The challenges of lifelong learning in biological and artificial systems

Sashank Pisupati

Yael Niv

Princeton Neuroscience Institute, Princeton University USA

Author Note
Sashank Pisupati (ORCID: https://orcid.org/0000-0003-0923-0585)
Yael Niv (ORCID: https://orcid.org/0000-0002-0259-8371)
Correspondence: pisupati@princeton.edu (Sashank Pisupati)

Abstract
How do biological systems learn continuously throughout their lifespans, adapting to change while retaining old knowledge, and how can these principles be applied to artificial learning systems? Here, we outline challenges and strategies of "lifelong learning" in biological and artificial systems, and argue that a collaborative study of each system's failure modes can benefit both.

Keywords: lifelong learning, continual learning, inductive biases, forgetting, reinforcement learning, Bayesian inference

The problem of lifelong learning

Humans and other biological learning systems display the astounding ability to continually accumulate knowledge over their entire lifetime. One remarkable feature of this form of learning is that even though experiences over a lifetime are sequentially sampled from changing (non-stationary) environments, organisms can adapt quickly to changes while retaining old knowledge. This contrasts with many contemporary machine-learning algorithms, which rely on independent, identically distributed samples from a stationary distribution to learn successfully, often failing to learn anew or forgetting old knowledge catastrophically when presented with incremental data from changing distributions [1]. What computational principles support effective learning and long-term knowledge retention in the face of change? The modern machine-learning subfields of *continual* and *lifelong learning* have taken inspiration from biological systems at multiple levels of abstraction to tackle this problem [1,2], and aim to rival human-level performance. The goal of such approaches is to learn continuously in non-stationary data regimes, avoid catastrophic forgetting and running out of capacity, and — more ambitiously — enable knowledge transfer and generalization between past and future tasks to improve learning efficiency.

Learning strategies observed in biological agents

Given the ever-changing natural world, it is perhaps no surprise that humans and other biological systems have evolved multiple mechanisms for lifelong learning. These include synaptic plasticity rules that protect previously learned associations, mechanisms that create new neural structures or representations when drastic changes are encountered, built-in neuromodulatory drives for persistent exploration, and architectural schemes that use multiple interacting learning and memory systems to balance generalization and segregation [2].

At a high level, these mechanisms can all be understood as *inductive biases*, or assumptions that learners bring to a problem, which shape learning and restrict the space of solutions. In particular, these mechanisms all embody the assumption that environments may not be stationary (hence, for example, the need to continuously explore new solutions), but may nevertheless have modular and/or recurring structure (i.e. changes may indicate a new task needs to be learned, but old knowledge should be retained as it may be useful for old tasks that resurface, or for generalizing to new problems that are similar).

One useful framework that has helped translate several of these mechanisms into a common language is the *contextual* or *latent-cause* framework. Here, the learner is assumed to constantly segment its experience into “contexts," “tasks" or “latent causes," discovering new contexts as well as inferring the recurring presence of old ones, and learning separate associations for each [3]. This reformulates the learning problem from one that involves using experiences to learn or track a fixed set of static or dynamic associative parameters (the typical assumption underlying Bayesian, gradient-based, or reinforcement-learning formulations; Figure 1a,b), to one that additionally involves partitioning experiences into (an unknown number of) contexts, each with their own set of learned parameters (Figure 1c).

Given an inferred partition, experiences can be used to selectively modify context-specific parameters, while these parameters are protected from erasure outside the relevant context. This means that new knowledge can be acquired quickly, and at the same time, old knowledge can be remembered and reused long into the future, if and when similar contexts are encountered (unlike, for example, standard reinforcement learning, where fast learning implies fast forgetting). Bayesian non-parametric models are a popular choice for the partitioning process since they allow flexible but judicious expansion of learnt latent structure. Such models avoid running out of capacity by using priors (such as the Chinese Restaurant Process prior) that constrain the number of latent causes inferred – a principle of parsimony that reflects humans' and animals' inductive biases across many different domains [3] and naturally emerges from efficient coding objectives [4].

