A comprehensive breakdown of the relationship between memory schemas and dreams using neural networks.

Avinash Dindial

1006020949

University of Toronto

Abstract

Sleep is well understood to have benefits to health such as muscle repair etc., however, the aspect of dreams remain ambiguous. REM sleep is the particular period of sleep where one dreams and is characterized by muscle atonia, rapid eye movements and brain wave levels that resemble that of a waking state. These brain waves can be measured using the EEG. The fact that other animals, not just humans, dream suggest it is an important part of life. With the exception of cetaceans, all other mammals dream to some degree (Siegel et. al., 2020). Sleep stages differ in their roles in memory. For instance, NREM sleep is more closely related to declarative memory, while REM sleep is mostly correlated with emotional memory and memory that participates in implicit learning (Zhang, 2016). The focus on this paper is to highlight the role of REM sleep and how is impacts memory schemas. This is done using a computational approach via a neural network that was developed to mimic the encoding, retrieval etc., of memory schemas.

The phenomenon of dreaming has escaped the understanding of humans for much of our existence. Sleep ensures our body rests and rebuilds muscle etc. but dreaming seems to be an elusive side-effect. The phenomenon of dreams has been often been explained as a way to aid us in learning. The brain makes use of schemas to avoid catastrophic forgetting and balances plasticity and stability (Hwu et. al., 2020). Plasticity refers to how the brain is plastic enough to form new memories but stable enough to make them last a possible lifetime. The role of memory schemas in learning is evident which makes the connection of dreams, schemas and learning prudent to understand since everybody dreams. Specifically in this paper, this notion of learning can be understood using the concept of overfitting. This approach uses a computational approach to understand psychological phenomena in neural networks and in humans. This computational perspective of the human brain has been proposed and developed since the early 20th century. This paper seeks to understand the role of dreams through the lens of memory schemas using neural networks. This paper is based on the article “A neural model of schemas and memory encoding” by Hwu & Krichmar, 2020. In this paper, a neural network was designed to mimic the neurobiological mechanisms behind tasks that involve memory schemas. In particular, the authors of this paper focused on the tasks by Tse et. al., 2007. In this paper, rats were trained for 20 days with the association of food well location and food flavour within a maze. On the 21st day, two food wells were removed and replaced with two new food wells with new foods, close to the locations of the original wells and referred to as the new paired associations within a pre-existing schema. The rats learned the locations of the two new food wells rapidly, suggesting that the presence of the wells within a familiar schema increased the rate of learning. This was then conducted for rats with HPC lesions and was shown that they were unable to learn the location of the new food wells but retained the original schema. Using all these variables, a neural network was then built that gave similar results to Tse et al. The neural network was consisted of two information streams: indexing and representation. The indexing stream mimics neural activity of the mPFC. Specifically, it mimics the way schemas are made in the mPFC using sensory input. In our case, we used a 2D grid with inputs of 1 if a food well existed in that grid location or 0 otherwise (Hwu et. al., 2020). The representation stream encodes the current schema. This is a multi-layered network in which weights are calculated using contrastive Hebbian learning. The model was validated by simulating the rat experiment explained above. Subsequently, two experiments were performed on this network. The goal of the first experiment is to show that new information matching a schema can be learned quickly and also that the HPC is necessary for learning. Recall in the original experiment, rats with lesioned HPC showed impaired learning. The purpose of the second experiment was to show that multiple schemas could be learnt by the same network and also that the HPC was necessary for this. The model was also tested under an 'HPC lesioned' condition which gave similar results to that of the rat experiment. The model retained information prior to being 'lesioned'. This showed that our biologically plausible neural network was able to learn schemas over time and quickly assimilate new information if it was consistent with a prior schema (Hwu et. al., 2020).

