Dopamine, Learning, and Production Rules: The Basal Ganglia and the Flexible Control of Information Transfer in the Brain

Andrea Stocco, Christian Lebiere, and John R. Anderson

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
stocco@cmu.edu; cl@cmu.edu; ja@cmu.edu

Abstract

One of the open issues in developing large-scale computational models of the brain is how the transfer of information between communicating cortical regions is controlled. Here, we present a model where the basal ganglia implement such a conditional information routing system. The basal ganglia are a set of subcortical nuclei that play a central role in cognition. Like a switchboard, the model basal ganglia direct the communication between cortical regions by alerting the destination regions to the presence of important signals coming from the source regions. This way, they can impose serial control on the massive parallel communication between cortical areas. The model also incorporates a possible mechanism by which subsequent transfers of information control the release of dopamine. This signal is used to produce novel stimulus-response associations by internalizing the representation being transferred in the striatum. We discuss how this neural circuit can be seen as a biological implementation of a production system. This comparison highlights the basal ganglia as bridge between computational models of small-size brain circuits and high-level characterizations of complex cognition, such as cognitive architectures.

Transfer of Information in the Brain

Many ambitious architectures of brain function have been proposed recently (e.g., Houk 2005; Hawkins and Blakesee 2004). These models differ widely from each other. One of the common problems they all have to solve is the transfer of information among brain regions. The simplest solution consists in hard-wiring the communication between brain regions as direct connections between layers in a network. However, in the human brain, cortico-cortical connections are estimated to make up more than 95% of all the inputs of a single brain region. Furthermore, about half of this amount is estimated to come from long-distance connections (Braitenberg and Schüz 1991). It is clear, therefore, that some organization needs to be overlaid over this massive set of connections.

A number of solutions to this problem have been proposed (e.g., Anderson 2007; van der Velde and de Kamps 2006). In this paper, we propose that the transmission of information along the cortico-cortical pathways is modulated by a subcortical circuit. This circuit comprises important structures such as the basal ganglia and the thalamus. By means of this circuit, organized behavior is imposed upon an otherwise uncoordinated flow of information within the cortex.

Our solution has been implemented as a connectionist computational model. It provides two additional advantages over other attempts. First, it shows how the subcortical circuit is functionally equivalent to a production system. This equivalence makes an important connection between the anatomy of the brain and a widely studied and adopted computational framework. Second, it provides a natural framework for skill acquisition and habit learning compatible with known biological constraints.

The Role of the Basal Ganglia

Before dealing with computational details, this section will review some evidence in favor of our hypothesis that the basal ganglia play an important role in coordinating the transfer of information between cortical areas. Three converging lines of research support this assumption.

Physiologically, pathologies of the basal ganglia in humans result in an increase in the amount of correlated activity between cortical regions (e.g., Stoffers et al. 2008). This fact can be interpreted by assuming that, under normal conditions, widespread cortico-cortical communication is limited by the control function of the basal ganglia.

A second line of evidence comes from studies of human working memory. A number of experiments have shown that the basal ganglia play an important role in gating new information to short-term memory. Neuroimaging data indicating basal ganglia involvement in preparation of

working memory updates (McNab and Klingberg 2008). Also, genetic differences in basal ganglia metabolism correlate with individual performance in working memory tests (Zhang et al. 2007). Finally, individual differences in the severity of dopamine depletion in Parkinson's Disease also correlate with decline of working memory functions.

A final hint of the importance of the basal ganglia in shaping cortico-cortical connectivity comes from research on learning. It is known that skill acquisition results in a dramatic reorganization of cortical connectivity. Moreover, animal studies have shown that lesions of the basal ganglia result in a profound impairment in skill acquisition. In animals, it prevents the acquisition of new stimulus-response associations. In humans, it has been proven to disrupt the acquisition of new sensory-motor skills.

