A Neural Network Model of Complementary Learning Systems

Mika Sarkin Jain* **(mjain4@stanford.edu)**
Department of Physics, Department of Computer Science, Stanford University
Stanford, CA 94305 USA

Jack Lindsey* **(jacklindsey@stanford.edu)**
Department of Mathematics, Department of Computer Science, Stanford University
Stanford, CA 94305 USA

Abstract

We introduce a computational model capturing the high-level features of the complementary learning systems (CLS) framework. In particular, we model the integration of episodic memory with statistical learning in an end-to-end trainable neural network architecture. We model episodic memory with a nonparametric module which can retrieve past observations in response to a given observation, and statistical learning with a parametric module which performs inference on the given observation. We demonstrate on vision and control tasks that our model is able to leverage the respective advantages of nonparametric and parametric learning strategies, and that its behavior aligns with a variety of behavioral and neural data. In particular, our model performs consistently with results indicating that episodic memory systems in the hippocampus aid early learning and transfer generalization. We also find qualitative results consistent with findings that neural traces of memories of similar events converge over time. Furthermore, without explicit instruction or incentive, the behavior of our model naturally aligns with results suggesting that the usage of episodic systems wanes over the course of learning. These results suggest that key features of the CLS framework emerge in a task-optimized model containing statistical and episodic learning components, supporting several hypotheses of the framework.

Introduction & Motivation

Complementary Learning Systems Framework

The complementary learning systems (CLS) framework hypothesizes that learning in the brain requires the integration of an episodic memory system and a statistical learning system. The CLS framework suggests that statistical learning primarily takes place in the neocortex and is necessary for powerful inference, while an episodic memory system is present in the hippocampus and is necessary for incorporating new observations quickly and robustly.

The CLS framework accounts for numerous observations about hippocampal and neocortical function, which it unifies into a single theoretical structure. The CLS framework also offers a resolution to a weakness in connectionist models of learning: that such models have difficulty incorporating observations from new domains quickly without interfering with previously acquired knowledge, a phenomenon known as catastrophic interference (McCloskey & Cohen, 1989). The CLS framework suggests that episodic memory allows for fast storage of new observations without disrupting existing knowledge (Burgess, Maguire, & O'Keefe, 2002), and that over time, structure is discovered in these observations and subsequently incorporated into a powerful statistical learning system (McClelland, McNaughton, & O'Reilly, 1995).

The CLS framework has undergone some revision in recent years. In particular, the REMERGE model suggests that recurrency enables activation of multiple episodic memories at once, allowing some degree of generalization through the retrieval of associated memories in the medial temporal lobe (Kumaran & McClelland, 2012). More recent versions of the theory also allow the hippocampus to manipulate natural statistics in its representations in a goal-dependent fashion (Kumaran, Hassabis, & McClelland, 2016).

Modeling Approach & Aims

We constructed a computational model of complementary learning systems by integrating a model of episodic memory with a model of statistical learning in an end-to-end trainable neural network architecture. We model episodic memory with a nearest neighbors-based memory module that can retrieve past observations, and statistical learning with a parametric module that can perform inference on a given observation. We call this hybrid a "semiparametric" model. Crucially, we do not constrain our model to leverage these components in a particular way, allowing us to explore the role of each over the course of learning.

To our knowledge, our approach is the first computational model to integrate both parametric and nonparametric learning components into a single end-to-end trainable model. Such an approach provides new avenues for modeling the CLS framework.

We sought to compare our model's behavior to relevant behavioral and neural results. We begin by outlining these results. Hippocampal lesions are known to impede new learning and induce temporally graded retrograde amnesia for recent experiences (Winocur, 1990; Squire, 1992; J. J. Kim & Fanselow, 1992). Other work suggests that damage to the hippocampus hinders ability to generalize and transfer knowledge across tasks, while basal ganglia lesions are detrimental to overall learning performance (Myers et al., 2003). Clinical studies on patients with Parkinson Disease reveal a similar functional separations between learning and transfer generalization capabilities (Herzallah, Moustafa, & Misk, 2010). Experiments with lidocaine injections in the hippocampus in rats support the hypothesis that it is crucial to useful situational

* Both authors contributed equally to this work.

generalization (Packard & McGaugh, 1996). Hippocampal lesions also prove detrimental to early learning, consistent with the notion that the hippocampus underlies the recognition of novel patterns (S. M. Kim & Frank, 2009).

