ELSEVIER

Cognitive Science 26 (2002) 503–520

COGNITIVE SCIENCE

http://www.elsevier.com/locate/cogsci

Brief report

Learning representations in a gated prefrontal cortex model of dynamic task switching

Nicolas P. Rougier, Randall C. O'Reilly*

Department of Psychology, University of Colorado Boulder, 345 UCB, Boulder, CO 80309-0345, USA

Received 30 November 2001; received in revised form 8 April 2002; accepted 18 April 2002

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. Furthermore, we show that a lesioned model produces perseverative errors like those seen in frontal patients.

Keywords: Working memory; Wisconsin card sorting task; Reinforcement learning; Computational modeling; Neural networks

1. 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

* Corresponding author. Tel.: +1-303-492-0054; fax: +1-303-492-2967.
E-mail addresses: rougier@psych.colorado.edu (N.P. Rougier), oreilly@psych.colorado.edu (R.C. O'Reilly).

0364-0213/02/$ – see front matter
PII: S0364-0213(02)00073-3

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 task for demonstrating the involvement of the PFC in task switching is the Wisconsin card sorting task (WCST; Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtiss, 1993; Milner, 1963), though the literature on this remains somewhat controversial (e.g., Mountain & Snow, 1993; Reitan & Wolfson, 1994; Stuss et al., 2000). Further evidence of PFC involvement in task switching comes from tasks related to the WCST (e.g., Dias, Robbins, & Roberts, 1997; Owen et al., 1993; Roberts, Robbins, & Everitt, 1988), including tasks demonstrating task switching impairments in children presumably due PFC immaturity (e.g., Munakata & Yerys, 2001; Zelazo, Frye, & Rapus, 1996) 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 (Braver & Cohen, 2000; Frank, Loughry, & O'Reilly, 2001; Miller & Cohen, 2001; O'Reilly, Braver, & Cohen, 1999; O'Reilly & Munakata, 2000; O'Reilly, Noelle, Braver, & Cohen, 2002). 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 & Munakata, 2000; O'Reilly et al., 2002).

The present model goes beyond earlier models of task switching involving a dynamically gated PFC system (O'Reilly & Munakata, 2000; O'Reilly et al., 2002) 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.

2. 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

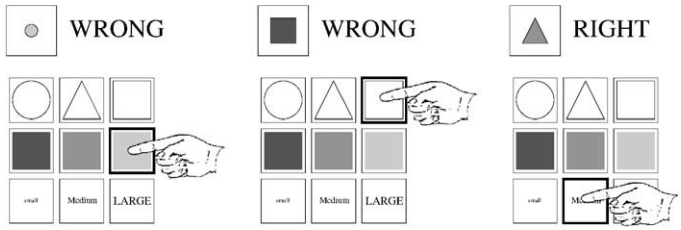

Fig. 1. The task. Each stimulus has three feature dimensions (e.g., shape, color, size) with three possible feature levels (circle/square/triangle, red/green/blue, small/medium/large) and the hidden dimension is the size of the presented sample stimulus. In the first two trials the subject chose the 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.

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 (Fig. 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.

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 (Fig. 1).

3. The model

Fig. 2 shows the structure of the model, which has the same basic elements as our previous task-switching models, and, like them, is implemented using the Leabra framework for the activation and learning mechanisms (O'Reilly, 1998; O'Reilly & Munakata, 2000) (see Appendix A for equations). 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. We tested a range of different hidden layer sizes from 16 to

