Goal-directed Learning with Reward Modulated Interaction between Striatal and Cerebellar systems

Sakyasingha Dasgupta, Florentin Wörgötter and Poramate Manoonpong

Bernstein Center for Computational Neuroscience, University of Göttingen,Germany.

Abstract

Goal-directed decision making in biological systems is broadly based on associations between conditional and unconditional stimuli. This can be further classified as classical conditioning (correlation based learning) and operand conditioning (reward-based learning). A number of computational and experimental studies have well established the role of the basal ganglia (striatal system) towards reward-based learning, where as the cerebellum evidently plays an important role in developing specific conditioned responses. Although, they are viewed as distinct learning systems [1], recent animal experiments point towards their complementary role in behavioral learning, and also show the existence of substantial two-way communication between the two structures [2]. Based on this notion of co-operative learning, in this work we hypothesize that the basal ganglia and cerebellar learning systems work in parallel and compete with each other (Figure 1). We envision such an interaction being driven by a simple reward modulated heterosynaptic plasticity (RMHP) rule [3], in order to guide the over all goal-directed behavior. Using a recurrent neural network actor-critic model of the basal ganglia and feed-forward correlation learning model of the cerebellum (input correlation learning-ICO) [4], we demonstrate that the RMHP rule can effectively combine the outcomes of the two learning systems. This is tested using simulated environments of increasing complexity with a four-wheeled animat in a dynamic foraging task. Although, they are modeled within a highly simplified level of biological abstraction, we clearly demonstrate that such a combined learning mechanism, leads to much stabler and faster learning of goal-directed behaviors in comparison to the individual systems.

Acknowledgements

This research was supported by the Emmy Noether Program (DFG, MA4464/3-1) and the Bernstein Center for Computational Neuroscience II Goettingen (01GQ1005A, project D1).

References

[1] K. Doya, "Complementary roles of basal ganglia and cerebellum in learning and motor control.," *Current opinion in neurobiology*, vol. 10, no. 6, pp. 732–9, 2000.

[2] A. C. Bostan, R. P. Dum, and P. L. Strick, "The basal ganglia communicate with the cerebellum.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 18, pp. 8452–6, 2010.

[3] S. Dasgupta, F. Worgotter, J. Morimoto, and P. Manoonpong, "Neural Combinatorial Learning of Goal-Directed Behavior with Reservoir Critic and Reward Modulated Hebbian Plasticity," in *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pp. 993–1000, 2013.

[4] B. Porr and F. Wörgötter, "Strongly improved stability and faster convergence of temporal sequence learning by utilising input correlations only," *Neural computation*, vol. 18, no. 6, pp. 1380–1412, 2006.

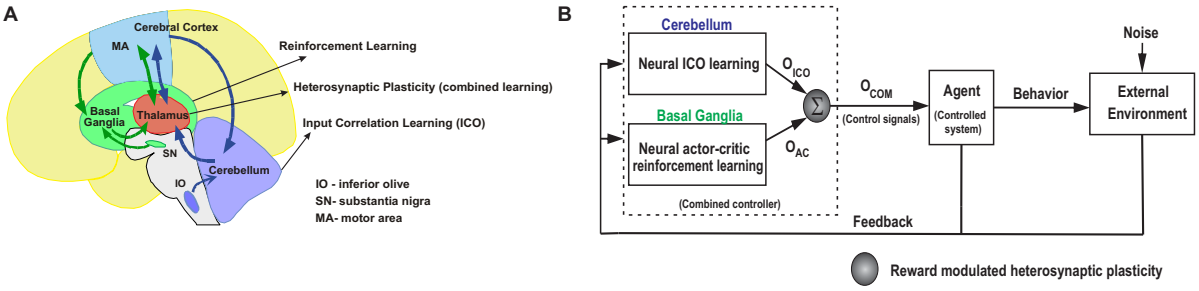

Figure 1: **(A)** Pictorial representation of the anatomical reciprocal connections between the basal ganglia, thalamus and cerebellum. Green arrows depict the cortico-striatal reward learning circuitry via the thalamus. Blue arrows depict the cortico-cerebellar recurrent loops for classically conditioned reflexive behaviors. **(B)** Combined learning framework with parallel combination of ICO learning and actor-critic reinforcement learning. Individual learning mechanisms adapt their weights independently and then their final weighted outputs (O_{ico} and O_{ac}) are combined into O_{com} using a reward modulated heterosynaptic plasticity rule (dotted arrows represent plastic synapses). O_{com} controls the agent behavior (policy) while sensory feedback from the agent is sent back to both the learning mechanisms in parallel.