Prefrontal Cortex: Delay-related activity

Spatial delayed-response task; Funahashi et al., 1989
Remember the SRN? (chap 6)

An Architecture for Sequence Learning

Simplified Recurrent Network (SRN):
Remember the SRN? (chap 6)

This is a *gating* network: context only updated at discrete timepoints.

An Architecture for Sequence Learning: Simple Recurrent Network (SRN)
Simple SRN story is not flawless

• How is hidden "copy" function implemented biologically?

• During settling, context must be actively maintained (ongoing hidden activity has no effect on context).

• What if distracting information presented in middle of sequence? Want to only hold on to relevant information in WM.

• Assumed all context is relevant. What if distracting information in WM.

• What if want to hold on to more than one piece of information in WM at a time? Or to selectively update one part of WM while continuing to robustly maintain others?

• What if the decision of whether or not to update information depends on currently internal WM state?

• And what if the decision of whether or not to update information actively maintained otherwise?
SZ patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.

Schizophrenia: Impaired PFC/BG Function
Schizophrenia: Impaired PFC/BG Function

• SZ Patients have deficits in the same cognitive tasks as those seen in PFC/BG damage.

But PFC patients don’t have delusions, hallucinations, psychoses. Also, some imaging studies show increase in PFC activity in SZ (e.g., Manoach, 2003; Callicott et al, 2003).
Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage. (Weinberger, O'Donnell, Greice, etc)

- But PFC patients don't have delusions, hallucinations, psychoses. Also, psychosis thought to stem from increased DA in BG (Weinberger, Callicott et al, 2003; Manoach, 2003). Some imaging studies show increased PFC activity in SZ (e.g. Weinberger, 2003).

- Patients with PFC damage also have deficits in the same cognitive tasks as those seen in SZ patients.
Schizophrenia: Impaired PFC/BG Function

- SZ Patients have deficits in the same cognitive tasks as those seen in patients with PFC damage.
- But PFC patients don't have delusions, hallucinations, psychosis. Also, some imaging studies show increased PFC activity in SZ (e.g., Manoach, 2003; Callicott et al., 2003).
- Psychosis thought to stem from increased DA in BG (Weinberger, et al., 2000; Remy et al., 2000).
- Other disorders with BG/DA dysfunction are associated with frontal-like cognitive deficits (PD, addiction, ADHD, etc); these deficits correlate with BG/DA dysfunction (Muller et al., 2000; Remy et al., 2000).
Widely Accepted Role of Prefrontal Cortex (PFC) is its robust maintenance of neural activity (working memory) and ability to rapidly switch state (e.g., task switching). PFC helps keep us on task, promotes cognitive flexibility via ability to maintain multiple tasks.
Widely Accepted Role of Prefrontal Cortex (PFC)

PFC helps keep us on task, promotes cognitive flexibility via ability to rapidly switch state (e.g., task switching):

- Robust maintenance of neural activity (working memory)
- Flexibility from adaptive gating (via Basal Ganglia, BG) that switches between maintenance and rapid updating
Widely Accepted Role of Prefrontal Cortex (PFC)

PFC helps keep us on task, promotes cognitive flexibility via ability to rapidly switch state (e.g., task switching):

- Extensive interconnectivity, allowing top-down biasing of task-relevant
- Flexibility from adaptive gating (via Basal Ganglia, BG) that switches
- Robust maintenance of neural activity (working memory)
- Widely Accepted Role of Prefrontal Cortex (PFC)
You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.

- Working memory: robust maintenance of information, but must also have ability to be rapidly updated — requires gating.

You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.

You’ve got to know when to hold ’em, know when to fold ’em.
Working Memory Demands: Updating & Maintenance

Hochreiter & Schmidhuber, 1997; Braver & Cohen, 1999; Frank et al, 2001

- Working memory: robust maintenance of information, but must also have ability to be rapidly updated — requires gating.
- You’ve got to know when to hold ’em, know when to fold ’em.
But who controls the controller?
BG damage $\Rightarrow$ deficits in motor, learning, motivation, working memory, cognitive control

Alexander, G. et al. (1986), Parallel Organization of Functionally Segregated Circuits Within Basal Ganglia 

Basal Ganglia Architecture: Cortically Based Loops
In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996).
In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996).

