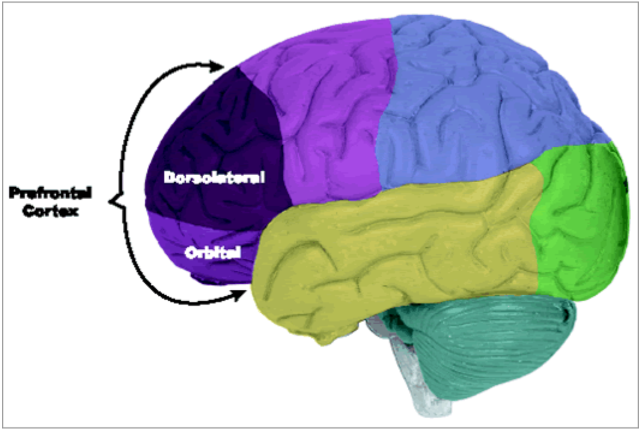
COSC451: Artificial Intelligence

Lecture 4: Planning and learning actions: the prefrontal cortex

Alistair Knott

Dept. of Computer Science, University of Otago

The prefrontal cortex



Other naming conventions divide the PFC into dorso- and ventrolateral. Orbitofrontal cortex is also called rostral prefrontal cortex.

Outline of the talk

1. PFC from a developmental/evolutionary perspective
2. General properties of PFC
3. Earl Miller's model of PFC
4. Action categories in the PFC

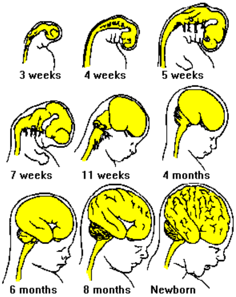
Outline of today's lecture

1. PFC from a developmental/evolutionary perspective
2. General properties of PFC
3. Earl Miller's model of PFC
4. Action categories in the PFC

PFC and brain development

Brain development has several different components:

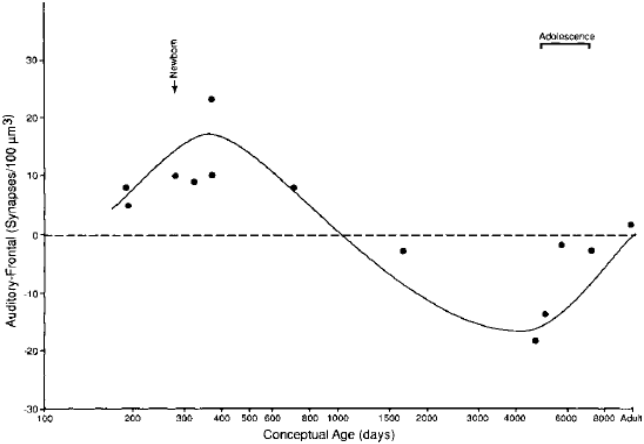
1. Neurogenesis (formation and migration of neurons)
2. Synaptogenesis (formation of synapses)
3. Dendritic growth
4. Myelination
5. Synapse elimination.



In humans, neurogenesis is largely complete by birth.
But the other processes occur for many years after birth.
(And so probably reflect *learning*).

PFC and brain development

In humans, processes 2-5 take longer in PFC than in primary sensorimotor cortices (see e.g. Huttenlocher and Dabholkar, 1997).

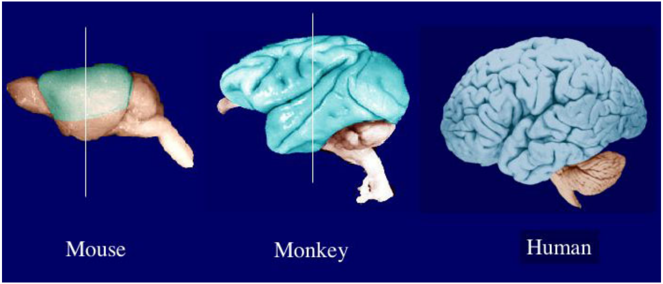


PFC and brain development

Sometimes I would look at students in frustration and think, ‘why don’t you act like an adult?’ The reality is they are not adults. The brain of a university student is very different to that of an adult, which is the source of both their charm and the chaos that they sometimes generate around the place. In particular, between ages 18 and 22 years, a large amount of brain development occurs in the prefrontal cortex. . . (Recent interview with Harlene Hayne, our new DVC)

PFC and evolution

The main difference between human brains and those of other animals is the relative size of the cortex.

