Unsupervised Temporal Learning During Sleep Supports Insight

Itamar Lerner (itamar.lerner@gmail.com)
Center for Molecular and Behavioral Neuroscience, Rutgers University – Newark,
197 University Ave., Newark, New Jersey 07102, U.S.A

Abstract:

Sleep, and particularly Slow Wave Sleep, is known to facilitate insight learning. The mechanism, however, remains unclear. Here, we suggest that sleep-dependent facilitation of insight is typically achieved in tasks requiring temporal pattern detection, and propose that compressed memory reactivations during sleep support such detection. We demonstrate this mechanism with a spiking neural network model of hippocampus-prefrontal cortex interactions. We show that supervised learning in the prefrontal cortex during wake benefits from prior sleep-dependent unsupervised learning in the hippocampus, allowing for quick extraction of temporal regularities that can be utilized to predict upcoming events. Without hippocampal learning, prediction of temporal patterns is very slow or fails completely. Thus, compressed memory replay at sleep gives a unique opportunity for developing insight into temporal patterns.

Keywords: sleep; insight; hippocampus; STDP; memory replay; memory consolidation

Introduction

During the last two decades, sleep, and especially Slow Wave Sleep (SWS), has been shown to facilitate a wide range of cognitive functions, including high-level cognition such as gist learning and insight into hidden rules (Rasch & born, 2013). One account for these findings is offered by the memory consolidation theory. It suggests that the hippocampus registers ongoing sequential experiences during waking and those encoded memories are reactivated and transferred to the prefrontal cortex (PFC) for permanent storage during SWS. This transfer occurs gradually and allows the PFC to extract regularities embedded in the encoded experiences and reorganize them in efficient patterns. Evidence for memory replay occurring at short bursts in the rat hippocampus during SWS supports the consolidation theory (Wilson & McNaughton, 1994). Moreover, replay of the most recent encoded experiences can also occur at wake before decision points, possibly supporting active planning (Rasch & Born, 2013). However, it is not clear how such replay facilitates insight learning.

We developed a novel 'Temporal Scaffolding' hypothesis of sleep-induced insight learning. This hypothesis is built on two observations: (1) The facilitatory effects of sleep on insight (defined here as rapid explicit recognition of new, unexpected patterns) in humans are most apparent when subjects are required to detect hidden ***temporal*** patterns in stimuli (Wagner et al., 2004; Fischer et al., 2006); and (2) Memory replay in the hippocampus during SWS tends to be time-compressed, such that encoded sequences are replayed in an accelerated manner (in rats, up to x20 faster than the original experience; Rasch & Born, 2013)

Building on these two observations, we suggest that if information encoded during wake contains unanticipated temporal regularities occurring over time gaps larger than seconds, it is difficult to extract online due to the limits of typical Hebbian-learning timescales (which require activity coincidence at the 50-100ms level). However, the time-compression characterizing replay of experiences during SWS bridges those gaps and allows the association of related bits of information that were experienced at disparate moments in time. Such temporal scaffolding could thus allow rapid detection of hidden temporal patterns in stored memories upon awakening, namely, insight.

Methods

To demonstrate the temporal scaffolding hypothesis, we developed a spiking neural network model of the hippocampus->PFC pathway. Bursts of spikes from the hippocampus (representing memory replay during wake that corresponds to the latest encoded sequential inputs from the environemnt) are transformed by the PFC into a prediction regarding the next upcoming environmental input in the sequence (Figure 1).

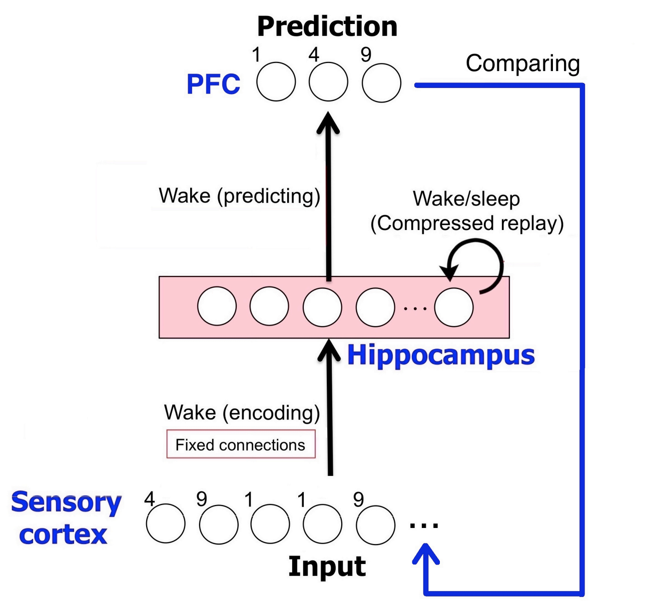

Figure 1: Core Network Architecture and Processes

Predictions are optimized through supervised learning based on the continuous environmental inputs, using the Tempotron algorithm (Gutig & Sompolinksy, 2006). If the prediction results in an error, for each pre-synapatic spike i occurring at time t_i within a replay, weights from the hippocampus to the PFC are updated according to:

$$\Delta w_i = \lambda \sum_{t_i < t_{\max}} < \boldsymbol{K}(\boldsymbol{t}_{\max} - \boldsymbol{t}_i)$$

With λ being the lerning rate, K being a standard function describing the normalized post-synaptic potential (PSP) of each spike, and t_{max} being the time at which the PSP is maximal following the replay event.

