Computational architecture for hierarchical cognitive processing (D3.13 – SGA3)

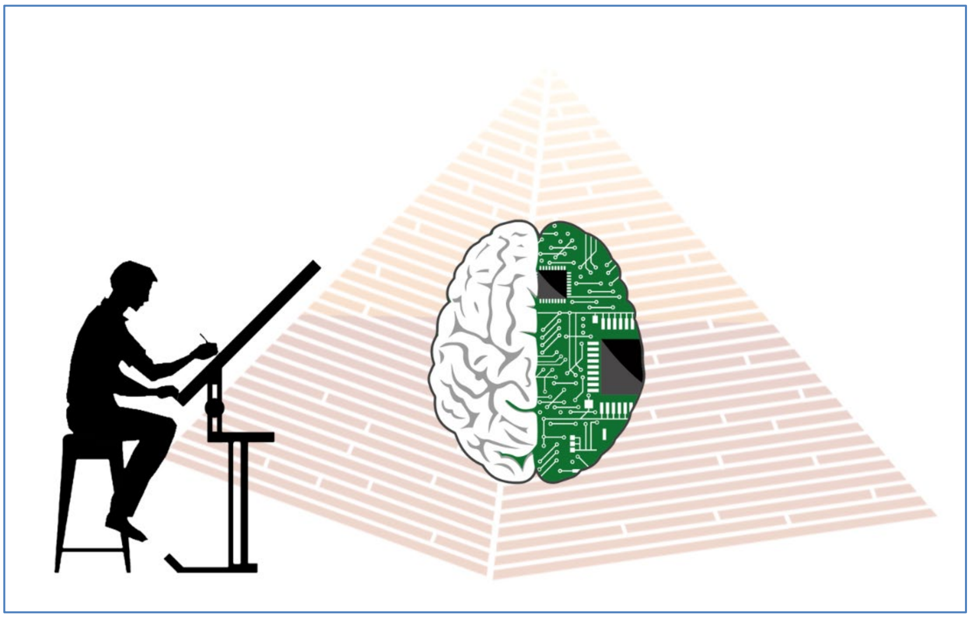

Figure 1: Architectures and learning methods for hierarchical cognitive processing

Cognition builds and exploits hierarchical structures in the real world to efficiently learn new tasks based on components learned in other tasks. We demonstrate such learning in the context of Learning-to-learn (Section 2.1) and Learning visual routines (Section 2.2).

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Computational architecture for hierarchical cognitive processing
Document Filename:	D3.13 (D82) SGA3 M42 RESUBMITTED 231208.docx
Deliverable Number:	SGA3 D3.13 (D82)
Deliverable Type:	Other
Dissemination Level:	PU
Planned Delivery Date:	SGA3 M42 / 30 SEP 2023
Actual Delivery Date:	SGA3 M42 / 15 SEP 2023 (resubmitted 08 Dec 2023)
Author(s):	Pieter ROELFSEMA, KNAW (P91)
Compiled by:	Alexandra VAN DEN BERG, CWI (P50)
Contributor(s):	Sander BOHTE, CWI (P50), contributed to Section 2.1 and Section 2.2 Alexandra VAN DEN BERG, CWI (P50), contributed to Section 2.1 Sami MOLLARD, KNAW (P91), contributed to Section 2.2 Mihai PETROVICI, UBERN (P71), contributed to Section 2.3
WP QC Review:	Anita VAN OERS MU (P117), Jeannette BOSCHMA MU (P117)
WP Leader / Deputy Leader Sign Off:	Rainer GOEBEL, MU (P117)
T7.4 QC Review:	Annemieke MICHELS, EBRAINS (P1)
Description in GA:	The release will make available to EBRAINS' user community the computational architectures developed to address how the brain can learn complex tasks that need to be broken down into multiple subtasks to be solved, necessitate the flexible memorisation and forgetting of task-relevant information, as well as tasks requiring the brain to learn an overarching rule spanning multiple sensory modalities. We put an emphasis on biologically plausible algorithms and learning rules and focused on the capacity of the models for generalization as a way to test the learnt algorithm and representations. The developed models are publicly available on the eBrains repository.
Abstract:	How can agents learn complex tasks that require the applications of subroutines (e.g. “opening the door”) while performing a larger task (e.g. “walking to the fridge), or tasks requiring the agent to learn an overarching rule spanning multiple sensory modalities (learning-to-learn, or meta-learning)? To resolve this question, we developed biologically plausible learning rules and neural network architectures trained on trial-and-error on complex tasks with multiple levels of abstraction and that can be applied to multiple tasks. Specifically, we developed learning rules that can cause attention shifts and create working memories for the relevant aspects of a task, exploiting memory mechanisms. One of the approaches implements biologically plausible learning in the form of contextual NMDA spikes. We also investigated how such learning rules can be extended to group atomic tasks into larger task sequences that act as visual subroutines. We provide access to developed network models as a release on the eBrains infrastructure.
Keywords:	memory, attention, local learning rules, learning-to-learn