We propose a model for a brain architecture where the basal ganglia have an overseeing role in directing and shaping cortico-cortical connectivity. The model we present here is a layered neural network that reflects several aspects of basal ganglia physiology. This network has two key properties. First, it can acquire new stimulus-response associations through practice. Second, its workings can be shown to be substantially similar to a production system. This provides a straightforward mapping between a well-established formalism for artificial intelligence and the biology of the brain.

A Switchboard Model for the Basal Ganglia

This architecture works as follows. Let us consider a collection of cortical areas $C = \{c_1 \ldots c_n\}$. For simplicity, let us assume they are all connected to each other. At each moment in time, each region receives signals from $n - 1$ other regions.

Essentially, the basal ganglia alert each region to attend to only a particular subset of "source" regions $S \subseteq C$. This process can be repeated for each region, providing a powerful system for prioritizing and simplifying the exchange of communication (Figure 1). Intuitively, there should be an optimal ratio of $|S|$ to $|C|$. If $|S|$ is too small, the communication between regions is eventually disrupted. If $|S|$ is too large, on the other end, each region receives too many competing signals.

One can consider a very simple model where the pattern held in each region is simply the sum of all the incoming signals from the other regions, $c_i = c_1 + c_2 + \ldots + c_n$.

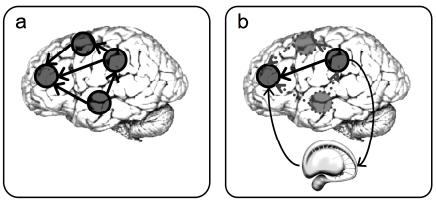

Figure 1: Comunication between massively connected cortical regions (left) can be organized by a cortico-subcortical circuit (right) that encompasses the basal ganglia (bottom right).

When the basal ganglia system is compromised, then $S \approx C$, and each region receives almost identical and largely overlapping signals. Therefore, *spatial* correlation between the regions increases. The *temporal* correlation increases as well, since updating events in one region are reflected in changes in larger group of connected regions.

Circuitry

In order to explain how the model works, one needs to introduce some biology.

The basal ganglia comprise a number of interconnected nuclei. They include the Striatum, the Internal (henceforth, GPi) and External (GPe) Globus Pallidus, the Substantia Nigra (SNr), and the Sub-Thalamic Nucleus (STN). The wiring among these nuclei is usually described in the following terms (Albin, Young, and Penney 1989). The striatum is the entry point of the circuit, receiving afferents from the entire cortex. The nuclei SNr and GPi constitute the system's output. These nuclei project mainly to the thalamus, and control the thalamic projections to the cortex. The Striatum and the SNr/GPI are connected by two pathways, which exert opposite effects. They are known as the *direct* and *indirect pathways*. The indirect pathway comprises the GPe and the STN (Figure 2).

A common interpretation, dating back to Albin, Young, and Penney (1989), is that the two pathways simply oppose each other. In particular, the direct pathway conveys excitatory signals to the cortex, while the indirect pathway contrasts this effect through direct inhibition. In the model, we expanded this interpretation as follows. The direct pathway carries a selection of *source* regions, whose representation has been chosen for transmission. The indirect pathway carries a selection of *destination* regions for each source region. In practice, the indirect pathway carries a *mask* that establishes which region each destination should be attending to.

The Striatum

The striatum is the largest nucleus of the circuit. The large majority of its cells are projection neurons (Graveland, Williams, and DiFiglia 1985). These neurons can be divided into two groups: Striatonigral (SN) cells, whose projections form the direct pathway, and Striatopallidal (SP) neurons, whose projections begin the indirect pathway (Figure 2).

In our model, SN and SP cells are organized into subdivisions. Each subdivision receives afferents from a single corresponding cortical region. Therefore, the striatal organization mirrors cortical topology. Subdivisions also possess a second-level, internal organization. Within a single subdivision, neurons are grouped into ensembles corresponding to the destination that the source region projects to.

Thus, the model striatal subdivisions reflect cortical topology at two levels. At a macro-level, they mirror the organization of cortex into specific regions. At a lower

level, each subdivision also reflects the cortical connectivity of the corresponding cortical region. This organization is compatible with some properties of corticostriatal projection distribution (e.g., Parthasarathy, Schall, and Graybiel 1992).

