

Neural Learning of Cognitive Control

Fred H. Hamker

Published online: 12 June 2012
© Springer-Verlag 2012

Abstract Our goal is to develop cognitive agents based on neuroscientific evidence. The efficiency of cognitive behavior depends on its capacity to select, represent and manipulate sufficient knowledge of the environment to achieve its goals. We designed a biologically motivated model of basal ganglia and particularly the prefrontal cortex and here review its foundations of neural learning and summarize our obtained results.

Keywords Learning · Cognitive agent · Working memory

1 The Challenges of Building Cognitive Systems

The aim to design cognitive agents has its roots already in the early days of artificial intelligence and a number of interesting architectures have been developed (for recent overviews see [9, 10]). Yet there has been no major breakthrough such that we now know how to build cognitive systems. Vernon, Metta and Sandini [17] distinguished between two major directions of cognitive systems research: Cognitivist models and emergent approaches, particularly enactive systems. Cognitivist models build internal, abstract, typically symbolic representations of the world and perform operations on these to determine behavior. Enaction has been put forward as an alternative to overcome limitations of cognitivist models such as symbol grounding and the strong dependency on the designer to find a suitable representation [18]. In enactive systems, the purpose of perception is to provide sensory information solely for guiding behavior

of an embodied agent. Thus, there are not necessarily states representing entities of the world but rather some form of self-organizing dynamics that fulfill the purpose of generating behavior. Most practical systems are cognitivist models with some form of emergence and self-organization. While there may be several possible research directions to address the challenges of building cognitive systems, neuroscience has revealed much data and insight during the last 20 years of research and the time appears ripe to design cognitive systems that closely match the functional principles of the brain. One of those challenges is flexibility, i.e. the ability to adapt to changes in the environment. Thus, neuro-cognitive systems should heavily build upon learning. Even relatively simple tasks such as working memory recruit much knowledge that has been acquired during development. For example, we are often not aware that in each moment in time we memorize only particular aspects from the seemingly endlessness of our visual input for each task at hand. This ability to “know” what is relevant and irrelevant for each given task should have been learned.

While enactive systems are very appealing [5, 18], they are difficult to implement without knowing how cognitive abilities develop. In the neurosciences, increasing evidence points towards the basal ganglia and the dopaminergic system to be crucially involved not only in motor control but also in learning, categorization and decision making by different parallel cortex-basal-ganglia loops. Learning such loops depends on the neurotransmitter dopamine which projects from neurons of nuclei in the midbrain to multiple brain areas [8]. Since the release of dopamine occurs in behaviorally relevant (unexpected or salient) events [14], learning takes particularly place in uncertain contexts. This learning will finally be selective for those neural activities that allow to improve the prediction so that this mechanism allows to sort out the relevant from the irrelevant. Indeed the

F.H. Hamker (✉)
Fakultät für Informatik, Technische Universität Chemnitz,
Straße der Nationen 62, 09107 Chemnitz, Germany
e-mail: fred.hamker@informatik.tu-chemnitz.de

machine learning concept of reinforcement learning shows several similarities to brain mechanisms mediated by basal ganglia and the dopaminergic system [19].

Different models have already been proposed for some cognitive functions in which the basal ganglia could be involved (see [4, 19] for an overview). Interestingly, a computational model of basal ganglia has recently put forward to replace the procedural module in the well known ACT-R cognitive architecture [15].

