Augmenting Machine Learning with Flexible Episodic Memory

Hugo Chateau-Laurent, Frédéric Alexandre

HAL Id: hal-03359384
https://inria.hal.science/hal-03359384

Submitted on 30 Sep 2021

Augmenting Machine Learning with Flexible Episodic Memory

Hugo Chateau-Laurent[1,2,3][a] and Frédéric Alexandre[1,2,3][b]

[1]*Inria Bordeaux Sud-Ouest, 200 Avenue de la Vieille Tour, 33405 Talence, France*

[2]*LaBRI, Université de Bordeaux, Bordeaux INP, CNRS, UMR 5800, Talence, France*

[3]*Institut des Maladies Neurodégénératives, Université de Bordeaux, CNRS, UMR 5293, Bordeaux, France*

{hugo.chateau-laurent, frederic.alexandre}@inria.fr

Keywords: Explicit Memory, Episodic Memory, Prospective Memory, Hippocampus, Prefrontal Cortex.

Abstract: A major cognitive function is often overlooked in artificial intelligence research: episodic memory. In this paper, we relate episodic memory to the more general need for explicit memory in intelligent processing. We describe its main mechanisms and its involvement in a variety of functions, ranging from concept learning to planning. We set the basis for a computational cognitive neuroscience approach that could result in improved machine learning models. More precisely, we argue that episodic memory mechanisms are crucial for contextual decision making, generalization through consolidation and prospective memory.

1 INTRODUCTION

Despite recent progress in machine learning, the field of artificial intelligence has not yet given rise to generally intelligent and autonomous agents endowed with autobiographical memory and capable of defining their own goals and planning their way to reach them. As a result, machine learning models are mainly confined to domain-specific processing. Complementary and interacting memories have been argued to be a key component of autonomous learning (Alexandre, 2016). We claim that current techniques in artificial intelligence mirror implicit memory, while topics related to building explicit memory systems remain largely unaddressed. The view we depict here is mostly consistent with the one proposed in (Botvinick et al., 2019) in that they both propose to draw inspiration on both the prefrontal cortex and the hippocampus to develop more powerful algorithms. Our work further expands this proposal by considering additional mechanisms inspired from anatomical and electrophysiological studies.

We first argue that models need complementary memories, then set the focus on the organization of episodic memory in biological organisms. Finally, we propose a roadmap to implement these mechanisms in artificial agents.

[a] https://orcid.org/0000-0002-2891-0503
[b] https://orcid.org/0000-0002-6113-1878

2 THE NEED FOR COMPLEMENTARY MEMORIES

Consider a task where you have different colored shapes displayed on a table and you can press one among a set of buttons to get rewards. There are hidden rules controlling reward delivery that you can discover by trial and error. An example rule would be that if a red square is on the left, pressing the left button is rewarded. This kind of task can be learned very easily and efficiently with classical machine learning techniques like reinforcement learning (RL). Similarly, the subtask of object identification is prototypical of gradient-based forward layered architectures. Through learning, these models build what is called in cognitive science an implicit memory. This corresponds to a slow and procedural learning of the underlying regularities. If information (like position or color of the shapes) is encoded in a topographical way, the model builds an overlapping representation, which enables it to generalize. In RL, this learning can be done with a model-free (MF) approach, meaning that the values of configurations are learned without explicit knowledge of the task structure (hence the term implicit memory).

These computational approaches have been paralleled with known principles in neuroscience. For instance, layered architectures of deep networks have been compared to the hierarchical structure of the sen-

sory posterior cortex. RL models have been very beneficial to better understand learning principles of the most ancient regions of the frontal cortex (i.e. the part of the cortex that is anterior to the central sulcus) that are also called agranular cortex, corresponding to the motor, premotor and lateral orbitofrontal cortices (Domenech and Koechlin, 2015). In association with the sensory cortex, these regions are involved in decision making via action selection from perceptual cues (associated to actions in motor and premotor cortex) and reward values (learnt in the orbitofrontal cortex). Loops associating these cortical regions with the basal ganglia and modulated by dopaminergic projections carrying reward prediction errors have been carefully studied, and corresponding models have been compared to MF-RL (Joel et al., 2002). It is also reported in cognitive neuroscience that, after a long period of repetition, this reward-based behavior is transformed into a purely sensorimotor behavior (also called habitual) simply associating the sensory and motor cortices, with no sensitivity to reward change (Boraud et al., 2018): after a long practice, you automatically press the left button when you see a red square on the left without even thinking of an anticipated reward.