Target Users/Readers:	Computational neuroscience community, computer scientists, neuroscientific community, researchers, scientific community, students

Table of Contents

1. Introduction 6
1.1 First release of network that can learn to attend the relevant information and store it in working memory 6
1.2 Computational architecture for multistep tasks 8
1.3 Multi-task learning with contextual NMDA spikes 10
2. References 11

Table of Figures

Figure 1: Architectures and learning methods for hierarchical cognitive processing 1
Figure 2: Architecture RECOLLECT 7
Figure 3: Multistep visual routines and their representation in the visual cortex of monkeys 8
Figure 4: Neural network for solving multistep visual routines 9
Figure 5: Contextual modulation of neurons in sensory processing pathways 10

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
15 SEP 2023	Deliverable submitted to EC
	Resubmission with specified changes requested in Review Report Main changes requested: • In this deliverable, the computational architecture for hierarchical cognitive processing has been presented, but it is not always clear which aspects of the developed architecture has already been released on EBRAINS and which developments still in progress, e.g. has RECOLLECT, or network for multistep visual routines been released yet? All the above should be clarified in the updated document, which should also report all relevant links to EBRAINS. • Moreover, the description is quite vague • and there are no cited references to SGA3 peer reviewed publications (which must be added in the revised document). • As regards the scientific content, the NMDA-based local learning as an alternative to backprop is hardly novel: can the consortium provide justification for this?
	Revised draft sent by WP to PCO. Main changes made, with indication where each change was made: • Change 1: at the end of Section 1.1, p7, we added the relevant public links ("The computational architecture and learning rule are released on the EBRAINS user portal (Output 3.18) and are also available on Github. "), and also at the end of Section 1.2, p10 ("We released the corresponding architecture as a publicly available repository on EBRAINS (Output 3.19. A paper is currently in review."). • Change 2: the text for the Field "Description in GA" on P1 has been elaborated. Additionally, some details have been added to the description of the results for the pro-/anti-saccade task (p7) ("We further show that the feature selectivity the network acquires, resembles that of neurons in primate cortex on the same task. "), ("Finally, we found that behaviour of the network during early stages of learning evolves in a similar manner as that of rodents learning the task. "). • Change 3: Cited references including SGA3 peer reviewed publications are added in Section 1.1 p6 ("such as Rombouts et al. (2015), Kruijne et al. (2021; P3194) and Zambrano et al. (2021; P3193)"); in Section 1.2, p8, we added a linked reference ("The neural network was trained using a biologically plausible learning rule (Pozzi et al. 2020; P2697)", and on p10, the link to the Wybo et al paper in PNAS was added ("(see Figure 5, Wybo et al., PNAS 2023)"). • Justification: the scientific content of the PNAS (2023) paper was not that we suggest NMDA-based local learning as an alternative to backprop. In fact, NMDA-dependent plasticity is on a different conceptual level than backprop. What we

	suggested in this paper, instead, is that a few dendritic NMDA spikes on the basal tree of pyramidal neurons can strongly change the processing of the whole network. We showed that the same feed-forward network with identical synaptic weights, but with additional dendritic NMDA spikes, can implement very different tasks. The performance of the NMDA-modulated network on the various tasks is almost as good is if the network would have been learned by backpropagation for each of the tasks individually.
	Revised version resubmitted to EC by PCO via SyGMa