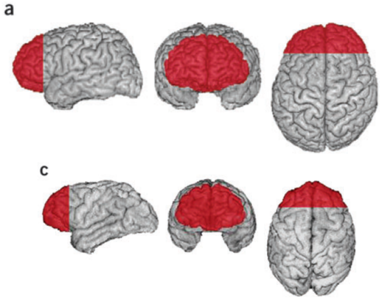


(Cortex is shown in blue.)

But there are also relative differences between regions of the cortex.

PFC and evolution

Proportionally, PFC has no more neurons in humans than in other primates.



human brain

chimp brain

However, it has proprtionally more *synapses* in humans than in other primates (Schoenemann *et al.*, 2005).

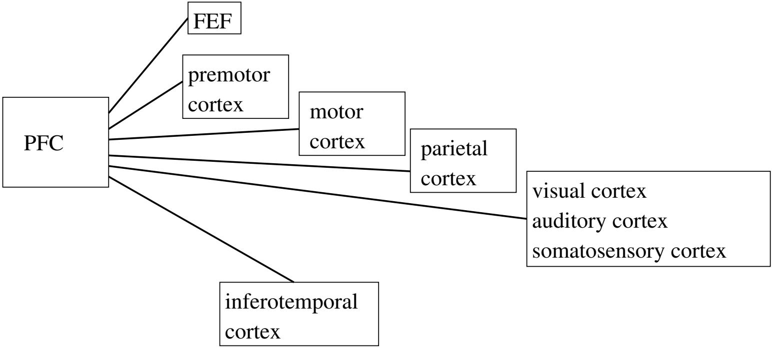
Maybe PFC is responsible for ‘distinctively human’ cognitive abilities?

Outline of today's lecture

1. PFC from a developmental/evolutionary perspective
2. General properties of PFC
3. Earl Miller's model of PFC
4. Action categories in the PFC

Connectivity of the PFC

The PFC is more widely connected than other cortical areas.



It can influence attentional actions (through FEF and IT) and motor actions (through sensory, parietal, premotor and motor cortices).

Behavioural consequences of damage to PFC

The consequences of damage to PFC are not immediately obvious.

- Perception and motor function seem fairly intact.

However, there are deficits in how perceptual stimuli are mapped to motor responses.

- Behaviour is 'stimulus-bound'; 'automatic'; 'inflexible'. . .

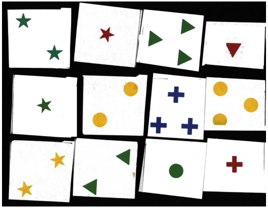
There are also difficulties in remembering sequences of stimuli, or of motor actions.

- These suggest a deficit in working memory.

Behavioural consequences of damage to PFC

A test for PFC damage is the Wisconsin card-sorting task.

Patients are given a stack of cards, and are told to sort these into groups, based on either number, shape or colour.



- PFC-damaged patients can do the first sorting task fine, but have difficulty switching from one sorting principle to another.

Behavioural consequences of damage to PFC

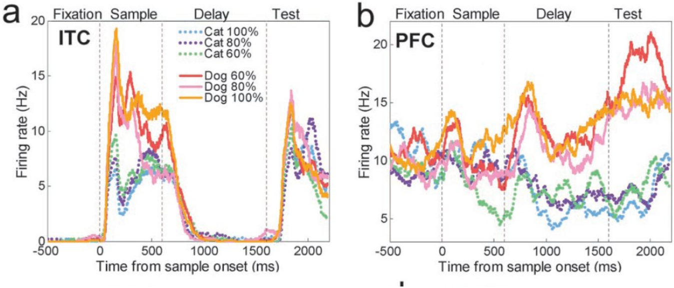
Another test for PFC damage is the Stroop task.

Patients are given a list of colour words, and must name the colour of each word.

- Everyone finds this hard!
- PFC-damaged patients find it particularly hard.

Working memory role of PFC neurons

Friedman *et al.* (2003) gave monkeys a delayed match-to-sample task (matching 'dog' and 'cat' stimuli), and recorded from PFC and IT cells.

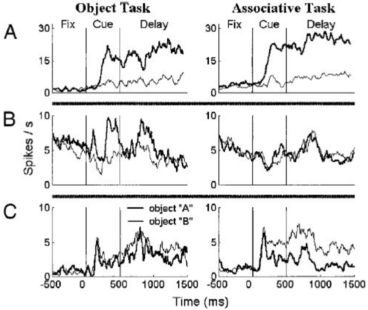


(a) a typical IT cell (b) a typical PFC cell

Task-specificity of PFC neuron activity

Asaad *et al.* (2000) gave monkeys a simple stimulus-response task, in two different conditions:

- ‘Object task’: respond to cue C with response R1
- ‘Associative task’: respond to cue C with response R2



A, B and C are PFC cells with different behaviours.