2.1   The need for episodic memory

The functions of these cortical regions are consequently tightly associated with implicit learning, as performed by the most famous and efficient models in artificial intelligence today, like deep networks and MF-RL agents. Nevertheless, these models suffer from a series of well-identified weaknesses. Layered architectures with overlapping representations are not good for learning data without structure or regularities. They learn very slowly, from global aspects to details and are sensitive to what is called catastrophic forgetting: if you first learn a relation (e.g. pressing the left button when seeing the red square on the left), then, when learning another one (e.g. pressing the left button when seeing the blue circle on the left), the previous association will be forgotten in the absence of regular recall.

This problem has been clearly identified and addressed in the framework of the complementary learning systems (CLS) (McClelland et al., 1995), introducing the role of the hippocampus (HPC). This neuronal structure has been originally studied in navigation tasks, demonstrating major learning capabilities and exhibiting place cells that respond to specific places in the environment (Stachenfeld et al., 2017). Later on, the HPC has been associated with more complex and general functions. It is thought to receive a summary sketch of the current cortical activation and perform functions like associative memory and recall. Thanks to its very sparse coding and an advanced function of pattern separation (Kassab and Alexandre, 2018), the HPC is able to learn very quickly, possibly in one shot, an arbitrary binding of any distributed cortical representation with a minimized risk of interference. Consequently, it cannot generalize well. However, it also performs pattern completion, namely the ability to rebuild a complete pattern from partial cues, and replay it in (i.e. send it back to) the cortex. Implemented in attractor networks, this function is useful for doing reinforcement learning in the real world with high-dimensional states and noisy observations (Hamid and Braun, 2017). It can consequently complement the cortex by learning rare patterns very quickly and binding a distributed cortical representation that would be hard to associate with a motor response at the cortical level. At each moment, when the HPC receives a cortical pattern, it has to decide whether it is novel and must be learned, or if it is reminiscent of a previously stored activation and must be completed (recalled) and sent back to the cortex. This is done by a clever mechanism described in the next section. It must also be stressed that what is manipulated by the HPC (originating from current cortical activation) does not represent a single event but rather, thanks to several mechanisms of temporal coding of the cortex evoked below, a full behavioral episode that can last several seconds. This faculty of storing and recalling specific episodes is termed episodic memory. It is a kind of explicit memory because information can here be explicitly questioned. In humans, this is often associated with a conscious declarative process.

We have now discussed principles of these complementary learning systems, with the cortex performing a slow statistical learning using distributed overlapping representations, and the HPC performing rapid arbitrary binding of cortical activation to store new episodes in a sparse associative memory and recall them later on. Interestingly, the capacity of the HPC to send back previously stored episodes to the cortex (Ólafsdóttir et al., 2018) is very useful because it can solve the problem of catastrophic forgetting. Thanks to the replay of virtual (i.e. not actually experienced) episodes, the HPC indeed has the capacity of gradually inserting examples of a new rule and interleaving them with examples of the previously stored rules not to forget them. This process of replay is also very useful for another learning mechanism called consolidation. We have explained above that it is difficult for the cortex to manipulate and for example associate by learning a concept which is distributed in distant regions of its surface. Conversely,

it is not possible for combinatorial reasons, to represent locally in associative regions of the cortex, any combination of elementary patterns that might constitute the premises of a behavioral rule. By this replay mechanism, if it appears that the HPC has to frequently send back an arbitrary binding of pattern to answer non tractable situations, the cortex will be able to slowly learn this combination of activities replayed by the HPC on its surface and to link units in an associative area to this distributed activation, thus forming locally a new concept to be employed in future behaviors, without the help of the HPC. Under this view, it can be said that the HPC is a supervisor of cortical learning. If the stored episodes represent a very rare case (or a particular case or an exception), it can continue and be treated by the HPC.

