A Gated Prefrontal Cortex Model of Dynamic Task Switching

Nicolas P. Rougier, Randall C. O'Reilly

Learning Representations in a Gated Prefrontal Cortex Model of Dynamic Task Switching

Nicolas P. Rougier
Department of Psychology
University of Colorado Boulder

Randall C. O'Reilly
Department of Psychology
University of Colorado Boulder

30th November 2001

Address correspondence to:
Randall C. O'Reilly
Department of Psychology
University of Colorado Boulder
345 UCB Boulder, CO 80309-0345
Tel: (303) 492-0054
Fax: (303) 492-2967
email: oreilly@psych.colorado.edu

Running head: A Gated Prefrontal Cortex Model of Dynamic Task Switching

Abstract

The prefrontal cortex is widely believed to play an important role in facilitating people's ability to switch performance between different tasks. We present a biologically-based computational model of prefrontal cortex (PFC) that explains its role in task switching in terms of the greater flexibility conferred by activation-based working memory representations in PFC, as compared with more slowly adapting weight-based memory mechanisms. Specifically we show that PFC representations can be rapidly updated when a task switches via a dynamic gating mechanism based on a temporal-differences reward-prediction learning mechanism. Unlike prior models of this type, the present model develops all of its internal representations via learning mechanisms as shaped by the demands of continuous periodic task switching. This advance opens up a new domain of research into the interactions between working memory task demands and the representations that develop to meet them. Results on a version of the Wisconsin Card Sorting task are presented for the full model and a number of comparison networks that test the importance of various model features.

Introduction

We've probably all had the experience of trying to pull open a door that should be pushed, or vice-versa. When it doesn't work, you have to switch your approach and try the opposite maneuver. Sometimes, however, you might catch yourself failing to switch, and retrying the incorrect maneuver again. This kind of *perseveration* behavior is a hallmark of patients with prefrontal cortex (PFC) damage. The classic example of the involvement of the PFC in task switching is the Wisconsin card sorting task (WCST), which typically but not exclusively (Stuss, Levine, Alexander, Hong, Palumbo, Hamer, Murphy, & Izukawa, 2000) shows impairments with frontal damage. Further evidence of PFC involvement in task switching comes from tasks related to the WCST (e.g., Dias, Robbins, & Roberts, 1997; Roberts, Robbins, & Everitt, 1988; Owen, Roberts, Hodges, Summers, Polkey, & Robbins, 1993) and other kinds of task switching paradigms (e.g., Burgess, Veitch, de Lacy Costello, & Shallice, 2000). Despite this evidence for the involvement of the PFC, the precise mechanistic role of this brain area in task switching remains unclear. In this paper we present a biologically-based model performing a WCST-like task switching task that helps to illuminate the mechanistic role that the PFC plays in task switching.

Specifically, our model is founded on the idea that the PFC is specialized for activation-based working memory (Miller & Cohen, 2001; Frank, Loughry, & O'Reilly, 2001; O'Reilly, Noelle, Braver, & Cohen, in press; O'Reilly & Munakata, 2000; O'Reilly, Braver, & Cohen, 1999). By representing information as maintained activation states, the PFC can contribute to task switching by rapidly updating these activation states in response to feedback. This activation switching can be much faster than the structural changes that underlie adaptation of connection strengths between neurons (as captured in standard neural network learning mechanisms). This rapid updating in PFC can be specifically triggered by a dynamic gating mechanism that controls the updating of activation-based working memories maintained in the PFC (O'Reilly et al., in press; O'Reilly & Munakata, 2000).

The present model goes beyond earlier models of task switching involving a dynamically gated PFC system (O'Reilly et al., in press; O'Reilly & Munakata, 2000) by developing PFC and other representations entirely from experience-driven learning mechanisms operating in the context of repeatedly switching among a set of tasks. Our previous models of PFC function have used hand-coded PFC representations, so this represents an important advance that opens up many new avenues of research into the nature of PFC representations as a function of task demands. Nevertheless, we build upon earlier work by employing a similar dynamic gating mechanism based on reward-prediction learning mechanisms. We begin with a summary of the task switching task, followed by a description of the model. We present results from both the dynamically-gated PFC model and a number of comparison models that systematically explore both the overall contribution of the PFC in the model, and the effects of various mechanisms that influence the dynamic gating process. We conclude with a discussion of the implications of having a dynamic gating model that can develop useful internal representations.