In parallel circuits, the BG may reinforce the updating of PFC working memory representations (Alexander et al., 1996; Frank, Loughry & O'Reilly, 2001).
In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996).

In parallel circuits, the BG may reinforce the updating of PFC working memory representations (Alexander et al., 1986; Frank, Loughry & O'Reilly, 2001).

Dopamine in PFC supports robust maintenance over time (Lewis & O'Donnell, 2000; Durstewitz & Seamans, 2002).

In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996).

BC Model Extension to Working Memory and Attention

BC
In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996). In parallel circuits, the BG may reinforce the updating of PFC working memory representations (Alexander et al., 1986; Frank, Loughry & O'Reilly, 2001). Dopamine in PFC supports robust maintenance over time (Lewis & O'Donnell, 2000; Durstewitz & Seamans, 2002). Phasic DA bursts thought to occur for task-relevant ("positive") information reinforce BG updating signals (O'Reilly & Frank, 2006). BC Model Extension to Working Memory and Attention
In the motor domain, the BG selectively facilitates one command while suppressing others (Mink, 1996).

• In parallel circuits, the BG may reinforce the updating of PFC working memory representations (Alexander et al., 1986; Frank, Loughry & O'Reilly, 2006).

• Dopamine in PFC supports robust maintenance over time (Lewis & O'Donnell, 2000; Durstewitz & Seemans, 2002).

• Phasic DA bursts thought to occur for task-relevant (“positive”) information, reinforcing BG updating signals (O'Reilly & Frank, 2006).

• Time course of DA activity: maintenance in PFC, updating thru BG.

• In parallel circuits, the BG selectively facilitates one command while suppressing others (Mink, 1996).
BG Gating of PFC Working Memory / Attention
PFC robustly maintains prior states.

Base state: Thalamus inhibited = Gate is closed

Go / NoGo

PFC = Gating of Working Memory / Attention

BC Disinhibition of PFC
PFC rapidly updated to maintain new information.

- Striatum fires: Thalamus disinhibited = Gate opened
  - PFC robustly maintains prior states.

- Base state: Thalamus inhibited = Gate is closed

\[
\text{PFC} \quad \text{Disinhibition of PFC} = \text{Gating of Working Memory / Attention}
\]
Parallel Stripes = Selective Gating.

PFC/BC loops form independent stripes = selective gating.

...PFC

...Posterior Cortex

thalamus

Striatum (matrix)

GP (tonic act)
Separate Maintenance vs. Output PFC / BG stripes
Superficial reflects inputs and maintenance
Maintenance via Thalamocortical loops, BG disinhibits

BC Gates flow: Superficial <- Deep PFC
Competition in striatum/GP between Maint vs Output and diff stripes

Multiple stripes, competition
Example Task: Specific Working Memory Demands

Must maintain outer loop (1, 2) while updating inner loop (A, X, ...).

Target (R):

\[ \text{L} = B - Y \]
\[ \text{L} = A - X \]
Example Task: Specific Working Memory Demands

Robust maintenance (over interfering stimuli; outer loop).

Must maintain outer loop (1, 2) while updating inner loop (A, X).

2 = B → Y
1 → A → X
Target (R):
Example Task: Specific Working Memory Demands

- Robust maintenance (over intervening stimuli; outer loop).
- Rapid and selective updating (keep 1, 2 while updating A, X).
Example Task: Specific Working Memory Demands

Many real-world examples: language, planning,...

- Robust maintenance (over intervening stimuli, outer loop).
- Rapid and selective updating (keep 1,2 while updating A,X).

