

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/373733297>

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

Conference Paper · January 2023

DOI: 10.32470/CCN.2023.1339-0

CITATIONS

0

READS

446

7 authors, including:



Qihong Lu

Columbia University

18 PUBLICATIONS 135 CITATIONS

SEE PROFILE



Tan Nguyen

Washington University in St. Louis

5 PUBLICATIONS 3 CITATIONS

SEE PROFILE



Jeffrey Zacks

Washington University in St. Louis

167 PUBLICATIONS 13,551 CITATIONS

SEE PROFILE

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

Qihong Lu¹ (qlu@princeton.edu)

Tan T. Nguyen³ (n.tan@wustl.edu)

Uri Hasson¹ (hasson@princeton.edu)

Thomas L. Griffiths^{1, 2} (tomg@princeton.edu)

Jeffrey M. Zacks³ (jzacks@wustl.edu)

Samuel J. Gershman⁴ (gershman@fas.harvard.edu)

Kenneth A. Norman¹ (knorman@princeton.edu)

Department of Psychology and Princeton Neuroscience Institute, Princeton University¹

Department of Computer Science, Princeton University²

Department of Psychological and Brain Sciences, Washington University in St. Louis³

Department of Psychology and Center for Brain Science, Harvard University⁴



Abstract

Humans spontaneously perceive a continuous stream of experience as discrete events. It has been hypothesized that this ability is supported by latent cause inference (LCI). We implemented this hypothesis using latent cause network (LCNet), a neural network model of LCI. LCNet interacts with a Bayesian LCI mechanism that activates a unique context vector for each inferred latent cause (LC). LCNet can also recall episodic memories of previously inferred LCs to avoid performing LCI all the time. These mechanisms make LCNet more neurally plausible and efficient than existing models. Across three simulations, we found that LCNet could 1) extract shared structure across LCs while avoiding catastrophic interference, 2) capture human data on curriculum effects on schema learning, and 3) infer the underlying event structure when processing naturalistic videos of daily activities. Our work provides a neurally plausible computational model that can operate in both laboratory experiment settings and naturalistic settings, opening up the possibility of providing a unified model of event cognition.

Keywords: event cognition; latent cause inference; episodic memory; neural network models; continual learning

Humans spontaneously perceive a continuous stream of experience as discrete events (Kurby & Zacks, 2008; Zacks & Tversky, 2001; Radvansky & Zacks, 2017). This seems effortless for humans, but it is challenging for standard neural network models. When viewing each event as a task, event perception can be viewed as a continual learning problem (Parisi, Kemker, Part, Kanan, & Wermter, 2019; Flesch, Saxe, & Summerfield, 2023) with a blocked curriculum, where both the ongoing task identity and the number of tasks are unknown.

The Structured Event Memory (SEM) model

In the human brain, it has been hypothesized that the process of grouping observations into events is supported by latent cause inference (LCI) (Gershman, Norman, & Niv, 2015; Franklin, Norman, Ranganath, Zacks, & Gershman, 2020; Niv, 2019; Shin & DuBrow, 2020) – recently, the Structured Event Memory (SEM) model has been proposed to explain how humans assign observations to latent causes (LCs) to facilitate learning, prediction, memory, and generalization in a context-dependent manner (Franklin et al., 2020). SEM uses a Bayesian non-parametric mechanism to perform LCI and uses separate neural networks to represent different LCs. Given the current observation x_t , SEM runs LCI to either 1) assign x_t to an existing LC and use the corresponding network to process x_t , or 2) create a new LC for x_t by generating a new network to process x_t . Such fully-separated representations can convert a continual learning problem into multiple single-task problems, circumventing the catastrophic interference that would otherwise occur with blocked learning (French, 1999; McClelland, McNaughton, & O’Reilly, 1995; McCloskey & Cohen, 1989). SEM can explain a wide range of human data, such as how event structure affects memory

reconstruction, and a variant of SEM (“SEM 2.0”; Bezdek, Nguyen, Gershman, et al., 2022) has recently been used to simulate data on how humans segment naturalistic videos of daily events (Bezdek, Nguyen, Hall, et al., 2022). However, using fully-separated representations is neurally implausible, and it impedes SEM’s ability to extract the shared structure across LCs, since a given observation is only used to train one network. Additionally, SEM performs LCI all the time, which is also implausible due to the heavy computational demands.