The Dynamic Naming Task

The task we use for testing the model is a modified version of the widely-studied Wisconsin card sorting task, where the inputs are cards having feature values along different stimulus dimensions (color, shape, number), and the essence of the task is to focus on one of these dimensions for a series of trials, and then switch to a different dimension. In our version, the output involves naming a feature instead of sorting cards into piles. Thus, we refer to this as the *dynamic naming task*. Specifically, each stimulus item has five feature dimensions with three possible feature levels in each of the dimensions. The goal of the task is to guess an unspoken *target* dimension, and report the feature value along that dimension for each item (figure 1). If the model outputs the correct feature value along the target dimension for the current stimulus, it is rewarded (otherwise not) — the patterns of reward can be used to guide the search for the unspoken dimension. The same target dimension (i.e., the subtask) is used for a period of n trials (typically $n = 50$), after which another target dimension is selected, requiring the model to switch to a new subtask.

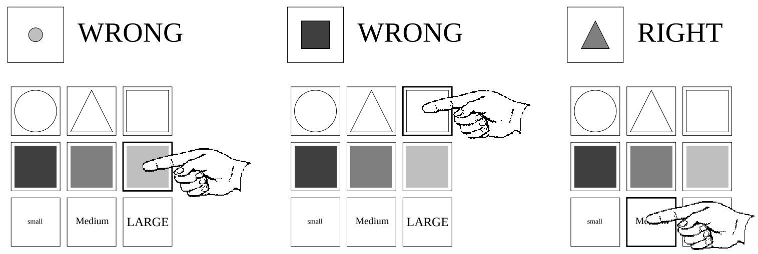

Figure 1: The task. In this simplified example, each stimulus has three feature dimensions (e.g. shape, color, size) with three possible feature levels (circle/square/triangle, red/green/blue, small/medium/large) and the hidden dimension is the size of the presented sample stimulus. In the first two trials the subject chose the sample matching level in the wrong dimension and consequently got negative feedback. In the last trial, the correct dimension and feature value is selected, leading to positive feedback.

The basic strategy to perform this task correctly is a simple win-stay, lose-shift type of rule:

- If a positive feedback is received, continue with the current subtask.
- If a negative feedback is received, switch to another subtask.

We show that this strategy emerges naturally from a reward-based gating mechanism that controls the maintenance of information in PFC working memory. The current stimulus dimension (subtask) is maintained in PFC until sufficient negative feedback results in the search for a different subtask.

The Model

Figure 2 shows the structure of the model, which is implemented using the Leabra framework (O'Reilly & Munakata, 2000; O'Reilly, 1998). The sample layer represents the input stimuli using localist representations of features within dimensions (i.e., the first row of 3 units represents shape, the second row represents color, etc.). This input is mapped via a hidden layer (representing posterior cortex) to an output layer, which represents the model's answer of a single active feature within one of the five dimensions. Thus,