2 Neural Learning Mechanisms

As it became clear from the outline above, the development of neuro-cognitive systems requires not only a systems level design but also a set of basic learning rules to be applied in the system. While recent research has revealed a dependence of learning on the exact timing of action potentials, so called spike-timing-dependent plasticity, learning in multiple brain areas can nevertheless be well approximated by firing rate dependent Hebbian learning (see [6] for a discussion). Thus, we modeled learning by dynamic Hebbian learning rules using a framework of rate coded neurons. According to this framework the membrane potential $m_i(t)$ of a neuron i is described by a differential equation:

$$\tau \cdot \frac{dm_i(t)}{dt} + m_i(t) = \sum_j w_{i,j}(t) \cdot u_j(t) + M + \epsilon_i(t) \quad (1)$$

where τ is the time constant of the neuron, $u_j(t)$ the firing rate of a presynaptic neuron j , $w_{i,j}(t)$ the weight between both neurons, M a baseline membrane potential and $\epsilon_i(t)$ a random noise term. The output of a neuron is computed from the membrane potential by a transfer function, typically just the positive part of the membrane potential $v_i(t) = (m_i(t))^+$, where $()^+$ denotes that negative values are set to zero.

Contrary to classical neural networks where learning usually takes place after each trial using static learning rules, we developed learning equations that allow continuous learning throughout the whole simulation period. Basically, learning refers to a change of the connectivity $w_{i,j}$ between a presynaptic neuron $u_j(t)$ and postsynaptic neuron $v_i(t)$ and relies on a variant of the Oja [12] learning rule using a covariance term:

$$\eta \cdot \frac{dw_{i,j}}{dt} = (u_j - \bar{u})^+ \cdot (v_i - \bar{v}) - \alpha \cdot (v_i - \bar{v})^2 \cdot w_{i,j} \quad (2)$$

The covariance term assigns each neuron as part of a neural population. The learning rule can be applied in each time step and states that the weight from the input $u_j(t)$ to neuron $v_i(t)$ is increased if the rate $u_j(t)$ is larger than the mean response of the input population $\bar{u}(t)$ and the postsynaptic

neuron $v_i(t)$ fires above the mean of its population $\bar{v}(t)$. A weak input $u_j < \bar{u}$ has no effect on learning this particular weight. A strong input $u_j > \bar{u}$, paired with weak postsynaptic response $v_j < \bar{v}$ leads to a decrease of the weight. Normalization by the second term ensures that the weight resource is limited and avoids an increase of the weight without bound. The exact definition of the covariance term can vary dependent on the goals of learning. This framework has also been used for learning receptive fields of neurons for object recognition [21].

However, this learning rule alone would not be sufficient for designing neuro-cognitive systems. First of all, learning does not easily ensure that the neural firing rate as well as the weights are in a meaningful range. Particularly learning across multiple hierarchical levels faces the problem that the rates permanently increase if the gain factor determined by the weights is too large or the neural responses die out if the weights are too small. This problem of stability in neural networks could be solved by a mechanism known as synaptic scaling [11]. Neurophysiological experiments revealed that neurons have the ability to multiplicatively scale the synaptic connections which has the advantage that the overall strength of synaptic transmission can be adjusted without destroying the encoded pattern which is memorized in the relative connection strength. Our studies [20] revealed that this physiological observation can be nicely accounted for by making the parameter α dependent on the firing rate of an individual neuron i so that

$$\tau_\alpha \cdot \frac{d\alpha_i(t)}{dt} + \alpha_i(t) = K_\alpha \cdot (v_i(t) - v^{\max})^+ \quad (3)$$

where v^{\max} is the target firing rate of the cell and K_α a constant that determines the speed of increases of α_i .

In addition, learning should not be solely unsupervised, but reinforced by success, i.e. it should be dependent on the prediction error in form of dopamine (DA) [19]. Thus, we introduced a third factor $f(\text{DA}(t))$ which is determined by a function f that considers the level of dopamine relative to its baseline. It leads to a weight increase if the obtained reward is larger than the actual received reward and to a decrease if the predicted reward is smaller than expected.