2.2 The need for cognitive control

Another series of weaknesses associated with implicit learning in the cortex is linked to the fact that, sometimes, and particularly in a complex and dynamic world, it is not sufficient for the behavior to be only reactive, stimulus driven, but it also need to be goal driven, guided by internal states like motivations, intentions or to obey external instructions providing a new rule. Consider cases where the rule associating the red square on the left with the left button changes for the right button and comes back to the left button (so-called reversal learning): a reactive stimulus-driven behavior would slowly unlearn the previous rule and learn the new one and unlearn it again, whereas it is rather expected to have a more flexible behavior including an internal deliberation to switch rapidly from one rule to the other. You can also consider cases (like in the Wisconsin Card Sorting Test (Bock and Alexandre, 2019)) where several rules are proposed but at a given moment only one is valid or cases where the new rule is given by instruction from a fellow creature. Here also you should give up immediately the previously learned rule and adopt a new one provided by internal process. You generally wait for errors (difference between anticipated and actual reward) to understand that rule has changed and that you have to find a new one. Sometimes, you can also discover that changes are associated with a contextual cue. In this case, you can learn contextual rules and change your behavior before making errors (Koechlin, 2014). Consider now cases where the rule becomes more complex and relies on non visible cues (for example press the left button if you see a red square on the left and if on the previous trial the red square was on the right). Here also a reactive behavior guided by observed cues is

not sufficient and should be completed by the possibility to decide from internal cues corresponding here to the memory of recently observed cues. Other internal cues to consider for biasing the decision could be also related to emotion or motivation associated to the present situation.

All these cases share common principles (Miller and Cohen, 2001). We are here in cases where the default behavior learned by a slow stimulus-driven process must be inhibited and replaced with a new behavior adapted to the present context. To ensure flexibility and reversibility, this process must be explicit and manipulate knowledge about the world, as in Model Based approaches in RL. Not only based on the perception of external stimuli, this system must rely on a internal source of activity to bias the default behavior suggested by stimuli and impose the new behavior in a top-down way. At the same time, it must be flexible but also robust to distractions from external stimuli.

It has been demonstrated by neuroscience that all these properties are ensured by the prefrontal cortex (PFC), a more recent (also called granular) region of the frontal cortex, particularly developed in primates (Fuster, 1989). This structure, organized according to the kind of task (Domenech and Koechlin, 2015) implements contextual rules which can ensure a top-down control of behavior. It has a modulatory role and does not replace reactive rules slowly acquired by implicit learning. Instead, it is able to inhibit the default (and not adapted here) behavior and trigger an attentional process to favor task-relevant stimuli and consequently elicit a previously learnt rule more adapted to the present context (O'Reilly et al., 2002). This modulatory role implies that, in case of a lesion, the behavior is not stopped but comes back to the (non appropriate) default case, yielding classical deficits of PFC damage like perseveration. All these mechanisms are implemented by the unique mechanism of working memory, where neuronal populations of the PFC exhibit a sustained activity to represent, maintain and impose this contextual activity. In addition to its role in learning, the role of the dopamine has been shown central to regulate the balance between maintenance and updating of the sustained activity (Braver and Cohen, 2000).

In contrast to the classical weight-based learning (as presented above with the HPC), what we describe here is an activity based control but if it is frequently exploited in certain context, it can become a default behavior by slow implicit learning.