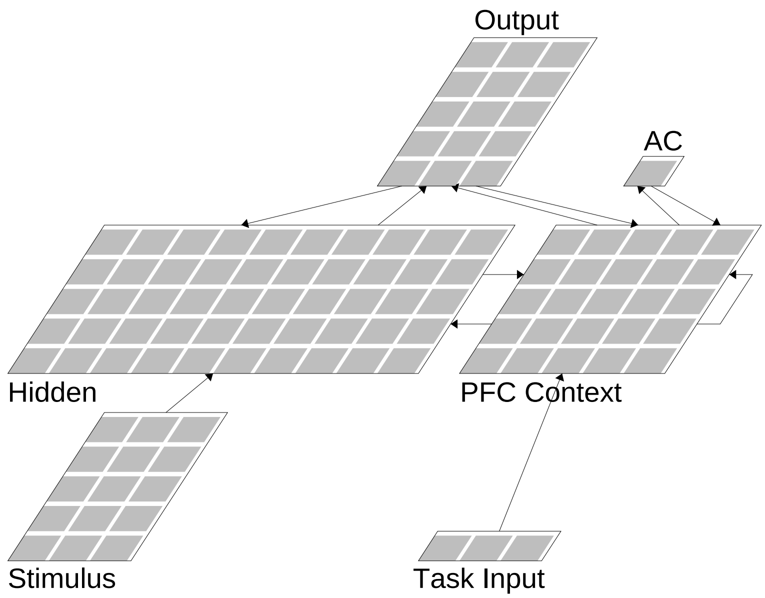

Figure 2: Dynamic task-switching model using active PFC context representations

the essential task that the network must perform is to select one dimension out of the five impinging on it from the input stimulus. The prefrontal cortex (PFC) layer facilitates this dimensional selection by providing top-down support or biasing (Cohen & O'Reilly, 1996; Cohen, Dunbar, & McClelland, 1990) for a given dimension. It facilitates task switching by being able to rapidly update to another dimensional representation, thereby providing different biases. Thus, as in our previous task-switching models (O'Reilly et al., in press; O'Reilly & Munakata, 2000), the PFC does not contribute directly to the input-output mapping task, but rather contributes by providing appropriate task-relevant biases. This contrasts with some other task-switching models (e.g, Dehaene & Changeux, 1991) as discussed in greater detail in O'Reilly et al. (in press).

The PFC layer maintains an activation-based working memory representation for the current subtask (i.e., stimulus dimension of relevance). To support this memory, and to encourage coherent task representations instead of blends of different task representations, the PFC has excitatory recurrent connections that enable units representing the same subtask to support each other. The updating of this layer is influenced by the reward-based dynamic gating mechanism as implemented in the AC layer, described next.

To enable exploration of "explicit" instruction in this task, we also included a task input layer where each unit represents a different subtask (i.e., dimension). During the naming task as described previously, this task input is uniformly activated so as to not provide any discriminant information about the relevant task. We have explored the interaction between this explicit signal and the intrinsic dynamics of the PFC representations during initial learning and subsequent task performance, but this is beyond the scope of the present paper.

The Adaptive Critic (AC) Dynamic Gating Mechanism

The dynamic gating mechanism in the model is based on the idea that working memory updating can be driven by changes in reward predictions (Braver & Cohen, 2000; O'Reilly et al., 1999), as formalized in the temporal-differences (TD) reinforcement learning mechanism (Sutton, 1988; Sutton & Barto, 1998). The TD algorithm employs an adaptive critic (AC) that attempts to predict future rewards, and it drives learning as a function of differences in these predicted rewards. The functional properties of the AC provide a good, if imperfect, fit to the firing properties of midbrain dopamine neurons in the ventral tegmental area (VTA). It has been shown that the VTA fires dopamine bursts for stimuli that are predictive of reward (e.g., Schultz, Apicella, & Ljungberg, 1993), in a way that is generally consistent with the AC mechanism (Montague, Dayan, & Sejnowski, 1996). If rewards are expected but not delivered (i.e., due to a behavioral error), the dopamine neurons exhibit reduced firing, corresponding to a *negative error signal*. Task-relevant information that should be maintained is a reliable predictor of reward, and should thus elicit dopamine firing, resulting in the updating of working memory (Braver & Cohen, 2000; O'Reilly et al., 1999), and the negative error signal should reset working memory representations. The net effect is to produce a form of *trial-and-error search* by activating and deactivating PFC representations (O'Reilly et al., in press; O'Reilly & Munakata, 2000).

In the context of the dynamic naming task, the AC unit learns to expect reward when the network is performing correctly, which stabilizes the PFC representations, and these expectations are disconfirmed when the task is switched and the network starts performing incorrectly, which destabilizes the PFC representations and allows a new pattern to be activated (i.e., a new task context). This stabilization and destabilization of PFC representations facilitates task switching. We implemented PFC active maintenance using a combination of recurrent excitatory connections and intracellular ionic conductances that provide persistent excitatory input to units that are active when the gating signal goes positive, enabling them to persist over time (see Frank et al., 2001 for details on this mechanism).

The next sections describe some additional mechanisms that we found to be important in improving the performance of the dynamic gating mechanism. Our full model includes these mechanisms, and we evaluate their contributions in the simulations described later in the results section.

