Model data; A, B) LCNet qualitatively replicated human behavior. C, D) LCI was more accurate in the blocked condition, and episodic memory reduced the need for full inference by 94%. E, F) LCI performance with full inference.

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.

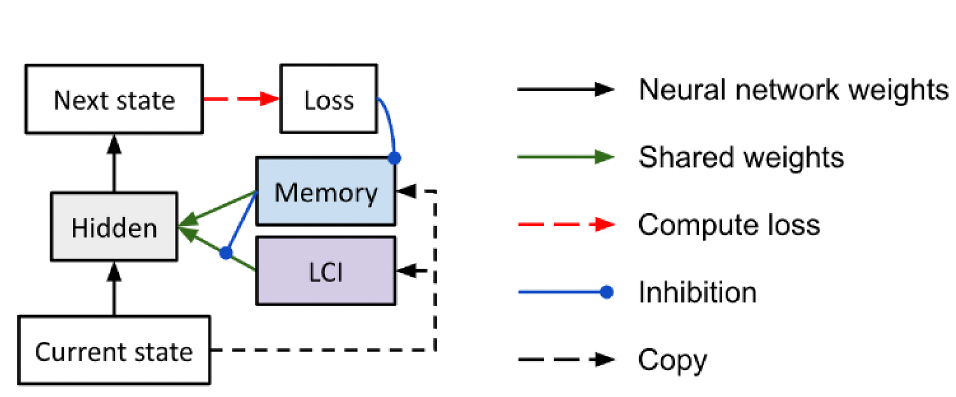

Figure 5:Use of episodic memory as a shortcut for full LCI – LCNet can recall previously used latent causes based on the current input state, instead of performing the laborious full LCI process. The memory shortcut is activated if it matches with the full inference result for a period of time; the full inference procedure is turned back on when the current loss is too high.

Inferring the event structure of naturalistic videos

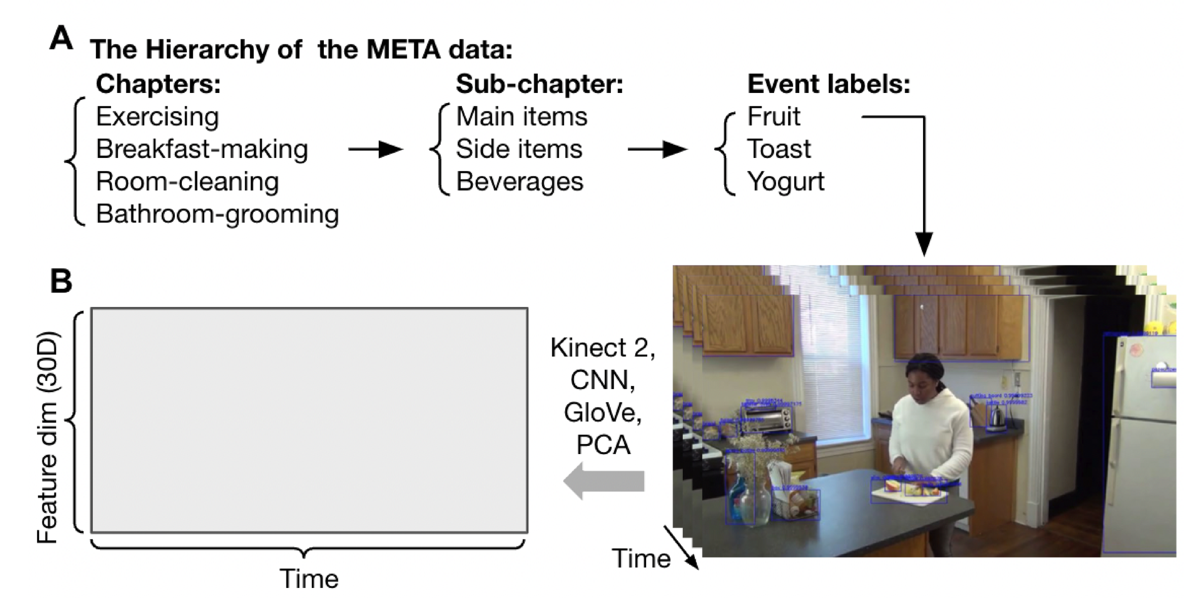

Figure 6:A) The hierarchical structure and B) preprocessing steps for the META dataset (Bezdek et al., 2022a), which consists of more than 100 hours of daily activities.

