Models of Conditioning

Peter Dayan
Gatsby Computational Neuroscience Unit

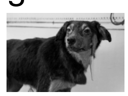
Nathaniel Daw   Sham Kakade   Read Montague   Yael Niv   Angela Yu

Conditioning

- Pavlovian (classical) conditioning
  - learning predictions of value
    - me: basic paradigms, simple Bayesian model
    - John: alternative `locally Bayesian' framework
    - Nathaniel: statistical models of representation, competition and cooperation
    - Bernard: neural substrates for predictions
    - me: models of value
    - Nathaniel: fMRI of value
- Instrumental (operant) conditioning
  - learning actions (RL)
    - Bernard: different neural systems for action control
    - me: models of habitual control
    - Nathaniel: models of competition

Pavlovian Conditioning

L → r



- predictive relationship
- stimulus (L: light),
- biologically significant outcome (r: reward or punishment)
- present L; monitor behaviour
  - preparatory (approach, withdrawal)
  - consummatory (eating)
- evolutionary `prior' – bias/variance

Pavlovian Conditioning

	Name	Set 1	Set 2	Test
1	Pavlovian		L → r	L ↛ r

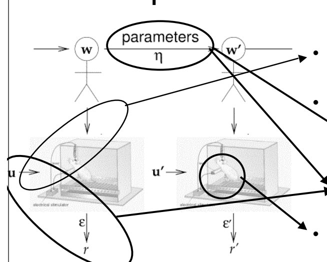
- competing/cooperating predictors
  - competition for learning/prediction
- initial, and sometimes continual, uncertainty
- conditional vs joint distribution (Daw)

Marrian Levels of Analysis

- computational
  - prior; likelihood; crank
- psychological
  - online learning rules
- neural
  - associabilities (amygdala); errors (VTA); representations (cortex)

$$\Delta w_T = \phi_T (r - V) u_T$$

Computational Conditioning

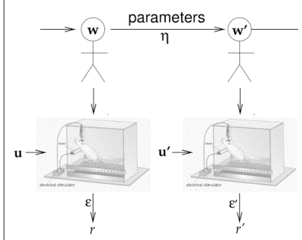


- output model
  - downwards unblocking
- update model
  - expected/unexpected uncertainty
  - uncertainty-dependent learning rates
  - backwards blocking
- decision model
  - action system competition
  - reinforcement learning

Psychological Intuitions

- contiguity (causality)
  - fine for L→ r
  - but blocking: L→ r ; L+T → r ; T → ·
- error correction Rescorla-Wagner $(r\text{-}V)u_T$
  - but ←blocking: L+T → r ; L→ r ; T → ·
- associability – learning competition
  - Mackintosh: GPGP
  - Pearce & Hall: AOAA
  - but L→ $2r$; L+T → r ; T → – r
- competitive prediction
  - Kruschke; Long & Dayan; Daw

Kalman Filter



expt $\mathbf{w}' = \mathbf{w} \quad + \eta$

reward given $r = \mathbf{w} \cdot \mathbf{u} + \epsilon$

allowable drift $\eta \sim N[0, \sigma^2 \mathbb{I}]$

output noise $\epsilon \sim N[0, \rho^2]$

- Markov random walk (or OU process)
- no punctate changes
- additive model of combination
- forward inference

Kalman Posterior

The Kalman filter maintains uncertainty:

$$P(V) = \mathcal{N}[\hat{\mathbf{w}} \cdot \mathbf{u}, \mathbf{u} \cdot \Sigma \cdot \mathbf{u}]$$

where

Assumed Density KF

Diagonal approx to $\Sigma = \text{diag}(\sigma_i^2)$

If $\mathbf{w} \sim \mathcal{N}\left[\hat{\mathbf{w}}, \text{diag}(\sigma_i^2)\right]$, then

$$\Delta \hat{w}_i = \frac{\sigma_i^2}{\sum_j \sigma_j^2 + \rho^2}(r - \mathbf{u} \cdot \hat{\mathbf{w}})u_i$$

- Rescorla-Wagner error correction
- competitive allocation of learning
  - P&H, M

