10.03: The PBWM Computational Model

The biological properties of the PFC/BG system that we reviewed above are captured in a computational model called PBWM (prefrontal cortex basal ganglia working memory) (O'Reilly & Frank, 2006; Hazy et al 2006, 2007) (Figure 10.11). The PFC neurons in this model are organized into separately-updatable stripes, and also into separate functional groups of maintenance and output gating (described more below). Furthermore, each PFC stripe is represented in terms of superficial layers (2,3) and deep layers (5,6) -- the deep layer neurons specifically have the ability to sustain firing over time through a variety of mechanisms, representing the effects of NMDA and mGluR channels and excitatory loops through the thalamus. The flow of activation from the superficial to deep layers of a given PFC stripe is dependent on BG gating signals, with the BG layers also organized into corresponding maintenance and output gating stripes. The Matrix layer of the BG (representing the matrisomes of the striatum) has separate Go and NoGo neurons that
project to a combined GPi and thalamus (GPIThal) layer with a single neuron per stripe that fires if the Go pathway is sufficiently stronger than the NoGo (this mechanism abstracts away from the detailed BG gating circuitry involving the GPe, GPi/SNr, STN and thalamus, as simulated in the motor chapter, and simply summarizes functionality in a single GPIThal layer). A GPIThal Go signal will update the PFC deep layer activations to reflect the current superficial layer activations, while a NoGo leaves the PFC alone to continue to maintain prior information (or nothing at all).

The PVLV phasic dopamine system drives learning of the BG Go and NoGo neurons, with positive DA bursts leading to facilitation of Go and depression of NoGo weights, and vice-versa for DA dips -- using the same reinforcement learning mechanisms described in the Motor chapter.

The main dynamics of behavior of the different PBWM components are illustrated in Figure 10.12 (Not Created Yet). Perhaps the single most important key for understanding how the system works is that it uses trial and error exploration of different gating strategies in the BG, with DA reinforcing those strategies that are associated with positive reward, and punishing those that are not. In the current version of the model, Matrix learning is driven exclusively by dopamine firing at the time of rewards, and it uses a synaptic-tag-based trace mechanism to reinforce/punish all prior gating actions that led up to this dopaminergic outcome. Specifically, when a given Matrix unit fires for a gated action, synapses with active input establish a synaptic tag, which persists until a subsequent phasic dopaminergic outcome signal. Extensive research has shown that these synaptic tags, based on actin fiber networks in the synapse, can persist for up to 90 minutes, and when a subsequent strong learning event occurs, the tagged synapses are also strongly potentiated (Redondo & Morris, 2011; Rudy, 2015; Bosch & Hayashi, 2012). This form of trace-based learning is very effective computationally, because it does not require any other mechanisms to enable learning about the reward implications of earlier gating events. In earlier versions of the PBWM model, we relied on CS (conditioned stimulus) based phasic dopamine to reinforce gating, but this scheme requires that the PFC maintained activations function as a kind of internal CS signal, and that the amygdala learn to decode these PFC activation states to determine if a useful item had been gated into memory. Compared to the trace-based mechanism, this CS-dopamine approach is much more complex and error-prone. Instead, in general, we assume that the CS’s that drive Matrix learning are more of the standard external type, which signal progress toward a desired outcome, and thus reinforce actions that led up to that intermediate state (i.e., the CS represents the achievement of a subgoal).

The presence of multiple stripes is typically important for the PBWM model to learn rapidly, because it allows different gating strategies to be explored in parallel, instead of having a single stripe sequentially explore all the different such strategies. As long as one stripe can hit upon a useful gating strategy, the system can succeed, and it quickly learns to focus on that useful stripe while ignoring the others. Multiple stripes are also critical when more than one piece of information has to be maintained and updated in the course of a task -- indeed, it is this demand that motivated the development of the original PBWM model to supersede earlier gating models, which used phasic dopamine signals to directly gate PFC representations but did not support multiple gating and hence was limited to a capacity of a single item. One interesting consequence of having these multiple stripes is that “superstitious” gating can occur in other stripes -- if that gating happens to reliably enough coincide with the gating signals that are actually useful, it too will get reinforced. Perhaps this may shed light on our proclivity for being superstitious?
Output Gating

A - Maintenance, pre-Output

B - Output gating

Figure (10.13): Schematic to illustrate the division of labor between maintenance-specialized stripes and corresponding output-specialized stripes. A - Maintenance stripe (left) in maintenance mode, with corticothalamocortical reverberant activity shown (red). Information from that stripe projects via layer Vb pyramids to a thalamic relay cell for the corresponding output stripe, but the BG gate is closed from tonic SNr/GPi inhibition so nothing happens (gray). B - Output gate opens due to 'Go'-signal generated disinhibition of SNr/GPi output (green), triggering burst firing in the thalamic relay cell, which in turn activates the corresponding cortical stripe representation for the appropriate output. Projection from output stripe's layer Vb pyramids then activates cortical and subcortical action/output areas, completing a handoff from maintenance to output. MD = mediodorsal nucleus of the thalamus; VP/VL = ventroposterior or ventrolateral (motor) thalamic nuclei.

As we saw in Figure 10.3, some PFC neurons exhibit delay-period (active maintenance) firing, while others exhibit output response firing. These populations do not appear to mix: a given neuron does not typically exhibit a combination of both types of firing. This is captured in the PBWM framework by having a separate set of PFC stripes that are output gated instead of maintenance gated, which means that maintained information can be subject to further gating to determine whether or not it should influence downstream processing (e.g., attention or motor response selection). We typically use a simple pairing of maintenance and output gating stripes, with direct one-to-one projections from maintenance to output PFC units, but there can be any form of relationship between these stripes. The output PFC units are only activated, however, when their corresponding stripe-level BG/GPiThal Go pathway fires. Thus, information can be maintained in an active but somewhat "offline" form, before being actively output to drive behavior. Figure 10.13 illustrates this division of labor between the maintenance side and the output side for gating and how a "handoff" can occur.

For more PBWM details, including further considerations for output gating, how maintained information is cleared when no longer needed (after output gating), and gating biases that can help improve learning, see PBWM details Subtopic, which also includes relevant equations and default parameters.