The latent-cause framework has been used to successfully account for previously puzzling features of learning in humans and animals, such as their ability to learn and remember multiple conflicting beliefs or behavioral policies, and enhanced learning speeds and accuracy when encountering successive new tasks or previously learned ones [3,5]. Neurally, this framework has been used to understand synaptic processes of memory modification and protection [6], circuit-level gating of learning updates by latent-cause representations in the prefrontal cortex and/or hippocampus [3], and global, long-term interactions between learning and episodic memory in replay, consolidation, and retrieval [7].

Biologically inspired strategies in artificial agents

To engineer artificial agents capable of continual learning, artificial intelligence research has taken inspiration from biological inductive biases at several levels. Particularly within deep learning, these efforts roughly fall under four approaches: gradient-based approaches inspired by synaptic plasticity rules, modular architectures that add capacity when new tasks are encountered, memory-based approaches that store and/or replay past experiences, and meta-learning approaches that attempt to learn useful inductive biases by optimizing an evolution-like "outer-loop" [1,2]. Many of these approaches leverage sparse, modular latent causal structure, with some solutions explicitly formulating this structure in probabilistic terms [8]. In particular, approaches that eschew expensive storage of past experiences in favor of a more compressed solution in the form of discrete, low dimensional "anchors" representing abstractions of past tasks [1] closely resemble latent-cause models of context-bound episodic memory and cortico-hippocampal interactions.

Failure modes of lifelong-learning agents

As with all inductive biases, those that allow for lifelong learning in changing natural environments induce particular failure modes when their assumptions are not met. Humans' and animals' naturally adaptive inductive biases about non-stationarity are perhaps no more evident than in unnaturally stationary laboratory tasks, where such implicit assumptions may hamper task performance, manifesting as sequential dependencies or persistent exploration [9-11].

Even in tasks that truly do involve a change in environmental statistics, the assumption of recurrence of old contexts may not be accurate, and the same lifelong-learning mechanisms that successfully protect old knowledge for future reuse may prove to be maladaptive. This is evident in fear-extinction paradigms, where expectations of outdated threats may resurface long after the threat has been removed, revealing intact associations bound to old latent causes [3].

Indeed, a number of psychiatric conditions are thought to involve mismatches between assumptions about distribution shifts and reality. For instance, certain forms of anxiety may reflect over-segmentation of threat experiences and/or over-enthusiastic protection of long outdated threat associations, making them stubbornly resistant to updating [12]. A similar over-active protection of early drug-related associations has been proposed to underlie addiction phenomenology such as relapse. Conversely, mechanisms that enable forward or backward transfer by generalizing knowledge between past and future contexts may inappropriately generalize negative experiences to neutral situations, giving rise to the widespread biased evaluations observed in post-traumatic stress disorder (PTSD) and depression [7, 13].

Such dramatic failure modes of powerful biological learning systems could offer valuable lessons for the safety analysis of new artificial systems capable of continual learning [14], and help anticipate situations that may lead to maladaptive behaviors before they occur, diagnose those that do occur, and perhaps even build in compensatory safety mechanisms to protect against them. In parallel, lifelong-learning algorithms could serve as powerful models of biological systems for the purposes of computational psychiatry, particularly for chronic conditions that recur over an individual's lifetime and reflect over-protection of old knowledge or over-generalization to new situations.

Looking ahead: a joint investigation of challenges?

Biological and artificial agents run up against similar challenges when attempting to remain adaptive throughout their lifetimes in non-stationary environments with recurring structure. These include the need to accurately recognize changes in one's environment, quickly update one's knowledge when this happens, judiciously protect old knowledge while doing so without running out of capacity, and appropriately reuse past learning to make smart generalizations about the future. Biological systems seem to be equipped with a number of inductive biases that help them solve this problem, and artificial systems have taken inspiration from these at various levels of granularity to arrive at algorithmic solutions. Beyond cross-inspiration from each system's successes, studying the inevitable failure modes of these inductive biases in both systems offers a fruitful avenue for potential future collaboration between the two fields, with mutual benefits: to the study of biological systems, it could offer a computational understanding of recurrent maladaptations commonly encountered in psychiatry, and inspire treatments that tame these mechanisms; and to the study of artificial systems, it could offer the ability to anticipate, diagnose and protect against maladaptive behaviors in lifelong learning agents.