A more neurally plausible model of LCI

Inference To address these limitations, we propose the Latent Cause Network (LCNet), a more biologically plausible neural network model of LCI. LCNet uses the same network to process all observations, and it uses different context vectors to achieve task/context-sensitivity (Figure 1A, G, Q), akin to the classic connectionist idea of context representation (Cohen, Dunbar, & McClelland, 1990; Rougier, Noelle, Braver, Cohen, & O’Reilly, 2005). Given the current observation x_t , it still performs LCI using the same Bayesian non-parametric mechanism. If LCI assigns x_t to an existing LC, the corresponding context vector is fed to the hidden layer of the network, or if LCI creates a new LC for x_t , a new context vector is generated. Importantly, context vectors are simply random vectors sampled from a Gaussian distribution. Since high dimensional random vectors are approximately orthogonal, they can reduce catastrophic interference across LCs.

Episodic memory LCNet has an episodic memory (EM) mechanism that retrieves previously inferred LCs instead of performing LCI on each time step. Concretely, EM is implemented as a lookup table with a narrow generalization gradient (McClelland, 2013; McClelland et al., 1995; O’Reilly & Norman, 2002; Norman & O’Reilly, 2003; Norman, Detre, & Polyn, 2008). Every time LCNet performs a full LCI, EM encodes the current observation and the inferred LC in its buffer. Later, given an observation x_t , if LCNet can retrieve an LC previously associated with x_t , it uses the retrieved LC instead of performing full inference (Figure 1G, Q). We found that having EM makes LCNet much more computationally efficient.

Function learning with shared structure

In Simulation 1, we compare the performance of a feedforward LCNet (Figure 1A), a feedforward version of SEM, and a regular feedforward network on learning four polynomial regression tasks. Given the input and a context-indicative signal (CIS) specifying which function is currently active, the model had to produce the corresponding output. Learning was fully blocked, and all models learned one function for every epoch. Unbeknownst to the model, each function (Figure 1C) is a sum of a component that is shared across all functions and an idiosyncratic component (Figure 1B). After learning, an LCI-lesioned LCNet reconstructs the shared component (Figure 1D), suggesting that the shared structure is encoded in the network weights (blue pathway in Figure 1A). We observed strong catastrophic interference for the regular network, al-