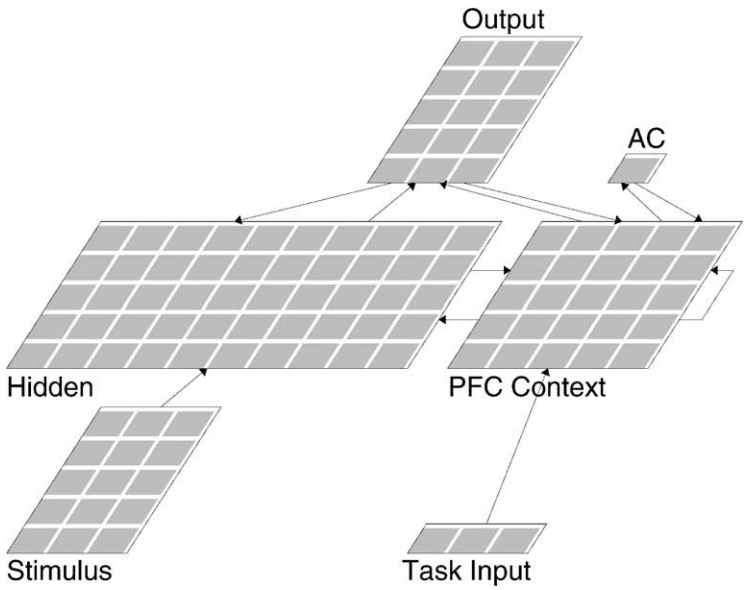

Fig. 2. Dynamic task-switching model using active PFC context representations.

50 with no noticeable changes in performance—this simple task is not particularly computationally demanding, but we have used this same basic model for a number of more complex tasks in our current research.

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 PFC layer facilitates this dimensional selection by providing top-down support or biasing (Cohen, Dunbar, & McClelland, 1990; Cohen & O'Reilly, 1996) 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 & Munakata, 2000; O'Reilly et al., 2002), 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. (2002).

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. As described in the next section, the updating of this layer is influenced by the reward-based dynamic gating mechanism as implemented in the adaptive critic (AC) layer, which acts on excitatory intracellular currents within the PFC units. Aside from these key specializations, the PFC units are otherwise the same as all the other units in the network.

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.

3.1. The adaptive critic 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 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 segmental 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 & Munakata, 2000; O'Reilly et al., 2002).

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.

3.2. 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 versus 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.

3.3. 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 Section 4, this mechanism is of a great help for stabilizing PFC representations and acts indeed as a very short term memory of the past.

3.4. Overview of model's task switching behavior

We can now illustrate the overall behavior of these mechanisms by considering the example given in Fig. 3. In this example, the model is currently performing the subtask of naming the feature values along the fifth dimension, while the dimension has just been switched to the fourth one. On the first trial following the switch and because of the context present within PFC,

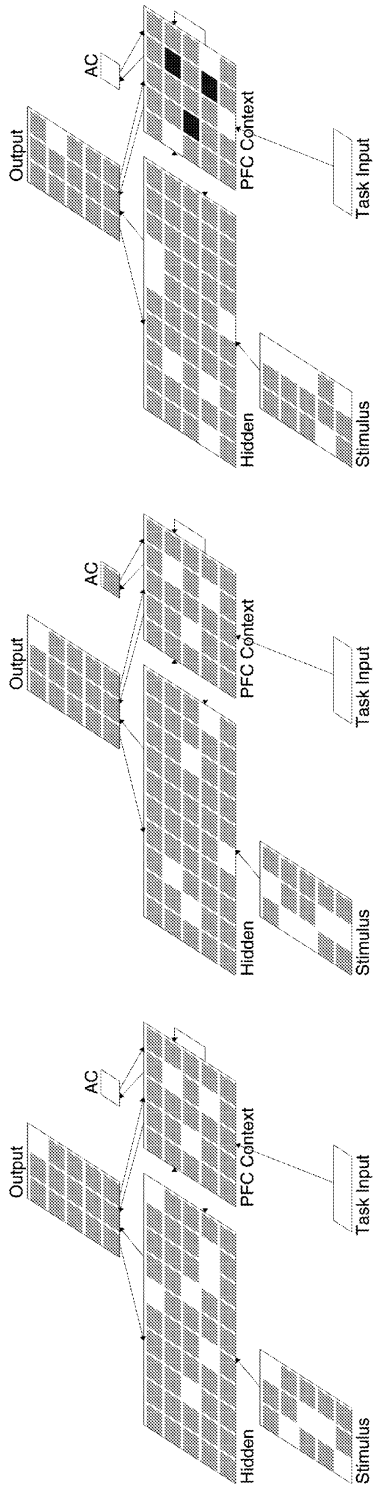

Fig. 3. Sequence of context activation within the PFC layer during a task switch. Grey squares represent inactive units, white squares represent active units and dark grey squares represents units with a negative bias activity.

the model continues to produce its answer based on the feature active in the fifth 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 fifth 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 fourth 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.

