# The neural and cognitive architecture for learning from a small sample

Aurelio Cortese[1], Benedetto De Martino[2,3] and Mitsuo Kawato[1,4]

Artificial intelligence algorithms are capable of fantastic exploits, yet they are still grossly inefficient compared with the brain's ability to learn from few exemplars or solve problems that have not been explicitly defined. What is the secret that the evolution of human intelligence has unlocked? Generalization is one answer, but there is more to it. The brain does not directly solve difficult problems, it is able to recast them into new and more tractable problems. Here, we propose a model whereby higher cognitive functions profoundly interact with reinforcement learning to drastically reduce the degrees of freedom of the search space, simplifying complex problems, and fostering more efficient learning.

**Addresses**
[1] Computational Neuroscience Laboratories, ATR Institute International, Kyoto, Japan
[2] Institute of Cognitive Neuroscience, University College of London, Alexandra House, 17-19 Queen Square, London WC1N 3AR, United Kingdom
[3] Wellcome Centre for Human Neuroimaging, University College London, WC1N 3BG London, United Kingdom
[4] RIKEN Center for Advanced Intelligence Project, ATR Institute International, Kyoto, Japan

## Introduction

Artificial Intelligence (AI) has come a long way since the summer of 1956, when it was first envisaged at the Dartmouth Summer Research Project on Artificial Intelligence. In the last ten years, we have witnessed how the principles of supervised and reinforcement learning, when embedded in neural networks composed of many hidden layers ('deep neural networks', or 'DNN'), can reach – and often surpass – human-level performances in visual object recognition, and in playing video-games and GO [1–3]. Despite DNN's massive computational capabilities, there are two aspects that temper these accomplishments: first, the number of training samples required to reach acceptable performances is huge - tens or hundreds of millions; second, these architectures show a limited ability to generalize to new tasks/settings that were not encountered during training. These limitations become largely evident in motor control as shown by the clumsy behavior of humanoid robots in the DARPA robotic challenge [4].

A second avenue of exciting progress in AI has come from probabilistic machine learning (a.k.a. Bayesian machine learning) [5], where agents can achieve impressive performance on one-shot learning or using a limited amount of examples [6,7]. This probabilistic approach resonates well with intuitive theories of human cognitive development and inductive reasoning [8]. The learning algorithm tries to find among all possible models the one that best explains the data (or, by extension, infer what causes the reality perceived through the sensorium). This approach, while conceptually appealing, is unlikely to provide a realistic model of how the brain operates. The main issue is that fully probabilistic inference might work well in simple and well constrained conditions, but becomes quickly computationally intractable for more complex and unconstrained scenarios. To exacerbate the problem, in order to function efficiently, probabilistic programs have to be endowed with ad-hoc definitions of the necessary representations [5]. Strictly speaking, in Bayesian inference we usually do not have a principled way to select initial priors. Generalization is thus limited to the class of problems, for which the program was designed for [6,7].

A considerable hurdle for artificial agents concerns generalization; how can machine learning algorithms deal with new and never-experienced scenarios? Humans and animals can easily and appropriately respond to new scenarios, mostly transferring knowledge acquired in loosely related contexts. What are the brain mechanisms that enable the human brain with its remarkable generalization capacity? Plain reinforcement learning is too slow, and hierarchical architectures [9], albeit ameliorating the algorithm by sub-dividing learning among multiple systems and meta-variables [10,11•], remain dependent on the need for ad-hoc definitions. Here, we suggest that brains do not simply solve supervised classification problems but transform them into different – and more tractable – problems. We propose that an adaptive role of higher cognition is to allow precisely this transformation to take place.

More specifically, we propose a model of how higher cognition is able to simultaneously operate the dimensionality

reduction and feature selection processes necessary for simplifying complex problems. Adjusting the degree of synchronization between neurons has been suggested as one possible way to control the degrees-of freedom of a neural system [12]. We draw on similarities with simulated annealing, exploring how different frequency modes of brain dynamics serve as inherent implementation channels to reduce degrees-of-freedom and reach optimal solutions.
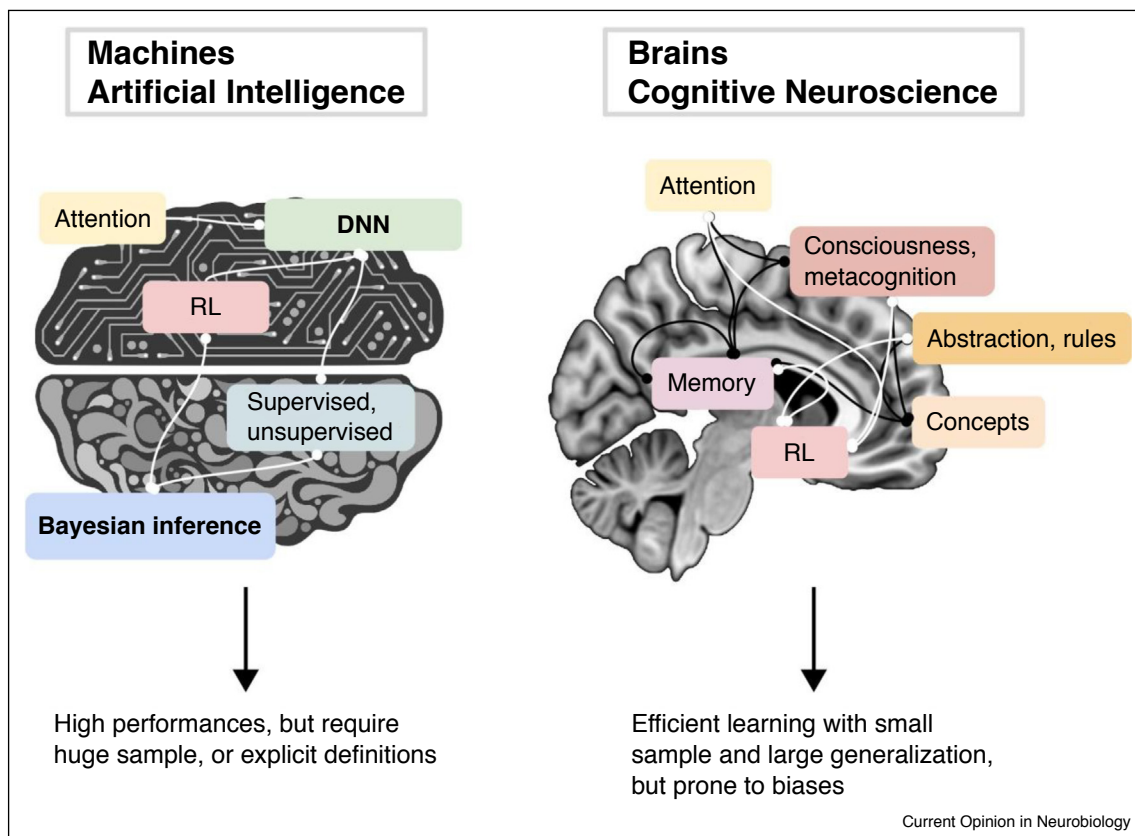
## Computational advantages of higher cognitive functions in learning