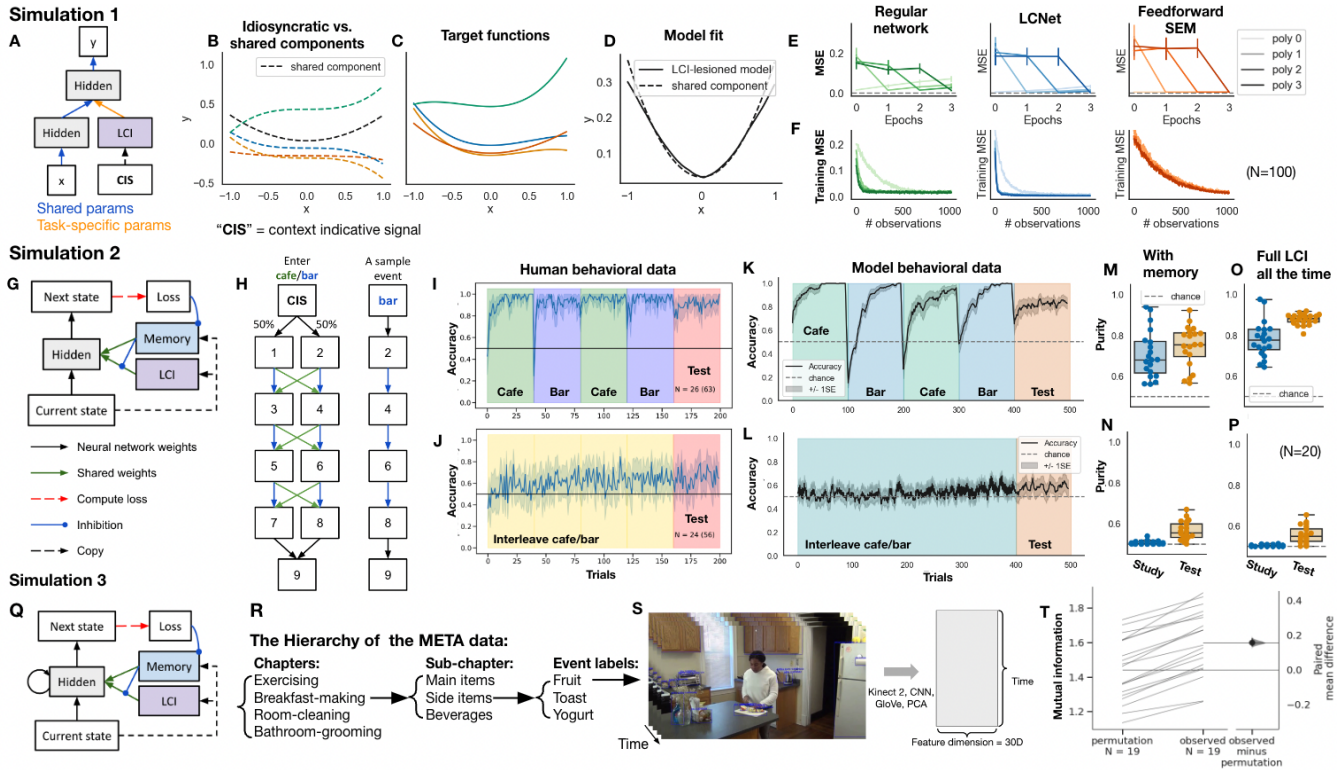


Figure 1: A) Model architecture for Simulation 1. B) The target functions that the model had to learn. Each function is a sum of a shared term and an idiosyncratic term shown in panel C. D) An LCI-lesioned LCNNet reconstructs the shared component, even though the shared component was never directly observed. E) Learning curves for each polynomial plotted separately over epochs – the model was only trained on the i -th polynomial at epoch i . F) Learning curves for each polynomial plotted separately over the number of samples. G) Model architecture for Simulation 2. H) The state-transition graph used in Beukers et al. (2023). I and J) human behavioral data in the blocked (I) vs. interleaved (J) conditions. K and L) model behavioral data in the blocked (K) vs. interleaved (L) conditions. M and N) the purity of the inferred LCs of the model relative to the ground truth in the blocked (M) vs. interleaved (N) conditions. O and P) the same as panel M and panel N, except that the memory module of the model was lesioned. Q) Model architecture for Simulation 3. R) The hierarchical structure of the META dataset (Bezdek, Nguyen, Hall, et al., 2022). S) Each frame in the META video recordings is compressed as a 30D feature vector. T) The LCs inferred by the model and the ground truth event labels shared a significant level of mutual information.

most no interference for LCNNet, and no interference for SEM (Figure 1E). Moreover, learning one function sped up new learning for both the regular network and LCNNet but not SEM (Figure 1F). These results show that LCNNet can extract shared structure and reuse it when learning new, related tasks while avoiding catastrophic interference.

Curriculum effect on schema learning

In Simulation 2, LCNNet (Figure 1G) was trained on a context-dependent sequential prediction task (Figure 1H). Given the current state, the model had to predict the upcoming state. Unbeknownst to the model/human, the first observation is context-indicative and determines the transition structure of the event graph (Figure 1H). Empirical results show that humans learned much better under the blocked curriculum (Figure 1I) than the interleaved curriculum (Figure 1J; Beukers et al., 2023). LCNNet qualitatively replicated this pattern (Figure 1K vs. 1L) as LCI was much more accurate in the blocked condition (Figure 1M vs. 1N), quantified by cluster purity. Importantly, although LCI was more accurate when full inference

was performed all the time (Figure 1O vs. Figure 1M), LCNNet with EM saved 94% of LCI while still being able to capture human data.