Models of PFC function

The role of PFC has been described in several different ways.

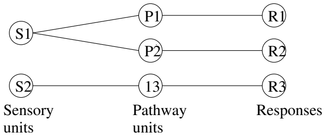
- PFC is where 'working memories' are stored
- PFC is where the agent's current 'task'/'goal'/'intention' is stored
- PFC is where 'planning' happens
- PFC is responsible for 'high-level cognitive control'.

Earl Miller's model of PFC provides a computational mechanism which suggests how it can have this range of functions.

Outline of today's lecture

Miller's model of the PFC

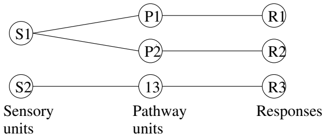
The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.



Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

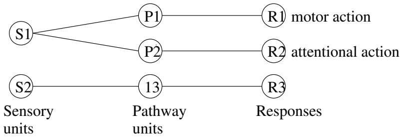
(We're talking *high-level* sensory representations, and *high-level* response representations.)



Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

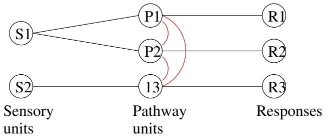


- Some responses are motor actions; others may be attentional actions.
- The pathways involve parietal/premotor cortex in either case.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

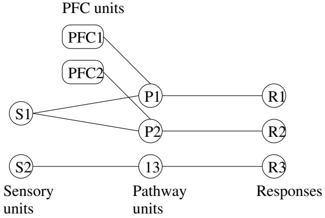


- Intermediate units in these pathways compete with one another.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

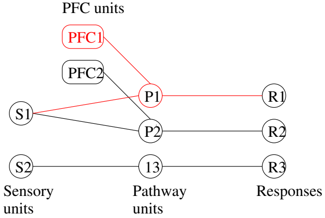


- Intermediate units in these pathways compete with one another.
- PFC units can bias this competition towards one pathway or another.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

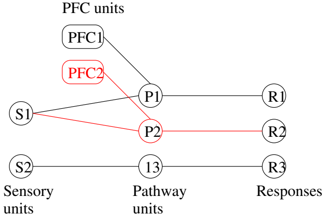


If the PFC1 assembly is active, this biases the agent towards responding to S1 with R1.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

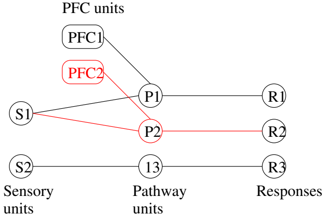


If the PFC2 assembly is active, this biases the agent towards responding to S1 with R2.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)

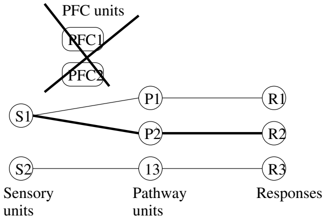


Since pathway units compete with each other, either bias makes the agent tend to *ignore* S2.

Miller's model of the PFC

The starting point for Miller's model is the idea of a neural pathway from stimuli to responses.

(We're talking *high-level* sensory representations, and *high-level* response representations.)



If PFC is damaged, you'll always follow the most common S-R pathways.

- So behavior is 'automatic', 'inflexible', 'stimulus-bound'.

Reinforcement learning

We use reinforcement learning to learn how to behave. Here's a simple model of this type of learning.

In many synapses, Hebbian learning is modulated by the presence of the neurotransmitter dopamine.

- The more dopamine is present, the more learning occurs.
- Dopamine is generated by pleasurable stimuli, and suppressed by unpleasant ones.

We reinforce S-R associations when nice things happen, and inhibit them when nasty things happen.

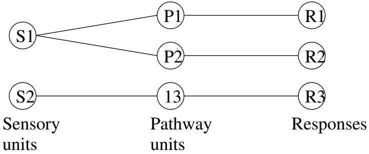
Learning in PFC

Miller's proposal: there are two levels at which reinforcement learning can apply.

- 'Simple learning':
  We learn which are the best S-R pathways, *in general*.
  This type of learning just involves strengthening connections in S-R pathways.
- 'Meta-level learning':
  We learn which S-R pathways are good in which circumstances.
  This is the kind of learning which involves the PFC.