Recent evidence has challenged a strict dichotomy between a purely episodic hippocampus and slowly adjusting statistical neocortex. Several experiments indicate a capacity to link related, recently experienced memories (Zeithamova, Schlichting, & Preston, 2012; Preston, Shrager, Dudukovic, & Gabrieli, 2004; Dusek & Eichenbaum, 1997). Though neural traces of episodic memories are initially quite distinct, as in the traditional view of the hippocampus, fMRI data indicate that traces of memories with shared features show overlap after one week (Tompary & Davachi, 2017). Detailed models of the hippocampus indeed suggest that some regularities in memories may be uncovered by the hippocampus itself, contradicting the most uncompromising theories of its role as housing uncorrelated episodic traces (Schapiro, Turk-Browne, & Botvinick, 2017).

Data show a decreasing dependence on episodic hippocampal representations as learning progresses. For instance, the behavior of rats on a water maze task reflects instance-based learned representations initially and neocortical parametric representations after a month of learning (Richards et al., 2014). Behavior consistent with this account is also observed in the aforementioned lidocaine experiment (Packard & McGaugh, 1996).

In light of CLS theory and the experimental literature outlined above, we set out to devise a computational model with the following high-level characteristics:

1. the model contains parallel episodic/nonparametric and statistical/parametric components,
2. the parametric component uses representations in the nonparametric component to form its own representations,
3. the episodic component may retrieve multiple related memories in response to a relevant stimulus,
4. the behavior of the episodic component may be learned to benefit task performance rather than directly reflect natural statistics.

We test our model's consistency with the following claims:

1. an episodic/nonparametric system aids in domain transfer generalization,
2. an episodic/nonparametric system aids in rapid learning from few examples,
3. a statistical/parametric system is important for attaining good performance on difficult vision and reinforcement learning tasks,
4. a model containing complementary learning systems will, with learning, exhibit increasing representational overlap for similar inputs,
5. such a model will rely increasingly on its statistical, parametric component as learning progresses.

Methods

Modeling CLS for Visual Recognition Tasks

We constructed a hybrid parametric and non-parametric deep learning model designed to perform image classification. This model consists of 1) a neural network that maps input to an embedding space 2) a fully differentiable nearest-neighbors-based classifier that operates on this embedding space and 3) a classifier network that operates on both the nearest-neighbors results and embedding space representation of an input image. This model architecture is diagrammed in Figure 1 and detailed in Algorithm 1.

To make our model end-to-end trainable, we use a differentiable nearest-neighbors calculation that operates on data-label pairs of the current batch during training. Given the embedding v_i of example i, we compute its squared distance $d_{ij} = (v_i - v_j)^2$ to the embedding of each example j, $j \neq i$, in the current batch. From these, we compute weights $w_{ij} = d_{ij}^{-\tau}$ where τ is a hyperparameter that modulates the emphasis on tight clustering in embedding space (we used $\tau = 2$). The calculation outputs (unnormalized) class probabilities $P_a(c_i = a) = \sum_{j \neq i} w_{ij} 1[c_j = a]$. These probabilities are concatenated with the embedding v_i of the current example to form the input to the final classifier network.

We trained our semiparametric model on the standard MNIST and CIFAR-10 datasets for image classification. We compared against a parametric baseline, identical to the semiparametric model but without the nearest-neighbor component, and a nonparametric baseline which does not include the parametric classifier network.