4. 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.
- To compare the performance of the intact and frontally-lesioned model to task-switching performance in intact and frontal patients.

4.1. Development of PFC representations

The most important result is that the full model as already described 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 Fig. 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 Fig. 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. These dimensional representations developed reliably across all runs of the model, and are in fact the kinds of representations we assumed should exist in the PFC in our earlier hand-coded models (O'Reilly & Munakata, 2000; O'Reilly et al., 2002). 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.

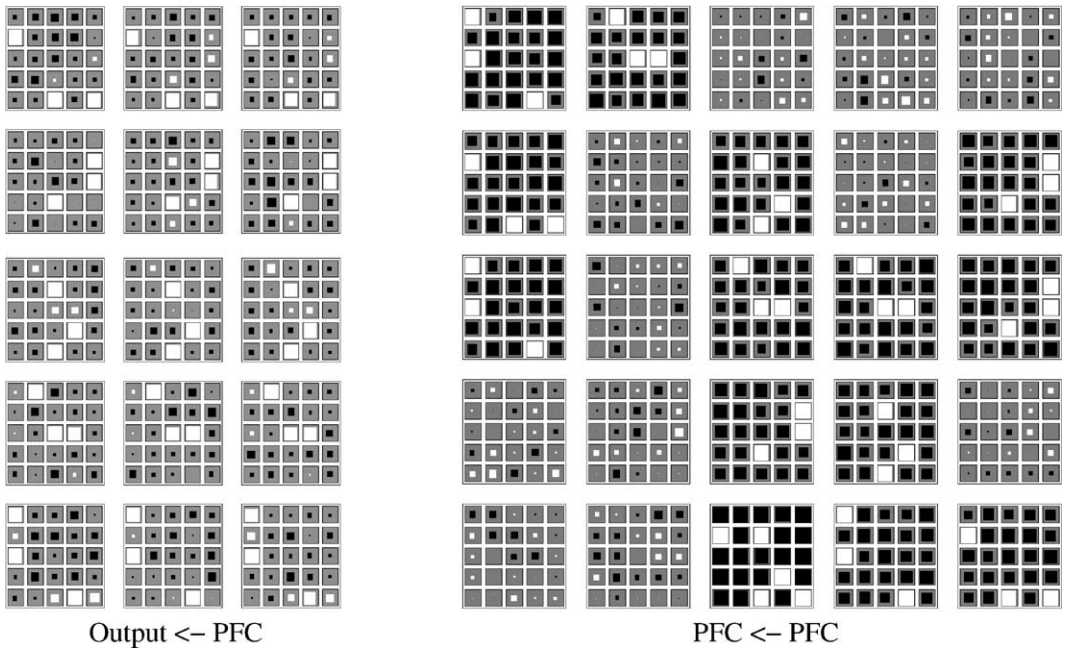

Fig. 4. Matrices of weights from PFC to output layer and from PFC to PFC layer. The large-scale grids represent the layout of output units (left panel) and PFC units (right panel), with the smaller grids for each unit showing weights from the 5×5 PFC layer. Large white squares represent very strong weights (near 1) and large black square represent very weak weights (near 0). The output layer is organized by row into the five 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.

4.2. 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) 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.