2.3 The need for prospective memory

It is now time to describe what could be called a side effect of the interactions between these two kinds of explicit memories, considering that the HPC and the PFC are interconnected (Simons and Spiers, 2003) and apply their respective functions one on the other. In short, the HPC can replay episodes of cortical activation, including prefrontal activation; reciprocally, the PFC is informed about contextual aspects of the situation by the HPC and can control the generation of adapted episodes and their storage. These episodes can be manipulated explicitly by the PFC, the same way it creates new rules adapted to the situation. Altogether, this gives rise to prospective memory (Schacter et al., 2007), also called memory of the future (Fuster, 1989), the unique capacity of imagination, of explicitly anticipating the future from an internal model of the world and of making a decision informed by the potential outcomes of possible future scenarios and deliberating on their respective interest. This is among the most powerful capabilities of human cognition, also associated to planning and reasoning, clearly going beyond reactive automatic behavior and enabled by the combination of the two kinds of explicit memory evoked above, episodic and working memory. An alternative computational account of how these memories interact has been proposed by (Zilli and Hasselmo, 2008). Yet, this framework is rather elementary and its limitations have been highlighted by (Dagar et al., 2021), who argued that more complex mechanisms are needed in realistic cases, in the same vein as what we propose here.

Models for implicit learning are today well mastered and employed in artificial intelligence. Models for cognitive control by the PFC have been proposed for some time in computational neuroscience (O'Reilly et al., 2002) and their principles begin to be transferred to machine learning (Wang et al., 2018). Episodic learning has also been considered recently for machine learning (Ritter et al., 2018) but for the moment, the corresponding models propose a rather superficial view of this cognitive function and of its multiple roles in cognitive control, consolidation and prospective memory, as we have evoked here. In the forthcoming section, we explain that integrating information about episodic learning from different sources in neuroscience can help understand and give a more precise computational account of this cognitive function in order to define an integrated explicit learning involving both the HPC and the PFC.
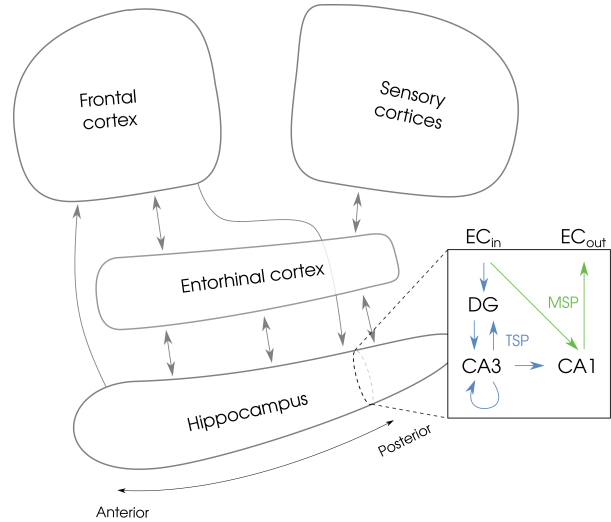
3 ORGANIZATION OF EPISODIC MEMORY

Figure 1: Connectivity within the hippocampus and between the hippocampus and the frontal, posterior and entorhinal cortices. Connections of the trisynaptic pathway are shown in blue while connections of the monosynaptic pathway are shown in green.

3.1 Memory storage in the HPC

As mentioned above, the HPC is able to store and replay sequences of experienced states, as well as construct novel ones. These sequences are embedded in two coupled oscillatory regimes: a theta rhythm (4-10 Hz) and faster gamma rhythm (20-80 Hz). According to their respective frequency, approximately seven gamma cycles are nested in each theta cycle. The entorhinal cortex (EC) constitutes the main input to the HPC by its superficial layers that we call ECs. They are thought to provide one state per theta cycle (Jensen and Lisman, 2005). The HPC then either engages its gamma regime in a retrospective code supporting short-term memory and episodic storage, or in a prospective code with which it tries to predict future states (Lisman and Otmakhova, 2001). In both cases, sequences are represented in a compressed way, with one state represented per gamma cycle. In the retrospective (or learning) code, the sequence of gamma states reflects the order at which past states have just been experienced, with the current state being represented at the end of the theta cycle. The main computational reason behind having such a buffer is that the gamma timescale is more appropriate at a biophysical level (relating to constants of time of molecular processes) for learning state transitions than the