Unsupervised learning within the hippocampus (assumed to reflect learning in lateral connections during spontaneous sleep-dependent replay of previously encoded environmental inputs) allows detecting patterns in the input prior to wake. Unsupervised learning is implemented with Spike-Timing-Synaptic Plasticity (STDP) combined with auto-encoding. For each two spikes i and j, occurring at times t_i and t_j during replay S, hippocampal weights are updated according to:

$$\Delta w_{ij} = \begin{cases} \boldsymbol{a}_+ \boldsymbol{e}^{\frac{-(t_i - t_j)}{\tau^+}} & , \quad \text{spike}_i \in \boldsymbol{S} \\ -\boldsymbol{a}_- \boldsymbol{e}^{\frac{-(t_i - t_j)}{\tau^+}} & , \quad \text{spike}_i \notin \boldsymbol{S} \end{cases}$$

With a_+, a_- and τ^+, τ^- being the learning rates and time constants for strengthening and weakening of weights, respectively.

We simulated the 'Number Reduction Task', an insight learning task shown in humans to be facilitated by SWS (Wagner et al., 2004). Inputs were sequences of digits that included a fixed temporal pattern (specifically, a mirror image; e.g., 4 9 1 1 9 4), arriving during wake at a speed of approximately 1 per second. The model was run (1) before sleep-dependent hippocampal unsupervised learning; (2) after sleep-dependent learning. Hippocampal replay was compressed by a factor of ~80.

Results

Before sleep, the hippocampus was not able to perform pattern completion of the full mirror-sequence when given its initial part (Figure 2A, upper row, middle). Moreover, the PFC typically failed to converge on correct predictions given partial replay as input (Figure 2A, upper row, left; 2C, blue curve). On some runs (depending on the random initialization of the weights) correct predictions were achieved based on the partial input, but only after many trials (Figure 2C, brown curve).

During sleep-dependent unsupervised learning, the hippocampus picked up the mirror-structure of the stored inputs (Figure 2B). Following sleep, pattern completion was present after partial replay (Figure 2A, upper row, middle). This time, the PFC consistently learned to predict the next input in the sequence, and it did it very quickly, demonstrating "insight" behavior (Figure 2C, green curve). PFC decision times were based on the completed replay pattern rather than its initial pattern (compare Figure 2A upper and lower rows, right).

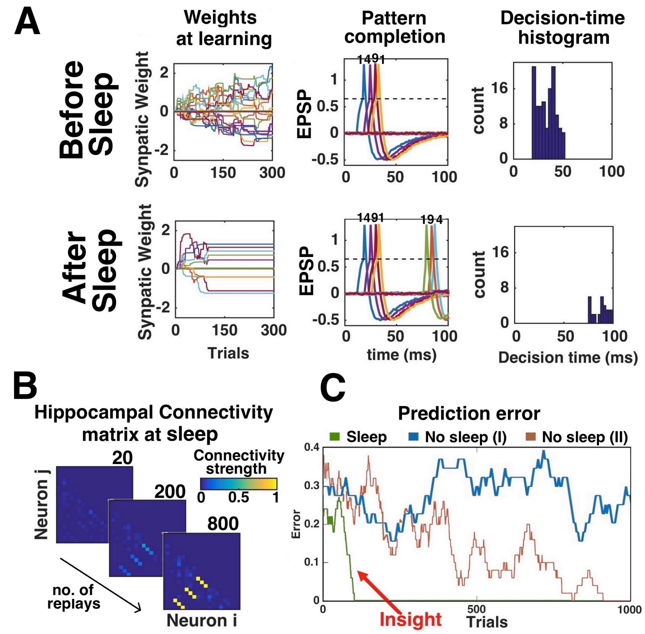

Figure 2: A. Learning before and after sleep. B. Hippocampal weights during sleep. C. Prediction errors at wake.

Conclusion

Our results demonstrate that temporal pattern detection during SWS could be the result of compressed memory replay, which allows regularities to fit into sufficiently short time periods to be picked on by Hebbian mechanisms. On the computational level, our work also demonstrates how initial unsupervised temporal learning could assist subsequent supervised temporal learning in transforming inputs into required outputs.

References

Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning–explicit knowing: a role for sleep in memory system interaction. *Journal of Cognitive Neuroscience, 18*(3), 311-319.

Gütig, R., & Sompolinksy, H. (2006). The tempotron: a neuron that learns spike timing-based decisions. *Nature Neuroscience, 9*(3), 420-428.

Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological reviews, 93*(2), 681-766.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature, 427*(6972), 352-355.

Wilson, M.A., & McNaughton, B.L. (1994). Reactivation of hippocampal ensemble memories during sleep. *Science, 265*(5172), 676-679.