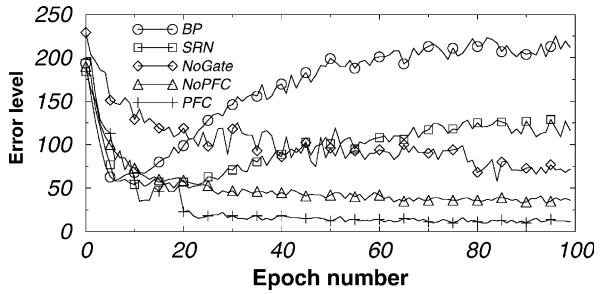

Fig. 5. Comparison of the learning performance (error count, with error defined as any unit on the wrong side of .5 for a given trial) of the full PFC model with a range of comparison models. Best possible error is 5 per epoch (one for each task switch). Clearly, the gated PFC is critical for good task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

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 51–100, dimension 3 for events 101–150, etc. The input stimulus was randomly generated for each event. Thus, the best possible error level would be five errors per epoch, one for each task switch.

The results from all of these models, and the full PFC model, are shown in Figs. 5 and 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

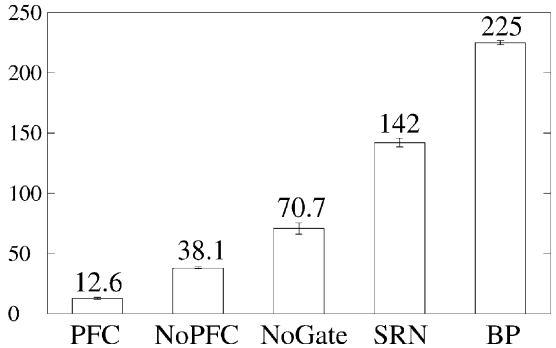

Fig. 6. Comparison of the learning performance between all different models. Presented results are the error count at epoch 100 averaged over 10 simulations.

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).

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 impair task switching because they confuse the network regarding the currently relevant dimension.

4.3. Contributions of additional gating mechanisms

To evaluate the contributions of the additional gating mechanisms already described, 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.

From the results presented in Fig. 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.

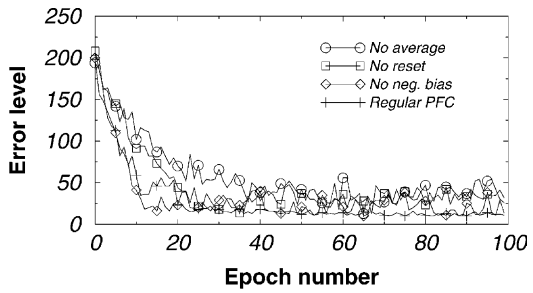

Fig. 7. Comparison results of the full PFC model with versions having a specific additional gating mechanism disabled. It is clear that the full set of mechanisms is necessary for optimal performance.

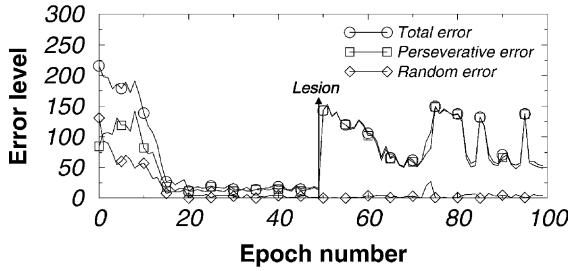

Fig. 8. The frontal model has been heavily lesioned at epoch 49 resulting in an immediate degradation in performance. There is a clear dissociation between random and perseverative errors as observed in the neuropsychological data.

4.4. Impact of lesions on perseveration

We now compare the performance of our model to the patterns of behavior seen in normal humans and brain damaged patients. Because there is considerable variability across studies, we will focus on the basic notion that frontal damage or immaturity (in children) leads to an increase in perseverative errors on the WCST task (e.g., Grant & Berg, 1948; Heaton et al., 1993; Milner, 1963; Munakata & Yerys, 2001; Zelazo et al., 1996), instead of trying to fit the data from any one study. A perseverative error occurs when the response is consistent with the previous sorting rule after the sorting rule has changed. It is clear that to the extent that the PFC facilitates rapid switching to the new rule, damage or immaturity (in children) of the PFC should produce perseverative errors. To account for these data, we observed the type of errors made by both an intact and a lesioned PFC model (Fig. 8). In early training (epoch 0–19), which is analogous to early development as it involves the development of the basic representations underlying task performance, the model is indifferently performing random and perseverative errors. After this "developmental" period, the model produces mainly perseverative errors. Once each task has been learned (epoch 19–49), there are still residual errors that occur at each task switch (not shown on Fig. 8). As explained before, these errors are necessary for the model to be able to switch, and they are perseveration errors. At epoch 49, we heavily lesioned the PFC layer (75% of units were lesioned), producing an immediate impact on performance. Consistent with the behavioral data, almost all of the errors produced by this lesion were perseverative.

