Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures

Frédéric Alexandre, Maxime Carrere

Modeling Neuromodulation as a Framework to Integrate Uncertainty in General Cognitive Architectures

Frédéric Alexandre[1,2,3] and Maxime Carrere[2,1,3]

[1]Inria Bordeaux Sud-Ouest, 200 Avenue de la Vieille Tour, 33405 Talence, France
[2]LaBRI, Université de Bordeaux, Bordeaux INP, CNRS, UMR 5800, Talence, France
[3]Institut des Maladies Neurodégénératives, Université de Bordeaux, CNRS, UMR 5293, Bordeaux, France
`frederic.alexandre@inria.fr`

Abstract. One of the most critical properties of a versatile intelligent agent is its capacity to adapt autonomously to any change in the environment without overly complexifying its cognitive architecture. In this paper, we propose that understanding the role of neuromodulation in the brain is of central interest for this purpose. More precisely, we propose that an accurate estimation of the nature of uncertainty present in the environment is performed by specific brain regions and broadcast throughout the cerebral network by neuromodulators, resulting in appropriate changes in cerebral functioning and learning modes. Better understanding the principles of these mechanisms in the brain might tremendously inspire the field of Artificial General Intelligence. The original contribution of this paper is to relate the four major neuromodulators to four fundamental dimensions of uncertainty.

Keywords: Neuromodulation, bio-inspiration, decision making

1 Introduction

Computational Neuroscience has contributed many models of cognitive functions [32], emphasizing their implementation in cerebral circuits or underlying neuronal mechanisms, notably related to learning. Nevertheless, from this huge amount of local models of cognitive functions, defining a general cognitive architecture, really autonomous and versatile is not just a matter of integration and a lot of work remains to be done to propose an effective framework of the brain, seen as a global cognitive model at this systemic level.

One of the more puzzling capacities of human cognition is our ability to adapt to any new circumstances without explicit labeling that the circumstances have changed and require a specific adaptation. This is particularly critical in our dynamic and stochastic world: Does an unusual event correspond to noise and can be neglected, or announce an important change and must be analyzed with

care ? Whereas we are rather good at answering such questions, they turn out to be difficult to integrate in modeling studies.

This is revealed in the stability-plasticity problem, studied in the domains of Artificial Neural Networks [19], Computational Neuroscience [23] and Lifelong Machine Learning [30]. In short, this problem originates from the fact that we would like at the same time to have a learning system able to adapt to any change (plastic) and preserving its past experience and not deleting it, in case any exception occurs (stable). A related dilemma is that of exploitation-exploration, also encountered in many domains [6]. On the one hand we might have the will to exploit our current knowledge of the present situation but on the other hand we might also want to explore other recipes and possibly find more efficient solutions.

This problem has been hardly studied in Cognitive Science. To our knowledge, [14] is one of the rare experimental studies to analyze human cerebral activations for such conflicts during decision making. Cohen and colleagues [11] underline that in reinforcement learning, formal solutions of optimal policies have been proposed only in simplified cases of this problem, particularly corresponding to discovering noisy but stationary processes with fixed probabilities of reward, but not in the case of unstationarity whereas changing environments are probably more realistic as far as human cognition is concerned.

The case of unstationary environments has been tackled in [17] formalizing them as a set of stationary environments and a certain probability to switch from an environment to the other. Accordingly, the very interesting MMBRL model has been proposed [17], as a multi-modules model, each specialized on an environment, learned by Q-learning, and predicting the current state. This model is very original because at each moment, all the modules participate in the decision and learn, but the number of environments must be fixed in advance and cannot evolve.

In the framework of decision making, the distinction between stationary and unstationary environments can be presented as follows: when you try to associate an action to a stimulus to get a reward, you might be willing to apply an associative rule that you have learned before, proposing, for a given stimulus, the best action to trigger to get a reward. Now, suppose that suddenly the rule does not work any longer (you get no reward): Which conclusion should you draw from this failure ? On the one hand, you might decide that the reason is that the environment is stationary but the rule is probabilistic (not always valid). In this case, the decison to make is only to revise its rate of validity and wait for the next trial (hopefully more rewarding). On the other hand, you might decide that the environment is unstationary and that the rule has suddenly changed. In this case, you will have to elaborate a new rule or re-use an old one. It is explained in [33] that these two interpretations are intricately linked. For example, if a rule is highly stochastic, you will be less prone to consider a failure as corresponding to unstationarity. Conversely a failure with a highly probable rule will be immediately considered as requiring a new rule.