behavioral timescale at which events naturally occur (Lisman and Otmakhova, 2001), especially given the fact that the most anterior parts of the HPC learn transitions between temporally distant events as we will discuss below. In the prospective (or recall) code, the item provided by the cortex is followed by one item per gamma cycle. The order and identity of these subsequent items reflect a prediction of upcoming states, potentially based on previous learning through retrospective coding. Alternation between these two mechanisms is presumably based on the ability (or inability) of the prospective code to accurately predict future states. Under this view, the HPC begins by considering the current episode as already known and attempts to predict the next states until errors indicate that the current episode is too different from stored situations and must consequently be considered novel and learned accordingly. In fact, a crucial part of the HPC called CA1 is often linked to novelty detection, namely computation of the mismatch between predictions and incoming sensory information (Duncan et al., 2012). When unpredicted sequences of stimuli are presented to the HPC, a dopaminergic signal is used to switch from the prospective code to the learning retrospective code in order to update the internal model (Lisman and Otmakhova, 2001).

Retrospective and prospective codes are expanded with mechanisms of pattern separation and pattern completion (Kassab and Alexandre, 2018). The HPC indeed tries to complete each gamma state (whether being actual or predicted) by relating it to states it has experienced in the past. This is the task of a recurrent region called CA3. Conversely, if the state does not match any stored pattern, its representation is modified in such a way that it becomes as dissimilar (orthogonal) as possible to any other pattern, to avoid interference in this very fast learning process. This is performed in the largest region of the HPC, the Dentate Gyrus (DG), thus encoding what makes each state unique, with a very sparse coding. In sum, the HPC performs a succession of auto-associations (encoding states) and hetero-associations (encoding their sequence, whether being retrospective or prospective) (Lisman and Otmakhova, 2001).

The mechanisms just described are thought to underlie the learning of individual experienced sequences of states (i.e. episodes). They mainly involve the trisynaptic pathway : ECs → DG → CA3 → CA1. CA1 also receives information from a more direct monosynaptic pathway: ECs → CA1 → ECd. While fast learning enabling the trisynaptic pathway to learn to link items represented in successive gamma cycles is obviously useful for one-shot sequence learning, slower learning deployed in the monosynaptic path-

way may serve additional roles. We have seen that what is encoded in DG and CA3 is not the full description of a state but rather what makes it different from the others, through orthogonalization processes. When the HPC sends information back to the cortex for reasons we will detail in the next part, this particular encoding cannot be used as is. In fact, the role of CA1 is to act as a translator between the sparse representation of the HPC and the dense overlapping representation used in the cortex (Schapiro et al., 2017). Learning mechanisms in the monosynaptic pathway ensure CA1 is tuned in such a way that what is reconstructed in deep layers of EC (ECd) corresponds to the original information sent by ECs. This tuning can be done at the moment of learning, when both the original information and its reconstruction are present. At the moment of prediction, the same comparison mechanism serves as a novelty proxy for switching between recall and learning regimes, like previously mentioned. A side effect of this slower learning mechanism is the involvement of CA1 in gathering statistics over multiple episodes (Schapiro et al., 2017), thus explaining why the activity of some hippocampal units like place cells reflect statistical regularities (Stachenfeld et al., 2017). Computer simulations reveal the importance of such a mechanism for discovering associations not directly experienced in for example transitive inference, namely the discovery that if A is associated to B and B is associated to C, A is indirectly associated to C (Schapiro et al., 2017). The role of episodic aggregation has also been described for building a kernel function in the context of episodic reinforcement learning (Gershman and Daw, 2017). Finally, these statistics might serve as an intermediate step for the transfer between episodic and cortical memory systems (Kumaran et al., 2016). We explain this in more detail in the next part.

3.2 Interfacing between the HPC and the cortex

A constant dialog between the HPC and the cortex is ensured by multiple pathways, including a major hub: the entorhinal cortex. This highly integrated cortical region receives inputs from most regions of the entire cortex in its superficial layers and buffers them in episodes to be sent to the HPC. Reciprocally, its deep layers receive replays from the HPC and distribute them back to the cortex. Other minor paths for the specific communication with the PFC will also be evoked below. While the precise description of these pathways is outside our scope, we would like to highlight a few key characteristics before addressing the computational role of this dialog.