Finally, note that the lesioned frontal model is roughly equivalent to the NoPFC model, we introduced earlier, and yet it produces worse overall performance. This is because it has to unlearn its prior reliance on the PFC layer, and learn to use its hidden layer in a new way. This provides an interesting illustration of possible complications in neural reorganization following brain damage.

4.5. Predictions

Because it embodies a number of very specific mechanisms and assumptions, our model naturally leads to a number of testable predictions. Here, we draw out one set of particularly interesting and perhaps counter-intuitive such predictions. The first prediction is that if one were

able to induce performance errors in people during a given block of same-task trials, one should be able to destabilize their PFC representations through the actions of the gating mechanism. To test such a prediction, one would need to find a way of inducing errors in performance other than the usual switching of the task—perhaps speeded responding and/or degraded stimuli could be used. In addition, it would be tricky to observe the behavioral signature of this destabilization over an above the actual behavioral errors produced by the manipulation—once the network has switched into a given task, the weights in the non-PFC portion reinforce performance on that task, making it less susceptible to switching to a new task. For these reasons, it might be easier to observe this PFC destabilization effect using neuroimaging techniques. If it were possible to observe this basic effect, then a number of other interesting predictions could be tested, for example regarding the number of errors in a row required for destabilization, and the effects of any inhibition of previously-active PFC representations.

5. 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 segmental area and its dopamine neuromodulatory outputs) can support more rapid task switching. We demonstrated that the model is consistent with the empirical 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 & Munakata, 2000; O'Reilly et al., 2002) 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 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) (e.g., Elman, 1990). 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.

Appendix A. Implementational details

The model was implemented using the Leabra framework, which is described in detail in O'Reilly and Munakata (2000), and summarized here. This framework has been used to simulate over 40 different models in O'Reilly and Munakata (2000), and a number of other research models. Thus, the model can be viewed as an instantiation of a systematic modeling framework using standardized mechanisms, instead of constructing new mechanisms for each model. The model can be obtained by e-mail at oreilly@psych.colorado.edu.

A.1. Pseudocode

The pseudocode for Leabra is given here, showing exactly how the pieces of the algorithm described in more detail in the subsequent sections fit together.

A.1.1. Outer loop

Iterate over events (trials) within an epoch. For each event:

1. Iterate over minus and plus phases of settling for each event.
   (a) At start of settling, for all units:
      i. Initialize all state variables (activation, v_m, etc.).
      ii. Apply external patterns (clamp input in minus, input and output in plus).
   (b) During each cycle of settling, for all non-clamped units:
      i. Compute excitatory net input ($g_e(t)$ or η_j, Eq. (A.2)).
      ii. Compute kWTA (k-Winners-Take-All) inhibition for each layer, based on $g^{\Theta;i}$ (Eq. (A.6)):
         A. Sort units into two groups based on $g^{\Theta;i}$: top k and remaining $k+1$ to n.
         B. If basic, find k and $k+1$th highest; if average-based, compute average of $1 \rightarrow k$ and $k+1 \rightarrow n$.
         C. Set inhibitory conductance g_i from $g^{\Theta;k}$ and $g^{\Theta;k+1}$ (Eq. (A.5)).
      iii. Compute point-neuron activation combining excitatory input and inhibition (Eq. (A.1)).
   (c) After settling, for all units:
      i. Record final settling activations as either minus or plus phase ($y^{-;j}$ or $y^{+;j}$).
2. After both phases update the weights (based on linear current weight values), for all connections:
   (a) Compute error-driven weight changes (Eq. (A.8)) with soft weight bounding (Eq. (A.9)).
   (b) Compute Hebbian weight changes from plus-phase activations (Eq. (A.7)).
   (c) Compute net weight change as weighted sum of error-driven and Hebbian (Eq. (A.10)).
   (d) Increment the weights according to net weight change.

