Classical Conditioning V:
TD learning in the brain & inhibitory conditioning

Plan for today

• Recap of TD learning & more neural implementation
• Inhibitory conditioning
• Opponent appetitive and aversive processes
Temporal-difference learning: an algorithm

1. start with some values for each of the states $V(S_i)$
   (these can be all 0, or random, or anything you want)

2. every time you encounter some state (eg: $S_3$) wait to see what happens:
   what reinforcement do I get, and what is the next state I end up with

3. then calculate the prediction error (if any), eg:
   \[ \delta = R_{next} + V(S_{next}) - V(S_3) \]

4. finally, update based on the prediction error:
   \[ V(S_3) \leftarrow V(S_3) + \eta \delta \]

Schultz, Schiller,及 others, 1997

Dopamine as a prediction error

\[ \delta_t = R \]
\[ \delta_t = V(S_t) \quad \delta_t = R - V(S_{t+1}) \]
\[ \delta_t = V(S_t) \quad \delta_t = 0 - V(S_{t+1}) \]
dopamine and synaptic plasticity

- prediction errors are for learning...
- cortico-striatal synapses show dopamine-dependent plasticity
- three-factor learning rule: need presynaptic + postsynaptic + dopamine to strengthen synapse

![Graph showing dopamine activity](image)

functional magnetic resonance imaging (fMRI)

- measure BOLD ("blood oxygenation level dependent") signal
- oxygenated vs de-oxygenated hemoglobin have different magnetic properties
- detected by big superconducting magnet

Idea:
- The brain is functionally modular
- Neural activity uses energy & oxygen
- Measure brain usage, not structure

- Spatial resolution: ~3mm 3D “voxels”
- temporal resolution: 5-10 seconds
imaging prediction errors in humans

4 stimuli:
- 40c
- 0/40c
- 20c
- 0c

You won
40 cents

0.5 sec
ITI

5 sec
ISI

0.5 sec

what would a prediction error look like?

searching for prediction error signals in humans

What would a prediction error look like?
searching for prediction error signals in humans

imaging prediction errors in humans

Wittmann et al., 2008
Gerstman et al., 2009
Glascher et al., 2010
(in my lab)

anatomical ROI

BOLD signal

model prediction
why is this useful?

if I can measure your dopamine when you see a stimulus, I can find out how much reinforcer you are expecting!

All models are wrong, some models are useful

summary so far

Thinking computationally about prediction learning

• The problem: prediction of future reward
• An algorithm: temporal difference learning
• Neural implementation: dopamine dependent learning in BG
  ⇒ Solves our puzzles: explains dopaminergic firing patterns, 2nd order conditioning
  ⇒ Compelling account for the role of dopamine in classical conditioning: prediction errors drive prediction learning