Connectivity with the cortex is not homogeneous throughout the HPC. In humans, the HPC runs along an anterior-posterior longitudinal axis. Along this axis, genetic, electrophysiological, and anatomical differences underlie differences in functionality, and in particular, decreasing degrees of spatio-temporal abstraction (see (Strange et al., 2014) for a review). In the spatial domain, the gradient of abstraction translates into place cells being sensitive to larger portions of the environment in the anterior HPC compared to place cells located in the posterior HPC (Jung et al., 1994). More generally, it is hypothesized that on the one hand, more posterior parts of the HPC which are mostly connected to posterior (and sensory) parts of the cortex are concerned with specifics (details of a scene, precise location,...). On the other hand, more anterior parts which are mostly interacting with anterior parts of the cortex (dealing with cognitive and behavioural control but also emotional and motivational aspects) represent more abstract and general information (context, larger scale location, semantic information, mood...). Intermediate portions are thought to mediate interactions between the two ends of the longitudinal axis using intermediate representations (Strange et al., 2014). Organisation in trisynaptic and monosynaptic pathways is preserved all along, although the dentate gyrus is more prominent in posterior HPC while CA1-3 are more prominent in its anterior part (Malykhin et al., 2010). Hence, the same mechanisms for learning detailed perceptual state transitions in the posterior HPC are used for learning transitions between more spatially and temporally distal cognitive cues in the anterior HPC. Interestingly, this also constitutes a potential substrate for transitive inference, since items that are not associated at the posterior scale of functioning can potentially be associated at the anterior scale of temporal abstraction (Strange et al., 2014).

Let us now discuss candidate functions of hippocampo-cortical dialog, beginning with topics involving communication through the entorhinal cortex. Firstly, hippocampal pattern completion can be taken advantage of for the recall of missing information in the cortex (Rolls, 2013). It is obviously helpful in the noisy domain of perception, in which objects must be identified while exhibiting a constantly changing aspect, for example dependent on light conditions in the case of vision. As mentioned before, similar mechanisms are used with more abstract representations in intermediate and anterior regions of the HPC. Consequently, pattern completion might also be useful for retrieving more abstract information related to context, emotional states, and even behavioral plans, a requirement for prospective memory. Secondly, a fur-

ther function of the dialog is to use knowledge about state transitions stored in the HPC. This has been shown to be useful for predictive coding in the visual cortex (Hindy et al., 2016) and cognitive anticipation (Gershman and Daw, 2017). Thirdly, the complementary learning systems theory suggests that the HPC acts as a temporary store for memories to be transferred to the cortex (McClelland et al., 1995). More precisely, dense representation and slow learning are deployed for the cortex to learn high-level generalizations while avoiding catastrophic forgetting respectively, resembling implicit memory of classical deep learning models. The other side of the coin is that the cortex is not able to learn from a single exposure to a piece of information. Conversely, the HPC deploys fast learning and sparse representation for enabling one-shot acquisition. Since sparse encoding is not suitable for generalization, the HPC alone is not sufficient either. In fact, it is really the conjunction and interaction of these two systems that enable humans to learn fast and exploit acquired knowledge efficiently. During resting periods, the HPC is thought to "teach" the cortex by sending recently acquired information interleaved with more established knowledge. This process is called interleaved learning and is used to incorporate new information into cortical knowledge without interference (McClelland et al., 1995). Note that the aforementioned hypothesis regarding the complementarity of the monosynaptic and trisynaptic pathways was introduced as an analogous and intermediate step to this consolidation process (Schapiro et al., 2017). Indeed, intermediate statistics discovered in the monosynaptic pathway might preferentially be sent to the cortex. There would therefore be a trisynaptic-monosynaptic-cortical gradient of learning abstraction, in the perpendicular direction of the anterior-posterior axis of representational abstraction. To our knowledge, understanding the differential contributions of these two axes to functions like transitive inference remains an open and unaddressed question. As a concluding remark, the fact that the monosynaptic pathway might play a more dominant role in the anterior HPC and the trisynaptic pathway a more dominant role in the posterior one suggests that these two axes might not be completely orthogonal.