Statistical learning theory of singular problems demonstrates that the generalization error is given by dividing the degrees-of-freedom ($d$) of the search space by the number of training sample ($n$): $e \propto d/(2n)$ [13,14•]. If brains ($d \sim 10^{11}$ neurons) need to solve arbitrary classification problems utilizing only a few hundred learning samples ($n \sim 10^2$), the generalization error would become huge, at least $10^{11}/10^2 = 10^9$. We postulate that brains transform these intractable learning problems into more feasible reinforcement learning problems with small degrees of freedom while being guided by reward and penalty. Higher cognitive functions such as attention, memory, concept formation, and metacognition might find low-dimensional manifolds

of meta-representations that are essential for learning from a small sample (Figure 1). Here, we will briefly review findings from attention, memory, concept formation, and metacognition, focusing on their role in facilitating learning. We are aware that these are vast and active areas of investigation and that it would be hard to do justice to all the relevant work that has been done. We have, therefore, decided to provide a snapshot of properties, modules and architectures that we believe are particularly relevant to inspire the development of new AI architectures. It is important to recognize that a recent work in the field of machine learning has also started to incorporate some of the intuitions discussed here.

Attention is the ability to direct computing resources toward relevant dimensions (stimulus attributes, spatial location, etc.) for focal processing, acting as a filter to amplify relevant information while dampening background clutter [15]. But how does an agent learn what to attend? Rewards and punishments serve to constrain attentional focus [16], and attending to specific features rather than to the whole improves versatility [17]. Essentially, we learn *what* to attend to at the same time as we are paying attention to what we are learning [18,19•]. In

**Figure 1**



General schematic, solutions to complex problems in artificial intelligence and nature (brains). The key insight discussed here is that higher cognitive functions continuously interact between them and with reinforcement learning to drive generalization and learning from small sample.

machine learning, a useful and efficient way to use attention mechanisms is to decompose tasks or questions into a series of simpler operations [20], or target specific parts of a query (e.g. particular words in a sentence) [21].

An intelligent agent must also be able to remember or even sometime forget past events. Accordingly, episodic memory plays a special role in goal-directed behavior and learning [22]. Reality is statistically structured however, and forms of gist-like memory can enhance reinforcement learning [23••]. Schematic memory still depends on episodes; it is by virtue of statistics over individual traces that summaries can be created. Not surprisingly, both persistence (remembering) and transience (forgetting) are essential ingredients to optimize decision-making [24]. Human-like memory processes are very different from what is usually considered in AI agents, where memory is often deterministic and non-sparse as well as storing all information. Linking neural networks to external buffer memory resources already produces impressive learning capabilities, unattainable by classic neural networks architectures [25,26•]. The development of predictive memory architectures, where memory formation itself is guided by a process of predictive inference [27], is one step toward systems storing only relevant information.

Concepts *are* abstractions, closely intertwined with schematic memories. Concepts can be created almost at will, and a key aspect is that they can be connected, creating conceptual maps [28]. Being highly hierarchical and compositional, more abstract concepts can be formed from existing ones. New concepts or conceptual maps can emerge from learning, but can also direct subsequent learning [29]. Concepts share obvious links with memory in their ability to represent schematized information, but work in AI has not yet capitalized on this approach. Conceptual representations in AI are currently restricted to simple visual domain examples that make use of the principles of hierarchy and compositionality [30].