1. Introduction

How animals learn to represent and memorise the key features of sensory stimuli for the guidance of actions, and how this learning can proceed by trial-and-error has been extensively studied. However, how agents can learn more complex tasks that require the applications of subroutines (e.g. “opening the door”) while performing a larger task (e.g. “walking to the fridge”), or tasks requiring the agent to learn an overarching rule spanning multiple problems (learning-to-learn, or meta-learning) is still understudied.

The goal of the research described in this Deliverable was to develop biologically plausible learning rules and neural network architectures trained on trial-and-error on complex tasks with multiple levels of abstraction. Those tasks require mechanisms of selection (i.e. attention) and storage of information into working memory. We developed learning rules that can cause attention shifts and create working memories for the relevant aspects of the task, and we showed how contextual modulation dynamically reshapes network function to solve new tasks. We furthermore exploited memory mechanisms and investigated how they can be used to group atomic actions into larger sequences that act as subroutines.

The development of learning rules and architectures that can learn hierarchical task structures represent a genuine breakthrough for (1) computational neuroscientists who build biologically inspired networks and (2) systems neuroscientist who can use the emerging computational framework to addresses the relevant fundamental questions in their experiments. As Figure 1 alludes to, we hope that this work, by modelling multistep cognitive routines and meta-reinforcement learning in the brain will help the neuroscience community gain a greater understanding on how the brain works. Conversely, we hope that by drawing inspiration from the brain, our work will be valuable to the AI community, by providing path to develop more robust and energy efficient artificial neural networks and learning rules.

Looking forward, in future work, the architectures and learning rules developed here will be extended towards deeper networks and applied onto more complex and naturalistic input data.

The work described here made a heavy use of the Fenix Infrastructure resources to train, test and run the models effectively and on a large-scale.

1.1 First release of network that can learn to attend the relevant information and store it in working memory

Contributing Partners: KNAW (P91), CWI (P50)

Humans can increase the speed at which they learn tasks when they have been exposed to similar problems in the past. This process is usually referred to as learning to learn. However, this can be challenging for neural network models given that new learning often causes networks to forget the tasks they previously acquired, rather than building upon these experiences.

Meta-reinforcement learning models have been designed that do have the ability of learning-to-learn. However, these typically rely on architectures that are not biologically plausible since they require information that is non-local to the synapse for training (e.g. backpropagation-through-time). An additional difficulty with these models is their computational complexity, which complicates their interpretability in relation to the brain. On the other hand, there are models, which use local learning rules to learn to represent information in memory, such as Rombouts et al. (2015), Kruijne et al. (2021; P3194) and Zambrano et al. (2021; P3193). A limitation here regards the inability to forget information, which is important for learning-to-learn. Therefore, the goal was to develop a model that 1) uses gating mechanisms to forget information efficiently, 2) is not unnecessarily complex, and 3) uses only local information for learning.

The resulting model is called “REinforCement learning of wOrking memory with bioLogically pLausible rECurrent uniTs”, or RECOLLECT. We used RECOLLECT to train networks composed of three layers (see Figure 2). The first layer is an input layer that represents the sensory environment. This layer projects towards a second layer with light-gated recurrent units (Light-GRU; Ravanelli et al., 2018). These units use candidate input units to process the sensory inputs. Before this information is directed towards the memory units, one gating unit per Light-GRU selects how sensitive the memory unit is to new input and how much of the memory is retained. As a result, information can be forgotten according to task requirements. The resulting memory units then connect to the output layer, where a unit estimates the Q-value of each action the network can take. This Q-value is the expectation of the (discounted) reward should this action be chosen.

In order to learn, RECOLLECT - as in AuGMEnT (Rombouts et al., 2015) - uses tags and traces. The tags constitute an attentional feedback signal that identify which output unit provided the winning action and should be updated. The traces store presynaptic activity over time. As such, the tags and traces provide spatial and temporal credit assignment, respectively, and enable RECOLLECT to learn using only local information.

