

The role of serotonin

Nisheet Patel

9th September, 2022

Abstract

Neuroscience has established at least one fairly concrete role for dopamine in the brain, namely, that it indicates prediction errors for reward signals. Serotonin, on the other hand, has eluded any such reductionism for several decades now. After hearing some of Zach's recent thoughts, I think there's very strong reason to believe not only that such a reduction might be possible, but also that it's likely within our reach. In this highly exploratory essay, we're going to explore the conjecture that serotonin reflects prediction errors for states. This begs the questions of what the nature of this prediction error might be, what functional purpose it might serve. Eventually, we're going to attempt to construct a framework to investigate these questions.

Musings of a theorist

Before we dive in and get our hands dirty on any sort of theoretical development, I think it's worth going through some thought experiments to examine our conjecture and evaluate its implications. As a first of these, I think it's worth shining the lamp of learning schemas to contextualize the problem a little better. Broadly speaking, machine learning researchers often classify learning into schemas of supervised, unsupervised, and reinforcement learning (RL). Of these, we have a half-decent understanding of how RL might be implemented in the brain and strong evidence for at least the temporal difference framework. We also have some inklings of unsupervised learning in the form of the predictive coding theory which people have proposed probably since the dawn of theoretical neuroscience. Supervised learning, though, is a thing of its own and probably not worth going into the details of. Among other reasons, it would require a "true" label in a world devoid of reality, which is just merely an illusion that we agree upon.

For the purpose of this essay, we will draw analogies between the role of dopamine for reinforcement learning and that of serotonin for unsupervised learning, perhaps implemented with the predictive coding framework. Broadly speaking, we hope to eventually validate or invalidate the broad conjecture that serotonin reflects temporal-difference (TD) state prediction errors (SPE) for the predictive coding framework, just like dopamine reflects temporal-difference (TD) reward prediction errors (RPE) for the reinforcement learning (RL) framework. Note that this framing in and of itself implies that serotonin might affect behavior during a reward-based task since the RL framework requires the notion of a state, but not the other way around, at least not directly, since the prediction coding framework does not require the notion of a reward.

Although the concrete predictions might depend on the details of the implementation, for now we're going to abstract away from (*what Pablo once described as*) the essential triad for a learning system - the learning rule, architecture of the system, as well as the optimization algorithm - and focus on the broad implications of the conjecture by leaning into the analogy and deriving inspiration for ideas from it.

DA ↔ RL

Consider the DA ↔ RL story. What we know is that dopaminergic neurons in the VTA reflect the temporal difference (TD) reward prediction error (RPE) in their firing rate. Once the RPE is in, RL may happen in many different ways. For instance, it may be used for TD(0) or TD(λ) with some fixed policy once the values are computed or for an Actor-Critic like setup with a potentially stochastic actor, both of which have been heavily explored by neuroscientists. Some recent ideas also propose that not all DA neurons may signal the same scalar error signal for the mean reward, but they may be involved in some effort to capture the whole distribution of the scalar reward signal, and/or that they may also be capturing different time-scales of the reward signal.

One line of thought that at least I've never carefully considered before is that the following. The error itself is reflected by the firing rate of the dopaminergic neurons in the VTA. Because they are dopaminergic neurons,

when they fire an action potential, they release dopamine as a neurotransmitter directly at the synaptic cleft to the next neuron(s) with a probability $p \approx 0.1$ (as per Sarah from Bellone lab). However, the RPE itself is only directly related to dopamine as a neurotransmitter and not to dopamine as a neuromodulator, as far as I understand. Thus, if any global changes result from it, they might be due to the well-established dopaminergic pathways, which I have no idea about.

A lot of this is still unclear to me, so I have noted some information in the box below to clarify some of these differences. And I will actively update it with new information that I find. I have also tried to include some potential links to ideas from machine learning for building intuition. Feel free to skip to the next section on Serotonin ↔ Predictive Coding if you feel like.

Understanding chemical signalling the brain

Neurotransmitters and neuromodulators are two types of chemical substances released by neurons in the brain. It turns out that both the protagonists of our story, serotonin and dopamine, can act as both neurotransmitters as well as neuromodulators! Unfortunately, the definitions are quite muddled and overlapping at least in my head, so I'm going to lay out the basics here to try and understand the mechanistic definitions of the two and differences before delving into reductionist theories of their global function.

Neurotransmitter vs neuromodulator

Generally speaking, neurotransmitters send signals directly to a target neuron via synapses much like weights in an artificial neural network, whereas neuromodulators are a subset of those neurotransmitter molecules which may be released in the fluid surrounding the neurons to coordinate a global change with many different neurons in many different regions to modulate some properties of their local synaptic transmission. One stark difference between the two is in the time-scale of their effects. While neurotransmitters work on relatively fast time-scales (seconds?), neuromodulators can only induce slow, long-lasting changes.

