

Behavior-State Dependent Modulation of Perception Based on a Model of Conditioning

Jordi-Ysard Puigbò^{1,2(✉)}, Miguel Ángel Gonzalez-Ballester^{2,3},
and Paul F.M.J. Verschure^{1,3,4}

¹ Laboratory of Synthetic, Perceptive, Emotive and Cognitive Science (SPECS),
DTIC, Universitat Pompeu Fabra (UPF), Barcelona, Spain

jordiyysard.puigbo@upf.edu

² Laboratory for Simulating, Imaging, and Modeling of Biological Systems
(SIMBIOsys), DTIC, Universitat Pompeu Fabra (UPF), Barcelona, Spain

³ Catalan Research Institute and Advanced Studies (ICREA), Barcelona, Spain

⁴ Catalan Institute of BioEngineering (IBEC), Barcelona, Spain

Abstract. The embodied mammalian brain evolved to adapt to an only partially known and knowable world. The adaptive labeling of the world is critically dependent on the neocortex which in turn is modulated by a range of subcortical systems such as the thalamus, ventral striatum and the amygdala. A particular case in point is the learning paradigm of classical conditioning where acquired representations of states of the world such as sounds and visual features are associated with predefined discrete behavioral responses such as eye blinks and freezing. Learning progresses in a very specific order, where the animal first identifies the features of the task that are predictive of a motivational state and then forms the association of the current sensory state with a particular action and shapes this action to the specific contingency. This adaptive feature selection has both attentional and memory components, i.e. a behaviorally relevant state must be detected while its representation must be stabilized to allow its interfacing to output systems. Here we present a computational model of the neocortical systems that underlie this feature detection process and its state dependent modulation mediated by the amygdala and its downstream target, the nucleus basalis of Meynert. Specifically, we analyze how amygdala driven cholinergic modulation these mechanisms through computational modeling and present a framework for rapid learning of behaviorally relevant perceptual representations.

1 Introduction

In the last years we have been a significant increase in the literature regarding topics the topics of artificial intelligence, supervised learning and reinforcement learning.

This increase is clearly due to the finding of workarounds to the limiting computational cost of certain algorithms, principally represented by shared weights

through convolutions that reduce the parameter spaces and increased generalisation on deeper networks.

In this sense, the field of machine learning is rapidly increasing the number of tasks that can be learned with combinations of algorithms from mainly reinforcement learning, recurrent neural networks and deep networks, as soon as there is enough computational power, time and data. Of special interest for robotics could be reinforcement learning algorithms, which allow virtual or real agents to learn action policies according to the current inputs. These algorithms have been recently used for playing atari games with better than human performance, simulating multi-agent interactions in solving foraging tasks, coordinating for goals and solving a wide range of tasks that require not just optimal action selection but also memory aided decision making (see [2] for critical review). All this being achieved within the last two years makes one wonder why we are still unable to apply this to real robots performing in real world tasks, but just in simulations. The answer is not really unknown to anyone: training times are extremely big and require lots of failures to converge into good performance, sometimes even better than humans. While time could eventually not be a mandatory problem, the current ability of robots to recover from failure is scarce, while some failures can be unrecoverable by themselves. Additionally, these algorithms require to be trained on the agent that will be performing the action at any moment.

The critical point, then, is that we need both robots able to recover from failure, like an animal that falls can stand up again, and algorithms that allow them to anticipate and avoid critical failure, as a child would learn to avoid a hot frying pan after the first burning contact. Learning in this conditions becomes hard, as you want to minimize the number of exposures or samples of that same event, going against the current trend of learning by having more and better data: in a critical failure event, you can't afford gathering more data and you might want to favor overfitting. Psychology theories on classical conditioning and its subsequent neuro-anatomical studies have provided insights of the circuitry behind this fast and drastic behavior. First, our bodies have reactive feedback systems that avoid additional harm once an aversive stimuli (unconditioned stimuli or US) is sensed. These systems have evolved through millions of generations and species to what we can find today in animals. On top of this, a precisely designed architecture learns to associate predictive cues to perfectly-timed reactions that will, eventually, avoid to sense the US again. The anatomy behind this behavior has been located numerous times in the cerebellum and has been extensively studied [1]. Finally, the most important point relies on acquiring accurate, stable representations of the events that predict the aversive USs. This suggests that the anticipatory behavior needs to be acquired in two steps: an initial fine tuning of the perceptions to detect cues predictive of aversive events, and then the precise association of these events with anticipatory, pre-defined actions. Several studies have highlighted that fear responses (mainly startles and fear-related neuromodulators release) are associated to the CS before the anticipatory response is correctly anticipated while neocortical neurons change their preferred responding stimuli both after CS-US pairing and

after pairing with neuromodulators like acetylcholine (ACh). We will revise what are the changes provoked in the cortical substrate by this neuromodulator and analyze how this neuromodulation can become relevant for quickly redefining perceptual representations and eventually to stabilize anticipatory behavior.