Self-monitoring processes, a more abstract class of cognitive functions, can encompass much richer representations. The ability to monitor one's thoughts is referred to as metacognition, and is linked to the psychological construct of confidence, that is, how good an agent is at keeping track of the probability of a choice being correct [31,32]. This aspect is very important for AI since it dovetails with a broad range of phenomena such as error monitoring and reality checking [33]. Of particular relevance to AI systems is the ability to explicitly track the evolution of the level of self-knowledge, which might provide biological agents with significant advantages when interacting with their environment [34–38]. Although metacognition and consciousness are intimately related, the question of what is the computational advantage of consciousness itself remains currently unanswered. Consciousness could represent the selection of information for global broadcasting within the system,

making it flexibly available for local (and distant) computational units [33]. In machine learning consciousness could also be interpreted as a powerful constraint on low-dimensional representations [39•]. Earlier efforts suggest that some forms of self-monitoring are computationally simple and can directly arise even in two-layer attractor networks [40]. Generative adversarial networks (GANs) are an exciting development in this direction: a generative model captures the data distribution, and a discriminative model, akin to metacognition, operates a reality check on new samples [41].
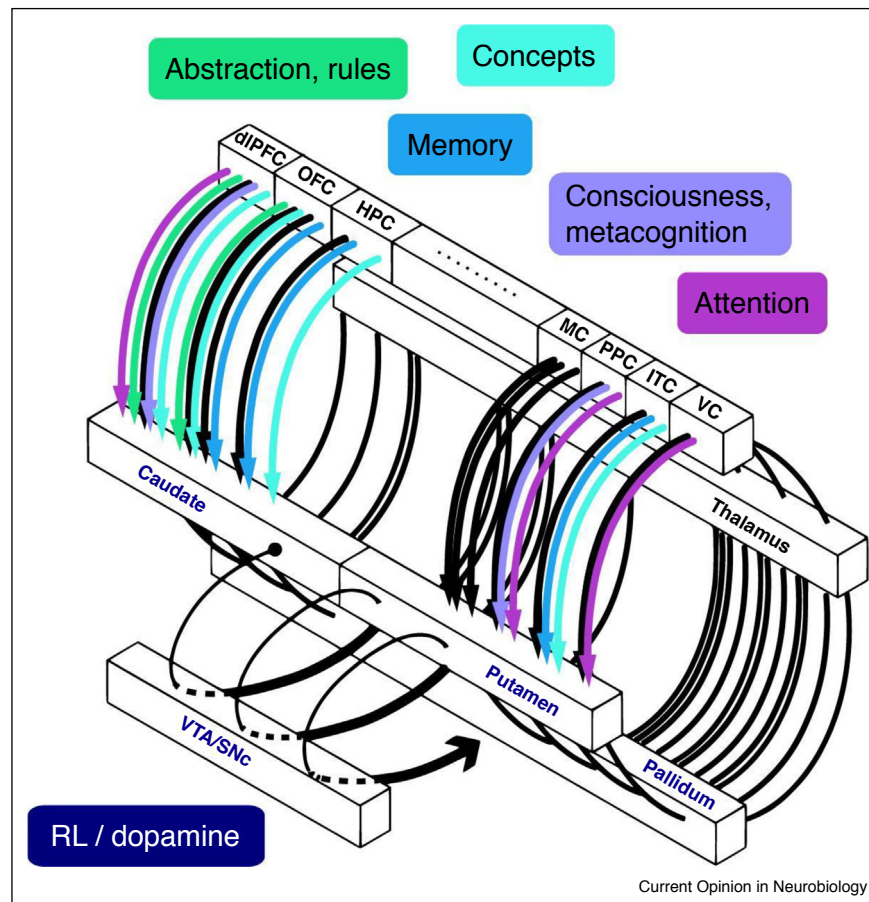
## Neural implementation of high level cognitive architecture

To generate solutions leading to efficient learning and flexible behaviors, nature had to solve a number of physical constraints. The brain cannot be equipped with *ad hoc* representations for every possible event in the world, since the horizon of possible states is practically infinite; moreover, it does not have unlimited computing resources. Understanding how the brain has overcome these constraints may be inspirational for developing new AI, yet it is important to keep in mind that some biological constraints (e.g. positive neuronal firing rates) may be bypassed by *in-silico* intelligent systems.

In its most basic interpretation, solving complex problems for the brain accounts to finding the relevant (hidden) states for RL. One solution to accelerate the search for hidden states is to capitalize on the brain's massively parallelized neural circuit architecture. Parallel searches are instantiated in multiple recurrent circuits linking basal ganglia with the cortex (Figure 2). These recurrent circuits effectively are information-transmitting loops (bi-directional and closed neural-circuit connections between basal ganglia and cortex): they can carry task-dependent explicit representations (stimuli, goals, etc.), abstract summaries, reward prediction errors (RPEs), and predicted states. Although parallel loops carry heterogeneous information, they do not function independently from each other. Rather, loops formed by sparse neural populations continuously interact at the synaptic level through cooperation and competition. Excitatory interactions (cooperation) appear between loops with similar, inclusive and related representations. In contrast, inhibitory interactions (competition) develop between loops with exclusive, different or unrelated representations. Because of the dynamic nature of the neural networks comprising these excitatory and inhibitory interactions, a winner-take-all scenario emerges [42,43]. That is, only the loop with the 'best' representation survives while other loops are suppressed. Here 'best' means the loop associated with the representation that minimizes RPE. Therefore, the selection of the best loop essentially corresponds to the automatic selection of relevant states for RL.

Excitatory and inhibitory interactions can occur virtually anywhere in the brain. However, the basal ganglia, a group of structures located deep within the cerebral

**Figure 2**



Winner-take-all parallel computing takes place in loops spanning basal ganglia and neocortex. Excitatory and inhibitory interactions between multiple loops formed by the basal ganglia, thalamus and cerebral cortical regions implement winner-take-all computations. The loop with the best representation for reinforcement learning and minimum reward prediction errors thus wins and all other loops are suppressed. These winner-take-all computations implement dimension reduction and feature selection accelerated by high cognitive functions as follows. Attention executes feature selection rather than dimension reduction, with relatively low abstraction. Episodic memory, among different kinds of memories, represents feature selection in the time domain, and its abstraction level is relatively low. Conceptualization executes dimension reduction rather than feature selection, and its abstraction level is high. Metacognition does both dimension reduction and feature selection and its abstraction level is very high. Consciousness has the highest abstraction level and results in pure dimension reduction. dlPFC, dorsolateral prefrontal cortex; OFC, orbitofrontal cortex; HPC, hippocampal formation; MC, motor cortex; PPC, posterior parietal cortex; ITC, inferior temporal cortex; VC, visual cortex; VTA/SNc, ventral tegmental area/substantia nigra; RL, reinforcement learning. Figure modified from Haruno and Kawato [67].

hemispheres, should play the most important role in these synaptic interactions for the following reasons: (1) multiple inhibitions and direct, indirect, as well as hyperdirect pathways link basal ganglia to cortex; therefore, winner-take-all computations can best be implemented in basal ganglia [43,44]. (2) RPEs are largely computed in basal ganglia [45], making these nuclei the ideal focal point for the comparison and selection of loops carrying the smallest RPEs.

