Oral presentation

Spike timing dependent plasticity implements reinforcement learning

Roberto A Santiago[2], Patrick D Roberts*[1] and Gerardo Lafferriere[2]

Address: [1]Neurological Sciences Institute, Oregon Health & Science University, Portland, OR, USA and [2]Department of Mathematics and Statistics, Portland State University, Portland, OR, USA

Email: Patrick D Roberts* - robertpa@ohsu.edu

* Corresponding author

An explanatory model is developed to show how synaptic learning mechanisms modeled through spike-timing dependent plasticity (STDP) can result in longer term adaptations consistent with reinforcement learning models. In particular, the reinforcement learning model known as temporal difference (TD) learning has been used to model neuronal behavior in the orbitofrontal cortex (OFC) and ventral tegmental area (VTA) of macaque monkey during reinforcement learning. While some research has observed, empirically, a connection between STDP and TD there is as yet no explanatory model directly connecting TD to STDP. Through analysis of the STDP rule, the connection between STDP and TD is explained. We further show that an STDP learning rule drives the spike probability of reward predicting neurons to a stable equilibrium. The equilibrium solution has an increasing slope where the steepness of the slope predicts the probability of the reward. This connection begins to shed light into more recent data gathered from VTA and OFC which are not well modeled by TD. We suggest that STDP provides the underlying mechanism for explaining reinforcement learning and other higher level perceptual and cognitive function.