When tested on a pro-/anti-saccade task, in which the model has to memorise a location and is cued to either report that location (pro-saccade) or the opposite location (anti-saccade), RECOLLECT can effectively learn to remember task-relevant information and forget the information again when the trial ends. We further show that the feature selectivity the network acquires, resembles that of neurons in primate cortex on the same task. Moreover, RECOLLECT can learn-to-learn on a reversal bandit task. In this task, two levers are presented with differing reward probabilities (e.g. 25% and 75%). After performing the task for 100 trials (i.e. an episode), the reward probabilities are reversed. This process is repeated across several episodes. The goal for the model is to learn-to-learn an effective exploration-exploitation policy that allows for efficient switching between actions upon episode reversals, rather than re-learning the task every time when reversals occur. Indeed, once training was completed and exploration was disabled, RECOLLECT quickly managed to discover the high-rewarding lever after reversals; indicating learning-to-learn. Finally, we found that behaviour of the network during early stages of learning evolves in a similar manner as that of rodents learning the task. In conclusion, RECOLLECT is a novel gated recurrent network that is trained using a local learning scheme based on LightGRU units that permit forgetting. RECOLLECT can learn-to-learn using exclusively local information in both space and time for network updates, therefore representing a biologically plausible alternative to long-short term memory networks in meta-reinforcement learning.

The computational architecture and learning rule are released on the EBRAINS user portal [(Output 3.18)]() and are also available on https://github.com/Alexandra-van-den-Berg/RECOLLECT. Moreover, a variant of the network that extends the method to deeper networks using BrainProp ([P2697]()) will also be released on this platform (Output 3.20).