Physiologically, the activity of projections neurons is controlled by interneurons (IN), which exert a powerful inhibitory pressure (Tepper and Bolam 2004). Because of the inhibition coming from interneurons, only a small number of ensembles of projection neurons contain active and firing cells. Active neurons in a striatonigral subdivision signal that the corresponding cortical region is a source region, and its contents have been picked up for routing. Active striatopallidal ensembles within a subdivision, on the other hand, signal destinations where the selected representations should not be transferred.

Source and destination information travel separately along the direct and the indirect pathway, and eventually combine in the output nuclei SNr/GPi. From the output nuclei, the signals reach the thalamus and, from there, come back to the cortex to enable the proper transfer path.

Relation to Production Systems

From a purely computational point of view, routing operations can be seen as a neural network analog to production rules in production systems. Production rules are control statements expressed in the form of condition-action clauses ("if… then…"). The similarity between the conditional routing model and a production system can be seen if one assumes the following mappings.

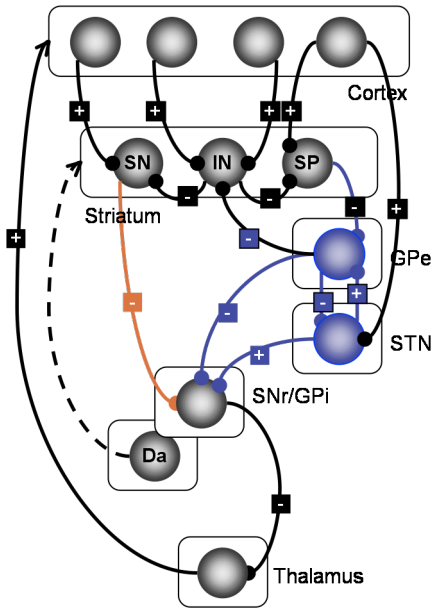

Figure 2: Outline of the basal ganglia circuit and connectivity. Blue arrows and nuclei represent the indirect pathway; Orange arrows represent the direct pathway. Dotted arrows represent dopamine (*Da*) projections.

A *rule* is embedded in the incoming and outgoing synaptic matrices of a set of striatal interneurons. The *condition* (i.e., Left-hand side) part of the rule is represented by incoming synapses to striatal interneurons. The synapses encode the specific cortical representation that will trigger the interneuron to fire. The *action* (i.e., Right-hand side) is encoded in the outgoing synapses to the striatal projection neurons. An action corresponds to the activation of particular ensembles of SN or SP projection neurons in the striatum. Their activation triggers the transmission of information from the source region of the cortex to the destination region.

Part of the flexibility of production systems originates from the use of variables in the production rules. However, variables are not easily dealt with in neural networks. To overcome this problem, a number of procedures have been proposed over the years (e.g., Touretzky and Hinton 1988; Smolensky 1990; Stewart and Eliasmith 2008).

In our model, a *variable* simply corresponds to a specific location in the cortex, i.e., a cortical region. During the execution of routing operations, moving content from a source region to a destination region corresponds to binding the variable to the contents of the source region and transferring the bound value to the destination region. Production rules can also specify constants. This corresponds to transferring a fixed content to a destination region. This particular content is not dependent on any particular region, and is embedded in the synaptic weights of the circuit. Such a case will be illustrated in the forthcoming sections on learning.

Dopamine

Dopamine is a neurotransmitter that affects neural plasticity, promoting long-term potentiation and long-term depression among neurons. Some properties of the dopamine signal in the basal ganglia have been successfully modeled as the error term in Sutton's (1988) Temporal Difference algorithm (Schultz 2002). That is, the dopamine signal reflects the error between two subsequence predictions of a specific state's value. Many models of dopamine function have been proposed (see Joel et al. 2002 for a review). However, some of the dopamine signal's features lay beyond pure reinforcement learning (Redgrave and Gurney 2006). More importantly, the connection between dopamine and procedural learning has been less investigated. It is easy to imagine that dopamine also underlies procedural learning. This is particularly important because the exact mechanisms by which new skills and habits can be acquired have seldom been modeled (See Ashby et al. 2007 for a notable exception).