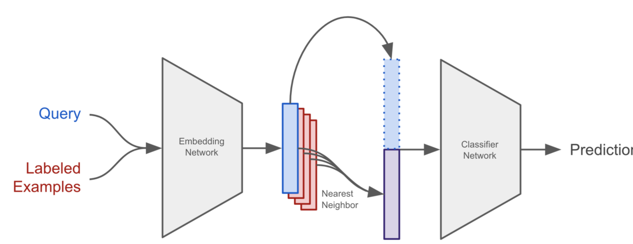

Figure 1: Architecture of our neural network model of complementary learning systems. Our model is designed to generate a prediction given an input query and labeled example queries. Out model incorporates nonparametric learning by performing nearest neighbor-based retrieval in a learned embedding space, and parametric learning by simultaneously incorporating inferences made on the embedding of the current input. Gradients flow through the entire architecture. Our model differs slightly when applied to control tasks, which is elaborated in-text.

Algorithm 1 Semiparametric Model of CLS

B: batch size.
(x_i, c_i): (example, label) pairs
$\tilde{c}$: one-hot encoding of class c
τ: cluster separation hyperparameter (set to 2)
for each batch **do**
  for $i = 1, 2, \ldots, B$ **do**
    Map example x_i into embedding v_i.
  end for
  for $i = 1, 2, \ldots, B$ **do**
    for $j = 1, 2, \ldots, B, j \neq i$ **do**
      $d_{ij} \leftarrow (v_i - v_j)^2$
      $w_{ij} \leftarrow (d_{ij})^{-\tau}$
    end for
    Estimate $\tilde{c}_i{}' = \sum_{j \neq i} w_{ij} c_j$
    Concatenate v_i and $\tilde{c}_i{}'$, map to prediction $\hat{\tilde{c}}_i$
    Train predicted label $\hat{\tilde{c}}_i$ on actual label c_i
  end for
end for

Modeling CLS for Control Tasks

We also constructed a hybrid parametric and non-parametric deep learning model for reinforcement learning tasks. The model is similar in most regards to our model for visual recognition, with minor changes made to suit the setting and improve performance. In particular, as in deep Q learning (Mnih et al., 2015), the model learns to estimate values of state-actions pairs (termed Q-values in the reinforcement learning literature). Our memory embedding space stores Q-values as well—in essence, these replace the role of class probabilities in the image classification setting.

As in the classification setting, our model begins with a trainable embedding network which maps raw inputs to an embedding space. Our episodic memory module and differentiable approximation to nearest neighbor-based classification takes after (Pritzel, Uria, Srinivasan, & Puigdomènech, 2017). Throughout training, we store N-step approximations to Q-values of observed state-action pairs in a dictionary, along with their embeddings at observation time. We subsequently find a specified number (fixed at $K = 50$ in our experiments) of neighbors with embeddings and stored Q-values. The retrieval-based Q-value estimate is computed as in the classification setting, but using only these nearest neighbors for computational efficiency.

We apply a fully parametric Q-value prediction network to the current observation. In the reinforcement learning setting we interpret it as a correction to the output of the nonparametric component; we find empirically that this better incorporates the nonparametric computation and improves performance.

We trained our model on the Atari games Venture, Bowling, H.E.R.O., and Enduro (Bellemare, Naddaf, Veness, & Bowling, 2013). We also trained it on a simple Unity-based ball-rolling task (in which the objective is to collect twelve fixed tokens in a square field under a time constraint) in order to observe the algorithm's behavior all the way through convergence on a task, given our resource constraints.

Results & Discussion

Visual Recognition Tasks

We first observed that the semiparametric model matches the final performance of the parametric baseline model on MNIST and CIFAR-10.

Next, we tested each model on a domain adaptation problem. We trained to convergence on a subset of MNIST containing half the available classes. Then we trained on n examples of unseen classes ("new domain"), varying n. We found (Figure 2) that the nonparametric model gave good performance most quickly on the new domain and that the semiparametric model captured some of this advantage. The same results hold when we include all classes in the second ("expanded") domain, demonstrating that the semiparametric model can adapt without catastrophic interference. On a more challenging dataset, CIFAR-10, we performed a similar adaptation experiment but allowing multiple iterations over the training data in the second domain. We find that the semiparametric model adapts quickly. On CIFAR-10, the parametric baseline fails to learn on the expanded domain even after 500 epochs, indicating that it struggles to learn new categories without interfering with existing knowledge.