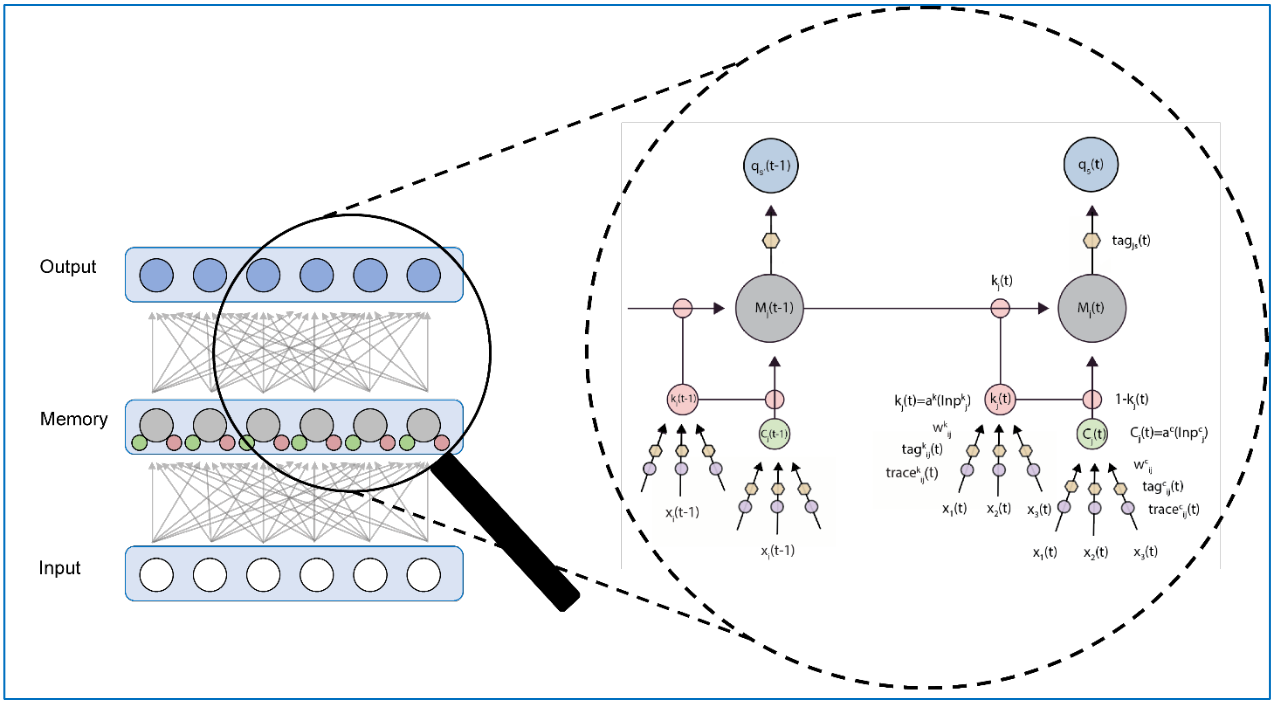

Figure 2: Architecture RECOLLECT

RECOLLECT can selectively represent sensory information in memory units and forget memory when required.

1.2 Computational architecture for multistep tasks

Contributing Partners: KNAW (P91), CWI (P50)

Numerous problems can be broken down into a series of smaller sub problems, which are then solved one after another in a sequential manner. During this process, it is crucial to store and pass on relevant intermediate results from one sub problem to the next until the main objective is achieved. Similarly, when it comes to visual tasks that can be divided into a sequence of basic visual operations, experimental evidence indicates that intermediate results are attended, represented by enhanced neuronal activity in the visual cortex. This heightened activity focus can then be utilised by subsequent subroutines. However, it remains unclear how such dynamics can emerge in neural networks that are trained solely based on rewards, as observed in animals.

To address this, we propose a novel recurrent architecture designed to solve complex visual tasks within a reinforcement-learning framework. We trained neural networks on visual tasks for which electrophysiological recordings of monkey's visual cortex are available: a search-then-trace task and a trace-then-search task. The neural network was trained using a biologically plausible learning rule (Pozzi et al. 2020; P2697)

In the search-then-trace task (see Figure 3) the trials starts when an agent (e.g. a monkey) directs gaze to a fixation point, which can be one of two colours. The agent had to trace the curve that was connected to a disk that had the same colour as the fixation point. This task thus requires the composition of two elemental operations: first a search operation to identify the beginning of the curve to trace, and then a trace operation. In the trace-then-search task (see Figure 3), the order of the two subroutines was inversed: the agent first had to trace the curve connected to the fixation point, register the colour of the end of that curve and make an eye movement toward another disk that had the same colour. Previous studies showed that the outcome of an elemental operation caused a focus of enhanced activity in the visual cortex. For example, in the search-then-trace task, the representation of the beginning of the target curve, which is the outcome of the search operation, is enhanced first in the visual cortex (159ms, Figure 3C), and then the tracing operation can begin (229ms, Figure 3C). In the trace-then-search task, the outcome of the trace operation, i.e. the disk at the end of the curve is enhanced first in the visual cortex (180ms, Figure 3F), before the search operation can take place (267ms, Figure 3F). Apparently, the outcome of an operation causes the tagging of the neuronal representation of a location in space with an enhancement of activity, which makes the outcome available as a focus of attention to be read as the input for the next elemental operation. Interestingly, the sequence of mental operations is associated with a sequence of attention shifts at a psychological level of description.