The paper by Yu and Dayan [33] had a big impact in the computational neuroscience community because it was also proposing that two neuromodulators in the brain, Acetylcholine (ACh) and Noradrenaline (or Norepinephrine NE) were respectively signaling these two kinds of uncertainty. This view is partly consistent with the wider view by Doya [16], proposing more generally a role for neuromodulators in reinforcement learning.

In the present paper, based on biological knowledge extracted from the literature and on models and experimental simulations including from our group, we propose to gather information about the main neuromodulators and insert them in a more systematic framework, particularly relating their role to the adaptation of the cerebral circuitry and the underlying cognitive functions, to the kind of uncertainty and to the part of the behavioral episode concerned with experienced uncertainty. Our goal is to show that neuromodulation in the brain is a very powerful way to drive it in different functioning and learning modes, at a lower cost regarding the complexification of the cognitive architecture. We will also explain that this is made possible because the nature of the information carried by neuromodulators is well adapted to such global modulations. We will finally propose that this information is not so heterogenous as classically reported but rather corresponds to adapting the cerebral network to different kinds of uncertainty.

2 The role of neuromodulation in neural processing

Information flows that feed a neural network can have different kinds of effects on its functioning and on its learning. Based on biological inspiration, three main flows can be considered. Friston has contrasted driver and modulator flows [18], corresponding respectively to sensory feed-forward afferent flows and to feed-back flows carrying expectations from higher regions of the network. Whereas the driver flows are directly integrated in the functioning rules of the neuronal units (classically in the weighted sum performed therein) and have a major (driving) impact on the resulting activation states, the modulator flows, as their name tells, have a more limited impact. They cannot generate an activation by themselves but can modulate an existing activation, by acting on some internal parameters like the gain of the activation rules. Relying on sparser and less focused connectivity, the feed-back flows can have consequently an attentional effect that can modify the level of activity in some regions. In addition, refering to the classical hebbian framework, it is also clear that both kinds of flows have distinct roles in the learning rules.

In natural and artificial networks, neuromodulation can also have a distinct effect on neural networks but is more rarely considered in modeling approaches. The main types of neuromodulators are monoamines (dopamine, serotonin, noradrenaline) and acetylcholine. The underlying mechanism [8, 15] is that, for each kind of neuromodulators, specific neurons gathered in a nucleus project to most regions of the brain. Accordingly, they can widely modify the functioning and learning modes of extended networks by modifying intrinsic properties of neu-

rons and synaptic weights. This is done thanks to specialized receptors types on different parts of neurons. This kind of information broadcasting is interesting in a distributed architecture like the brain, where the anatomy of interconnections is rather stable and cannot be systematic but where some important information has to be known in many regions to modulate their functioning and learning modes [15]. Basically, neuromodulatory neurons can have two kinds of activity, a phasic activity corresponding to a reaction synchronized to a specific event, typically the reception of a reward, and a tonic activity corresponding to a sustained release of the neuromodulator. Whereas phasic activity of neuromodulatory neurons can have dramatic roles in brain functioning (cf. for example the phasic dopamine, often described as representing the reward prediction error [29]), their tonic activity is more consistent with the modulatory role that we discuss here and will be only considered in the remaining of the paper.

In order to introduce the influence of neuromodulators on the normal functioning of the cerebral structures, let us first rapidly define what we call a 'normal' (or nominal) functioning of the cognitive architecture. In short, an important part of our deliberative cognition consists in analyzing current exteroception and interoception to detect respectively important information in the external world (sensory cues) and in the internal world (needs) and to apply corresponding sensorimotor rules to satisfy currents goals [1]. Here, sensorimotor rules must be understood as relations encoded in the prefrontal cortex associating perception to the result obtained when a decision is executed, which can correspond to trigger a motor action in the external world or to trigger more internal decisions. In the ideal case, the world is perfectly known and the corresponding important sensory cues and rules have already been extracted by learning. But the world is not ideal but uncertain, as it can be detected by reward prediction errors that monitor uncertainty.