We experimented with training all models on 1000-image subsets of MNIST and CIFAR-10. As shown in Figure 4A,

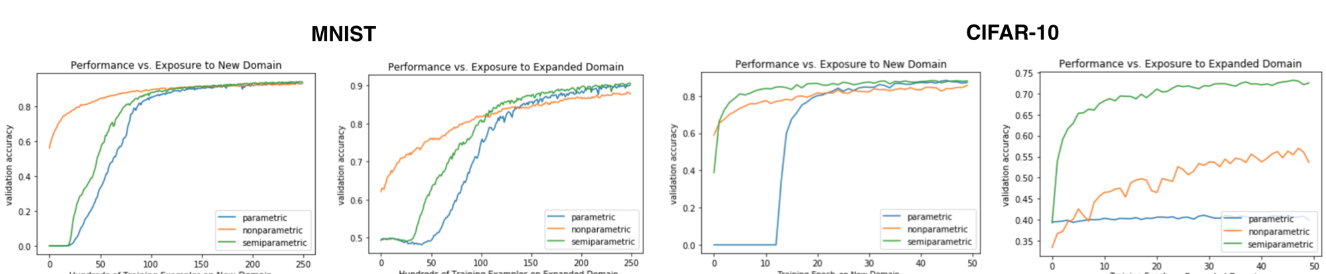

Figure 2: Model validation accuracy after exposure to training examples from a new domain (unseen classes) or expanded domain (unseen and previously seen classes).

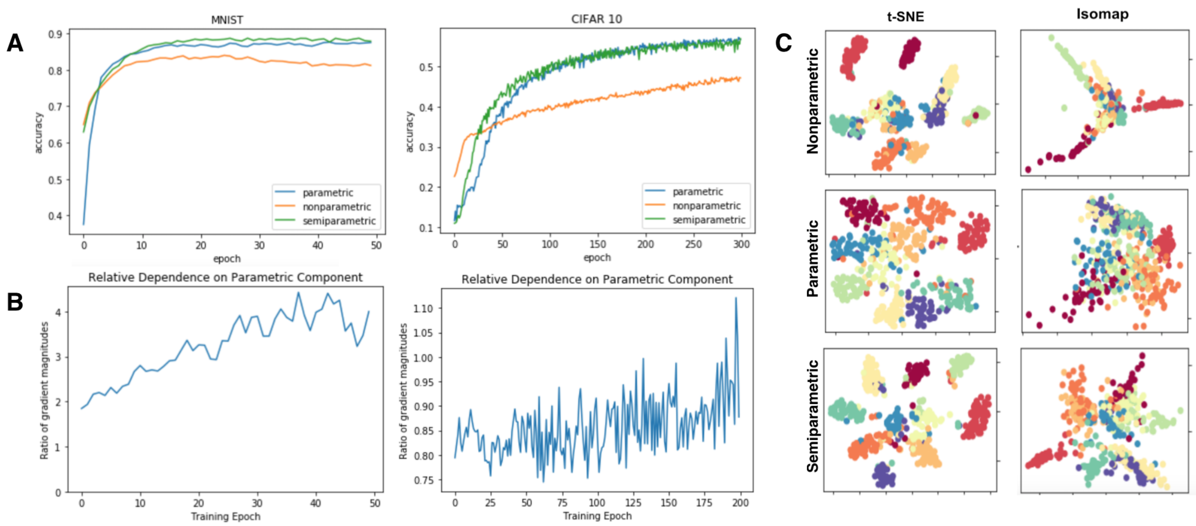

Figure 3: (A) Image classification accuracies on MNIST and CIFAR-10 validation sets. (B) The model's relative dependence on nearest-neighbors-based information, as measured by the ratio of gradient magnitudes with respect to the model's nearest-neighbor and parametric components. (C) T-SNE and Isomap plots of the embedding space of each model after training, color-coded by class.

the nonparametric baseline learns quickly, but not asymptotically well; the parametric model has asymptotically better performance but initially requires more training time to learn. Our semiparametric model, on the other hand, learns both quickly and asymptotically well.