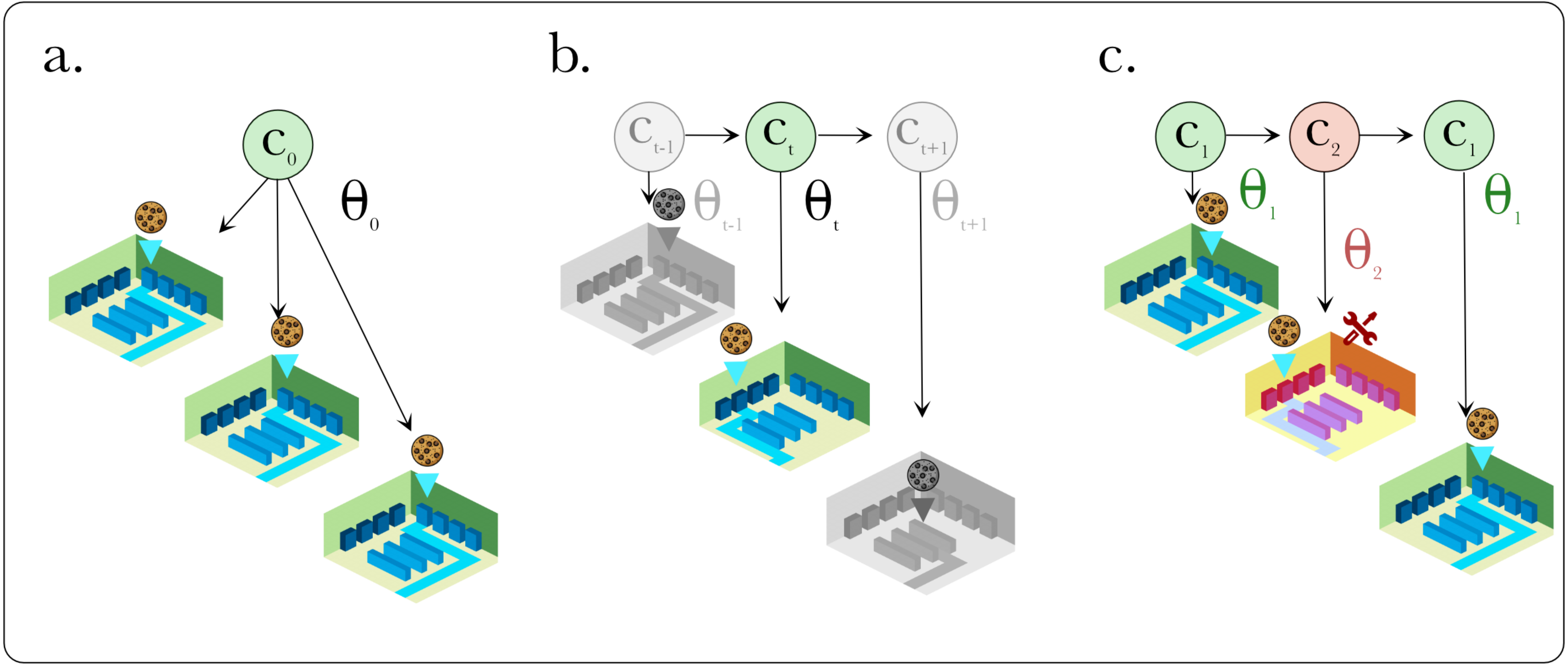
Acknowledgements

Special thanks to Nora Xiao for inspiring the figure, conceptualizing the continual learning example and helpful discussions. We would like to thank Angela Langdon, Gili Karni, Dan Mirea, Nadav Amir, Oded Bein, Rachel Bedder, Isabel Berwian and Stephanie Groman for helpful discussions and feedback on the work. This work was funded by NIDA grant R01DA042065, NIMH grant R01MH119511 and a grant from the Templeton Foundation.