Concerning specific paths of communication between the PFC and the HPC, several principles must be mentioned. In short, a unidirectional loop of communication has been described (Eichenbaum, 2017). The anterior HPC is hypothesized to send contextual information about the retrieved episode to the PFC, which uses it to better select appropriate behavioral rules. Reciprocally, the PFC controls episodic recall and encoding, either directly by modulating ac-

tivity in the posterior HPC (Rajasethupathy et al., 2015) or indirectly by biasing the activity of the sensory and entorhinal cortices. As a result, hippocampal recall and its prospective code are oriented towards task-relevant episodes, whereas encoding and its retrospective code are modulated to control memory organization. The PFC would therefore need a model of what is stored in the HPC, a function that is sometimes referred to as metamemory. The direction of these information flows are presumably controlled by an additional communication pathway linking the PFC and the HPC through the thalamus, the mechanisms of which are still to be understood in more details. To better understand the crucial role of these recently discovered and yet to model interactions, let us consider the example task of deciding how to dress. While learning to wear warm clothes when perceiving rain and lighter clothes when perceiving sunlight is sufficient in some situations, elements of controlled episodic memory can help us making a better decision. The HPC might be able to recall and send information to the PFC concerning a weather forecast inconsistent with your current perception. The PFC could then predict that the weather is likely to change in the near future. It might in turn inhibit the automatic recall of an alternative but fictional weather forecast you saw in a movie. This situation illustrates how interactions between the HPC and the PFC harness memory of the past to make decisions in the present by predicting the future.

4 ROADMAP FOR THE FUTURE

In this paper, we have shown that important limitations in today's machine learning are due to a lack of explicit memory and we have reported that, in the brain, the HPC and the PFC play a fundamental role in creating and managing this memory. We have also explained that this form of memory plays a central role in concept learning, planning, reasoning and other major cognitive functions. Although this global view begins to be understood and exploited in machine learning, the specific role of episodic memory and its substrate in these functions is still not well understood and corresponding working models could be augmented with additional mechanisms.

We attempted to report results and hypotheses coming from different domains of neuroscience in order to provide an increasingly precise view of hippocampal circuitry and its relationship with posterior and prefrontal cortices. On the one hand, we think that the role of the HPC in supervising the sensory cortex to allow for consolidation without interference could be investigated in more details and give rise to better balanced neural networks capable of forming new conceptual knowledge online. On the other hand, even if several pieces of information are still missing, preliminary studies about the involvement of interactions between the HPC and the PFC in forming a prospective memory should also be considered for improving decision making algorithms. Finally, we believe that explicit cognition is a key ingredient for creating artificial agents capable of deliberating and explaining their actions.

REFERENCES

Alexandre, F. (2016). Beyond machine learning: Autonomous learning. In *8th International Conference on Neural Computation Theory and Applications (NCTA)*, pages 97–101.

Bock, P. and Alexandre, F. (2019). [Re] The Wisconsin Card Sorting Test: Theoretical analysis and modeling in a neuronal network. *ReScience C*, 5(3):#3.

Boraud, T., Leblois, A., and Rougier, N. P. (2018). A natural history of skills. *Progress in Neurobiology*, 171:114–124.

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., and Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in cognitive sciences*, 23(5):408–422.

Braver, T. S. and Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working. In *MIT Press. Making Working Memory Work*, pages 551–581.

Dagar, S., Alexandre, F., and Rougier, N. (2021). Deciphering the contributions of episodic and working memories in increasingly complex decision tasks. In *Proceedings of the International Joint Conference on Neural Networks, IJCNN 18-22 July 2021, Shenzhen, China*. IEEE.

Domenech, P. and Koechlin, E. (2015). Executive control and decision-making in the prefrontal cortex. *Current Opinion in Behavioral Sciences*, 1:101–106.

Duncan, K., Ketz, N., Inati, S. J., and Davachi, L. (2012). Evidence for area ca1 as a match/mismatch detector: A high-resolution fmri study of the human hippocampus. *Hippocampus*, 22(3):389–398.

Eichenbaum, H. (2017). Memory: Organization and Control. *Annual Review of Psychology*, 68(1):19–45.