Finally, reward is typically delivered after the events that are relevant for obtaining reward. Thus, some form of memory is required. While some approaches relied on working memory [4], alternative neural mechanisms of learning could provide an essential contribution as well. At least two different mechanisms could provide a synaptic memory for minutes up to a few hours. In most synapses, learning critically depends on the influx of Ca^{2+} which itself is dependent on the pre- and postsynaptic response. Ca^{2+} has no long memory trace but it triggers Ca^{2+} -dependent kinases, such

as CaMKII that leads to synaptic changes by protein synthesis during minutes to hours [1] and thus provides a mechanism of short term synaptic memory. Moreover, it has recently been discovered that activated synapses can be tagged and only later, when dopamine has been released, they undergo long-term changes [3, 7]. Thus, we expanded Hebbian learning rules to include Calcium eligibility traces $Ca_{i,j}(t)$ [13].

$$\eta^{Ca} \cdot \frac{dCa_{i,j}}{dt} + Ca_{i,j} = (v_i - \bar{v})^+ \cdot (u_j - \bar{u}) \quad (4)$$

$$\eta^{Ca} = \begin{cases} \eta^{inc} & \text{if } dCa_{i,j} > 0 \\ \eta^{dec} & \text{else.} \end{cases} \quad (5)$$

η^{Ca} is the time constant of the calcium trace, η^{inc} a parameter controlling the speed of calcium level increase and η^{dec} a parameter controlling the speed of calcium level decline. By using a small value for η^{inc} , the corresponding calcium level increases rapidly. As η^{dec} is much larger than η^{inc} , the calcium level declines rather smoothly and thus leaves a synaptic trace of its previous activation, so that these calcium eligibility traces determine the amount of long term changes (learning) at the time of DA release [13]:

$$\eta \cdot \frac{dw_{i,j}(t)}{dt} = f(DA(t)) \cdot Ca_{i,j}(t) - \alpha_i(t) \cdot (v_i(t) - \bar{v}(t))^2 \cdot w_{i,j}(t) \quad (6)$$

3 Learning Working Memory

I now illustrate how cognitive control can be learned by the example of working memory (WM). Working memory describes the ability of a system to temporally store information for its later usage [2]. By maintaining information in WM, an organism can detach its responses from its immediate sensory environment and exert deliberate control over its actions. Working memory requires to determine when to store a representation (of a stimulus) and to preserve it from overwriting as long as it is required. Most importantly, working memory has to be learned specifically for each task, since it requires the knowledge which of the myriad of stimuli are required for future use. Based on anatomical evidence [8] of a brain structure known as Basal Ganglia (BG), we designed a hierarchical model where a set of (prefrontal) cortico-BG loops are involved in cognitive control and provide context information for decisions learned in a parallel motor cortico-BG-thalamic loop (Fig. 1). This architecture of parallel and hierarchically interconnected loops provides a potential anatomic substrate for how visual, cognitive and motor information can be maintained for extended periods of time (closed loops) and for how cognitive processes can bias response selection [13]. While I here focus

on WM, such closed loops could also provide a neural systems substrate for conscious processes [16]. Neurons in prefrontal cortex provide “high-level” context information for the “lower-level” motor loop. Thus, the motor loop does not have to deal with those details that can be solved already at the higher level and the learning problem becomes easier.

As demonstrated in [13] the model can learn complex working memory tasks such as the 1-2-AX task. In this task, a long sequence of letters and numbers from eight possible stimuli (1, 2, A, B, C, X, Y and Z) is shown and a button press is required after each stimulus. A right button press is required after the sequence 1-A-X or 2-B-Y. In all other cases, a left-button press is correct. Thus, the solution is not memorizing all stimuli shown but to selectively memorize a stimulus given the prior presentation of stimuli. For example, the occurrence of ‘A’ is only relevant if at some time earlier ‘1’ has been shown but not ‘2’. Moreover, this task also requires the selective resetting of memory. The task has been trained in a three-step shaping procedure according to which the model is initially confronted with the numbers 1 and 2 and a delayed response such that it has to implement a memory of the numbers. If the model reliably learned the first step, the letters A, B and C are added to the sequence and finally the full task is presented. During the shaping procedure relearning of responses is required as each shaping step is not simply a component of the final task. In additional experiments we have shown that once a task has been learned, additional learning or relearning is facilitated.

