Oral presentation

Solving the distal reward problem through linkage of STDP and dopamine signaling

Eugene M Izhikevich*

Address: The Neurosciences Institute, San Diego, CA, USA

Email: Eugene M Izhikevich* - Eugene.Izhikevich@nsi.edu

* Corresponding author

Learning the associations between cues and rewards (classical or Pavlovian conditioning) or between cues, actions, and rewards (instrumental or operant conditioning) involves reinforcement of neuronal activity by rewards or punishments. Typically, the reward comes seconds after reward-predicting cues or reward-triggering actions, creating an explanatory conundrum known in the behavioral literature as the *distal reward problem* and in the reinforcement learning literature as the *credit assignment problem*. Indeed, how does the animal know which of the many cues and actions preceding the reward should be credited for the reward? In neural terms, in which sensory cues and motor actions correspond to neuronal firings, *how does the brain know what firing patterns, out of an unlimited repertoire of all possible patterns, are responsible for the reward if the patterns are no longer there when the reward arrives*? How does it know which spikes of which neurons result in the reward if *many* neurons fire during the waiting period to the reward? Finally, how does the common reinforcement signal in the form of the neuromodulator dopamine (DA) influence the right synapses at the right time, if DA is released globally to many synapses? Here, I show how the credit assignment problem could be solved in a network of cortical spiking neurons with DA-modulated plasticity.

The model is based on the experimental findings that DA modulates synaptic plasticity by enhancing long-term potentiation (LTP) and long-term depression (LTD): For example, in hippocampus, dopamine D1 receptor agonists enhance tetanus-induced LTP, but the effect disap-

pears if the agonist arrives at the synapses 15–25 seconds after the tetanus, thereby suggesting the existence of a short window of opportunity for the enhancement. My major hypothesis is that DA acts the same way on the spike-timing dependent synaptic plasticity (STDP). That is, a particular order of firing induces a synaptic change (positive or negative), which is enhanced if extracellular DA is present during the critical window of a few seconds.

I show that DA modulation of STDP has a built-in property of instrumental conditioning: It can reinforce firing patterns occurring on a millisecond time scale even when they are followed by rewards that are delayed by seconds. This property relies on the existence of slow synaptic processes that act as "synaptic eligibility traces" or "synaptic tags". These processes are triggered by nearly-coincident spiking patterns, but due to a short temporal window of STDP, they are not affected by random firing during the waiting period to the reward. This "insensitivity" of the synaptic tags to the random ongoing activity during the waiting period is the key feature that distinguishes my approach from previous studies, which require that the network be quiet during the waiting period or that the patterns are preserved as a sustained response. I also discuss why *this mechanism works only when precise firing patterns are embedded into the sea of noise* and why it fails in the mean firing rate models. I also present a spiking network implementation of the most important aspect of the temporal difference (TD) reinforcement learning rule – the

shift of reward-triggered release of DA from unconditional stimuli to reward-predicting conditional stimuli.

This study emphasizes the importance of precise firing patterns in brain dynamics and suggests how a global diffusive reinforcement signal in the form of DA can selectively influence the right synapses at the right time. The model provides a testable prediction on the action of DA on STDP, which will be tested by G. Bi (Pittsburgh University) and R. Froemke (UCSF) (personal communications).

References

1.	Izhikevich EM: **Solving the distal reward problem through linkage of STDP and dopamine signaling.** *Cerebral Cortex* 2007. DOI: 10.1093/cercor/bhl152