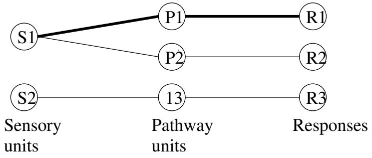
Simple reinforcement learning

1. You see a stimulus $S1$.
2. You choose a response $R1$—initially at random.
3. If you get a reward, you strengthen the pathway from $S1$ to $R1$.



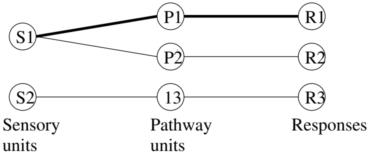
Simple reinforcement learning

1. You see a stimulus $S1$.
2. You choose a response $R1$—initially at random.
3. If you get a reward, you strengthen the pathway from $S1$ to $R1$.



Simple reinforcement learning

1. You see a stimulus $S1$.
2. You choose a response $R1$—initially at random.
3. If you get a reward, you strengthen the pathway from $S1$ to $R1$.



That works okay if $R1$ occurs while $S1$ is still active. But...

- What if there's a delay between $S1$ and $R1$?
- What if the stimulus signals *which S-R mappings are rewarded*?

PFC-based reinforcement learning

How about this reward schedule:

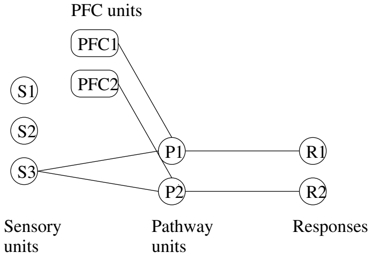
- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1/S2$ and suitable PFC units.

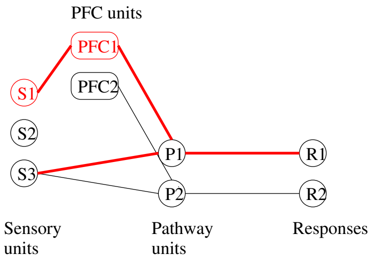


PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1/S2$ and suitable PFC units.

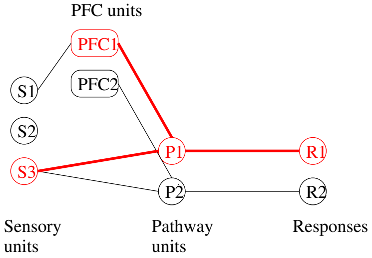


PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1/S2$ and suitable PFC units.

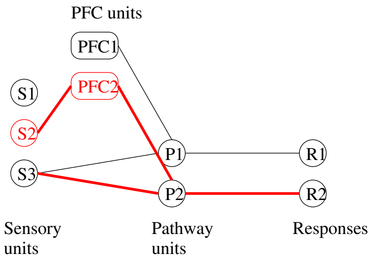


PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1$/$S2$ and suitable PFC units.

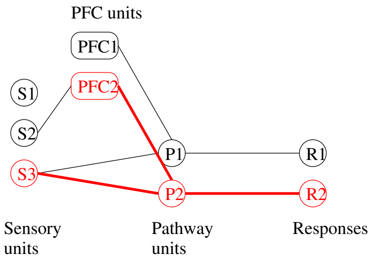


PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1/S2$ and suitable PFC units.

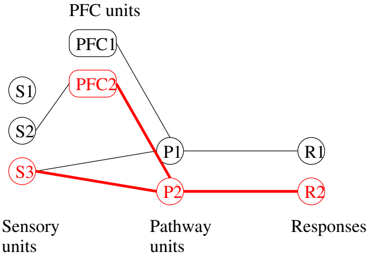


PFC-based reinforcement learning

How about this reward schedule:

- If you see $S1$, then respond to any subsequent $S3$ with $R1$
- If you see $S2$, then respond to any subsequent $S3$ with $R2$.

What we want here is to learn a mapping between $S1/S2$ and suitable PFC units.

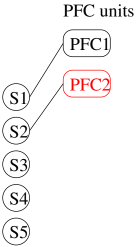


Note: we also want to make sure PFC units remain active after their 'triggering' stimulus disappears.

PFC-based reinforcement learning

The connection from sensory stimuli to the PFC is quite special.

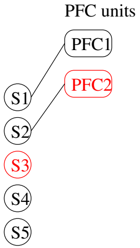
- Most of the time, PFC units don't respond to changes in perceptual stimuli.
  (PFC units hold 'plans', which don't get changed willy-nilly.)