An important characteristic of the neuromodulatory systems is that they are triggered by the central nucleus of the amygdala, a limbic structure known to be active when an error of (negative or positive) reward prediction is detected. As explained below, other evaluations are made in other cerebral regions to differentiate the kind of uncertainty and inhibit non relevant neuromodulation.

3   Four kinds of neuromodulators

Acetylcholine (ACh) is released by basal forebrain nuclei like the nucleus basilis. Its role has been related to expected uncertainty [33], corresponding to the stochastic case evoked above, when the rule has not changed but is not always valid. The observed effects of ACh in the sensory cortex are attentional and correspond to increase the signal-to-noise ratio [27], resulting in promoting feed-forward sensory information against feed-back expectations, when the environment is judged highly stochastic. As modeled in [10], another observed effect of ACh in conditioning experiments [9] is to promote learning about the context and not about (noisy) sensory cues.

Noradrenaline (or Norepinephrine NE) is released by the Locus Coeruleus and has been associated to unexpected uncertainty [33] when the rules have changed, like it is the case during reversal [2]. Accordingly, it is also associated to an increase of exploration, to extract new contingencies and elaborate new rules [3]. NE has been reported to have the same kinds of attentional effects as ACh to favor exploration of new sensory cues [33] The identification of the circumstances as corresponding to unexpected uncertainty has been modeled in [24] by a signal depending on the reward rate and on measures of response conflict estimated from two windows encompassing long term and short term history of activity.

Concerning the other two neuromodulators considered here, Serotonin (or 5HT) released from the dorsal raphe nucleus and Dopamine (DA) released from the Ventral Tegmental Area in the midbrain, the situation is more controversial. Old models like [13] propose a dual role for these neuromodulators, with tonic DA for average appetitive reward prediction and 5HT for average aversive reward (punishment) prediction. More recent papers acknowledge a role for 5HT in aversive processing [12] but mainly for behavioral inhibition and passive avoidance. Accordingly, low levels of 5HT are also associated to impulsivity. This is consistent with the view expressed in [16] relating 5HT to time discounting, corresponding to the degree of preference between immediate over delayed rewards, higher levels of 5HT corresponding to greater patience. This kind of impulsivity can also be linked to the concept of risk, another major theory for the role of 5HT [5], where a greater variance in rewards might be prefered to more stable payoffs, even if disavantageous on the total reward received. Defining a utility function as the trade-off bewteen the expected reward and its variance is proposed as a model for the level of tonic 5HT [5]. Concerning tonic DA, recent papers generally agree for a role to set the trade-off between exploration and exploitation [21], which has of course to be confronted to the hypothetized role of NE mentioned above. The evoked model of tonic DA implements the corresponding mechanisms in the output functions of the basal ganglia, a motor region responsible for action selection, with a probability distribution that can be modulated by tonic DA from flat (to favor exploration) to sharp (for exploitation) [21] shapes.

4   Towards a common framework of interpretation

We propose that setting the focus on the concept of action is a good way to make clear the ambiguities evoked above. This idea has also been proposed by Y. Niv [25] to better explain the differences between the pavlovian and instrumental conditioning paradigms. Instead of evoking mainly rewards, actions should be considered as representing the main difference between both paradigms: The new dimension brought by instrumental conditioning (and related decision making) over pavlovian conditioning (and related reward value) is that the animal or human agent is responsible for deciding to trigger an action and more precisely for its frequency, yielding the rapidity or the vigor of this action, but also a certain

cost corresponding to the energy necessary to trigger the action. Consequently the vigor of the response (and the corresponding energetic price to pay for it) is a good indicator of the motivation to have a reward and, similarly, the cost of a delay in this procedure is more meaningful.

We propose to exploit the same duality between the rate of the action and the reward as a way to disambiguate the respective roles of the neuromodulators considered here and propose accordingly to go deeper in the specification of different kinds of uncertainty. We have evoked several times above that uncertainty manifests itself by the acknowledgement that the sensorimotor rule which has been applied doesn't bring the expected reward. The next question would be to know what is the cause of this problem: Is it because the selected sensory cues is not longer valid (unexpected sensory uncertainty) or just because it is only noisy (expected sensory uncertainty) ? Or is it because the action to associate to the (still correct) sensory cue has changed (unexpected motor uncertainty) or just because it is only noisy and not always working (expected motor uncertainty). We propose here that these cases respectively trigger the release of Noradrenaline, Acetylcholine, Dopamine and Serotonin.