The Computation of Reward

The computation of reward plays a critical role in the model because it drives the AC unit behavior. The most straightforward solution would be to send the reward signal directly to the AC unit, but this solution would cause instability within the network because of intra-task errors. That is, the model, and actual subjects, always have a low level of background errors in task performance due to small weight changes interacting with interactive activation dynamics (O'Reilly, 1996). Consider the case where the model produces an intra-task error: with a direct reward signal, the error causes the AC unit to immediately destabilize the PFC context representation, leading to a search among other possible subtasks. Before switching back to the subtask the model was just in (which was in fact the right subtask), it will then produce additional errors while trying other contexts (which are wrong at this time). The reward computation therefore must be fault-tolerant.

We adopt a fault-tolerant solution by averaging the reward across a period of n steps, and setting a threshold on this average for positive vs. negative reward (i.e., if the average reward is above-threshold, a positive reward is given, otherwise a negative reward is given). Thus, if isolated errors are produced, the average remains high and the AC unit nonetheless gets a positive reward signal and does not induce context switching. With the consistent errors associated with task switching, the average will go below threshold and a negative reward will be given. Note, however, that this has a drawback in that the model has to produce more errors before being able to switch context, degrading overall performance. The choice of n is then a compromise between stability and performance. We used a value of 2.

One further optimization can be made. Consider the situation when the model just switched to a new context. This switch occurred because the model produced wrong answers, and because we average the reward over time, we carry forward errors from the past when switching to a new context. These errors will incorrectly penalize the new context, which should be tested before deciding it might be the wrong one. Therefore, we reset the average reward value after it goes below threshold and results in an actual error signal communicated to the AC unit. These additional assumptions are important for the model's behavior (as we demonstrate later) and thus stand as testable predictions about how reward signals are filtered through to the midbrain dopamine systems in the brain.

Inhibition of Prior Task Representations

One final mechanism that we added prevents the immediate reactivation of previously active task representations as the network searches for a new task context. This clearly makes sense because when the task switches, the previous task context should not be considered among the options for the new task context. The situation is analogous to the well-known inhibition of return phenomenon in visual search, and there may indeed be a common underlying biological mechanism of neural fatigue or synaptic depression. In the model, we used a negative bias weight learning mechanism that rapidly builds up a negative bias weight in response to negative changes in activation states (as when a PFC unit is deactivated). This negative bias then makes the unit unlikely to be reactivated. It then decays steadily back toward zero to release the inhibition over time. As we will show in the results section, this mechanism is of a great help for stabilizing PFC representations and acts indeed as a very short term memory of the past.

Overview of Model's Task Switching Behavior

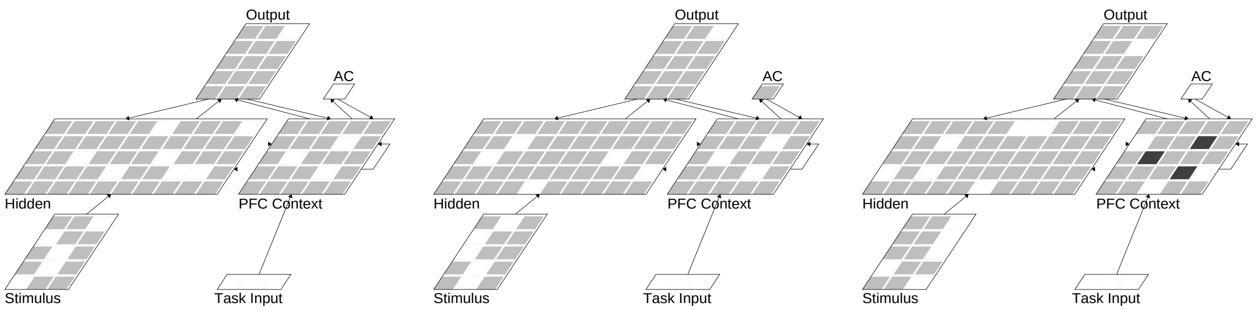

Figure 3: Sequence of context activation within the PFC layer during a task switch. Grey squares represent inactive units, white squares represent active units and dark grey squares represents units with a negative bias activity.