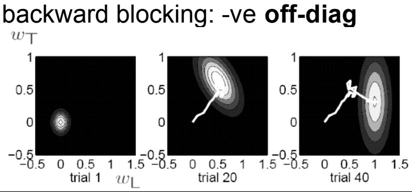
Blocking

forward	L→r	L+T→r	T→ ·
backward	L+T→r	L→r	$T \to \cdot$

- forward blocking: error correction

$$\cdot(r - \mathbf{u} \cdot \hat{\mathbf{w}})$$

- backward blocking: -ve **off-diag** $\Sigma_{LT} < 0$



Mackintosh vs P&H

- under diagonal approximation:

$$E(r - \mathbf{u} \cdot \hat{\mathbf{w}})^2 = \rho^2 + \textstyle\sum_j \sigma_j^2 u_i^2$$

- for slow learning,

σ_j^2 changes with correlation of $(r - V)$ and u_i

  - effect like Mackintosh

Summary

- Kalman filter models many standard conditioning paradigms
- elements of RW, Mackintosh, P&H
- but:
  - downwards unblocking
    $L \to r\Delta r \quad L+T \to r \quad T \not\to \pm r$
    predictor competition
  - negative patterning L→r; T→r; L+T→·

stimulus/correlation rerepresentation (Daw)

  - recency vs primacy (Kruschke)

Uncertainty (Yu)

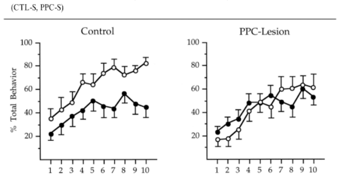
- expected uncertainty - ignorance
  $$\mathbf{w} \sim \mathcal{N}\left[\widehat{\mathbf{w}}, \mathrm{diag}(\sigma_i^2)\right]$$
  - amygdala, cholinergic basal forebrain for conditioning
  - basal forebrain for top-down attentional allocation?
- unexpected uncertainty – `set' change
  - noradrenergic locus coeruleus
- unexpected uncertainty – 'state' change
  - noradrenergic locus coeruleus
- part opponent; part synergistic interaction

ACh in Conditioning

- Given ***uncertainty***, **ACh** b***oosts*** learning to stimuli of uncertain consequences

Table 1. Outline of procedures for Experiment 1

Treatment condition (groups)	Phase 1: consistent L–T relation	Phase 2: experimental change in L–T relation	Phase 3: test of conditioning to L
Consistent (CTL-C, PPC-C)	L → T → food; L → T	L → T → food; L → T	L → food
Shift (CTL-S, PPC-S)	L → T → food; L → T	L → T → food; L	L → food

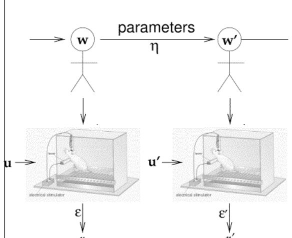


(Bucci, Holland, & Galllagher, 1998)

Pavlovian Conditioning

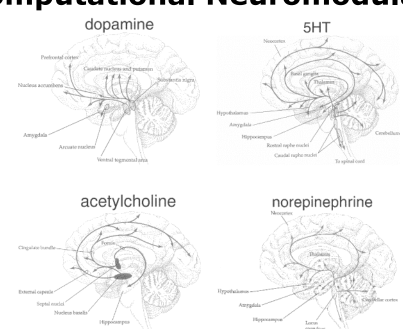
- evolutionary prior
  - amoebae don't learn to absorb food!
- Pavlovian/instrumental competition
  - +ve : negative automaintenance; chicks in a looking glass world
  - -ve : depressive realism

Conditioning



- output model
  - striatum/PFC?
- update model
  - unexpected uncertainty for change: (NE)
  - uncertainty-dependent learning rates: (ACh)
  - backwards blocking: (?)
- decision model
  - action system competition: (ACh?)
  - reinforcement learning (DA, 5HT)

Computational Neuromodulation



- general: excitability, signal/noise ratios
- specific: prediction errors, uncertainty signals