2 Neocortex and Acetylcholine

Acetylcholine (ACh) is a neuromodulator that mediates the detection, selection and further processing of stimuli [5] in the neocortex. ACh is typically associated with attention and synaptic plasticity [6]. The main source of cortical acetylcholine is the Nucleus Basalis of Meynert (NBM) in the basal forebrain. Cholinergic neurons in the NBM receive their main excitatory projections from the amygdala and from ascending activating systems. The amygdala is known to be involved in building associations that are key for fear conditioning. While projections from the ascending activating systems are generally related to bottom-up mechanisms for maintaining arousal and wakfulness, other excitatory projections from prefrontal and insular cortex convey top-down control of acetylcholine release. The principal targets of cholinergic afferents from the NBM are in the neocortex. These projection preserve a topological organization with the cortex and the amygdala. Cholinergic modulation of the neocortex depends principally on two distinct cholinergic receptor types: metabotropic Muscarinic and ionotropic Nicotinic receptors. The activation of muscarinic receptors in the neocortex of the rat has a fast, global excitatory effect on both excitatory and inhibitory neurons. In contrast, binding with nicotinic receptors have a slower disinhibitory response on the neural substrate. Hence, the global effects of acetylcholine in a population of cortical neurons must depend on the balance and distribution of these two opponent types of cholinergic receptors. Here we will investigate this relationship in the context of sensory processing in classical conditioning. In particular, we will investigate the role of the differential drive of cortical interneurons by acetylcholine on the behavioral state dependent gating and representation of sensory states.

2.1 Two Degrees of Multistability in the Neocortex are Mediated by Inhibitory Interneurons

The complex structure of the neocortex is comprised by 80% excitatory and 20% inhibitory neurons organized in six differentiable layers. Among the 20% of inhibitory interneurons, two major interneuron subtypes are immunoreactive to either Parvalbumin or Somatostatin [5], but not the other, accounting for most of the GABAergic interneurons accross cortical layers. These two interneuron subtypes are also different in terms of dendritic arborization (e.g. basket cells, Martinotti cells, chandelier cells), spiking patterns (e.g. fast spiking, regular spiking). PV expressing (+) interneurons (PVi+) in the neocortex have fast spiking dynamics where basket cells are the classical example. SST expressing interneurons (SSTi+) instead have more regular spiking dynamics and broader