We can now illustrate the overall behavior of these mechanisms by considering the example given in (figure 3). In this example, the model is currently performing the subtask of naming the feature values along the 5^{th} dimension, while the dimension has just been switched to the 4^{th} one. On the first trial following the switch and because of the context present within PFC, the model continues to produce its answer based on the feature active in the 5^{th} dimension, which now is the wrong one. Nonetheless, because of the fault tolerant mechanism we use for computing reward, the model maintains the current context and consequently, in the second trial, the model gives again its answer in this same 5^{th} dimension. Having produced two errors in a row, the average reward is below threshold and a negative reward signal is sent to the AC unit, which destabilizes the PFC representations. This destabilization allows a new context representation to become active within the PFC layer, which is the right one in our example. Because this context biases the hidden layer to produce an answer in the 4^{th} dimension, it naturally leads to positive rewards which in turn stabilizes the PFC context layer until the task switches again. Furthermore, it is to be noted units participating in the previous context representation receives a negative bias activity preventing then to be immediately reactivated.

Results

The objectives of the following simulations are as follows:

- To determine if useful PFC representations can develop on their own (from random initial weights) in the context of repeatedly performing the dynamic naming task, with periodic task switching.
- To evaluate the contribution of the PFC and dynamic gating mechanisms in our model in comparison to other models lacking these mechanisms.
- To evaluate the importance of various features of the dynamic gating mechanism and reward computation mechanisms as described earlier.

Development of PFC Representations

The most important result is that the full model as described above can indeed learn useful task representations from random initial weights through the process of performing the dynamic naming task. The first line of evidence is that the network learns to solve the task quite well in terms of asymptotic error levels (see figure 5, PFC data). Second, we examined the PFC representations that developed to see if there were distinct representations for each task context (stimulus dimension). As shown in figure 4, the weight matrices between PFC layer and output are clearly organized along stimulus dimension. All features within a given dimension (i.e., each row of the output layer) share the same set of weights from the PFC layer. Furthermore, these same set of PFC units are strongly interconnected via their reciprocal connections. These results clearly show that the model has developed abstract representations of stimulus dimensions in this context. We are currently leveraging this basic finding into a large-scale investigation of how such PFC representations might facilitate flexible behavior in other kinds of task contexts.