Must maintain outer loop (1,2) while updating inner loop (A,X).

\[
\begin{align*}
2 &= B \land \\
1 &= A \land \\
\text{Target} (R) &:
\end{align*}
\]
PFC/BG Model of I-2-AX

(Frank, Loughry & O'Reilly, 2001)
FC/BG Model of 1-2-AX

(Frank, Loughry & O’Reilly, 2001)
FC/BG Model of 1-2-AX

(Frank, Loughry & O'Reilly, 2001)
But How do BG "Know" What to Update/Ignore?
But how do BG “know” what to update/ignore?
DA effects on BG Updating Of PFC

Same mechanism as in basic motor circuitry! (Frank, 05; Gerfen, 00)

\[ \text{Dip} = \text{Release of D2 inhibition on NoGo} \]

\[ \text{Burst} = \text{Excitatory D1 on Go} \]

* DA bursts/dips modulate Go vs. NoGo firing and learning.

- DA bursts/dips modulate Go vs. NoGo firing and learning.
PFC/BG Model, Learns 12-AX and other WM tasks

(O’Reilly & Frank, 2006)
Phonological Loop Task

Store

Recall

input bg pic

Phonological Loop Task

r

1

a

b

c

s
go

a

b

c

r

2

a

b

c

s
go

a

b

c

r

3

a

b

c

s
go

a

b

c
Phonological Loop Task

PBWMLSTMRBPSRN

Algorithm

Epochs to Criterion

Phono Loop Training Time

Phono Loop Generalization

Generalization Error %

Generalization Error %

Phonological Loop Task

PBWM = Prefrontal Basal-ganglia Working Memory; LSTM = Long Short Term Memory

(P.B. Schmidhuber et al.); RBP = Recurrent Backpropagation; SRN = Simple Recurrent Network.

(O’Reilly & Frank, 2006)
Simulating BG DA increases in SZ:
Effects on 12AX Working Memory Performance

Training Needed to Learn

Accuracy

12AX Working Memory Performance

Epochs to Criterion

% Networks Solving Task

Effects on 12AX Working Memory Performance
Simulating BG DA increases in SZ:
Evidence for BG gating of PFC

Instruction cue signals whether subsequent yellow information is distracting (and should not be stored in working memory).

McNab & Klingberg (2008), Nature Neuroscience
Evidence for BG gating of PFC: Lesion patients
Evidence for BG gating of PFC: Neuroimaging

Distractor versus no-distractor contrast: NoGo!

McNab & Klingberg (2008), Nature Neuroscience
Both BG and PFC activity correlate with WM capacity. Better WM function correlates with PFC (middle frontal gyrus) and BG activity in Dist vs NoDist contrast. Globus pallidus = output of BG; stronger ↔ more NoGo

McNab & Klingberg (2008), Nature Neuroscience
(spatial storage; this parietal region is sensitive to memory load).

More GP activity = Less “unnecessary storage” of distracting info in parietal cortex.

Evidence for BG gating of PFC: Neuroimaging.
McNab & Klingberg (2008), *Nature Neuroscience*

No such correlation in PFC, suggesting that BG = filter (gate).

Spatial storage: this parietal region is sensitive to memory load.

More GP activity = less unnecessary storage of distracting info in parietal cortex.

Evidence for BG gating of PFC: Neuroimaging
Testing the Model: AX-CPT Task

Cohen et al. (1997); Barch et al. (2001); Frank & O'Reilly, 2006

• Ignore all distractors (5, 7)

• Non-target sequences A - Y, B - Y; press L button.

• Target sequence is A - X; press R button.

The target sequence is A - X; press R button.

Cohen et al. (1997); Barch et al. (2001); Frank & O'Reilly, 2006
Testing the Model: AX-CPT Task

Cohen et al (1997); Barch et al (2001); Frank & O'Reilly, 2006

- Target sequence is A - X; press R button.
- Ignore all distractors (5,7).
- Tests gating and maintenance components of working memory / attention.

• Target sequence is A - X; press R button.
• Non-target sequences A - Y, B - X, B - Y; press L button.
• Ignore all distractors (5,7).

—Testing the Model: AX-CPT Task—
ta
আ
१०
A-X target sequence on 70% of trials.

B-Y is control condition, not dependent on working memory.