A.2. Point neuron activation function

Leabra uses a *point neuron* activation function that models the electrophysiological properties of real neurons, while simplifying their geometry to a single point. This function is

nearly as simple computationally as the standard sigmoidal activation function, but the more biologically-based implementation makes it considerably easier to model inhibitory competition, as described. Further, using this function enables cognitive models to be more easily related to more physiologically detailed simulations, thereby facilitating bridge-building between biology and cognition.

The membrane potential V_m is updated as a function of ionic conductances g with reversal (driving) potentials E as follows:

$$\Delta V_m(t) = \tau \sum_c g_c(t) \bar{g}_c (E_c - V_m(t)) \tag{A.1}$$

with three channels (c) corresponding to e, the excitatory input; l leak current; and i inhibitory input. Following electrophysiological convention, the overall conductance is decomposed into a time-varying component $g_c(t)$ computed as a function of the dynamic state of the network, and a constant g_c that controls the relative influence of the different conductances.

The excitatory net input/conductance $g_e(t)$ or η_j is computed as the proportion of open excitatory channels as a function of sending activations times the weight values:

$$\eta_j = g_e(t) = \langle x_i w_{ij} \rangle = \frac{1}{n} \sum_i x_i w_{ij} \tag{A.2}$$

The inhibitory conductance is computed via the kWTA function described in the next section, and leak is a constant.

Activation communicated to other cells (y_j) is a threshold (Θ) sigmoidal function of the membrane potential with gain parameter γ:

$$y_j(t) = \frac{1}{1 + 1/(\gamma [V_m(t) - \Theta]_+)} \tag{A.3}$$

where $[x]_+$ is a threshold function that returns 0 if $x < 0$ and x if $X > 0$. Note that if it returns 0, we assume $y_j(t) = 0$, to avoid dividing by 0. As it is, this function has a very sharp threshold, which interferes with graded learning mechanisms (e.g., gradient descent). To produce a less discontinuous deterministic function with a softer threshold, the function is convolved with a Gaussian noise kernel ($\mu = 0$, $\sigma = .005$), which reflects the intrinsic processing noise of biological neurons:

$$y_j^*(x) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\Pi}\sigma} e^{-z^2/2\sigma^2} y_j(z - x)\, dz \tag{A.4}$$

where x represents the $[V_m(t) - \Theta]_+$ value, and $y^{*;j}(x)$ is the noise-convolved activation for that value. In the simulation, this function is implemented using a numerical lookup table.

A.3. k-Winners-Take-All inhibition

Leabra uses a kWTA function to achieve inhibitory competition among units within a layer (area). The kWTA function computes a uniform level of inhibitory current for all units in the layer, such that the $k+1$th most excited unit within a layer is below its firing threshold, while the

*k*th is above threshold. Activation dynamics similar to those produced by the kWTA function have been shown to result from simulated inhibitory interneurons that project both feedforward and feedback inhibition (O'Reilly & Munakata, 2000). Thus, although the kWTA function is somewhat biologically implausible in its implementation (e.g., requiring global information about activation states and using sorting mechanisms), it provides a computationally effective approximation to biologically plausible inhibitory dynamics.

kWTA is computed via a uniform level of inhibitory current for all units in the layer as follows:

$$g_i = g^{\Theta}_{k+1} + q(g^{\Theta}_k - g^{\Theta}_{k+1}) \tag{A.5}$$

where $0 < q < 1$ (.25 default used here) is a parameter for setting the inhibition between the upper bound of $g^{\Theta;k}$ and the lower bound of $g^{\Theta;k+1}$. These boundary inhibition values are computed as a function of the level of inhibition necessary to keep a unit right at threshold:

$$g^{\Theta}_i = \frac{g^*_e \bar{g}(E_e - \Theta) + g_l \bar{g}_l (E_l - \Theta)}{\Theta - E_i} \tag{A.6}$$

where $g^{*;e}$ is the excitatory net input without the bias weight contribution—this allows the bias weights to override the kWTA constraint.