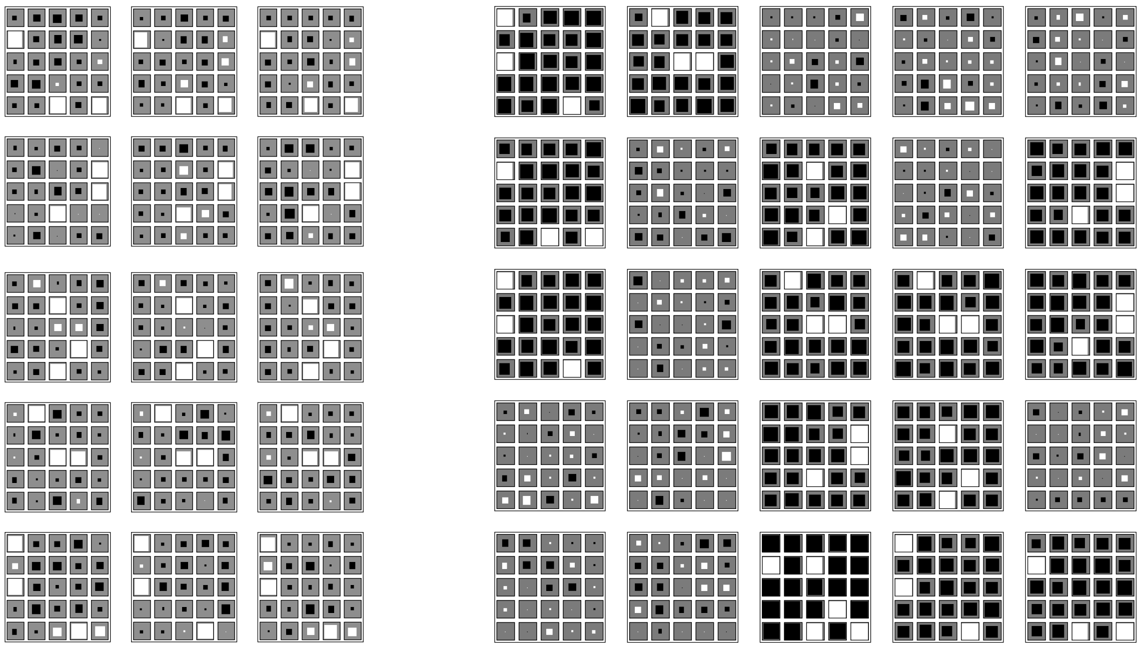

Figure 4: Matrices of weights from PFC to output layer and from PFC to PFC layer. The large-scale grids represent the layout of Output units (left panel) and PFC units (right panel), with the smaller grids for each unit showing weights from the 5x5 PFC layer. Large white squares represent very strong weights (near 1) and large black square represent very weak weights (near 0). The output layer is organized by row into the 5 different stimulus dimensions — the key finding is that each row receives the same pattern of strong weights from a subset of PFC units that thus represent the entire dimension. The right panel shows that these same PFC units are strongly interconnected with each other.

The Contribution of the Dynamically Gated PFC

Our next objective was to determine how important the dynamically gated PFC mechanism is to successful task performance. We did this by comparing the performance of the full PFC model with the following comparison models:

NoGate A model with the identical connectivity as the full PFC model but without the AC dynamic gating mechanism. This reveals the importance of the gating mechanism in comparison to the full model.

NoPFC A model lacking both the PFC and its dynamic gating mechanism — this model must rely exclusively on weight-based learning mechanisms and thus reveals the importance of the activation-based working memory mechanisms supported by the PFC.

SRN A simple recurrent network (SRN) network that has a context layer updated as a copy of the previous hidden layer activations (Elman, 1990). This provides a form of memory functionally similar to the PFC mechanism, but it lacks the specialized mechanisms of the dynamic gating mechanism.

BP A basic three-layer backpropagation network for comparison with the features of the Leabra algorithm (used in all the above networks).

All models had the same number of hidden units as the standard model, and all other parameters were the same. The BP model used cross-entropy error with an error tolerance of .01, learning rate of .1, and no momentum. Each model was tested for 100 epochs of 250 events each. These 250 events are sequentially organized along the five possible dimensions, that is: the hidden dimension is dimension 1 for the first 50 events, dimension 2 for events 51 to 100, dimension 3 for events 101 to 150, etc. The input stimulus was randomly generated for each event. Thus, the best possible error level would be 5 errors per epoch, one for each task switch.

The results from all of these models, and the full PFC model, are shown in figure 5 and figure 6. These results show clearly that this task cannot be solved by standard networks like simple recurrent networks or backpropagation networks. Even though they both appear to partially solve the task initially, performance deteriorates over time, presumably due to a build up of interference from repeated task switching. In the case of the BP network, we observed that it was producing a blend of the different possible output values, as has been observed previously in cases where, as in this task, the same input leads to different outputs on different trials (e.g., Movellan & McClelland, 1993).

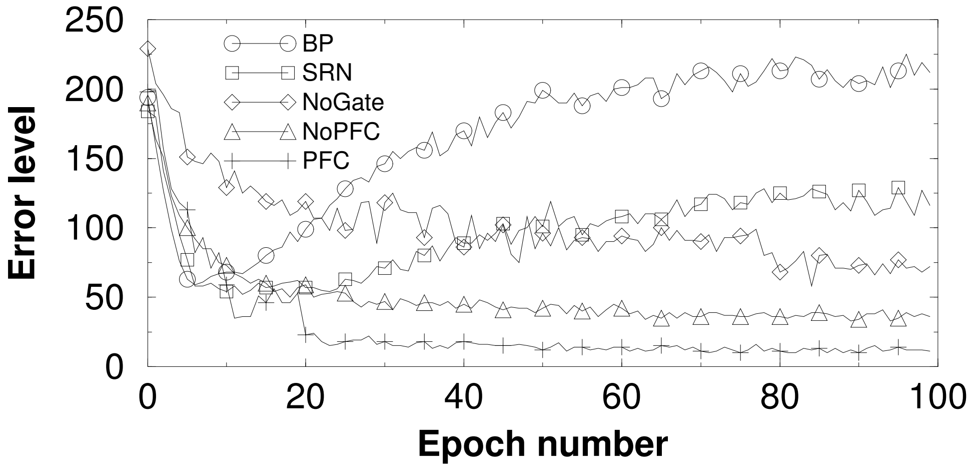

Figure 5: Comparison of the learning performance (error count, with error defined as any unit on the wrong side of .5 for a given trial) of the full PFC model with a range of comparison models. Best possible error is 5 per epoch (one for each task switch). Clearly, the gated PFC is critical for good task switching performance.