Our empirical results show that the semiparametric model mimics advantageous properties of the nonparametric model early in training but gradually converges toward the performance of the parametric model. We show directly that this phenomenon is due to initial reliance on the nonparametric component of the model which wanes over time. To measure this reliance, we compute the magnitude of the gradient of the model's output with respect to the embedding space, as a fraction of the magnitude of the gradient with respect to the nearest-neighbors step. This metric serves as a first-order approximation to the model's relative dependence on nearest-neighbors retrieval. Figure 4B provides empirical evidence that this dependence increases over time, consistent with the analogy to the psychological theory of CLS.

We analyzed the learned representations of each model, employing the low-dimensional embedding techniques t-SNE and Isomap. (Figure 4C). The embeddings for each trained model map examples of the same class into local clusters. However, it appears that the nonparametric model exhibits the tightest clustering, followed by the semiparametric model.

Control Tasks

As our method was designed to capture the benefits of a nonparametric nearest neighbors-based approach as well as those of traditional powerful deep parametric models, we compare our model to high-performing models in either category as baselines: a nonparametric variant of the Neural Episodic Control model (NEC) (Pritzel et al., 2017) and Double DQN with rank-based prioritized experience replay (which we will refer to as DDQN+) (Schaul, Quan, Antonoglou, & Silver, 2015). As the parametric baseline model trains slowly, we took asymptotic performance figures from published results and indicated them with dashed lines in Figure 3.

In early learning stages on the models tested, the semiparametric model matched or exceeded the baselines. We conclude that the semiparametric model captures and sometimes enhances the early-learning advantage of nonparametric methods. This result is reasonable, as the decision in the nonparametric baseline to weight neighbors according to inverse distance in embedding space from the current state-action is rather arbitrary. A more complex function of these distances, or one dependent on the current state embedding, might more accurately estimate the Q-function. Our method appears to provide this functionality.

We found that semiparametric learning matches or exceeds the asymptotic performance of NEC (Figure 4A). This result extended to games in which nonparametric learning gave poor results even in early stages. On Enduro, for instance, where NEC fails to learn, semiparametric learning did not suffer the same issue. The semiparametric model exhibited the same advantage on the Roll-a-ball task (Figure 4B).

We investigated our model's dependence on nearest-neighbors data relative to parametric learning over the course of learning. We quantified this dependency with two approaches: 1) calculating the magnitude of the correction to the nearest neighbors output over the course of training, and 2) by measuring a gradient magnitude ratio as in the classification setting. In all tasks measured, including the Roll-a-ball