Key conditions: B-X, A-Y

A-X target sequence on 70% of trials.

Working Memory Demands: AX-CPT Task
B-X: ↑ Working memory, ↑ performance
B-X: Working memory ↓ performance

Sensory Input

Working Memory

Gating open → closed

b) Maintain B

5

a) Update B

Input

Gating

Working Memory

B
B-X: \( \downarrow \) Working memory ↓ performance
B-X: Working memory ↓ performance

B: B-B \( \uparrow \) Working memory, \( \uparrow \) performance

Sensory Input

Gating

open

closed

Response

Motor

- b) Maintain B
- a) Update B
- c) Prepare Non-Targ Resp.

OK!
B-X: Working memory ↑ performance ↑
B-X: ↓ Working memory ↓ performance

Sensory Input

Working Memory

Gating

open

a) Update B

b) Update 1

B

5

B

5

open open
B-X: Working memory ↓ performance

Sensory Input

Gating

Working Memory

a) Update B

b) Update

c) No Relevant Context.

5

Motor Response

5

OOPS!
Working memory, ↓ performance

Sensory Input

Gating

open

Working Memory

(4) Update A

A-Y: ↓ Working memory, ↑ performance
A-Y: Working memory ↓ performance

a) Update A
b) Maintain A

Sensory Input
Working Memory
Gating

(open, closed)

Working memory, ↓ performance

Sensory Input

Working Memory

Gating

open

closed

a) Update A
b) Maintain A
c) Prepare Target Resp

Motor Response

A-Y: Working memory ↑ performance
A-Y:

Working memory, performance

Sensory Input

Working Memory

Gating

Open

Closed

OOPS!

Motor

Response

Prepare Target Resp

Maintain A

Update A

A-Y: Working memory↑, performance↓
A-Y: Working memory ↑ performance ↓
A-Y: Working memory \(\uparrow\) performance

\[
\begin{align*}
\text{Sensory Input} & \quad \text{Gating} \\
\text{Working Memory} & \quad \text{Update A}
\end{align*}
\]
A-Y: Working memory \( \uparrow \) performance

- Sensory Input
- Motor Response
- Working Memory
- Gating

1. No Relevant Context: A
2. Update 1: open
3. Update A: open

\( A \rightarrow \) Working memory \( \uparrow \) performance
A-Y: Working memory ↓ performance

OK!

Motor Response

(a) No Relevant Context:

(b) Update 1

(c) Update A

Working Memory

Sensory Input

Gating
Working Memory Demands: AX-CPT Task

• More active maintenance of task-relevant info

  More A-Y false alarms, Less B-X false alarms
WM context index = BX - AX accuracy

Barch et al (2001)

SZ: Impaired Working Memory

WM Context Index

AX-CPT Working Memory

Delay

Long

Short

Schiz

Controls
Double blind, within-subjects design (N=28).

• Cabergoline and Haloperidol: D2 agonist and D2 antagonist

• D2 agents: preferential action in BG

Frank & O'Reilly, 2006

Psychopharmacological Studies

• Arsten et al., 1994; Seemans et al, 2004
• Camps et al., 1987; Moghaddam & Bunney, 1990

Cabergoline and Haloperidol: D2 agonist and D2 antagonist • Double blind, within-subjects design (N=28).

Frank & O'Reilly, 2006

Psychopharmacological Studies
Working memory gating task
Frank et al. 2001, O'Reilly & Frank 2006, Hazy et al. 2007...
DA drug effects on working memory gating

see Moustafa et al. 08 and Frank et al. 07 for similar drug effects in Parkinson's and ADHD.
Analogous drug effects on learning and working memory.
How do multiple BG-FC circuits interact in motivated behavior?
Only some maintained PFC reps relevant for processing during intermediate stages of processing.
Reduces number of S-R mappings needed to be learned by motor circuit
Contextual dependencies for output gating (e.g., arithmetic); see also LSTM (Hochreiter & Schmidhuber 97)
Hierarchical interactions among BG-FC circuits

- BG-FG circuits learn reward probabilities via DA prediction error signals
- Anterior PFC influence posterior circuits via BG output gating
- BG gates frontal „actions“ (motor, working memory, context)
Hierarchical interactions in BG-FC circuits:

PFC influences on BG learning

Collins & Frank 2013, Psych Rev; Frank & Badre 2012
Why does motor control develop so slowly in humans?