Habit Learning and Production Systems

The connection between production systems and our model of basal ganglia function can be applied to the domain of habit learning. One easy way to form new habits from a chain of production rules is to create new operations that

simply associate the last action's response with the initial action's conditions. Another strategy consists in detecting intermediate steps that are only needed temporarily to transform some kind of representation. The processed representation can be eventually encoded directly into the production rule, and the intermediate step can be removed. A successful version of this idea was proposed by Taatgen and Lee (2003).

Such intermediate steps can be detected by dopamine neurons in the substantia nigra (SNc). These neurons receive two sets of inputs from the striatum, from the direct and the indirect pathway. Many reinforcement learning models assume that the signal travelling along the indirect is delayed, and can be used to obtain an temporal error term (Joel et al. 2002). In our model, the two signals can be used to compare the current set of source regions with the previous set of destination regions. When a region belongs to both, dopamine neurons receive an extra boost of activation. Since these neurons project to the striatum, the net result is that striatal cells in that region receive an additional amount of dopamine. In turn, the presence of additional dopamine modifies the activation of projection neurons and interneurons, and the synapses between them. As a result, a compressed copy of the cortical representation becomes permanently stored within the striatum, and can be used even in absence of the original cortical contents.

Dopamine can have long-term effects of potentiating or depressing a synapse, depending on the types of pre- and post-synaptic neurons. These two effects were straightforwardly modeled as Hebbian (1) and anti-Hebbian (2) rules. In our model, Hebbian learning is applied to the excitatory cortico-striatal connections. Anti-Hebbian learning, on the other hand, takes place in the inhibitory synapses within the striatum. The two rules are implemented as follows:

$$\Delta w_{i,j} = \eta\,(\theta - w_{i,j})\,x_i\,x_j \qquad \text{if} \quad w_{i,j} > 0 \qquad (1)$$

$$\Delta w_{i,j} = -\eta\,(\theta + w_{i,j})(x_i - x_j) \qquad \text{if} \quad w_{i,j} < 0 \qquad (2)$$

where $w_{i,j}$ is the weight of the synapses between neurons i and j; x_i is the activation of neuron i; and η is the learning rate. The term θ is used to bound the weight growth, which is otherwise exponential and unstable.

Simulations

This section describes a series of simulations that illustrate the model's performance and its learning capabilities. In the simulations, the model was trained to perform a simple aural discrimination task. This task was originally used as part of series of multi-tasking experiments (Hazeltine, Teague, and Ivry 2002). In this task, participants responded to the presentation of a tone. Tones could have three different pitches (440, 880 and 3520 Hz), to which participants had to respond "one", "two", or "three", respectively.

This paradigm was later modeled by Anderson, Taatgen, and Byrne (2005). In their model, the task requires four steps: (1) The tone is encoded in an aural buffer; (2) The tone is used as a retrieval cue and matched against previously memorized stimulus-response associations in long-term memory; (3) A tone-response association that matches the attended stimulus is retrieved; (4) The response is used in a vocal command. The model also provides a mapping between functional steps and three brain regions: one in the temporal lobe (responsible for Step 1), one in the left prefrontal cortex (Steps 2-3) and one in the posterior frontal cortex (Step 4).

A simple cortico-basal ganglia circuit is generated for the simulation. The circuit contains only the three cortical regions required by the task. Correspondingly, the striatum only contains three main subdivisions. It is further assumed that each region is connected to the other two. To simulate memory retrieval, the prefrontal memory region is connected to a data structure (perhaps corresponding to the hippocampus) that can hold a series of associated patterns. The prefrontal region sends its internal representations to this structure, and receives back the pattern that is associated with the best-matching input representations. Each cortical region contains 100 artificial neurons, and each striatal subdivision is made of 10 units.