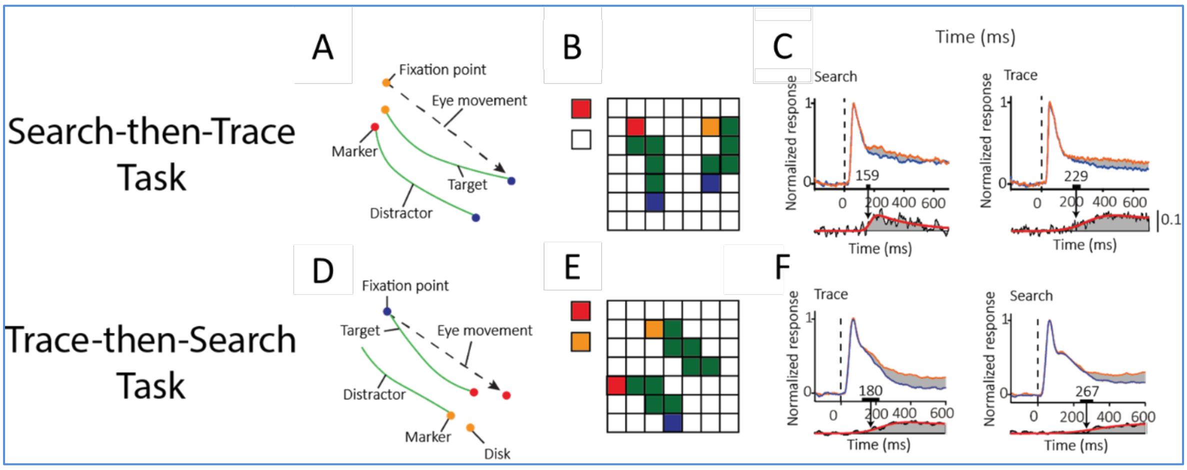

Figure 3: Multistep visual routines and their representation in the visual cortex of monkeys

Monkeys and neural networks were trained on a search-then-trace and a trace-then-search task. In monkeys and networks, elements relevant to transfer between subroutines are tagged by an enhancement of activity.

Although there is substantial evidence that visual routines are implemented in the visual cortex through the propagation of enhanced activity during a recurrent processing phase, it remains unknown which architectures and learning rules can learn such routines. To examine learning of visual routines, we tested neural networks with feedforward, feedback and lateral connections. We asked whether the network (1) would learn multiple elemental operations, (2) whether it would learn to execute the elemental operations sequentially and in the correct order and (3) how the network ensured the transfer of information of one elemental operations to the next. We trained convolutional neural networks with an input layer; two hidden layers and an output layer (see Figure 4). In each layer, there was a group of feedforward units that only propagated information to the next layer and to recurrent neurons with the same receptive field. The other, recurrent group propagated information to both the higher and lower layer and to units with nearby receptive fields in the same layer. As a result, neurons in the feedback group could be modulated by activity outside the pixel that defined their receptive field. However, the activity of neurons in the feedback group was gated by neurons in the feedforward group that had the same receptive field so that this modulatory effect could not occur if the feedforward unit with the same receptive field was not active. We trained the networks with RELEARNN, which is a biologically plausible implementation of Q-learning that can be broken down in three phases. Upon presentation of the stimulus, the activity of the input neurons remains constant and activity propagates through the recurrent connections of the network. If the activity of neurons between two consecutive time steps was constant, we considered that a stable state was reached and the action was selected, but we did not wait longer than 50 time steps if the activity was not yet stable. When an action is selected, an "attentional" signal originating from the winning action is propagated through an accessory network to determine the influence of each neuron on the selected action. The network then gets a reward from the environment and computes a reward prediction error.