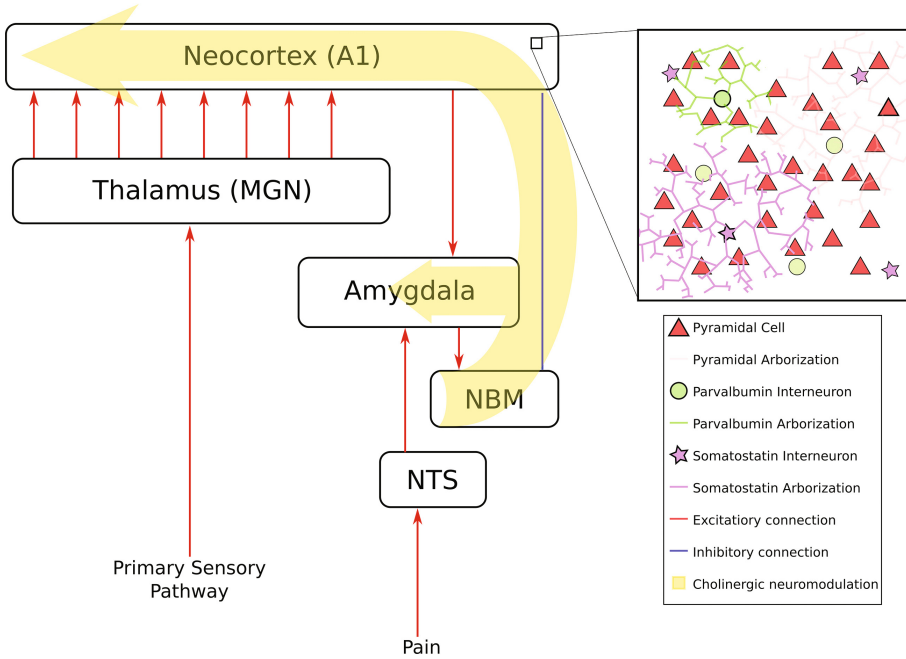


Fig. 1. Anatomical model of conditioning and cholinergic modulation. An auditory stimulus is pre-processed in the auditory pathway through Superior Olive (SO), Inferior Colliculus (IC) and the Medial Geniculate Nucleus of the thalamus (MGN) until it reaches the primary auditory cortex (A1). Our model of a small fraction of the cortex (top-right) is composed of 3 cell populations, one excitatory containing 80% of the cells and 2 inhibitory containing the remaining 20%. These inhibitory populations correspond to the PVi+ and SSTi+ responsive inhibitory interneurons. Both inhibitory populations receive excitatory input principally from the excitatory population, composed of pyramidal cells. Finally, they project back to it with inhibitory connections. Moreover, unconditioned stimuli (US), indicative of surprising or aversive events are relayed through the Nucleus Tractus Solitarius (NTS) to the Amygdala (Am). The amygdala, in turn, uses cortical information to predict future USs. Either the predictions or the US itself stimulate the Nucleus Basalis of Meynert. The NBM releases ACh in the neocortex and the amygdala (Am), therefore promoting the acquisition of new sensory features and its predictive component. Finally, the amygdala, which is receiving contextual information in the typical form of a conditioned stimulus (CS), predictive of the US, learns the association between the cortical predictive components of and the NBM stimulation, facilitating future learning events, now cued by the CS.

arborizations usually in the form of Martinotti and chandelier cells. Together with the less studied interneurons expressing Serotonin, these 3 immunohistochemically separable interneuron classes account for the majority of the GABAergic neurons in the neocortex [4].

In order to understand how different inhibitory populations affect the dynamics of the network we realized a series of computational simulations changing the

axonal range and the gain of the inhibitory neurons. We used rate based leaky linear units, defined by:

$$x_j(t) = \sum_i W_{ij} x_i(t-1) \quad (1)$$

Where x_j and x_i are the post-synaptic and pre-synaptic neurons and W is the a connectivity matrix representing the conductance between each neuron in the population. The connectivity matrix W is constructed following a pseudo-random rule based on exponential distance, where the probability of forming a connection (C_{ij}) between two neurons x_i and x_j separated by a distance d_{ij} is given by:

$$P(C_{ij}|d_{ij}) = ce^{-bd_{ij}} \quad (2)$$

Figure 2a shows how the parameters of the distance rule affect the stability of the network. Low ranges of inhibition decrease the number of stable modes as seen in the left and bottom edges. Figure 2 shows how the decrease in size of the network (i.e. lowering the number of observed neurons from the simulation) reduce the number of unstable modes (real eigenvalues > 0) until the critical point of zero. At this point the observed subnetwork is completely stable (Fig. 2b). In contrast, when we have 2 populations, one with double the range of the first, while modifying the gain the neural activity is compressed, highlighting some perceptions and inhibiting some others (Fig. 2c), although there are no direct effects in the stability of the network.

All together, the modulation of the proportions between both inhibitory populations could provide a mechanism to increase or uncover the number of stable points in the network, while temporarily compressing signals, in favor of less common sensory representations. If we considered that the equilibrium points in a neocortical population correspond to the stable representations that the network has learned, we could find the modulation of the different inhibitory populations convenient for escaping local minima and finding more relevant equilibrium points. Conveniently, ACh seems to be a neuromodulator intrinsically related with the regulation of inhibitory activity, although probably not the only one.

2.2 Acetylcholine Can Drive Towards Metastable States

Recent work has shown how PVi+ preferentially express the M1 muscarinic acetylcholine receptor (mAChR), while SSTi+ express the nicotinic acetylcholine receptors (nAChR) [3]. Both receptors are greatly involved in depolarizing both excitatory and inhibitory cells, mAChR having a faster response and nAChR being slower, globally disinhibiting the excitatory population of the neocortex after ACh release, through inhibition of PV+ cells. Together with the changes in network stability reported in the previous section, this evidence suggests a general role for acetylcholine in the regulation of both gain and stability in the neocortex. From the one side, we have shown that the modulation of short- and