4 Conclusion and Future Work

The review provides the present state of the art in the neural learning of cognitive control inspired by the function of the human brain. As illustrated above the usage of biologically realistic learning rules and homeostatic principles enables the learning of cognitive functions even if learning takes place in parallel at multiple sites in the model. In particular, this model provides an approach that allows to learn internal memory states and thus it could be further developed as a basis for goal directed systems. Presently, we scale this model up to large input and action states. Moreover, we apply this concept to acquire learning of visual categories. In the long run, this model could provide an alternative to more classical, machine-learning approaches of reinforcement learning for robots.

Acknowledgements This project report provided an overview of recent research on neural learning of cognitive control in our lab which involved contributions from multiple researchers, particularly from Henning Schroll, Julien Vitay and Jan Wiltchut.

This work has been funded by DFG HA2630/4-1 and HA2630/4-2 as well as by the EU FP7-ICT program “Eyeshots: Heterogeneous 3-D Perception across Visual Fragments”.

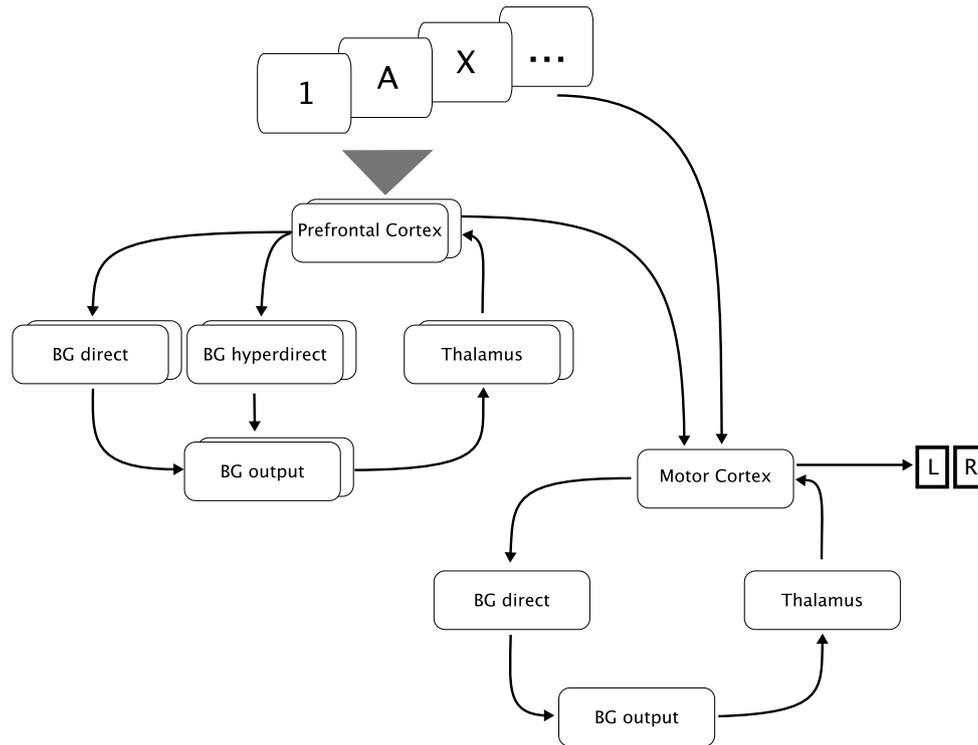


Fig. 1 Cognitive task and model layout for WM. A stimulus sequence is presented to the model and the model has to respond after each stimulus with a left (*L*) or right (*R*) button press. By learning it has to discover the underlying rules of the task. The task requires that prefrontal cortico-BG-thalamic loops flexibly control WM and guide a motor loop to choose between a set of possible responses. While the general layout of prefrontal and motor loops is the same, the prefrontal

loops have a direct and hyperdirect pathway and the motor loop is simplified and only relies on the direct pathway. *Boxes* represent the different layers of the model, arrows the connections between them. ‘*Double*’ boxes represent dual prefrontal circuits. Most of the connections are subject to learning using the concept of a prediction error for the DA signal in Eq. (6). A more detailed figure about the pathways can be found in [13]

