BMC Neuroscience

Poster presentation

A spiking temporal-difference learning model based on dopamine-modulated plasticity

Wiebke Potjansu*[1,2], Abigail Morrison[1] and Markus Diesmann[1,3,4]

Address: [1]Theoretical Neuroscience Group, RIKEN Brain Science Institute, Wako City, Saitama, Japan, [2]Institute of Neurosciences and Medicine, Research Center Jülich, Jülich, Germany, [3]Brain and Neural Systems Team, RIKEN Computational Science Research Program, Wako City, Saitama, Japan and [4]Bernstein Center for Computational Neuroscience, Albert-Ludwigs-University, Freiburg, Germany

Email: Abigail Morrison - abigail@brain.riken.jp

* Corresponding author

Making predictions about future rewards and adapting the behavior accordingly is crucial for any higher organism. One theory specialized for prediction problems is temporal-difference (TD) learning. Experimental findings suggest that TD learning is implemented by the mammalian brain. In particular, the resemblance of dopaminergic activity to the TD error signal [1] and the modulation of corticostriatal plasticity by dopamine [2] lend support to this hypothesis. We recently proposed the first spiking neural network model to implement actor-critic TD learning [3], enabling it to solve a complex task with sparse rewards. However, this model calculates an approximation of the TD error signal in each synapse, rather than utilizing a neuromodulatory system.

Here, we propose a spiking neural network model which dynamically generates a dopamine signal based on the actor-critic architecture proposed by Houk [4]. This signal modulates as a third factor the plasticity of the synapses encoding value function and policy. The proposed model simultaneously accounts for multiple experimental results, such as the generation of a TD-like dopaminergic signal with realistic firing rates in conditioning protocols [1], and the role of presynaptic activity, postsynaptic activity and dopamine in the plasticity of corticostriatal synapses [5]. The excellent agreement between the predictions of our synaptic plasticity rules and the experimental findings is particularly noteworthy, as the update

rules were postulated employing a purely top-down approach.

We performed simulations in NEST [6] to test the learning behavior of the model in a two dimensional grid-world task with a single rewarded state. The network learns to evaluate the states with respect to its reward proximity and adapt its policy accordingly. The learning speed and equilibrium performance are comparable to those of a discrete time algorithmic TD learning implementation.

The proposed model paves the way for investigations of the role of the dynamics of the dopaminergic system in reward-based learning. For example, we can use lesion studies to analyze the effects of dopamine treatment in Parkinson's patients. Finally, the experimentally constrained model can be used as the centerpiece of closed-loop functional models.

Acknowledgements

References

1. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward.** *Science* 1997, **275:**1593-1599.
2. Reynolds JN, Hyland BI, Wickens JR: **A cellular mechanism of reward-related learning.** *Nature* 2001, **413:**67-70.

3. Potjans W, Morrison A, Diesmann M: **A spiking neural network model of an actor-critic learning agent.** *Neural Computation* 2009, **21:**301-339.
4. Houk JC, Adams JL, Barto AG: *A model of how the basal ganglia generate and use neural signals that predict reinforcement MIT Press, Cambridge, MA;* 1995.
5. Reynolds JN, Hyland BI, Wickens JR: **Dopamine-dependent plasticity of corticostriatal synapses.** *Neural Networks* 2002, **15:**507-521.
6. Gewaltig M-O, Diesmann M: **NEST (neural simulation tool).** *Scholarpedia* 2007, **2:**1430.