So far, we have discussed a relatively simplified model that is amenable to clearly delineate the theory. The reality of the brain is nevertheless more intricate. Several brain areas are

likely interacting to orchestrate an efficient search and ensure convergence to task-relevant low-dimensional manifolds. Prefrontal, sensorimotor, hippocampal cortices, as well as cerebellum, thalamus, and basal ganglia all share recurrent connections. Above this automatic machinery, what is the role of higher cognitive functions? How can they further accelerate learning computations? Metacognition, attention and memory synchronize abstract representations in prefrontal cortex (PFC) or hippocampal formation (HPC) with concrete representations in sensorimotor areas. Recurrent connections between these cortico-striatal regions connect reinforcement learning mechanisms with representational and abstraction engines that makes for an ideal candidate circuitry (Figure 2).

Dopamine inputs to the HPC ensure that learned or partially learned rules are conceptualized and stored in memory [46]. But HPC function stretches far beyond memories, and one influential idea is that it plays a key role in building cognitive maps for spatial [47] and conceptual navigation [28]. HPC neurons functionality probably extends so that they take the role of predictive units extracting structure and low-dimensional bases of the world [48,49•]. These discoveries are in line with a recent proposal that during decision making, inferable state-to-state transitions represented in the cortex keep track of the evolving hidden space to accelerate learning [50,51]. More specifically, the PFC is thought to hold an exclusive position along the hierarchy of representations as the substrate forging meta-representations [11•,31] and abstraction processes [10,52,53]. Furthermore, the PFC oscillatory frequencies act as mediators of abstraction: the synchronization frequency helps shuffle the abstraction level encoded in different regions of the PFC [54••]. In fact, this extends well beyond the PFC; oscillatory frequencies effectively form communication channels throughout the brain [54••,55–57].

At the neuronal level, one way to discover correlations between meta-representations and RPEs is to characterize these representations as statistical distributions in neural population activities. The cortex could perform probabilistic inference on such distributions either by sampling over neural populations [58], or by weighting correlations between neurons [59].
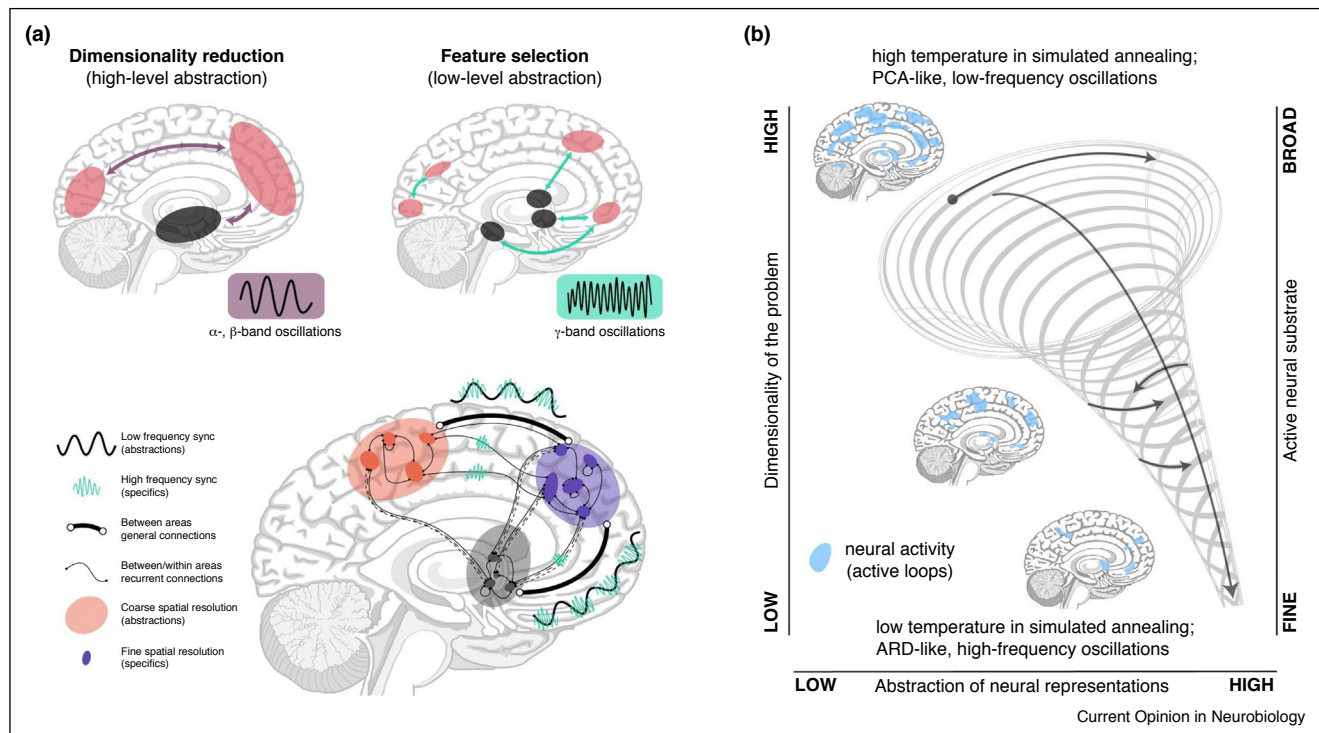
Learning new problems should invariably start with a consistent scenario: ignition of myriads of parallel loops resulting in widespread neural activity over extensive cortico-striatal networks. The selection of RL states starts with broad sweeps to evolve in a fine search. Initially, broad brain regions are equally activated and participate in the search. Next, dimension reduction and feature selection begin, in parallel, to drop-out less activated loops, a process accelerated by higher cognition such as attention, memory, metacognition (Figure 2). In the end, only a small number of loops will remain activated and neural activity should be concentrated to the few cortical locations carrying the most relevant representations and the corresponding parts of basal ganglia. A useful analogy for this search from broad to fine neural substrates (and representations) is simulated annealing or Gibbs sampling, optimization techniques to approximate global solutions in large search spaces [60,61]. For example, annealing starts by first using high temperatures causing large changes in the objective function, then iteratively descending to lower temperatures causing ever smaller rearrangements — until convergence.