Broader speculations:

"Fourth trimester"

Standard story: Infants born early due to large head, small birth canal.
Broader speculations:

Why does motor control develop so slowly in humans?

- Standard story: Infants born early due to large head, small birth canal.
- ‘Fourth trimester’
- But 3 month old infants still pretty incompetent (from babycenter.com):
Broader speculations:

Why does motor control develop so slowly in humans?

- Broader speculations:

  • Standard story: Infants born early due to large head, small birth canal.

  • "Fourth trimester" concept: Infants still fairly incompetent at birth.

  • But 3-month-old infants still pretty incompetent (from babycenter.com): "You no longer need to support his head. When he's on his stomach, he can lift his head and chest. He can open and close his hands."

- Standard story: Infants born early due to large head, small birth canal.
Why does motor control develop so slowly in humans?

Broader speculations:

Why does motor control develop so slowly in humans?

Standard story: Infants born early due to large head, small birth canal.

Fourth trimester:

But 3 month old infants still pretty incompetent (from babycenter.com):

• You no longer need to support his head. When he’s on his stomach, he can lift his head and chest. He can open and close his hands.

• Hypothesis: Human brain is wired to discover generalizable structure... which is initially inefficient.
Task-sets (TS)
Ts as abstract rule objects

Latent task-set space

Abstracting Task-set rules

Wooldridge et al. 2011
Reverberi et al. 2011
Abstracting Task-sets rules
Abstracting Task-sets Rules

\[ A_i \leftarrow S_i \]

\[ T_{S_i} \]

\[ C_7, C_6, C_5, C_4, C_3, C_2, C_1 \]
Abstracting Task-sets Rules
Abstracting Task-sets rules

Latent task-set space: Unknown size
see also Gershman et al 2010

- Chinese restaurant process: Jordan, Blei, Teh 2005
- $\alpha > 0$: clustering parameter

\[
\frac{\alpha}{(\alpha + 1)^2} \frac{d\pi}{d\nu} = \frac{(1 + u |e|)^2 S L = * S L d^* = \nu \neq \eta A \right)}{(1 + u |e|)^2 S L = * S L d^*} = (1 + u |e|) \cdot * S L d
\]

- Prior prob on TS space: Given a new C:

Prior prob on TS space is revised in new contexts and can be across contexts and can be clustered.

C-TS model
Model mimicry: C-TS and Hierarchical BG-PFC network

- Sparseness of context-PFC connectivity matrix is linked to α clustering
- fMRI evidence for Hierarchical PFC-BG mechanisms (Badre et al. 2010, Badre & Frank 2012)
- Both models are approximations of the same process: building TS structure

Collins & Frank 2013 Psych Rev
Predicts positive, negative transfer. The network learns efficiently unsupervised.

Neural Network - Results
Hierarchical learning and clustering: First, in adults

Experiment:
Hierarchical learning and clustering: First, in adults
Clustering affords faster learning within existing rule sets...
But initial (phase 1) clustering is inefficient and slow.
Neural model predictions and EEG
Neural model predictions and EEG
Now, in Infants!

Werchan et al., 2015; in prep
Learning Curves
Rostro-caudal axis and Hierarchical control

- 2nd level of hierarchical control (from Badre, Koechlin)

- Activity predicts learning in 2nd level hierarchy condition, declines in flat condition
How do BG-PFC circuits interact to discover hierarchical structure?