This view does not modify the acknowledged role of ACh on sensory processing and is also consistent with information given above for tonic DA and NE roles. DA would be for exploration/exploitation trade-off for action selection whereas NE would be concerned with the same trade-off for selective attention of the sensory cues. This is consistent with the fact that DA effect is mainly reported in the basal ganglia and the frontal cortex (known to be responsible for the organization of action) and NE effect in the posterior parietal cortex [4]. The role of NE to modify sensory processing in the thalamus and the cortex is also acknowledged [28]. Similarly, DA has a clear role for wanting in the motor pole and not for liking on the sensory side [7]. This separation of roles between NE and DA has also been proposed in modeling studies [21].

Similarly, relating the role of 5HT to the estimation of risk and variance of reward can rather be seen as a matter of noise or probabilistic distribution of effects, consistently with an interpretation of expected uncertainty related to actions. Even older interpretations of DA and 5HT as respectively related to reward and punishment can be re-examined here: It can be understood that when the risk is high, this should promote behavioral inhibition and passive avoidance, often related to aversive processing [12], whereas exploration due to high level of tonic DA might be misinterpreted as behavioral activation to get a reward.

5   Discussion

In this paper we have mainly discussed about the role of neuromodulators in cognition. Structurally, they have been presented [15] as a way to broadcast information in a sparsely connected network like it is the case for the brain. Functionally, it can be also argued that this kind of information passing is preferable for certain types of information, orthogonal to the information processed locally

and rapidly changing. This is the case with uncertainty, that we have proposed to be the key topic of neuromodulation. More precisely, we have proposed a double set of criteria to qualify uncertainty. In addition to uncertainty announcing a radical change in the underlying rules or simply reporting stochasticity in the environment, as already proposed in [33], we suggest, as the major contribution of the present paper, that uncertainty can also refer to the sensory cues or to the actions. In this view, each of the resulting four kinds of uncertainty would require fundamentally different modifications in the cerebral network and consequently different neuromodulators.

Altogether, the contribution of the present view to the framework of Artificial General Intelligence can be synthesized as follows: Deliberative decision-making is often summarized as learning and exploiting sensorimotor rules, associating the sensory context to the best actions to trigger to get rewards and reach goals. Such a framework is very classical in machine learning, particularly corresponding to reinforcement learning [31], and in cognitive science [1]. It is attractive because of its simplicity and because several computational frameworks have been developed to express and learn such rules, using symbols [1], neurons [31] and bayesian formalism [33] and proposing accordingly effective means to implement artificial intelligent agents. This simplicity can also be seen as a weakness, reducing cognition to simple sensorimotor rules. One argument is that adapting to the unknown and changing world is more complex than learning simple contingencies between situations and optimal decision to make, particularly because the world is uncertain, including dynamic and stochastic aspects. A really autonomous agent, in the perspective of Artificial General Intelligence, should be able to detect by itself such uncertainties and adapt to them.

The first lesson that we propose to be drawn by inspiration from neuromodulation in the brain is that, in order to take uncertainties into account in a cognitive architecture, there is no need to give up the framework of sensorimotor rules and look for a completely different or radically more complex framework. Alternatively, it can be sufficient to build additional modules to detect the kind of experienced uncertainty and to modulate or increment the set of sensorimotor rules. This is interesting because existing systems based on sensorimotor rules could be simply extended (and not fundamentally modified) and remain effective in an uncertain world, corresponding to an important improvement in autonomy for the corresponding agent.

The second element that we introduce here, corresponding to the most original contribution of this paper, is about the different kinds of uncertainties that can be experienced. In addition to the distinction between expected and unexpected uncertainties already proposed in [33], we propose here that the fact that uncertainty applies either to sensory or to motor dimensions is of prime importance. In our view, the resulting four kinds of uncertainty should be signaled by the four major neuromodulators in the brain, allowing for distinct modulation of the cerebral system and, accordingly, of the existing set of sensorimotor rules. Although we have given above some hints from existing models about the kind

of modulation performed by each neuromodulator in the cerebral system, a more precise specification of these effects should be the topic of future work.