Now that we have established how the neural network works in parallel with human cognition on this task, we can assume that this would generalize to other tasks of learning. The neurobiological implications of our model, our work could have practical applications to a range of tasks in artificial intelligence (Hwu et. al., 2020). In order for a neural network to become good at replicating human cognition, it must be fed information and allowed for this information to be replayed and manipulated etc. Neural networks operate in two phases: waking phase where information is actively being inputted, and a sleeping phase where no new information is being inputted. During the sleeping phase, the neural network ‘dreams’ scenarios using information in the waking phase. This allows for the neural network to reorganize information learnt during the waking phase into a more cohesive and generalizable schema. According to the Overfitted Brain Hypothesis proposed by Erik Hoel in 2017, the human brain undergoes a similar process when developing schemas. The reason dreams are so fantastical is because our brain is replaying input taken during the day and developing memory schemas which is then consolidated. Indeed, suppression of sleep in humans is linked to inhibited learning. The process of dreaming is akin to introducing noise in a neural network; this allows for connections and neural activity within the brain to reorganize themselves. This in turn strengthens prior memories or encodes new memories depending on what is inputted. This opponent processing model of forming schemas is well established in current literature. Schemas are dynamic structures constantly evolving with new experiences and memories through processes dubbed assimilation and accommodation (Gilboa et. al., 2017). To add validity to this claim, there exists the memory consolidation theory of dreaming. This states dreams serve to store or create new memories using assimilation and accommodation. However, this phenomenon also takes place during NREM sleep. For example, there is evidence that learning a new task leads to a greater activation during both REM and slow wave sleep in the task-relevant cortical areas, which indicates there is no preferential consolidation during dreaming (Hoel, 2021). However, the argument can be made that dreaming does not simply consolidate memory, it allows for the generation of predictive, generalizable schemas. This uses a Top-down processing approach where the fantastical nature of dreams is bootstrapped in the high-level nodes of the brain, thereby affecting the input of lower-level nodes. Dreams are often more common and potent in younger children as opposed to adults. Using this hypothesis, this makes sense as children are required to form new schemas due to their limited exposure to the world. The demand for new schemas to help them navigate the world is provided by dreams to help them anticipate the world. For instance, if a young child perceives a dog as a four-legged creature with fur that barks etc., the next time they see a similar animal, they can anticipate its behaviour. This has multiple benefits especially to children, primarily it helps keep them safe by using previous schemas to anticipate the world.

The previous arguments have all converged on the possibility that dreams enforce the creation or manipulation of memory schemas to attempt to predict the world. Essentially, dreams compress data taken in during the day into a 'best fit' schema. You can think of all the random information during the day as random dots on a graph and a line of best fit as the schema produced during dreaming. This gives us the best possible interpretation of the data. This is done by introducing noise in the brain in the form of dreams, then allowed to bootstrap in higher layers of the brain. The schema produced from this is then fed into lower-level nodes. The same result is given in neural networks who are tunes to mimic neurons in the brain. When noise is injected into a neural network, it reorganizes nodes and makes connections more efficient. This parallel in results converge on the Overfitted Brain Hypothesis and its implications for memory schema and dreams.

References

Asaf Gilboa, Hannah Marlatte. Neurobiology of Schemas and Schema-Mediated Memory. Volume 21, Issue 8. 618-631 (2017). https://doi.org/10.1016/j.tics.2017.04.013.

Heaton, Kristin & Hill, Clara & Hess, Shirley & Leotta, Carol & Hoffman, Mary Ann. (1998). Assimilation in therapy involving interpretation of recurrent and nonrecurrent dreams. Psychotherapy: Theory, Research, Practice, Training. 35. 147-162.

Hoel, E. (2021). The overfitted brain: Dreams evolved to assist generalization. *Patterns (New York, N.Y.)*, *2*(5), 100244–100244. https://doi.org/10.1016/j.patter.2021.100244.

Hwu, T., Krichmar, J.L. A neural model of schemas and memory encoding. *Biol Cybern* 114, 169–186 (2020). https://doi.org/10.1007/s00422-019-00808-7

Manger, P. R., & Siegel, J. M. (2020). Do all mammals dream?. *The Journal of comparative neurology*, *528*(17), 3198–3204. https://doi.org/10.1002/cne.24860

Zhang, W. (2016). A supplement to self-organization theory of dreaming. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00332