Abstract

A computationally explicit model of the basal ganglia as a reward-based learning system

Fernando A. Brücher
Psychiatry and Human Behavior
University of California Irvine

The basal ganglia (BG), a collection of brain structures generally unknown by the public, it's believed to be key to reward-based learning, a type of learning of major importance for animals as well as humans. We present an effort to synthesize all relevant evidence on the BG into a mathematical model that presents the BG as a reward-based learning system. The model in turn allows for predictions (simulation) and analysis. Using anatomical and cytoarchitectonical data, distribution of receptor types and neurotransmitters, and types of membrane currents; we define the particularities of each of the cell types that conform the BG and associated structures; their electro-chemical behavior, their roles and the pathways that connect them. We define the BG as having several segregated pathways. Two of these affect the target structures of BG, which in turn promote or inhibit action, while other two are solely devoted to reward/punishment prediction. We also consider the available evidence for plasticity within the system to define the cortico-striatal projection as the only site of plasticity within the BG, and the key role that dopamine plays on it. The presence of spines in synapses where plasticity is suspected to take place is an important feature of the model and related to the concept of eligibility traces in the reinforcement learning (RL) field. Finally we use existing evidence (lesion, animal training) on BG's contributions to behavior to further support the proposed model. Once a thorough structural model of the BG and related structures is presented, we assume the goal of learning within the system is to: maximize the expected discounted sum of future rewards (a common goal of in the RL field), and use a (stochastic) gradient based approach to derive the learning rules that should be implemented by the BG to attain that goal. We show the existing plasticity features of BG are indeed consistent with the assumed goal of the BG system. We also show the necessity for the learning rules to work of having part of the system devoted to producing behavior while a separate part of it focused on reward/punishment prediction (an arrangement known in the RL literature as actor-critic). The contribution of this work is two sided, on one end it helps to understand the BG's role as part of the brain, and on the other it proposes a biology based reward based learning method with similarities but also novel features from existing RL schemes.