Scene prediction on naturalistic video stimuli

In Simulation 3, we test LCNNet (Figure 1Q) on naturalistic video recordings of daily activities (Figure 1R, 1S) (Bezdek, Nguyen, Hall, et al., 2022). Since this dataset was generated in a controlled manner, the ground truth event labels are known (Figure 1R). We found that LCs inferred by LCNNet share a significant amount of mutual information with the event label (Figure 1T) even though 35% of LCI was saved by using EM. Being able to process naturalistic data demonstrates the generality of our framework. In the future, we will test LCNNet's ability to predict human event segmentation (comparing its performance to SEM 2.0; Bezdek, Nguyen, Gershman, et al., 2022), and we will use uncertainty to guide when to perform LCI (as in the CCN 2023 submission by Nguyen et al.). The model also generates time series predictions of when episodic retrieval happens, which we will test via human fMRI.

Acknowledgments

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

References

- Beukers, A. O., Collin, S. H. P., Kempner, R., Franklin, N. T., Gershman, S. J., & Norman, K. A. (2023). Blocked training facilitates learning of multiple schemas. *PsyArXiv*.
- Bezdek, M., Nguyen, T., Gershman, S. J., Bobick, A., Braver, T. S., & Zacks, J. M. (2022, September). Error-based updating of event representations enables prediction of human activity at human scale. *PsyArXiv*.
- Bezdek, M., Nguyen, T. T., Hall, C. S., Braver, T. S., Bobick, A. F., & Zacks, J. M. (2022, October). The multi-angle extended three-dimensional activities (META) stimulus set: A tool for studying event cognition. *Behavior Research Methods*.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990, July). On the control of automatic processes: a parallel distributed processing account of the stroop effect. *Psychological Review*, 97(3), 332–361.
- Flesch, T., Saxe, A., & Summerfield, C. (2023, March). Continual task learning in natural and artificial agents. *Trends in Neurosciences*, 46(3), 199–210.
- Franklin, N. T., Norman, K. A., Ranganath, C., Zacks, J. M., & Gershman, S. J. (2020, April). Structured event memory: A neuro-symbolic model of event cognition. *Psychological Review*, 127(3), 327–361.
- French, R. M. (1999, April). Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*, 3(4), 128–135.
- Gershman, S. J., Norman, K. A., & Niv, Y. (2015, October). Discovering latent causes in reinforcement learning. *Current Opinion in Behavioral Sciences*, 5, 43–50.
- Kurby, C. A., & Zacks, J. M. (2008, February). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2), 72–79.
- McClelland, J. L. (2013, November). Incorporating rapid neocortical learning of new schema-consistent information into complementary learning systems theory. *Journal of Experimental Psychology: General*, 142(4), 1190–1210.
- McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995, July). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3), 419–457.
- McCloskey, M., & Cohen, N. J. (1989, January). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 24, pp. 109–165). Academic Press.
- Niv, Y. (2019, October). Learning task-state representations. *Nature Neuroscience*, 22(10), 1544–1553.
- Norman, K. A., Detre, G., & Polyn, S. M. (2008). Computational models of episodic memory. In R. Sun (Ed.), *The Cambridge handbook of computational psychology* (Vol. 753, pp. 189–225). New York, NY, US: Cambridge University Press, xii.
- Norman, K. A., & O'Reilly, R. C. (2003, October). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review*, 110(4), 611–646.
- O'Reilly, R. C., & Norman, K. A. (2002, December). Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *Trends in Cognitive Sciences*, 6(12), 505–510.
- Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., & Wermter, S. (2019, May). Continual lifelong learning with neural networks: A review. *Neural Networks: the official journal of the International Neural Network Society*, 113, 54–71.
- Radvansky, G. A., & Zacks, J. M. (2017, October). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, 17, 133–140.
- Rougier, N. P., Noelle, D. C., Braver, T. S., Cohen, J. D., & O'Reilly, R. C. (2005, May). Prefrontal cortex and flexible cognitive control: rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, 102(20), 7338–7343.
- Shin, Y. S., & DuBrow, S. (2020, May). Structuring memory through Inference-Based event segmentation. *Topics in Cognitive Science*.
- Zacks, J. M., & Tversky, B. (2001, January). Event structure in perception and conception. *Psychological Bulletin*, 127(1), 3–21.