Furthermore, to extend this analogy, high temperatures are depicted as a form of dimension reduction, while low temperatures as akin to feature selection. We suggest that

dimension reduction relates to abstraction, operating at low oscillatory frequency modes with low spatial resolution and using large neural populations, while feature selection relates to specific content utilizing high frequency modes and sparse neural ensembles (Figure 3a). By frequency modes here we mean specific bands in neural oscillation frequencies – that is, the frequency at which neuronal populations show synchronized activity. Low frequency synchronization delineates the horizon of relevant *dimensions* so that high frequency-based *feature* selection can happen. To note, the direction of search from broad to fine holds in terms of brain areas/neural networks involved; but in terms of abstraction, the search directionality of the model discussed here is unconstrained. The key intuition is that in the brain the processes of dimension reduction and feature selection should take place *simultaneously*, accelerating interaction and winner-take-all convergence of loop drop-out (Figure 3b). A recent work has elegantly linked the brain's structural connectivity (particularly the thalamo-cortical system) with neural activity patterns and dynamics, providing a formal basis for harmonic patterns of certain frequencies [62••]. The authors of this work demonstrate that these connectome-specific harmonics patterns self-organize through the interplay of neural excitation and inhibition in coupled dynamical systems.

We can now delineate how cognitive functions may affect and expedite learning processes. We have proposed a system composed of massively parallelized modules centered around an RL machinery, where communication frequency determines the abstraction level of representations, and where cognitive functions have the ability to synchronize representations at different abstraction levels. At the very outset the search is characterized by activity over broad areas of the brain, but the system is typified by low abstraction and little synchronization. RPEs may be tied to any aspects of the task, with most RPEs being unspecific and irrelevant. Conceptualization or integration of (task, instruction) rules, together with visuospatial attention [63], can generate the first optimal low-frequency modes, leading an initial dimensionality reduction to define the search horizon where specific features can be selected. Attention and episodic memory also play important roles from the initial stages for the selection of relevant features (i.e. loops) [64••] through synchronization in high spatiotemporal-frequency channels [55,65], nested within low-frequency modes. As learning progresses, high-frequency feature-specific RPEs become predominant [64••,66], and the number of activated loops greatly decreases. The abstraction level is maintained high because feature-specific RPEs can be represented in summarized fashion, hence further reducing the dimensionality and complexity of the problem. Hidden states in reinforcement learning can now be discovered more readily because the search domain has shrunk. Importantly, the degree of certainty or uncertainty on neural meta-representations can provide a fast

**Figure 3**



Different frequency modes in synchronization of neural activity represent broad and fine dimension reduction and feature selection. **(a)** Neural oscillations are rhythmic repetitions of patterned neural activity. These oscillations can cover frequencies from approximately 0.05 Hz (low spectrum) to 500 Hz (high spectrum). While low-frequency spatio-temporal modes contain many neurons and connections, high frequency modes contain small numbers of neurons and connections. Nonlinear dynamic interactions between low and high-frequency modes provide the computational means for fast parallel search of the optimal metarepresentation, corresponding smallest reward prediction errors (RPE), neurons and connections. When a low frequency mode is selected first, all high frequency modes contained within it are generally activated because low and high frequency modes share common neurons and connections. Among the activated high-frequency modes, those with the highest correlations between meta-representations and RPE are further activated, and an optimal mode is thus selected. Consequently, dimension reduction with low-frequency synchronization and feature selection with high-frequency synchronization proceed together by closely interacting. **(b)** Illustration of the search implemented by the model. Low frequency mode corresponds to dimension reduction such as principal component analysis (PCA) and high temperature in annealing. High frequency mode corresponds to feature selection such as L1-norm regularization or automatic relevance determination, and low temperature in annealing. Real or simulated annealing takes long time but brains cannot afford that. There exists no external control of temperature in the proposed interaction between different modes; nonlinear brain dynamics directly implements simulated annealing. With low signal to noise ratio, a common issue in most learning problems, first an optimal low-frequency mode is activated because it contains many areas, neurons and connections. This increases the chances of correlation computations surviving high noise conditions. Then, high-frequency modes contained in it are generally activated, and correlations can be more reliably computed by constrained domains and general excitatory inputs to them. The selection of the optimal high-frequency mode can be executed more robustly. This interaction between low and high-frequency modes roughly implements annealing. The key difference with annealing though is that activation and inhibition of all loops start simultaneously in the brain. But convergence starts from large dimension and low abstraction (top left, black dot on top arrow) and proceeds to small dimension and high abstraction (bottom right). This is analogous to simulated annealing (direction of oblique arrow).

track to which states are relevant or irrelevant in reinforcement learning, by virtue of metacognition self-monitoring properties [33]. Finally, consciousness itself may lead learning to the highest level of abstraction. Conscious representation of the relevant states or decision policies can be interpreted as a maximally abstract summary, a tensor with very low dimensionality that nevertheless carries all the fundamental information [39•]. These meta-representation vectors are extremely useful for generalization because they can be easily applied to new, previously unexperienced, problems.