The neuromodulator doesn't really have an analogue in machine learning, but I've heard science-daddy Terry Sejnowski proposing this idea of transforming a weight vector \vec{w} by some function f to get $f(\vec{w})$ to allow exploiting the same network for multiple use-cases. This idea is super interesting and could be thought as one way to induce and learn some useful invariances that are useful for tasks and that make these circuits modular and resource-efficient. Perhaps, one example of such an invariance is that of mean-invariance, which is implemented in the brain by what we call divisive normalization and in artificial neural networks by what we call layer-norm. This analogy breaks down when we look at the time-scale of action, but then again, we can leave these speculations as food for thought for our future selves. For now, we will simply end this with the note that one of the major proposed role of neuromodulators is in facilitating adaptation to a rapidly changing environment.

Dopamine

Dopamine is commonly known as one of the brain's neurotransmitters, the fast-acting chemical messengers that are shuttled between neurons. But in the new work (by Assad et al.), dopamine is acting as a neuromodulator. It's a term for chemical messengers that slightly alter neurons to cause longer-lasting effects, including making a neuron more or less likely to electrically communicate with other neurons. This neuromodulatory tuning mechanism is perfect for helping to coordinate the activity of large populations of neurons, as dopamine is likely doing to help the motor system decide precisely when to make a movement.

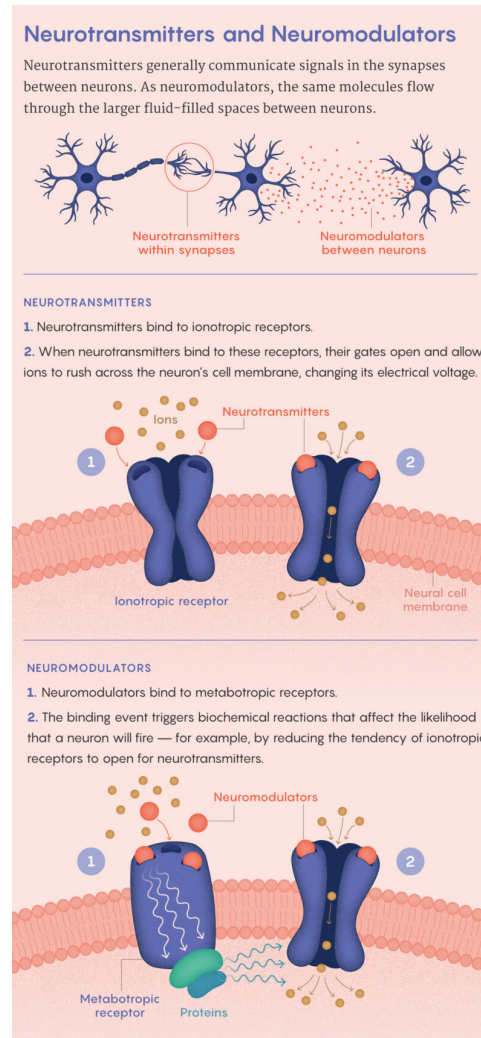
[Source](#)

Serotonin

*The neurotransmitter serotonin (5-HT), widely distributed in the central nervous system (CNS), is involved in a large variety of physiological functions. In several brain regions 5-HT is diffusely released by volume transmission and behaves as **a neuromodulator** rather than as a "classical" neurotransmitter. The modulatory action of 5-HT affect glutamate and γ-amino-butyric acid (GABA), which are the principal neurotransmitters*

mediating respectively excitatory and inhibitory signals in the CNS

[Source](#)



[Source](#)

Contextualizing this to refine our conjecture

With this newfound knowledge, it seems imperative to refine the loose and hitherto unknown (at least to the naive theorist) parts of our conjecture:

The firing rate of serotonergic neurons in the dorsal raphe reflects temporal-difference (TD) state prediction errors (SPEs) for the predictive coding framework, just like the firing rate of dopaminergic neurons in the VTA reflects temporal-difference (TD) reward prediction errors (RPEs) for the reinforcement learning framework.