References

- Abraham WC (2008) Metaplasticity: tuning synapses and networks for plasticity. *Nat Rev, Neurosci* 9:387–399
- Baddeley A (2003) Working memory: looking back and looking forward. *Nat Rev, Neurosci* 4:829–839
- Clopath C, Ziegler L, Vasilaki E, Buesing L, Gerstner W (2008) Tag-trigger-consolidation: a model of early and late long-term potentiation and depression. *PLoS Comput Biolol* 4:e1000248
- Cohen MX, Frank MJ (2009) Neurocomputational models of basal ganglia function in learning, memory and choice. *Behav Brain Res* 199(1):141–156
- Eggert J, Wersing H (2009) Approaches and challenges for cognitive vision systems. In: *Creating brain-like intelligence: from basic principles to complex intelligent systems*. Lecture notes in computer science, vol 5436. Springer, Berlin, pp 215–247
- Frégnac Y et al (2010) A re-examination of Hebbian-covariance rules and spike timing-dependent plasticity in cat visual cortex in vivo. *Front Synaptic Neurosci* 2:147
- Frey U, Morris R (1997) Synaptic tagging and long-term potentiation. *Nature* 385:533–536
- Haber SN (2003) The primate basal ganglia: parallel and integrative networks. *J Chem Neuroanat* 26:317–330
- Helie S, Sun R (2010) Incubation, insight, and creative problem solving: a unified theory and a connectionist model. *Psychol Rev* 117:994–1024
- Langley P, Laird JE, Rogers S (2009) Cognitive architectures: research issues and challenges. *Cogn Syst Res* 10:141–160
- Turrigiano GG, Nelson SB (2004) Homeostatic plasticity in the developing nervous system. *Nat Rev, Neurosci* 5:97–107
- Oja E (1982) A simplified neuron model as a principal component analyzer. *J Math Biol* 15:267–273
- Schroll H, Vitay J, Hamker FH (2012) Working memory and response selection: a computational account of interactions among cortico-basal ganglia-thalamic loops. *Neural Netw* 26:59–74
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27
- Stocco A, Lebiere C, Anderson JR (2010) Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol Rev* 117:541–574
- Trapp S, Schroll H, Hamker FH (2012) Open and closed loops: a computational approach to attention and consciousness. *Adv Cogn Psychol* 8:1–8
- Vernon D, Metta G, Sandini G (2007) A survey of artificial cognitive systems: implications for the autonomous development of mental capabilities in computational agents. *IEEE Trans Evol Comput* 11:151–180
- Vernon D, Metta G, Sandini G (2010) Enaction as a conceptual framework for developmental cognitive robotics. *Paladyn J Behav Robot* 1:89–98
- Vitay J, Fix J, Beuth F, Schroll H, Hamker FH (2009) Biological models of reinforcement learning. *Künstl Intell* 3:12–18

20. Vitay J, Hamker FH (2010) A computational model of the influence of basal ganglia on memory retrieval in rewarded visual memory tasks. *Front Comput Neurosci* 4:13
21. Wilschut J, Hamker FH (2009) Efficient coding correlates with spatial frequency tuning in a model of V1 receptive field organization. *Vis Neurosci* 26:21–34



Fred H. Hamker is professor of Artificial Intelligence at the Department of Computer Science of the Chemnitz University of Technology. He received his diploma in Electrical Engineering from the University of Paderborn in 1994 and his Ph.D. in Computer Science at the TU Ilmenau in 1999. He was a Post-Doc at the J.W. Goethe University of Frankfurt and the California Institute of Technology (Pasadena, USA). In 2008 he received his *venia legendi* from the Department of Psychology at the Westf. Wilhelms-University Münster.