## Conclusions

Fruitful interactions between neuroscience and AI have opened up a new exciting era beyond DNN, which require huge training samples. Brains utilize higher cognitive functions such as attention, memory, concept formation, and metacognition to transform seemingly intractable supervised learning problems with astronomical degrees-of-freedom state spaces and small samples, into reasonable reinforcement learning problems within a low-dimension meta-representation manifold.

We postulated that the neuronal mechanism implementing this transformation of computational problems is likely comprises parallel search of low-dimensional meta-representations via synchronization of multiple loops formed by the cerebral cortex, HPC and basal ganglia. Brain connectivity and nonlinear neural dynamics provide harmonic modes spanning from low to high spatiotemporal frequencies. Interactions between different modes may provide dimension reduction and feature selection analogous to simulated annealing, albeit much faster. That is, low frequency mode could allow for dimension reduction analogous to high temperature in annealing, while high frequency modes could select a small number of features analogous to low temperature. Furthermore, real-time interactions between high and low frequency modes may enable fast parallel searches to quickly determine the reinforcement learning search domain. Attention and episodic memory are presumed mechanisms operating feature selection, while conceptualization mainly takes the form of dimension reduction. Discovery of relevant hidden states may be greatly accelerated by metacognition through synchronization of meta-representations.

Taken together these cognitive modules, acquired over millions of years by natural selection, might inspire a new generation of AI architectures that will take us one step closer to human level intelligence.

## Conflict of interest statement
Nothing declared.

## Acknowledgements

## References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. He K, Zhang X, Ren S, Sun J: **Delving deep into rectifiers: surpassing human-level performance on imagenet classification**. In *Proceedings of the IEEE International Conference on Computer Vision*. 2015:1026-1034.

2. LeCun Y, Bengio Y, Hinton G: **Deep learning**. *Nature* 2015, **521**:436-444.

3. Silver D, Huang A, Maddison C, Guez A, Sifre L, Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M *et al.*: **Mastering the game of Go with deep neural networks and tree search**. *Nature* 2016, **529**:484-489.

4. Atkeson CG, Benzun PWB, Banerjee N, Berenson D, Bove CP, Cui X, DeDonato M, Du R, Feng S, Franklin P *et al.*: **What happened at the DARPA robotics challenge finals**. In *The DARPA Robotics Challenge Finals: Humanoid Robots to the Rescue*. Edited by Spenko M, Buerger S, Iagnemma K. Springer International Publishing; 2018:667-684.

5. Ghahramani Z: **Probabilistic machine learning and artificial intelligence**. *Nature* 2015, **521**:452-459.

6. Lake B, Salakhutdinov R, Tenenbaum J: **Human-level concept learning through probabilistic program induction**. *Science* 2015, **350**:1332-1338.

7. George D, Lehrach W, Kansky K, Lázaro-Gredilla M, Laan C, Marthi B, Lou X, Meng Z, Liu Y, Wang H *et al.*: **A generative vision model that trains with high data efficiency and breaks text-based CAPTCHAs**. *Science* 2017, **358**:eaag2612.

8. Lake BM, Ullman TD, Tenenbaum JB, Gershman S: **Building machines that learn and think like people**. *Behav Brain Sci* 2017, **40**:e253.

9. Kawato M, Samejima K: **Efficient reinforcement learning: computational theories, neuroscience and robotics**. *Curr Opin Neurobiol* 2007, **17**:205-212.

10. Buschman TJ, Miller EK: **Goal-direction and top-down control**. *Philos Trans R Soc Lond B Biol Sci* 2014, **369** 20130471–20130471.

11. Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H,
•   Leibo JZ, Hassabis D, Botvinick M: **Prefrontal cortex as a meta-reinforcement learning system**. *Nat Neurosci* 2018, **21**:860-868 http://dx.doi.org/10.1038/s41593-018-0147-8.
Stems from previous intuitions that basal ganglia and prefrontal cortex are two independent learning system. The paper proposes that dopamine from the basal ganglia learning system is used to train an independent, slower but more general and efficient (meta-)learning system in the prefrontal cortex.

12. Tokuda IT, Hoang H, Kawato M: **New insights into olivo-cerebellar circuits for learning from a small training sample**. *Curr Opin Neurobiol* 2017, **46**:58-67.

13. Watanabe S: *Algebraic Geometry and Statistical Learning Theory*. Cambridge University Press; 2009.

14. Yamazaki K: **Asymptotic accuracy of distribution-based
•   estimation of latent variables**. *J Mach Learn Res* 2014, **15**:3541-3562.
This is a theoretical paper that introduces a formal framework for error bounds in statistical learning theory. The author dealt with the generalization error of latent variables in singular cases of machine learning.

15. Xu K, Ba J, Kiros R, Cho K, Courville A, Salakhutdinov R, Zemel R, Bengio Y: **Show, attend and tell: neural image caption generation with visual attention**. *arXiv [csLG]* 2015. International conference on machine learning, 2048-2057.

16. Lee J, Shomstein S: **The differential effects of reward on space- and object-based attentional allocation**. *J Neurosci* 2013, **33**:10625-10633.