This begs the question of what dopaminergic neurons in the VTA do, what are the circuits involved, and what are these circuits currently implicated in? And the same for serotonergic neurons in the dorsal raphe. *Questions I never thought I would ask myself; fuck me.*

Serotonin ↔ Predictive coding

Now, consider the serotonin (5HT) ↔ predictive coding story. Analogous to the DA story, what would the nature/form of the state prediction error (SPE) be? And could we perhaps also loosely speculate on how downstream areas would use it? To answer these questions, note that one of the striking difference between the two frameworks is that reward is a scalar signal whereas state is a vector. This leads to the first open question of whether the serotonergic neurons in the dorsal raphe represent a scalar signal or a vector.

On the nature of the SPE

Scalar SPE

If the signal reflected by serotonergic neurons were to be a scalar, then it would need to be squashed down, for instance, with a norm. The exact function required for mapping the vector signal to a scalar would need careful consideration and perhaps a lot of experimental evidence, but let's leave that aside for now as a note for our future selves.

Vector SPE

Alternatively, one could try to argue that, similar to story of the diverse responses of DA neurons trying to facilitate distributional RL (albeit for a scalar reward signal), the many serotonergic neurons in the dorsal raphe code a state prediction error signal where the state itself is multi-dimensional. At first glance, this seems totally ridiculous. Why would the function of such a global signal of releasing serotonin would be able to capture the complexity of a signal as high-dimensional as the state of the world, which comes from countless sets of sensors from diverse sources? However, there are at least two possible counter-arguments that make this scenario worth considering.

First, it is at least possible to at least think otherwise. One could argue that predictive coding could be done with a variational autoencoder that has an immensely compressed latent representation. In such a scenario, it is possible that the state prediction error signal that serotonergic neurons in the dorsal raphe reflect is for this latent layer and hence it is low-dimensional.

Second, in light of my new understanding of the differences between neurotransmitters and neuromodulators, I note that serotonergic neurons in the dorsal raphe also act as neurotransmitters, communicating their information locally to the neurons they are directly connected to. So while there might be well-established pathways that broadcast this signal to potentially the whole brain, the signal in and of itself need not be scalar! As far as I understand, it is not serotonin, the molecule itself, that carries a binary signal of absence or presence all over the brain wherever serotonergic receptors are to be found. The signal itself is carried via a whole set of neurons involved in the relevant pathways, it's just that the neurons happen to be serotonergic. Hence, the signal itself can be multi-dimensional.

On the function of the SPE

The next key question is whether it would at all be useful to have a squashed global signal, and if it would be, what are the possibilities. It seems easier to imagine how it would be useful for other systems to know about the notion of a state and any errors in its conception, for instance, the reward system relies on it to associate a reward with it, and the motor system relies on it to know whether to continue with an action plan or to switch (unless that is a part of the state as well). One such inter-connection that we have discussed and that has also been part of the proposal for the ERC grant is that the reward system might want to pause or slow down learning if the representation of the state itself is highly uncertain. However, in the interest of investigating the conjecture itself in its pure form, we need to look at how the state prediction error signal would be used for the predictive coding framework independently of its influence on other systems.

Continuous representations of state

First, it seems reasonable to think, as it almost follows by definition, that regardless of the dimensionality of the signal, the state prediction error is an indicator of uncertainty (in the relevant dimension). However, consider the dilemma of representing states as continuous or discrete variables. For continuous representations, the idea of uncertainty makes a lot of sense. For instance, if a visual-only creature manages to predict away every temporal slice of visual input (*I'm trying hard not to use the word "pixel"*) correctly at all moments in time, the job of predictive coding system is done.

Discrete representations of state

However, we know that many animals and at least humans form abstractions. They cluster abstract representations as concepts, give them labels, and incessantly argue about their definitions using language. These abstractions, as I see it, are an attempt to discretize the otherwise continuous state representation.

As opposed to continuous representations of states, discrete representations come with a whole lot of complications for our story. For instance, could it be that by indicating SPEs, serotonergic neurons trigger discrete state transitions. In this scenario, if a discrete representation of a state does not exist, could it even trigger the creation of such a new state? One possible idea to investigate this is by looking at Dirichlet processes.

Interim remarks and notes for future

- Ideas for modelling a state prediction system:
 - PredNet
 - Variational autoencoder
 - Ones proposed in RL:
 - successor representation
 - first-occupancy representation
- How should we handle discrete states, and state transitions and creation?
 - Get inspiration from COIN model (Lengyel, Wolpert)
- Look at interactions between reward and state systems

Wilder speculations

Partly based on Alex's whacky claims. There are other broad learning frameworks than the ones we're considering here, for instance, competitive or adversarial learning, where one component is trying to fool the other by generating fake signals and the other is trying to detect whether the source of the signals is the impostor or the external world (some other brain region). This may be useful for self-organization in clusters. Could it be that some other neurotransmitter(s)/modulator(s) facilitate(s) this?