PFC-based reinforcement learning

The connection from sensory stimuli to the PFC is quite special.

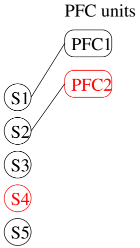
- Most of the time, PFC units don't respond to changes in perceptual stimuli.
  (PFC units hold 'plans', which don't get changed willy-nilly.)



PFC-based reinforcement learning

The connection from sensory stimuli to the PFC is quite special.

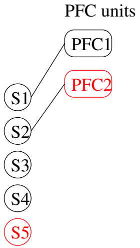
- Most of the time, PFC units don't respond to changes in perceptual stimuli.
  (PFC units hold 'plans', which don't get changed willy-nilly.)



PFC-based reinforcement learning

The connection from sensory stimuli to the PFC is quite special.

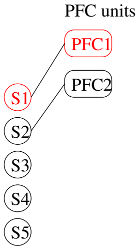
- Most of the time, PFC units don't respond to changes in perceptual stimuli.
  (PFC units hold 'plans', which don't get changed willy-nilly.)



PFC-based reinforcement learning

The connection from sensory stimuli to the PFC is quite special.

- Most of the time, PFC units don't respond to changes in perceptual stimuli.
  (PFC units hold 'plans', which don't get changed willy-nilly.)



However, there are special perceptual stimuli which cause the agent to *change its plan*. These stimuli update PFC.

PFC-based reinforcement learning

How do we learn when to update our plan?

One interesting idea relates to how a dopamine response changes over experience.

- To start with, a dopamine burst happens when the reward occurs.
- Over time, the dopamine burst moves progressively earlier, to the earliest stimulus which *predicts the reward*.

PFC-based reinforcement learning

How do we learn when to update our plan?

One interesting idea relates to how a dopamine response changes over experience.

- To start with, a dopamine burst happens when the reward occurs.
- Over time, the dopamine burst moves progressively earlier, to the earliest stimulus which *predicts the reward*.

Braver and Cohen (2000): perhaps connections from stimuli to PFC are only enabled if there's a dopamine burst.

- So PFC is updated whenever an unexpected reward is predicted. And the new PFC rep can help the agent learn the right S-R pathways to bias in this situation.

PFC-based reinforcement learning

For example: say you see a cake. You're hungry, but you don't know what to do with cakes.

- By random exploration, you sometimes pick up objects you see and put them in your mouth.
- When you first put a cake in your mouth, you get a DA burst.
- After a bit of experience, you get a (small) DA burst when you *see* the cake (because it's a weak predictor of reward).
- At this point, the 'cake' visual stimulus also triggers a (weak) update of your PFC rep.
- After that, when you see a cake and put it in your mouth, you learn an association between a PFC rep and a particular S-R pathway.
- This causes the 'cake' stimulus to become a *better* predictor of reward…

Outline of today's lecture

A repertoire of hand/arm actions

Consider an agent looking at a target object.

We have just been thinking about a single action: reach-to-grasp.
But of course there are many alternatives:

- 'hit', 'snatch', 'squash', 'push', 'put'...

Two questions:

- How are these different actions defined?
- How does the agent decide which action to do?

PFC and action categories

Ultimately, actions are defined *functionally*, by their effects.

- The effect is what gets a reward. (In some situations, slapping is rewarded; in others, squashing is rewarded, etc.)

However, the motor system must have a lower-level representation of individual actions.

At a lower level, hand actions vary along two dimensions:

- The hand follows different trajectories onto the target.
- The hand undergoes different preshape sequences.

(Perhaps a hand trajectory is specified as a sequence of biases on the movement vector, applied during execution.)

PFC and action categories

If PFC is where sequences of motor movements are ‘planned’, then PFC is one place where action categories could be defined.

- But damage to PFC doesn’t tend to impair our low-level repertoire of actions.

There are other parietal or premotor areas which are perhaps more plausible.

- e.g. the supplementary motor area (SMA)
- e.g. the pre-supplementary motor area (pre-SMA)
- e.g. the anterior intraparietal cortex (AIP).

However, we can assume that PFC *schedules* the actions which are defined in these areas.

Summary

PFC provides top-down input to help the agent decide what action to do next.

- The action can be an attentional action or a motor action.
- Input comes in the form of biases on S-R pathways.

PFC learns what to do next using reinforcement.

- PFC learns plans, and when to adopt them.

Action categories (e.g. 'squash', 'slap') are probably represented as planned motor sequences.

- They're probably stored in a premotor area rather than in PFC.
- PFC probably holds higher-level plans.