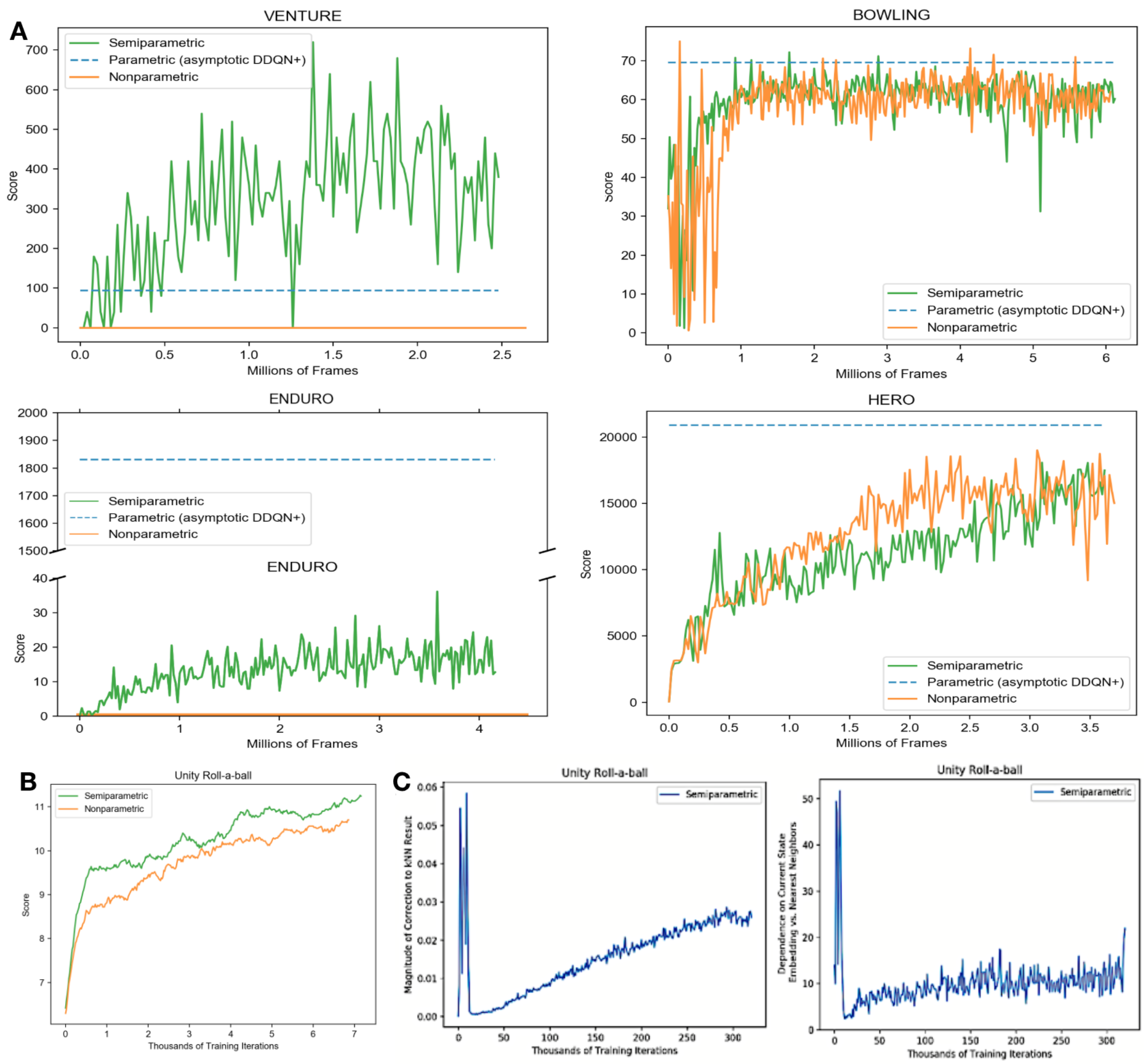

Figure 4: Reinforcement learning results on (A) Atari games and (B) the Roll-a-ball task. (C) The model's relative dependence on its parametric component during training on Roll-a-ball, as measured by the magnitude of corrections to the nonparametric Q-value estimates (left) and the ratio of gradient magnitudes with respect to the model's parametric and nearest-neighbor components (right).

task (Figure 4C), semiparametric learning appears to rely increasingly on its parametric corrections to the nearest neighbor results, as in the image classification setting. This is consistent with the paradigm that purely nonparametric, nearest neighbor-based methods become less advantageous as training progresses.

Conclusion

We introduced a neural network model of complementary learning systems. Our model integrates nonparametric and parametric learning computations, reflecting the broad roles in the CLS framework of episodic memory in the hippocampus and statistical learning in the neocortex. Our model is end-to-end differentiable, allowing the embeddings of observations to be manipulated to optimize task performance, as in modern CLS theory. Crucially, we make few assumptions about how and when our model incorporates parametric versus nonparametric computations during learning by allowing the details of this process to be learned by a neural network.

We have demonstrated that a model with these computational components exhibits properties consistent with neural and behavioral data. We find, for instance, that the model's nonparametric component is crucial to its ability to generalize across domain transfer and to learn rapidly from few observations, while its parametric component provides long-term learning power. These results mirror observations of animal and human subjects with impairments in the relevant brain regions. Moreover, our model learns to depend increasingly on parametric representations as learning progresses, consisting with behavioral studies and with the principles of CLS theory. We also observe, even in the nonparametric module, increasing representational overlap between qualitatively similar observations with learning, as has been observed in neural data.