The left part of Figure 3 illustrates how the three regions jointly perform the experimental task. In the figure, time flows vertically. The different vertical tracks detail the execution steps performed by each region. The boxes on the tracks illustrate the representation being currently held and processed by that region during one of the task steps. The arrows connecting the boxes denote transfers of information between regions. Two transfer operations are required to perform this task. They are represented in the two right panels of Figure 3. These two panels illustrate the state of the model thalamus in the circuit. For convenience, thalamic neurons have been re-arranged into groups within a 3 by 3 matrix. In this matrix, rows represent the three source regions and columns represent the same regions as possible destinations. For instance, the presence of activity in a matrix cell in the first row and second column corresponds to a signal that tell the second region to attend to the contents of the first. This is exactly what happens in the basal ganglia between Steps 1 and 2. Activation of these thalamic terminals determines the transition to Step 2. When the prefrontal region has received the auditory cue, it responds by producing in Step 3 a pattern corresponding to the response associated to the tone. Note that this transition occurs within the cortex, and the basal ganglia are not involved. The second routing operation (illustrated in the bottom right panel) is triggered at this point, and routes the retrieved response to the vocal region (Region 3), where it can be transformed into a vocal program. This corresponds to the final step of the task.

Effects of Dopamine on Task Performance

In the learning simulation, dopamine is triggered after the execution of Step 3 (see Figure 3). Learning is simulated as

a one-shot process by setting $\eta = 3.0$ and $\theta = 1.0$. Note that the same results would have been obtained by lowering η and having the model perform the task a number of consecutive times. Figure 4 illustrates the very same task after learning has occurred. A new operation has been produced that directly encodes part of the original cortical pattern and is able to transmit it to the thalamus. It is interesting to note that the new operation can possibly occur at the same time as the original response to the stimulus. This is explicitly represented in the right panel of Figure 4. The new activation pattern includes the responses that were previously found in two separate routing operations (compare to right panels of Figure 3). Therefore, learning had the effect of "shifting back" in time what originally was the second operation, and applying it in advance. This fact is consistent with the reorganization of firing patterns in the striatum following habit learning, where neurons responding originally at specific task events begin responding earlier after practice (Jog, Kubota, Connolly, Hillegaart, and Graybiel 1999).

Conclusions

This paper has presented a connectionist model of the basal ganglia. The model is based on the idea that this subcortical circuit oversees and controls the transfer of information among cortical brain regions. Because of the large underlying amount of cortico-cortical connection, this circuit plays a fundamental role in organizing the flow of information within the brain by selecting proper source and destination regions. The model can be seen as a neural implementation of a production system, where production rules correspond to routing operations among brain regions. This equivalence is important for two reasons. The first is that it provides a biological substrate for a powerful and well-known computational framework. Second, this equivalence provides a means to understand the neural basis of intelligence and flexible behavior, and bridge the gap between low-level and high-level computational descriptions of the brain.