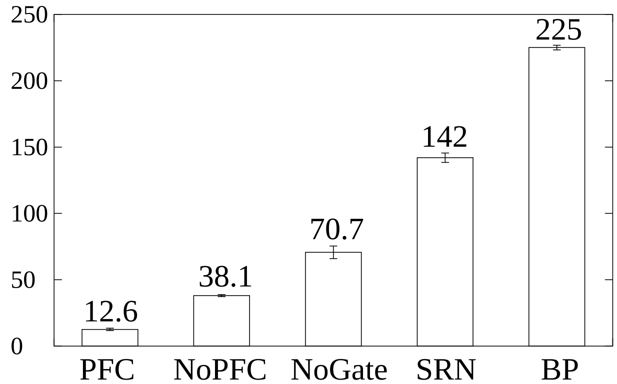

Figure 6: Comparison of the learning performance between all different models. Presented results are the error count at epoch 100 averaged over 10 simulations.

Unlike the feedforward BP network, the NoPFC Leabra model can exhibit attractor dynamics via bidirectional connectivity (between the hidden and output layers). These attractor dynamics, coupled with inhibitory competition in the Leabra algorithm, enable the network to more rapidly learn to settle into different output states for the same inputs (O'Reilly & Munakata, 2000). This explains the better performance of the NoPFC model relative to BP.

Finally, the NoGate model clearly shows that the dynamic gating mechanism is critical for making effective use of the PFC context representations in task switching. Without it, the model tends to develop several PFC context representations that are disconnected from the dimension information. These representations actually impairs task switching because they confuse the network regarding the relevant dimension at one time.

Contributions of Additional Gating Mechanisms

To evaluate the contributions of the additional gating mechanisms described above, we compared the full PFC model to PFC model variations where a specific mechanism was disabled:

No average The computation of reward is computed along one time step (instead of two).

No reset The average reward value is not reset after it goes below threshold.

No negative bias The negative bias mechanism is disabled.

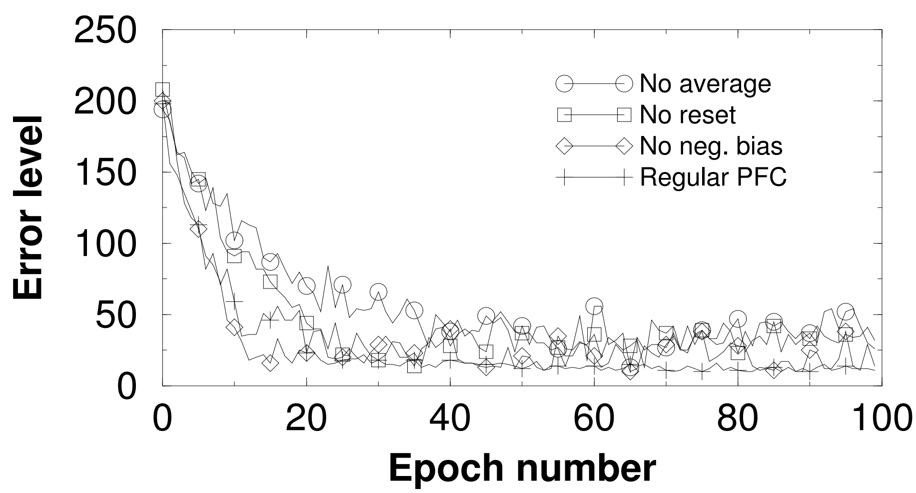

Figure 7: Comparison results of the full PFC model with versions having a specific additional gating mechanism disabled. It is clear that the full set of mechanisms is necessary for optimal performance.

From the results presented in figure 7, it is clear that each additional gating mechanism plays an important role in stabilizing PFC representations, consistent with the motivations provided when these mechanisms were introduced. Nevertheless, even with these mechanisms disabled the network is still generally performing better than the alternative networks explored in the previous section. Thus, these mechanisms can be considered a fine tuning of the overall gating process.