Striatum implicated in hierarchical rule learning.
Hierarchical interactions among BG-FC circuits
Anterior PFC affects posterior circuits via BG output gating

- Reduces # of stimulus-response mappings needed to be learned in motor circuit

Application to Badre task
Application to Bade task
Hierarchical structure exists (stay tuned) •

This is due to reward prediction error signals punishing gating of prePMD when no DA-based RL in gating networks needed for rapid learning in Hier cond •

As in fMRI data, model prePMD activity decreases in Flat relative to Hier •

Model prePMD activity by condition

Hierarchical Task

Modeling BG-PFC Hierarchical Learning
Modeling individual learners: abstract account

• Neural model makes plausible links to biology, but unsuitable for quantitative fits to individual subject behavior
• We developed abstract version of hierarchical rule learning using Bayesian mixture of experts (MoE), motivated by neural model

Modeling individual learners: abstract account
Modeling individual learners: abstract account

• Separate experts learn $P(\text{Rew} | \text{Resp}, \text{Shape})$, $P(\text{Rew} | \text{Resp}, \text{Orient})$, etc.
Modeling individual learners: abstract account

- Separate experts learn $P(\text{Rew} | \text{Resp, Shape})$, $P(\text{Rew} | \text{Resp, Orient})$, etc.
- Hierarchical experts separately learn statistics about each dimension contingent on a candidate higher order feature (cf. prePMD context for output gating).
- Separate experts learn $P(\text{Rew} | \text{Resp, Shape})$, $P(\text{Rew} | \text{Resp, Orient})$, etc.

Modeling individual learners: abstract account
Modeling individual learners: abstract account

• Separate experts learn $P(\text{Rew} | \text{Resp}, \text{Shape})$, $P(\text{Rew} | \text{Resp}, \text{Orient})$, etc.

• Hierarchical experts separately learn statistics about each dimension contingent on a candidate higher order feature (cf. prePMD context for output gating).

• Credit assignment mechanism learns probability that each expert contributes to observed rewards.

• Attentional mechanism selects among experts based on learned probability of success $P(\text{Rew} | e)$, $\forall e$.

• Separate experts learn $P(\text{Rew} | \text{Resp}, \text{Shape})$, $P(\text{Rew} | \text{Resp}, \text{Orient})$, etc.
weight each expert in proportion to learned probability of expert success, $P(\text{Rew} | e)$, $\forall e$. 
Goal: Infer latent hypotheses by observing sequence of responses.

... rewards, maximizing likelihood of choices with few free params for each subject.

Model:
- FLAT and HIER layers
- Connections between layers
- Observed responses and rewards
Params: priors to attend to each dimension, to attend to conjunctions, to consider hierarchy. $3$ softmax slopes (motor, within experts, between flat/hier).
MoE fits to humans
Example attentional weights: MoE fit to human
MoE fit to BG-PFC model

MoE fit to BG-PFC Net

All Nets: Attention to Hierarchy

Attentional Wts
Individual differences in attention to Hierarchy

All subjects: Attention to Hierarchy
Individual differences in attention to Hierarchy

Median split on Attention vs. Attention → more prePMD in Hier condition
Model-based fMRI: reward prediction errors tracked by entire striatum (cf. O'Doherty et al., 2004)
Prediction errors and attention to Hier vs Flat experts.
PE's predict PPMD decline in Flat condition

DECLINE in PREPMD should be predicted by sensitivity of BG to PE's associated with hierarchical rule
PE's predict PPMD decline in Flat condition

Decline in prePMD should be predicted by sensitivity of BG to PE's associated with hierarchical rule

PE's predict PPMD decline in Flat condition
Activations can be maintained in two different ways:

- Recurrent excitation:
- Intrinsic bistability: selectively activated ion channels.

Problem with recurrent: strongest activations at any point determine what is maintained – not necessarily the best stimulus or worst stimulus.

Activations can be maintained in two different ways:
Computational/Biological Details

Our proposal:

• Thalamic disinhibition activates layer 4 FC.

• Convergence of layer 4 and cortico-cortical 2/3 projections on either 2/3 or 5/6 neurons triggers maintenance ion channels.

Other proposals:

• Thalamic-cortical loops themselves drive maintenance (enough thalamic neurons to specify what is maintained?)

• Just recurrent excitation within FC.

(unsustainable, but useful complement to ion channels)