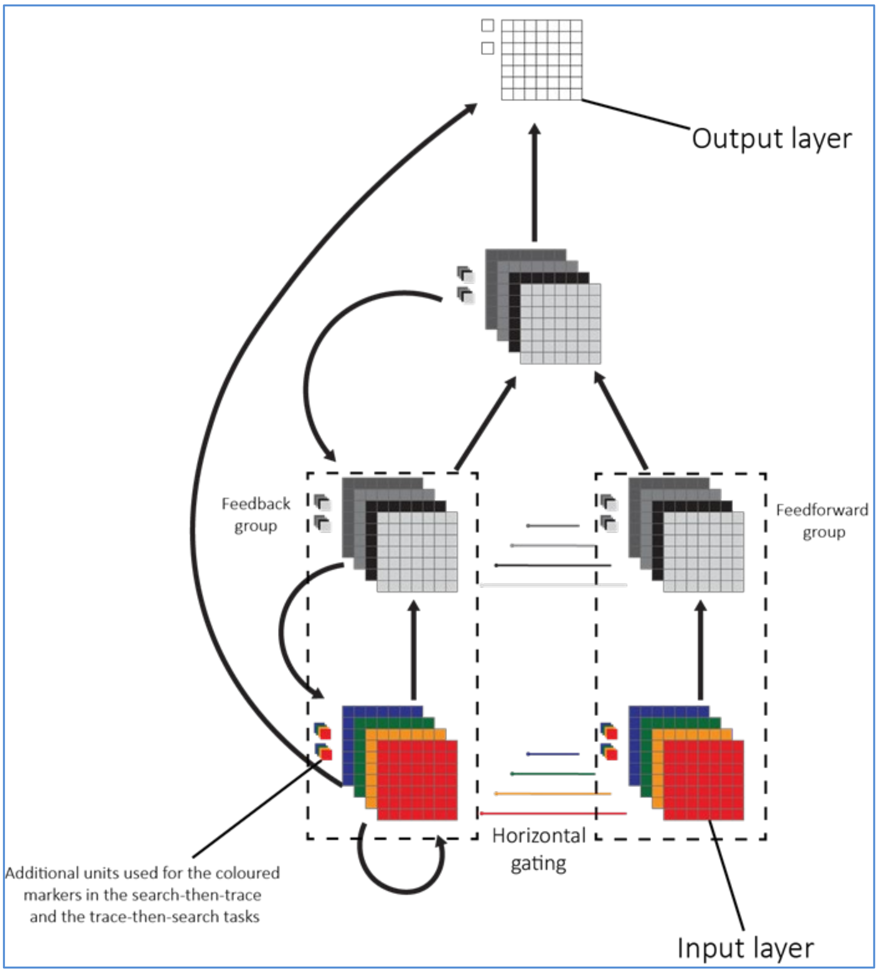

Figure 4: Neural network for solving multistep visual routines

Information is processed through a recurrent group (left) and a feedforward group (right). Units in the recurrent group are gated by units in the feedforward group

Networks were able to learn the tasks that within in a few thousand trials. After training, we examined the activation of spatially selective neurons, and that of neurons tuned to colour. We also compare these activity profiles to the activity of neurons in the visual cortex. Critically, we found that the artificial neural networks spontaneously learnt to develop the same response modulations that have been observed in the visual cortex of monkeys. During the trace operation, neurons with receptive fields on the target curve enhanced their activity sequentially, starting from the beginning of the curve. During the search operation, neurons representing the target colour enhanced their activity. Information is also transferred between the subroutines as a focus of enhanced activity in the network, so that the output of an elemental operation can be read as the input of the next one. As such, the networks were able to transfer the strategies used for one task or subtask (curve tracing for instance), to other tasks (like a search-then-trace task or a trace-then-search task, Output 3.19). We also showed how relevant information that needs to be transferred between subroutines is stored in working memory as an enhancement of activity, until the overall task has been completed. We released the corresponding architecture as a publicly available repository on EBRAINS (https://search.kg.ebrains.eu/?category=Model#8c62f509-5876-4101-8cfb-6db39c90ff98; Output 3.19). A paper is currently in review (https://plus.humanbrainproject.eu/publications/4066).

1.3 Multi-task learning with contextual NMDA spikes

Contributing Partners: UBERN (P71), FZI (P52)

Learning in the brain, as well in modern artificial intelligence, is typically considered to rely on synaptic modifications. However, every change in synaptic strength does not just affect the functionality of the network in the very particular context in which it occurs, but rather carries implications for all of the functions in which a network is involved. The dendritic morphology of cortical neurons offers another road to re-use the same neurons in new contexts and new tasks, without actually changing the synaptic strengths. Dendrites may receive context-dependent ‘top-down’ input that individually modulates the excitability of the neurons, so that they can be reused in the new context without adapting the synaptic strengths.

In deep learning, the standard approach to accommodate changing task demands is to train new output layers on top of a common trunk network, and, if needed, to relearn synapses throughout the whole network. However, the brain appears to take a radically different strategy, as neurons in all processing layers are modulated by contextual information. We showed that context-dependent dendritic afferents can powerfully modulate the neuronal output and that this modulation dynamically reshapes network function to solve new tasks, without adapting any feedforward synapses (see Figure 5, Wybo et al., PNAS 2023). We furthermore showed that these dendritic modulations could underlie self-supervised learning of deep networks, without relying on the backpropagation of errors across the layers of the network.