References

1. Hadsell, R., Rao, D., Rusu, A. A., & Pascanu, R. (2020). Embracing change: Continual learning in deep neural networks. *Trends in cognitive sciences*, *24* (12), 1028–1040.
2. Kudithipudi, D., Aguilar-Simon, M., Babb, J., Bazhenov, M., Blackiston, D., Bongard, J., Brna, A. P., Chakravarthi Raja, S., Cheney, N., Clune, J., et al. (2022). Biological underpinnings for lifelong learning machines. *Nature Machine Intelligence*, *4* (3), 196–210.
3. Gershman, S. J., Norman, K. A., & Niv, Y. (2015). Discovering latent causes I reinforcement learning. *Current Opinion in Behavioral Sciences*, *5*, 43–50.
4. Dasgupta, I., & Griffiths, T. L. (2022). Clustering and the efficient use of cognitive resources. *Journal of Mathematical Psychology*, *109*, 102675.
5. Heald, J. B., Lengyel, M., & Wolpert, D. M. (2021). Contextual inference underlies the learning of sensorimotor repertoires. *Nature*, *600* (7889), 489–493.
6. Gershman, S. J., Monfils, M.-H., Norman, K. A., & Niv, Y. (2017). The computational nature of memory modification. *Elife*, *6*, e23763.
7. Cohen, R. T., & Kahana, M. J. (2022). A memory-based theory of emotional disorders. *Psychological Review*, 129(4), 742-776.
8. Xie, A., Harrison, J., & Finn, C. (2020). Deep reinforcement learning amidst lifelong non-stationarity. *arXiv preprint arXiv:2006.10701*.
9. Dubey, R., Griffiths, T. L., & Dayan, P. (2022). The pursuit of happiness: A reinforcement learning perspective on habituation and comparisons. *PLoS Computational biology,* 18(8), e1010316
10. Ebitz, R. B., Sleezer, B. J., Jedema, H. P., Bradberry, C. W., & Hayden, B. Y. (2019). Tonic exploration governs both flexibility and lapses. *PLoS computational biology*, *15* (11), e1007475.
11. Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., & Churchland, A. K. (2021). Lapses in perceptual decisions reflect exploration. *Elife*, *10*, e55490.
12. Zika, O., Wiech, K., Reinecke, A., Browning, M., & Schuck, N. W. (2022). Trait anxiety is associated with hidden state inference during aversive reversal learning. *bioRxiv*. DOI: 10.1101/2022.04.01.483303
13. Norbury, A., Brinkman, H., Kowalchyk, M., Monti, E., Pietrzak, R. H., Schiller, D., & Feder, A. (2021). Latent cause inference during extinction learning in trauma-exposed individuals with and without ptsd. *Psychological Medicine*, 1–12.
14. Schulz, E., & Dayan, P. (2020). Computational psychiatry for computers. *Iscience*, *23* (12), 101772.

Figure 1. Three commonly encountered online learning settings. a. Static setting, the standard for most optimization problems: A stationary environment (variously referred to as a task, context, or latent cause) with a fixed set of learnable parameters produces an independently and identically distributed set of samples, even when encountered sequentially. Illustrated is a grocery store environment c_0 with a cookie reward in one location, where the learnable parameters θ_0 determine the optimal behavioral policy required to arrive at the cookie (shortest path from entrance, denoted in the image below), and the learner's task is to acquire an estimate of these parameters that converges to the true values. **b.** Dynamic setting with coupled learning and forgetting: A non-stationary environment c_t with a single set of parameters θ_t that undergoes (possibly unsignalled) changes. In this scenario, the cookie's location changes every so often (illustrated: different batches of episodes), with only its current location relevant for optimal performance. The learner must therefore track only the current value of the parameters, and may forget previous values since they are no longer relevant. **c.** Continual learning setting, requiring learning without forgetting: A non-stationary environment with recurrent structure. Here, past contexts may reappear in the future, or be relevant to future generalization. This requires the learner to partition their experiences into the appropriate contexts c_i, entertain the possibility of new contexts appearing, and maintain context-specific parameters θ_i in memory, implying a clustering problem with an unknown number of clusters. Failure to segregate learning in this way will result in forgetting of old knowledge, necessitating relearning if the old context resurfaces. Appropriately clustering the data can improve learning efficiency, and allows knowledge transfer between past and future tasks. Depicted is a case where the cookie has been moved due to construction in the cookie aisle (a second latent cause c_2). By protecting learning acquired in c_1, the learner can instantly reuse it when the construction is over.



a.
b.
c.
c_0
θ_0
c_{t-1}
c_t
c_{t+1}
θ_{t-1}
θ_t
θ_{t+1}
c_1
c_2
c_1
θ_1
θ_2
θ_1