We believe that analysis of such models can provide insight into the role between statistical and episodic learning systems. One might, for instance, examine which episodes are evoked by the memory retrieval process in response to a given input and examine how they are incorporated into decision-making. The model's graceful performance in response to domain transfer permits investigation into how observations from an unseen distribution can be incorporated in a connectionist model without catastrophic interference. Furthermore, our model's performance in comparison to traditional parametric and nonparametric models may make it of interest to the machine learning community. Future work might update our model to more concretely reflect neuroscientific understanding of memory consolidation.

Acknowledgments

We would like to thank Jay McClelland for his insights into CLS, and to Emma Brunskill and Amir Zamir for guidance on the computational component of this work.

References

Bellemare, M. G., Naddaf, Y., Veness, J., & Bowling, M. (2013). The arcade learning environment: An evaluation platform for general agents. *J. Artif. Intell. Res.(JAIR)*, *47*, 253–279.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625–641.

Dusek, J. A., & Eichenbaum, H. (1997). The hippocampus and memory for orderly stimulus relations. *PNAS*, *94*(13), 7109–7114.

Herzallah, M. M., Moustafa, A. A., & Misk, A. J. e. a. (2010). Depression impairs learning whereas anticholinergics impair transfer generalization in parkinson patients tested on dopaminergic medications. *Cog. and Behav. Neuro.*, *23*(2), 98–105.

Jain, M. S., & Lindsey, J. (2018a). Deep semiparametric learning. *In progress*.

Jain, M. S., & Lindsey, J. (2018b). Semiparametric reinforcement learning. *ICLR 2018 Workshop Track*.

Kim, J. J., & Fanselow, M. S. (1992). Modality-specific retrograde amnesia of fear. *Science*, *256*(5057), 675–677.

Kim, S. M., & Frank, L. M. (2009). Hippocampal lesions impair rapid learning of a continuous spatial alternation task. *PLoS One*, *4*(5), e5494.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? complementary learning systems theory updated. *Trends in cognitive sciences*, *20*(7), 512–534.

Kumaran, D., & McClelland, J. L. (2012). Generalization through the recurrent interaction of episodic memories: a model of the hippocampal system. *Psychological review*, *119*(3), 573.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). 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.

McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In *Psychology of learning and motivation* (Vol. 24, pp. 109–165). Elsevier.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... Ostrovski, G. e. a. (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529–533.

Myers, C. E., Shohamy, D., Gluck, M. A., Grossman, S., Kluger, A., Ferris, S., ... Schwartz, R. (2003). Dissociating hippocampal versus basal ganglia contributions to learning and transfer. *Journal of Cognitive Neuroscience*, *15*(2), 185–193.

Packard, M. G., & McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiology of learning and memory*, *65*(1), 65–72.

Preston, A. R., Shrager, Y., Dudukovic, N. M., & Gabrieli, J. D. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus*, *14*(2), 148–152.

Pritzel, A., Uria, B., Srinivasan, S., & Puigdomènech, A. e. a. (2017). Neural episodic control. *arXiv*.

Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature neuroscience*, *17*(7), 981.

Schapiro, A. C., Turk-Browne, N. B., & Botvinick, M. e. a. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Phil. Trans. R. Soc. B*, *372*(1711), 20160049.

Schaul, T., Quan, J., Antonoglou, I., & Silver, D. (2015). Prioritized experience replay. *arXiv*.

Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological review*, *99*(2), 195.

Tompary, A., & Davachi, L. (2017). Consolidation promotes the emergence of representational overlap in the hippocampus and medial prefrontal cortex. *Neuron*, *96*(1), 228–241.

Winocur, G. (1990). Anterograde and retrograde amnesia in rats with dorsal hippocampal or dorsomedial thalamic lesions. *Behavioural brain research*, *38*(2), 145–154.

Zeithamova, D., Schlichting, M. L., & Preston, A. R. (2012). The hippocampus and inferential reasoning: building memories to navigate future decisions. *Frontiers in human neuroscience*, *6*, 70.