17. Farashahi S, Rowe K, Aslami Z, Lee D, Soltani A: **Feature-based learning improves adaptability without compromising precision**. *Nat Commun* 2017, **8**:1768.

18. Yang SC-H, Wolpert DM, Lengyel M: **Theoretical perspectives on active sensing**. *Curr Opin Behav Sci* 2016, **11**:100-108.

19. Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y: **Dynamic
•   interaction between reinforcement learning and attention in multidimensional environments**. *Neuron* 2017, **93**:451-463.
The authors used model-based fMRI and eye tracking to follow trial-by-trial fluctuations in attention while participants performed a decision-making task in a multidimensional environment. Results indicate a bidirectional interaction between attention and learning: attention constraints learning to relevant dimensions, while the question of 'what to attend' is driven by reinforcement learning.

20. Hudson DA, Manning CD: **Compositional attention networks for machine reasoning**. *arXiv* 2018. 1803.03067.

21. Luong M-T, Pham H, Manning CD: **Effective approaches to attention-based neural machine translation**. *arXiv* 2015. 1508.04025.

22. Lengyel M, Dayan P: **Hippocampal contributions to control: the third way**. In *Advances in Neural Information Processing Systems*

*20.* Edited by Platt JC, Koller D, Singer Y, Roweis ST. Curran Associates, Inc.; 2008:889-896.

23. Santoro A, Frankland P, Richards B: **Memory transformation**
•• **enhances reinforcement learning in dynamic environments**. *J Neurosci* 2016, **36**:12228-12242.
This modeling study shows that switching from detailed episodic memory to more generalized and schematic memory is a useful strategy for reinforcement learning in changing environments. Using gist-like memories optimizes behavior over several timescales.

24. Richards B, Frankland P: **The persistence and transience of memory**. *Neuron* 2017, **94**:1071-1084.

25. Graves A, Wayne G, Danihelka I: **Neural turing machines**. *arXiv [csNE]* 2014. 1410.5401.

26. Graves A, Wayne G, Reynolds M, Harley T, Danihelka I, Grabska-
• Barwińska A, Colmenarejo SG, Grefenstette E, Ramalho T, Agapiou J *et al.*: **Hybrid computing using a neural network with dynamic external memory**. *Nature* 2016, **538**:471-476.
Building directly from neural turing machines, this paper introduces a machine learning model that can read from and write to an external memory buffer (like a standard computer), but learns to do so with data. The proposed architecture shows improved ability to solve complex and structured tasks.

27. Wayne G, Hung C-C, Amos D, Mirza M, Ahuja A, Grabska-Barwinska A, Rae J, Mirowski P, Leibo JZ, Santoro A *et al.*: **Unsupervised predictive memory in a goal-directed agent**. *arXiv [csLG]* 2018. 1803.10760.

28. Constantinescu AO, O'reilly JX, Behrens TEJ: **Organizing conceptual knowledge in humans with a gridlike code**. *Science* 2016, **352**:1464-1468.

29. Mack ML, Love BC, Preston AR: **Dynamic updating of hippocampal object representations reflects new conceptual knowledge**. *Proc Natl Acad Sci U S A* 2016, **113**:13203-13208.

30. Higgins I, Sonnerat N, Matthey L, Pal A, Burgess CP, Botvinick M, Hassabis D, Lerchner A: **SCAN: learning abstract hierarchical compositional visual concepts**. *arXiv [statML]* 2017. 1707.03389.

31. Cortese A, Amano K, Koizumi A, Kawato M, Lau H: **Multivoxel neurofeedback selectively modulates confidence without changing perceptual performance**. *Nat Commun* 2016, **7**:13669.

32. Pouget A, Drugowitsch J, Kepecs A: **Confidence and certainty: distinct probabilistic quantities for different goals**. *Nat Neurosci* 2016, **19**:366-374.

33. Dehaene S, Lau H, Kouider S: **What is consciousness, and could machines have it?** *Science* 2017, **358**:486-492.

34. Friston K, Lin M, Frith C, Pezzulo G, Hobson J, Ondobaka S: **Active inference, curiosity and insight**. *Neural Comput* 2017, **29**:2633-2683.

35. Nassar MR, Wilson RC, Heasly B, Gold JI: **An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment**. *J Neurosci* 2010, **30**:12366-12378.

36. McGuire JT, Nassar MR, Gold JI, Kable JW: **Functionally dissociable influences on learning rate in a dynamic environment**. *Neuron* 2014, **84**:870-881.

37. Vaghi MM, Luyckx F, Sule A, Fineberg NA, Robbins TW, De Martino B: **Compulsivity reveals a novel dissociation between action and confidence**. *Neuron* 2017, **96**:348-354.e4.

38. Guggenmos M, Wilbertz G, Hebart M, Sterzer P: **Mesolimbic confidence signals guide perceptual learning in the absence of external feedback**. *eLife* 2016, **5**.

39. Bengio Y: **The consciousness prior**. *arXiv [csLG]* 2017.
• 1709.08568.
A formal proposal to model consciousness in machine learning. The approach is simple and amenable to current systems, considering consciousness as a vector containing simple statements drawn from summaries of the feature space.

40. Insabato A, Pannunzi M, Rolls ET, Deco G: **Confidence-related decision making**. *J Neurophysiol* 2010, **104**:539-547.

41. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y: **Generative adversarial nets**. *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc.; 2014.

42. Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE: **Adaptive mixtures of local experts**. *Neural Comput* 1991, **3**:79-87.

43. Grossberg S: **Contour enhancement, short term memory, and constancies in reverberating neural networks**. *Stud Appl Math* 1973, **52**:213-257.