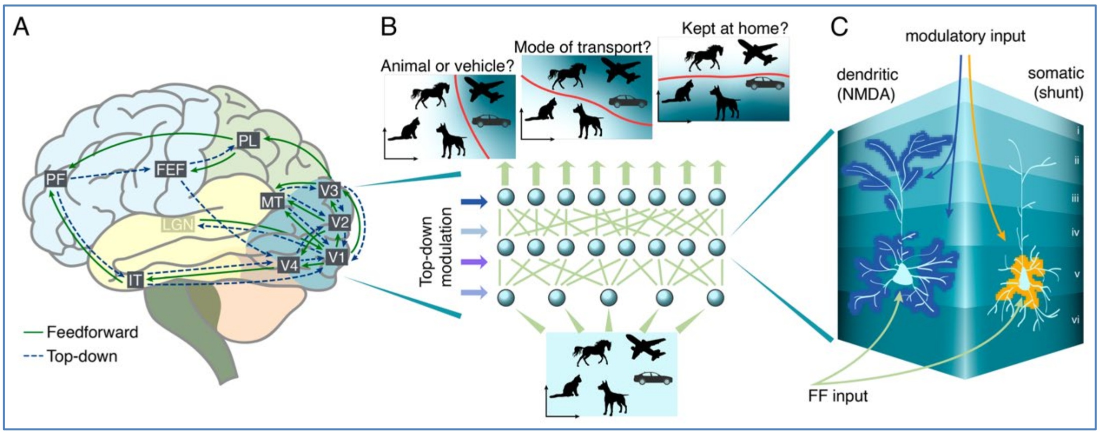

Figure 5: Contextual modulation of neurons in sensory processing pathways

Adapted from [P3781], Wybo et al., PNAS 2023. (A) Top-down connections from prefrontal and motor areas relay high-level information to early sensory processing neurons [adapted from Gilbert et al., 2013), LGN: lateral geniculate nucleus of the thalamus, V1-4: visual area 1-4, MT: medial temporal area, IT: inferior temporal cortex, PL: parietal lobe, FEF: frontal eye field, PF: prefrontal cortex]. (B) We hypothesize that high-level information from prefrontal

and motor areas modulates the activity of early sensory neurons, enhancing response properties of neurons with task-relevant receptive fields. These modulations induce a task-dependent functional remapping of sensory processing pathways built on fixed, task-agnostic feedforward connectivity. (C) At the biophysical level, we investigate two plausible candidate mechanisms that could implement quasi-tonic neuron-specific modulations: somatic shunting inhibition and dendritic NMDA spikes.

2. References

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. Nature Reviews Neuroscience, 14(5), 350-363.

Kruijne, W., Bohte, S. M., Roelfsema, P. R., & Olivers, C. N. L. (2021). Flexible working memory through selective gating and attentional tagging. Neural Computation, 33(1), 1-40.

Pozzi, I., Bohte, S. M., & Roelfsema, P. R. (2020). Attention-gated brain propagation: How the brain can implement reward-based error backpropagation. 34th Conference on Neural Information Processing Systems (NeurIPS 2020).

Ravanelli, M., Brakel, P., Omologo, M., & Bengio, Y. (2018). Light gated recurrent units for speech recognition. IEEE Transactions on Emerging Topics in Computational Intelligence, 2(2), 92-102.

Rombouts, J. O., Bohte, S. M., & Roelfsema, P. R. (2015). How attention can create synaptic tags for the learning of working memories in sequential tasks. *PLoS computational biology*, *11*(3), e1004060.

Wybo, WAM, Tsai MC, Tran VAK, Illing, B, Jordan, J, Morrison, A and Senn, W (2023). NMDA-driven dendritic modulation enables multitask representation learning in hierarchical sensory processing pathways. Proceedings of the National Academy of Sciences 120 (32), e2300558120

Zambrano, D., Roelfsema, P. R., & Bohte, S. M. (2021). Learning continuous-time working memory tasks with on-policy neural reinforcement learning. Neurocomputing, 461, 635-656.