Conclusions

We have shown that the specialized activation-based mechanisms that we hypothesize are supported by the prefrontal cortex and associated subcortical neural systems (in this case the ventral tegmental area and its dopamine neuromodulatory outputs) can support more rapid task switching. This is consistent with cognitive neuroscience data (Burgess et al., 2000), and with the broader literature showing the involvement of the prefrontal cortex in the Wisconsin Card Sorting Task, which our task is a simpler variant of. This model extends earlier models demonstrating similar points (O'Reilly et al., in press; O'Reilly & Munakata, 2000) by showing that the network can develop its representations strictly through learning mechanisms in the process of repeatedly switching among a set of tasks over an extended period. These are advances help to establish the general importance of these mechanisms.

The importance of having working-memory representations that can develop on their own in response to task constraints can be highlighted by contrast with the most commonly used models for temporally-extended tasks, the simple recurrent network (SRN). In the SRN, the context representation is simply a copy of the hidden layer, and thus does not enable the network to develop different representations in the context. In contrast, we saw that the present model was capable of developing PFC context representations that abstracted out the notion of a stimulus dimension, because such an abstraction was critical for task switching performance. The hidden layer did not develop such an abstraction because it needs to produce a specific input/output mapping at the level of stimulus features, not dimensions. We think that this example may be representative of general distinctions between posterior cortex and PFC representations (i.e., posterior is more embedded and diffuse while PFC is more discrete and systematic), and that this may have important implications for understanding the unique contributions that the PFC makes in human cognition. We are currently exploring this possibility using the present model on a range of other tasks.

Acknowledgments

This work was supported by ONR grant N00014-00-1-0246 and NSF grant IBN-9873492.

References

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell, & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 713–737). Cambridge, MA: MIT Press.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*, 848–863.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review*, *97*(3), 332–361.

Cohen, J. D., & O'Reilly, R. C. (1996). A preliminary theory of the interactions between prefrontal cortex and hippocampus that contribute to planning and prospective memory. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp. 267–296). Mahwah, New Jersey: Erlbaum.

Dehaene, S., & Changeux, J. P. (1991). The Wisconsin Card Sorting Test: Theoretical analysis and modeling in a neuronal network. *Cerebral Cortex*, *1*, 62–79.

Dias, R., Robbins, T. W., & Roberts, A. C. (1997). Dissociable forms of inhibitory control within prefrontal cortex with an analog of the Wisconsin Card Sort Test: Restriction to novel situations and independence from "on-line" processing. *Journal of Neuroscience*, *17*, 9285–9297.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between the frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, and Behavioral Neuroscience*, *1*, 137–160.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167–202.

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *Journal of Neuroscience*, *16*, 1936–1947.

Movellan, J. R., & McClelland, J. L. (1993). Learning continuous probability distributions with symmetric diffusion networks. *Cognitive Science*, *17*, 463–496.

O'Reilly, R. C. (1996). Biologically plausible error-driven learning using local activation differences: The generalized recirculation algorithm. *Neural Computation*, *8*(5), 895–938.

O'Reilly, R. C. (1998). Six principles for biologically-based computational models of cortical cognition. *Trends in Cognitive Sciences*, *2*(11), 455–462.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1999). A biologically based computational model of working memory. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control.* (pp. 375–411). New York: Cambridge University Press.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.

O'Reilly, R. C., Noelle, D., Braver, T. S., & Cohen, J. D. (in press). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral Cortex*.

Owen, A. M., Roberts, A. C., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1993). Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain*, *116*, 1159–1175.

Roberts, A. C., Robbins, T. W., & Everitt, B. J. (1988). The effects of intradimensional and extradimensional shifts on visual discrimination learning in humans and non-human primates. *Quarterly Journal of Experimental Psychology*, *40*, 321–341.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, *13*, 900–913.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., Murphy, K. J., & Izukawa, D. (2000). Wisconsin card sorting test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*, 388–402.

Sutton, R. S. (1988). Learning to predict by the method of temporal diferences. *Machine Learning*, *3*, 9–44.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction.* Cambridge, MA: MIT Press.