44. Graybiel AM, Aosaki T, Flaherty AW, Kimura M: **The basal ganglia and adaptive motor control**. *Science* 1994, **265**.

45. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward**. *Science* 1997, **275**:1593-1599.

46. Du H, Deng W, Aimone J, Ge M, Parylak S, Walch K, Zhang W, Cook J, Song H, Wang L *et al.*: **Dopaminergic inputs in the dentate gyrus direct the choice of memory encoding**. *Proc Natl Acad Sci U S A* 2016, **113**:E5501-E5510.

47. O'Keefe J, Dostrovsky J: **The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat**. *Brain Res* 1971, **34**:171-175.

48. Stachenfeld K, Botvinick M, Gershman S: **The hippocampus as a predictive map**. *Nat Neurosci* 2017, **20**:1643-1653 http://dx.doi.org/10.1038/nn.4650.

49. Kumaran D, Hassabis D, McClelland J: **What learning systems**
• **do intelligent agents need? Complementary learning systems theory updated**. *Trends Cogn Sci* 2016, **20**:512-534.
A review of the complementary learning systems theory that integrates recent advances in neuroscience to draw a thorough picture of how hippocampus and neocortex support efficient learning and generalization. Furthermore, core principles of the theory and new findings are discussed in the machine learning context, with the prospect of designing new agents in artificial intelligence.

50. Funamizu A, Kuhn B, Doya K: **Neural substrate of dynamic Bayesian inference in the cerebral cortex**. *Nat Neurosci* 2016, **19**:1682-1689.

51. Schuck NW, Cai MB, Wilson RC, Niv Y: **Human orbitofrontal cortex represents a cognitive map of state space**. *Neuron* 2016, **91**:1402-1412.

52. Sarma A, Masse N, Wang X-J, Freedman D: **Task-specific versus generalized mnemonic representations in parietal and prefrontal cortices**. *Nat Neurosci* 2015, **19** nn.4168.

53. Bowman C, Zeithamova D: **Abstract memory representations in the ventromedial prefrontal cortex and hippocampus support concept generalization**. *J Neurosci* 2018, **38**:2605-2614.

54. Wutz A, Loonis R, Roy J, Donoghue J, Miller E: **Different levels of**
•• **category abstraction by different dynamics in different prefrontal areas**. *Neuron* 2018, **97**:716-726.
Different levels of category abstraction engage different oscillatory dynamics in different PFC areas. Lower-level categories are first extracted from bottom-up inputs to the ventrolateral PFC reflected in gamma power and spiking. Dorsolateral PFC encode at lower frequencies more abstract levels, depending on top-down processing.

55. Bastos AM, Vezoli J, Bosman CA, Schoffelen J-M, Oostenveld R, Dowdall JR, De Weerd P, Kennedy H, Fries P: **Visual areas exert feedforward and feedback influences through distinct frequency channels**. *Neuron* 2015, **85**:390-401.

56. Saleem AB, Lien AD, Krumin M, Haider B, Rosón MR, Ayaz A, Reinhold K, Busse L, Carandini M, Harris KD: **Subcortical source and modulation of the narrowband gamma oscillation in mouse visual cortex**. *Neuron* 2017, **93**:315-322.

57. Buzsáki G, Draguhn A: **Neuronal oscillations in cortical networks**. *Science* 2004, **304**:1926-1929.

58. Orbán G, Berkes P, Fiser J, Lengyel M: **Neural variability and sampling-based probabilistic representations in the visual cortex**. *Neuron* 2016, **92**:530-543.

59. Rubén M-B, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A: **Information-limiting correlations**. *Nat Neurosci* 2014, **17**:1410-1417.

60. Kirkpatrick S, Gelatt CD, Vecchi MP: **Optimization by simulated annealing**. *Science* 1983, **220**:671-680.

61. Geman S, Geman D: **Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images**. *IEEE Trans Pattern Anal Mach Intell* 1984, **6**:721-741.

62. Atasoy S, Donnelly I, Pearson J: **Human brain networks function**
•• **in connectome-specific harmonic waves**. *Nat Commun* 2016, **7**:10340.
Seminal work showing that harmonic frequency modes naturally emerge from the brain's structural connectivity and excitatory-inhibitory networks dynamics. Different harmonic modes contain either low or high frequency both in space and time. These harmonic modes reproduce resting-state fMRI networks.

63. Lobier M, Palva JM, Palva S: **High-alpha band synchronization across frontal, parietal and visual cortex mediates behavioral and neuronal effects of visuospatial attention**. *Neuroimage* 2018, **165**:222-237.

64. Oemisch M, Westendorff S, Azimi M, Hassani SA, Ardid S,
•• Tiesinga P, Womelsdorf T: **Feature specific prediction errors and surprise across macaque fronto-striatal circuits during attention and learning**. *Nat Commun* 2019, **10**:176 http://dx.doi.org/10.1101/266205.
Macaque fronto-striatal loops are the direct neural correlates of feature-specific prediction errors in multidimensional spaces. Learning reduces the number of loops involved, converging to specific traces representing the most important features.

65. Staresina BP, Michelmann S, Bonnefond M, Jensen O, Axmacher N, Fell J: **Hippocampal pattern completion is linked to gamma power increases and alpha power decreases during recollection**. *eLife* 2016, **5**:e17397.

66. Saez I, Lin J, Stolk A, Chang E, Parvizi J, Schalk G, Knight RT, Hsu M: **Encoding of multiple reward-related computations in transient and sustained high-frequency activity in human OFC**. *Curr Biol* 2018, **28**:2889-2899.e3.

67. Haruno M, Kawato M: **Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning**. *Neural Netw* 2006, **19**:1242-1254.