A.4. Hebbian and error-driven learning

For learning, Leabra uses a combination of error-driven and Hebbian learning. The error-driven component is the symmetric midpoint version of the GeneRec algorithm (O'Reilly, 1996), which is functionally equivalent to the deterministic Boltzmann machine and contrastive Hebbian learning (CHL). The network settles in two phases, an expectation (minus) phase where the network's actual output is produced, and an outcome (plus) phase where the target output is experienced, and then computes a simple difference of a pre- and post-synaptic activation product across these two phases. For Hebbian learning, Leabra uses essentially the same learning rule used in competitive learning or mixtures-of-Gaussians which can be seen as a variant of the Oja normalization. The error-driven and Hebbian learning components are combined additively at each connection to produce a net weight change.

The equation for the Hebbian weight change is:

$$\Delta_{\text{hebb}} w_{ij} = x^+_i y^+_j - y^+_j w_{ij} = y^+_j (x^+_i - w_{ij}) \tag{A.7}$$

and for error-driven learning using CHL:

$$\Delta_{\text{err}} w_{ij} = (x^+_i y^+_j) - (x^-_i y^-_j) \tag{A.8}$$

which is subject to a soft-weight bounding to keep within the 01 range:

$$\Delta_{\text{sberr}} w_{ij} = [\Delta_{\text{err}}]_+ (1 - w_{ij}) + [\Delta_{\text{err}}]_- w_{ij} \tag{A.9}$$

The two terms are then combined additively with a normalized mixing constant k_{hebb}:

$$\Delta w_{ij} = \varepsilon[k_{\text{hebb}}(\Delta_{\text{hebb}}) + (1 - k_{\text{hebb}})(\Delta_{\text{sberr}})] \tag{A.10}$$

A.5. Temporal differences and adaptive critic gating mechanisms

To implement the temporal differences (TD) algorithm in Leabra, the AC unit in the VTA layer has plus–minus phase states that correspond to the expected reward at the previous time step (minus) and the current time step (plus). The difference between these two states is the TD error δ, which is essentially equivalent to the more standard kinds of error signals computed by the error-driven learning component of Leabra, except that it represents an error of prediction over time, instead of an instantaneous error in the network output. This δ value then modulates the strength of an excitatory ionic current (labeled *m* here) that helps to maintain PFC activations:

$$g_{\mathrm{m}}(t-1) = 0, \quad \text{if } |\delta(t)| > r \tag{A.11}$$

$$g_{\mathrm{m}}(t_j) = g_{\mathrm{m}}(t-1) + \delta(t) y_j(t) \tag{A.12}$$

Thus, a positive δ increases this maintenance current for active units, and a negative δ decreases it. Furthermore, if δ is sufficiently large in magnitude (greater than the reset threshold *r*), it resets any existing currents.

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* (Vol. 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, NJ: Lawrence 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.

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl type card sorting problem. *Journal of Experimental Psychology*, *38*, 404–411.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin card sorting test manual: Revised and expanded*. Odessa, FL: Psychological Assessment Resources Inc.

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

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology*, *9*, 90–100.

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.

Mountain, M. A., & Snow, W. G. (1993). Wisconsin card sorting test as a measure of frontal pathology: A review. *The Clinical Neuropsychologist*, *7*, 108–118.

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

Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. *Psychological Science*, *12*, 335–337.

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. (2002). Prefrontal cortex and dynamic categorization tasks: Representational organization and neuromodulatory control. *Cerebral Cortex*, *12*, 246–257.

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.

Reitan, R. M., & Wolfson, D. (1994). A selective and critical review of neuropsychological deficits and the frontal lobes. *Neuropsychology Review*, *4*, 161–198.

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.

Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development*, *11*, 37–63.