Fuster, J. (1989). *The prefontal cortex. Anatomy, physiology and neurophysiology of the frontal lobe*. Raven Press, New-York.

Gershman, S. J. and Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual review of psychology*, 68:101–128.

Hamid, O. H. and Braun, J. (2017). Reinforcement learning and attractor neural network models of associative

learning. In *International Joint Conference on Computational Intelligence*, pages 327–349. Springer.

Hindy, N. C., Ng, F. Y., and Turk-Browne, N. B. (2016). Linking pattern completion in the hippocampus to predictive coding in visual cortex. *Nature neuroscience*, 19(5):665–667.

Jensen, O. and Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in neurosciences*, 28(2):67–72.

Joel, D., Niv, Y., and Ruppin, E. (2002). Actor–critic models of the basal ganglia: new anatomical and computational perspectives. *Neural Networks*, 15(4-6):535–547.

Jung, M. W., Wiener, S. I., and McNaughton, B. L. (1994). Comparison of spatial firing characteristics of units in dorsal and ventral hippocampus of the rat. *Journal of Neuroscience*, 14(12):7347–7356.

Kassab, R. and Alexandre, F. (2018). Pattern separation in the hippocampus: distinct circuits under different conditions. *Brain Structure & Function*, 223(6):2785–2808.

Koechlin, E. (2014). An evolutionary computational theory of prefrontal executive function in decision-making. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 369(1655):20130474+.

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

Lisman, J. E. and Otmakhova, N. A. (2001). Storage, recall, and novelty detection of sequences by the hippocampus: elaborating on the socratic model to account for normal and aberrant effects of dopamine. *Hippocampus*, 11(5):551–568.

Malykhin, N., Lebel, R. M., Coupland, N., Wilman, A. H., and Carter, R. (2010). In vivo quantification of hippocampal subfields using 4.7 t fast spin echo imaging. *Neuroimage*, 49(2):1224–1230.

McClelland, J. L., McNaughton, B. L., and 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):419–457.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, 24:167–202.

O'Reilly, R. C., Noelle, D. C., Braver, T. S., and Cohen, J. D. (2002). Prefrontal cortex and dynamic categorization tasks: representational organization and neuromodulatory control. *Cereb Cortex*, 12(3):246–257.

Rajasethupathy, P., Sankaran, S., Marshel, J. H., Kim, C. K., Ferenczi, E., Lee, S. Y., Berndt, A., Ramakrishnan, C., Jaffe, A., Lo, M., et al. (2015). Projections from neocortex mediate top-down control of memory retrieval. *Nature*, 526(7575):653–659.

Ritter, S., Wang, J., Kurth-Nelson, Z., and Botvinick, M. (2018). Episodic Control as Meta-Reinforcement Learning. preprint, Neuroscience.

Rolls, E. (2013). The mechanisms for pattern completion and pattern separation in the hippocampus. *Frontiers in systems neuroscience*, 7:74.

Schacter, D. L., Addis, D. R., and Buckner, R. L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8(9):657–661.

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., and Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711):20160049.

Simons, J. S. and Spiers, H. J. (2003). Prefrontal and medial temporal lobe interactions in long-term memory. *Nature Reviews Neuroscience*, 4(8):637–648.

Stachenfeld, K. L., Botvinick, M. M., and Gershman, S. J. (2017). The hippocampus as a predictive map. *Nature neuroscience*, 20(11):1643.

Strange, B. A., Witter, M. P., Lein, E. S., and Moser, E. I. (2014). Functional organization of the hippocampal longitudinal axis. *Nature Reviews Neuroscience*, 15(10):655–669.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., and Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, 21(6):860–868. Number: 6 Publisher: Nature Publishing Group.

Zilli, E. A. and Hasselmo, M. E. (2008). Modeling the role of working memory and episodic memory in behavioral tasks. *Hippocampus*, 18(2):193–209.

Ólafsdóttir, H. F., Bush, D., and Barry, C. (2018). The Role of Hippocampal Replay in Memory and Planning. *Current Biology*, 28(1):R37–R50.