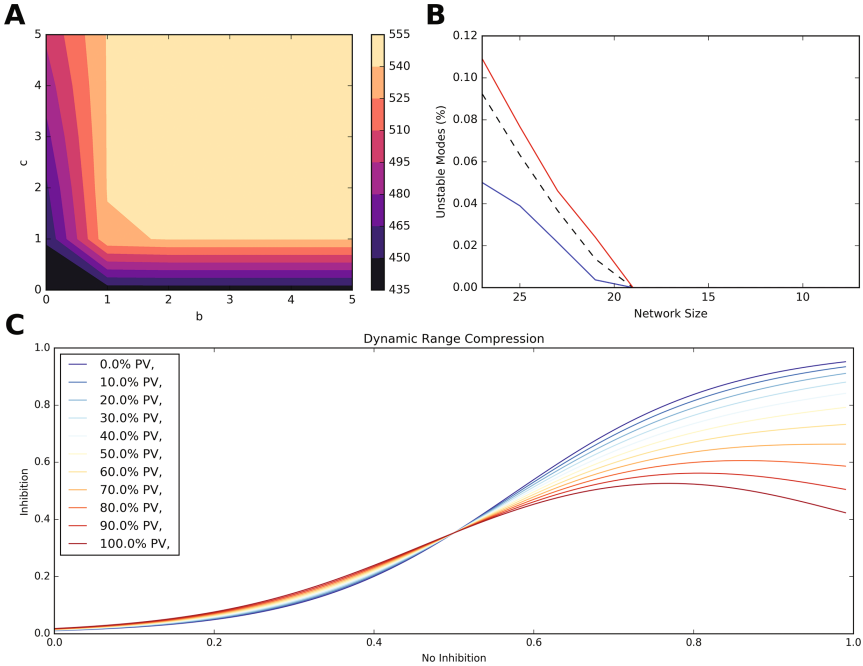


Fig. 2. Characterization of cortical inhibition. Figure A visualizes the number of stable eigenvalues (< 0) found in function of the parameters of the distance rule. B shows the decrease in the number of eigenvalues when studying the stability of reduced portions of the network. C shows how inhibition affects the excitatory dynamics, this is how biasing the inhibition synaptic gains towards decreased global inhibition (SST), strongly reduces the firing rate of the most active neuron, compressing the data and allowing to see better other stimuli.

long-range inhibitory populations affects the stability of a neural network. From the other side, ACh neuromodulation rapidly makes unstabilizes local attractors, to later stabilise them. This can also be understood as a mechanism that dynamically unstabilizes the attractor state from local to global features, with a potential role in context switching. This could then further increase learning speed by setting the brain in an unstable state that promotes switching from pre-learned attractor states to new potential representations.

3 Conclusions and Future Work

In this study we tested this hypothesis by analysing how the stability of a linear neural network is affected by different kinds of inhibition. We further discuss how this mechanism can affect learning speed in critical situations where acetylcholine is typically released, i.e. during dangerous or surprising events. In this paper we have introduced a preliminary analysis of the role of cholinergic modulation of

the inhibitory substrate of the neocortex. We have found that the modulation produced by acetylcholine favors the exploration of the perceptual space by compressing the neural signals and opening more equilibria to encapsulate novel perceptions.

We argue that this mechanism is key for speeding up learning, as it switches between two behavior-state dependent modes: the more frequent, stable mode that provides a coherent representation of the word, and a rare, unstable mode that rapidly embeds new perceptions or even knowledge during critical events. Moreover, physiological experiments show acetylcholine release not only in fearful events, but also in the presence of *unexpected* rewarding or aversive stimuli and during sustained attention. The conclusions presented in this short dissertation support that the mechanisms for switching between stable, exploitative mind states and more exploratory states would be useful to maintain attentive states and respond to surprising, unexpected events.

The inclusion of this mechanism in robotic or other artificial agents, will provide the substrate for autonomous learning, by allowing to detect future potential dangers or errors with enough time to react or stop. Following studies will aim to test the presented hypothesis in a computational model of cholinergic-based learning in a robotic platform.

References

1. Herreros, I., Verschure, P.F.: Nucleo-olivary inhibition balances the interaction between the reactive and adaptive layers in motor control. *Neural Netw.* **47**, 64–71 (2013)
2. Lake, B.M., et al.: Building machines that learn and think like people. arXiv preprint [arXiv:1604.00289](https://arxiv.org/abs/1604.00289) (2016)
3. Disney, A.A., Aoki, C., Hawken, M.J.: Gain modulation by nicotine in macaque V1. *Neuron* **56**, 701–713 (2007). doi:[10.1016/j.neuron.2007.09.034](https://doi.org/10.1016/j.neuron.2007.09.034)
4. Rudy, B., Fishell, G., Lee, S., Hjerling-leffler, J.: Three Groups of Interneurons Account for Nearly 100% of Neocortical GABAergic Neurons (2010). doi:[10.1002/dneu.20853](https://doi.org/10.1002/dneu.20853)
5. Kawaguchi, Y., Kubota, Y.: GABAergic cell subtypes and their synaptic connections in rat frontal cortex. *Cereb. Cortex* **7**, 476–486 (1997). doi:[10.1093/cercor/7.6.476](https://doi.org/10.1093/cercor/7.6.476)
6. Klinkenberg, I., Sambeth, A., Blokland, A.: Acetylcholine and attention. *Behav. Brain Res.* **221**, 430–442 (2011)