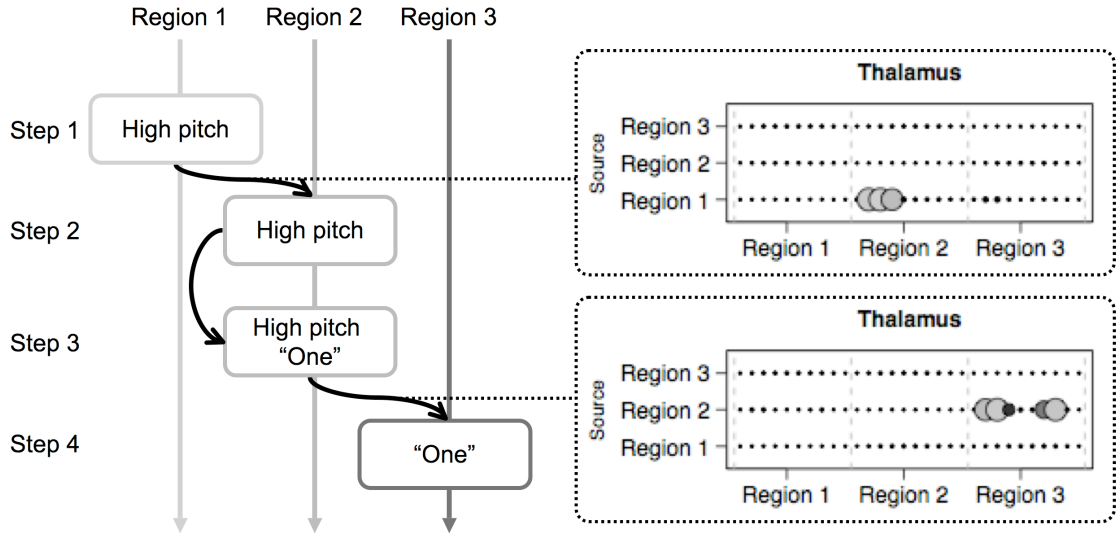

Figure 3. Performance of the Aural Discrimination task by the model. *Left*: The sequence of operations necessary to perform the task, as executed by the different regions. *Right:* Thalamic units activation during information transfers across regions.

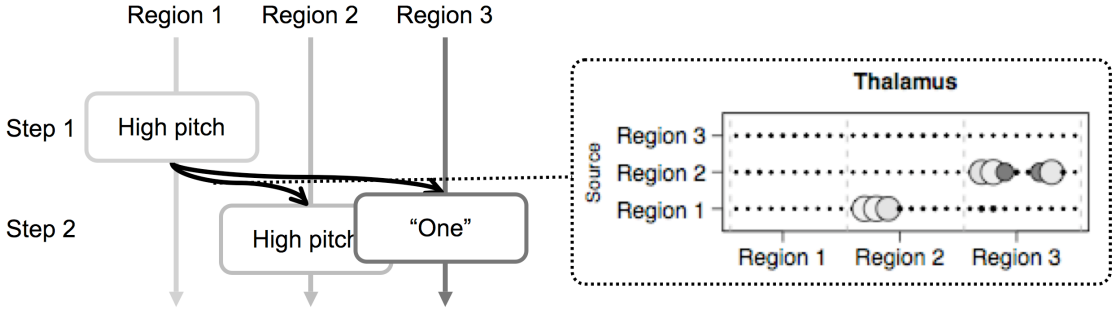

Figure 4: Performance of the same task after dopamine-triggered learning. Notice that the two operations are now triggered at the same time. Furthermore, the response representation has been encoded in the striatum, and can be transferred even in absence of a suitable representation in the source region.

Furthermore, the connection between production systems and the basal ganglia circuit provides a new perspective on the neural basis of practice and skill acquisition. This connection was exploited to adapt the principles of production system practice effects to the neural circuit. The simulation results show that habit learning can be implemented by using the very same mechanisms that have been hypothesized for dopamine-based reinforcement learning. Additionally, these results are qualitatively consistent with electrophysiological data.

Acknowledgments

This research was supported by award FA9550-08-1-0404 from the Air Force Office of Scientific Research (AFOSR) to John Anderson, Randall C. O'Reilly, and Christian Lebiere.

References

Albin, R. L.; Young, A. B.; and Penney, J. B. 1989. The Functional Anatomy of Basal Ganglia Disorders. *Trends in Neurosciences* 12: 366—375.

Anderson, J. R. 2007. *How Can the Human Mind Occur in the Physical Universe?* New York, NY: Oxford University Press,

Anderson, J. R.; Taatgen, N. A.; and Byrne, M. D. 2005. Learning to Achieve Perfect Timesharing: Architectural Implications of Hazeltine, Teague, and Ivry (2002). *Journal of Experimental Psychology: Human Perception and Performance* 31: 749—761.

Ashby, F. G.; Ennis, J. M.; and Spiering, B. J. 2007. A Neurobiological Theory of Automaticity in Perceptual Categorization. *Psychological Review* 114: 632—656.