Another key topic hardly evoked here, is about the identification of the kind of uncertainty experienced by the agent and consequently of the kind of neuromodulator to release. We have evoked models performing this identification from history of response conflicts and reward errors [24]. The way these elements are estimated and determination of their cerebral encoding is the topic of ongoing research in the team. It can be also mentioned that metalearning occurs at this level to help determine as fast and accurately as possible this critical information, particularly depending on the context in which it occurs [26].

This tentative explanation for the role of neuromodulators in cognition needs to be confirmed by a deeper anchoring in neuroscience. Additional clues might be particularly searched in the nucleus accumbens, known to be the gateway between sensory limbic and motor sides of cognition [22] and more generally in the basal ganglia concerning the respective roles of DA and 5HT in the balance between inhibition and excitation of behavior [20]. Finally, an important issue to consider in forthcoming works in about interneuromodulatory interactions: It has been shown that many interactions exist between neuromodulators resulting in more intricate roles than presented here [15].

References

1. Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychol Rev 111(4), 1036–1060 (Oct 2004), `http://dx.doi.org/10.1037/0033-295x.111.4.1036`

2. Aston-Jones, G., Rajkowski, J., Kubiak, P.: Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. Neuroscience 80(3), 697–715 (Jul 1997), `http://dx.doi.org/10.1016/s0306-4522(97)00060-2`

3. Aston-Jones, G., Cohen, J.D.: An integrative theory of Locus Coeruleus-Norepinephrine function: Adaptive Gain and Optimal Performance. Annual Review of Neuroscience 28(1), 403–450 (2005), `http://dx.doi.org/10.1146/annurev.neuro.28.061604.135709`

4. Aston-Jones, G., Rajkowski, J., Cohen, J.: Role of locus coeruleus in attention and behavioral flexibility. Biological Psychiatry 46(9), 1309–1320 (Nov 1999), `http://dx.doi.org/10.1016/s0006-3223(99)00140-7`

5. Balasubramani, P.P., Chakravarthy, V.S., Ravindran, B., Moustafa, A.A.: An extended reinforcement learning model of basal ganglia to understand the contributions of serotonin and dopamine in risk-based decision making, reward prediction, and punishment learning. Frontiers in computational neuroscience 8 (2014)

6. Berger-Tal, O., Nathan, J., Meron, E., Saltz, D.: The exploration-exploitation dilemma: A multidisciplinary framework. PLoS ONE 9(4), 1–8 (04 2014)

7. Berridge, K.C., Robinson, T.E.: What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research Reviews 28(3), 309–369 (Dec 1998), `http://dx.doi.org/10.1016/s0165-0173(98)00019-8`

8. Bouret, S., Sara, S.J.: Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends in Neurosciences 28(11), 574–582 (2005), `http://dx.doi.org/10.1016/j.tins.2005.09.002`

9. Calandreau, L., Trifilieff, P., Mons, N., Costes, L., Marien, M., Marighetto, A., Micheau, J., Jaffard, R., Desmedt, A.: Extracellular hippocampal acetylcholine level controls amygdala function and promotes adaptive conditioned emotional response. The Journal of neuroscience : the official journal of the Society for Neuroscience 26(52), 13556–13566 (Dec 2006)

10. Carrere, M., Alexandre, F.: A pavlovian model of the amygdala and its influence within the medial temporal lobe. Frontiers in Systems Neuroscience 9(41) (2015)

11. Cohen, J.D., McClure, S.M., Yu, A.J.: Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 362(1481), 933–942 (May 2007), `http://dx.doi.org/10.1098/rstb.2007.2098`

12. Cools, R., Nakamura, K., Daw, N.D.: Serotonin and dopamine: unifying affective, activational, and decision functions. Neuropsychopharmacology 36(1), 98–113 (Jan 2011), `http://dx.doi.org/10.1038/npp.2010.121`

13. Daw, N.D., Kakade, S., Dayan, P.: Opponent interactions between serotonin and dopamine. Neural Networks 15(4-6), 603–616 (Jun 2002), `http://dx.doi.org/10.1016/s0893-6080(02)00052-7`

14. Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J.: Cortical substrates for exploratory decisions in humans. Nature 441(7095), 876–879 (Jun 2006), `http://dx.doi.org/10.1038/nature04766`

15. Dayan, P.: Twenty-Five Lessons from Computational Neuromodulation. Neuron 76(1), 240–256 (Oct 2012), `http://dx.doi.org/10.1016/j.neuron.2012.09.027`

16. Doya, K.: Metalearning and neuromodulation. Neural Networks 15(4-6), 495–506 (Jun 2002), `http://dx.doi.org/10.1016/s0893-6080(02)00044-8`
17. Doya, K., Samejima, K., Katagiri, K.i., Kawato, M.: Multiple Model-Based Reinforcement Learning. Neural Comp. 14(6), 1347–1369 (2002), `http://neco.mitpress.org/cgi/content/abstract/14/6/1347`
18. Friston, K.: Functional integration and inference in the brain. Progress in neurobiology 68(2), 113–143 (Oct 2002), `http://view.ncbi.nlm.nih.gov/pubmed/12450490`
19. Grossberg, S.: Adaptive Resonance Theory: How a brain learns to consciously attend, learn, and recognize a changing world. Neural Networks 37, 1–47 (Jan 2013), `http://dx.doi.org/10.1016/j.neunet.2012.09.017`
20. Haber, S., Fudge, J., McFarland, N.: Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. The Journal of Neuroscience 20(6), 2369–2382 (2000)
21. Humphries, M.D., Khamassi, M., Gurney, K.: Dopaminergic control of the exploration-exploitation trade-off via the basal ganglia. Frontiers in Neuroscience 6(9) (2012)
22. Mannella, F., Gurney, K., Baldassarre, G.: The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. Frontiers in behavioral neuroscience 7 (2013)
23. McClelland, J.L., McNaughton, B.L., O'Reilly, R.C.: Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. Psychological review 102(3), 419–457 (1995)
24. McClure, S., Gilzenrat, M., Cohen, J.: An exploration-exploitation model based on norepinepherine and dopamine activity. In: Weiss, Y., Schölkopf, B., Platt, J. (eds.) Advances in Neural Information Processing Systems 18, pp. 867–874. MIT Press (2006), `http://www.csbmb.princeton.edu/\~{}smcclure/pdf/MGC\_NIPS.pdf`
25. Niv, Y.: Cost, Benefit, Tonic, Phasic: What Do Response Rates Tell Us about Dopamine and Motivation. Annals of the New York Academy of Sciences 1104(1), 357–376 (May 2007), `http://dx.doi.org/10.1196/annals.1390.018`
26. Pauli, W.M., Hazy, T.E., O'Reilly, R.C.: Expectancy, Ambiguity, and Behavioral Flexibility: Separable and Complementary Roles of the Orbital Frontal Cortex and Amygdala in Processing Reward Expectancies. Journal of Cognitive Neuroscience 24(2), 351–366 (2011), `http://dx.doi.org/10.1162/jocn\_a\_00155`
27. Pauli, W.M., O'Reilly, R.C.: Attentional control of associative learning–a possible role of the central cholinergic system. Brain Research 1202, 43–53 (Apr 2008)
28. Sara, S.J., Bouret, S.: Orienting and Reorienting: The Locus Coeruleus Mediates Cognition through Arousal. Neuron 76(1), 130–141 (Oct 2012), `http://dx.doi.org/10.1016/j.neuron.2012.09.011`
29. Schultz, W.: Predictive reward signal of dopamine neurons. Journal of Neurophysiology 80(1), 1–27 (1998), `http://jn.physiology.org/content/80/1/1`
30. Silver, D., Yang, Q., Li, L.: Lifelong machine learning systems: Beyond learning algorithms. In: AAAI Spring Symposium Series (2013)
31. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An introduction. MIT Press (1998)
32. Trappenberg, T.P.: Fundamentals of Computational Neuroscience (2. ed.). Oxford University Press (2009)
33. Yu, A.J., Dayan, P.: Uncertainty, Neuromodulation, and Attention. Neuron 46(4) (2005)