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).

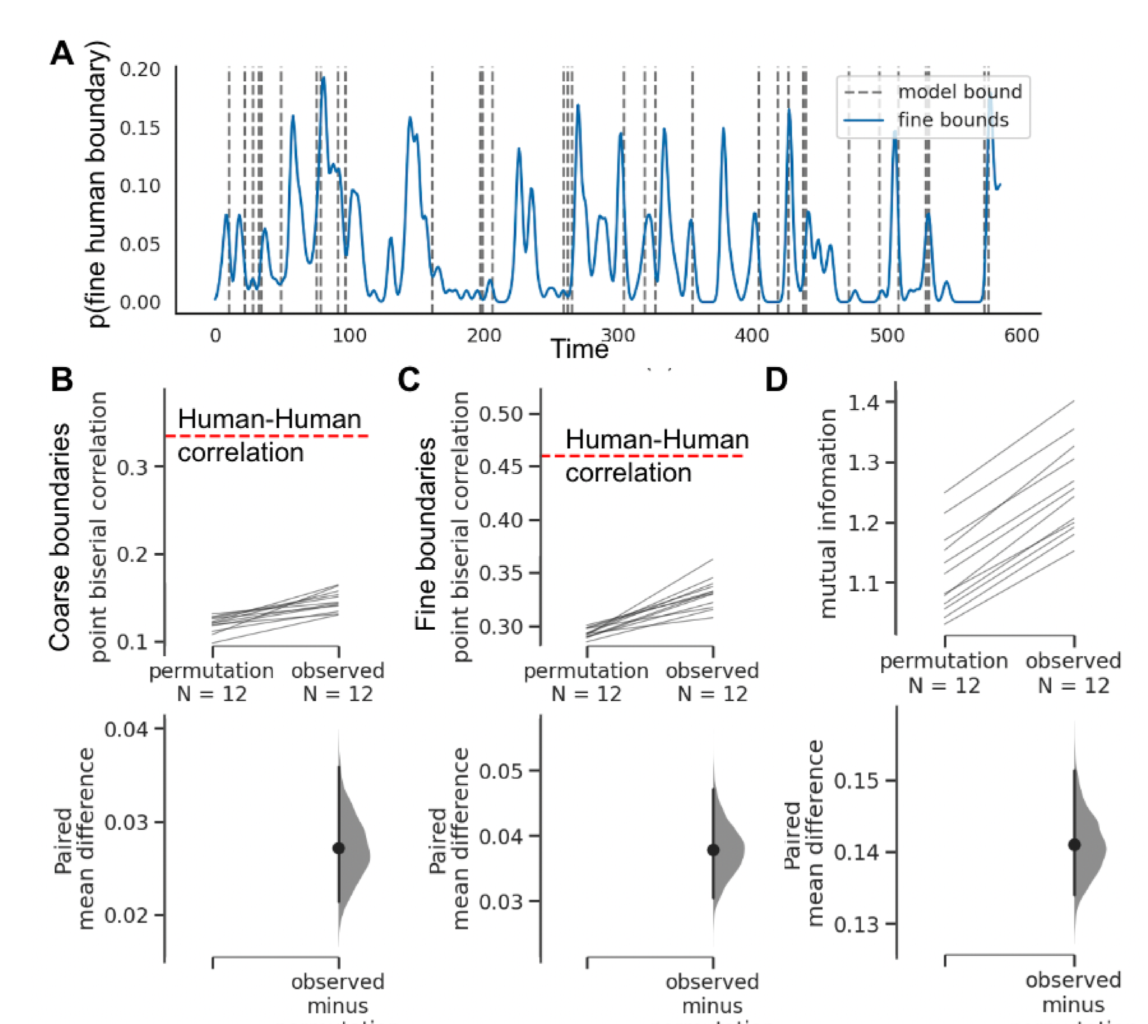

Figure 7:A) Model boundaries vs human boundaries for an example video. B, C) The event boundaries extracted from the models are significantly correlated with human coarse/fine boundaries. D) The LCs inferred by the model and the ground truth event labels shared significant mutual information.

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`