Braitenberg, V., and Schüz, A. 1991. *Anatomy of the Cortex. Statistics and Geometry*. Heidelberg, Germany: Springer Verlag,

Graveland, G. A.; Williams, R. S.; and DiFiglia, M. 1985. A Golgi Study of the Human Neostriatum: Neurons and Afferent Fibers. *Journal of Comparative Neurology* 234: 317—333.

Hawkins, J., and Blakeslee, S. 2004. *On Intelligence*. New York, NY: Times Books.

Hazeltine, E.; Teague, D.; and Ivry, R. B. 2002. Simultaneous Dual-Task Performance Reveals Parallel Response Selection After Practice. *Journal of Experimental Psychology: Human Perception and Performance* 28*:* 527—545.

Houk, J. C. 2005. Agents of the Mind. *Biological Cybernetics* 92: 427—437.

Joel, D.; Niv, Y.; and Ruppin, E. 2002. Actor-Critic Models of the Basal Ganglia: New Anatomical and Computational Perspectives. *Neural Networks* 15: 535—547.

Jog, M. S.; Kubota, Y.; Connolly, C. I.; Hillegaart, V.; and Graybiel, A. M. 1999. Building Neural Representations of Habits. *Science* 286: 1745—1749.

McNab, F., and Klingberg, T. 2008. Prefrontal Cortex and Basal Ganglia Control Access to Working Memory. *Nature Neuroscience* 11: 103—107.

Packard, M. G., and Knowlton, B. J. 2002. Learning and Memory Functions of the Basal Ganglia. *Annual Review of Neuroscience* 25: 563—593.

Parthasarathy, H. B.; Schall, J. D.; and Graybiel, A. M. 1992. Distributed but Convergent Ordering of Corticostriatal Projections: Analysis of the Frontal Eye Field and the Supplementary Eye Field in the Macaque Monkey. *Journal of Neuroscience* 12: 4468—4488.

Redgrave, P., and Gurney, K. 2006. The Short-Latency Dopamine Signal: A Role in Discovering Novel Actions? *Nature Reviews Neuroscience* 7: 967—975.

Schultz, W. 2002. Getting Formal With Dopamine and Reward. *Neuron* 36: 241—263.

Smolensky, P. 1990. Tensor Product Variable Binding and the Representation of Symbolic Structures in Connectionist Systems. *Artificial Intelligence* 46: 159—216.

Stewart, T. C., and Eliasmith, C. 2008. Building Production Systems with Realistic Spiking Neurons. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Stoffers, D.; Bosboom, J. L.; Deijen, J. B.; Wolters, E.; Stam, C. J.; and Berendse, H. W. 2008. Increased Cortico-Cortical Functional Connectivity in Early-Stage Parkinson's Disease: An MEG Study. *Neuroimage* 41: 212—222.

Sutton, R. S. 1988. Learning to Predict by the Methods of Temporal Differences. *Machine Learning* 3: 9—44.

Taatgen, N. A., and Lee, F. J. 2003. Production Compilation: A Simple Mechanism to Model Complex Skill Acquisition. *Human Factors* 45: 61—76.

Tepper, J. M., and Bolam, J. P. 2004. Functional Diversity and Specificity of Neostriatal Interneurons. *Current Opinion in Neurobiology* 14: 685—692.

Touretzky, D., and Hinton, G. H. 1988. A distributed connectionist production system. *Cognitive Science* 12: 423—466.

van der Velde, F., and de Kamps, M. 2006. Neural blackboard architectures of combinatorial structures in cognition. *Behavioral and Brain Sciences* 29: 37—70.

Zhang, Y.; Bertolino, A.; Fazio, L.; Blasi, G.; Rampino, A.; Romano, R.; Lee, M.-L. T.; Xiao, T.; Papp, A.; Wang, D.; and Sadée, W. 2007. Polymorphisms in Human Dopamine D2 Receptor Gene Affect Gene Expression, Splicing, and Neuronal Activity During Working Memory. *PNAS* 104: 20552—20557.