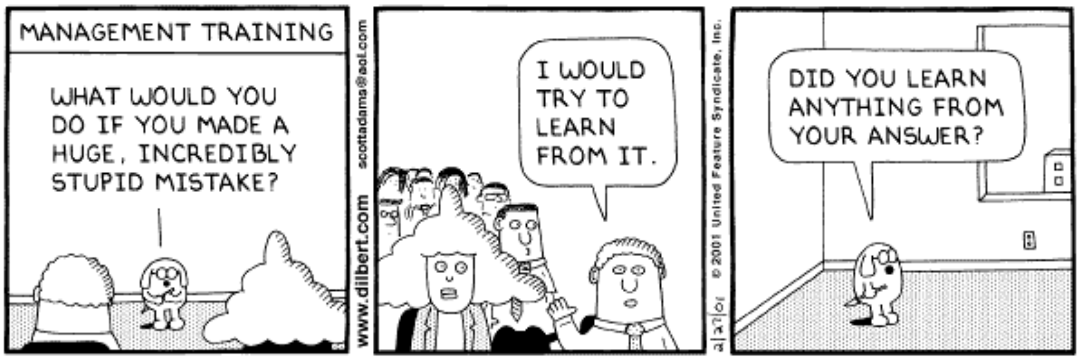
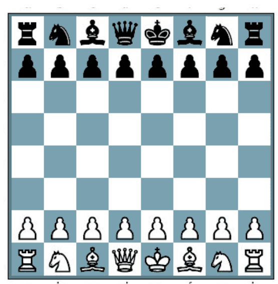
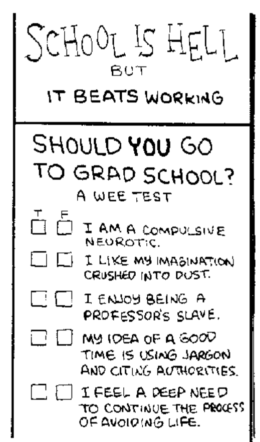
Classical Conditioning II: Learning from prediction errors



PSY/NEU338: Animal learning and decision making:
Psychological, computational and neural perspectives

where were we?



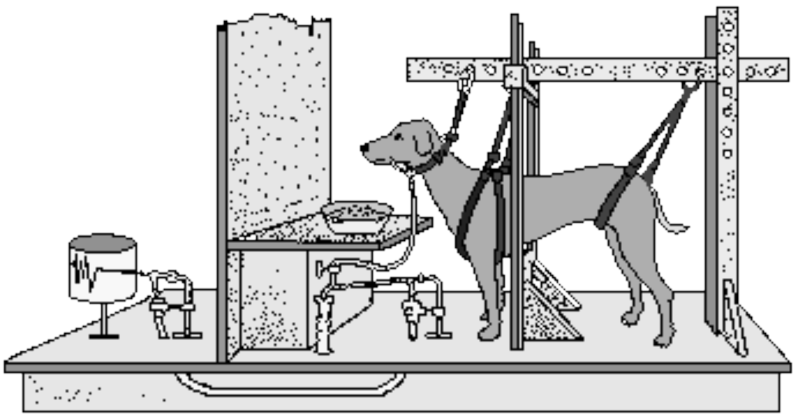


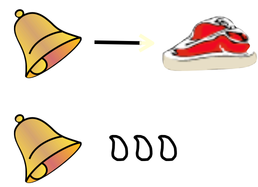
decision making is hard

- Reward/punishment may be delayed
- Outcomes may depend on a series of actions

⇒ “credit assignment problem” (Sutton, 1978)

predicting the future can help...
...and animals learn predictions!





examples of classical conditioning from daily life

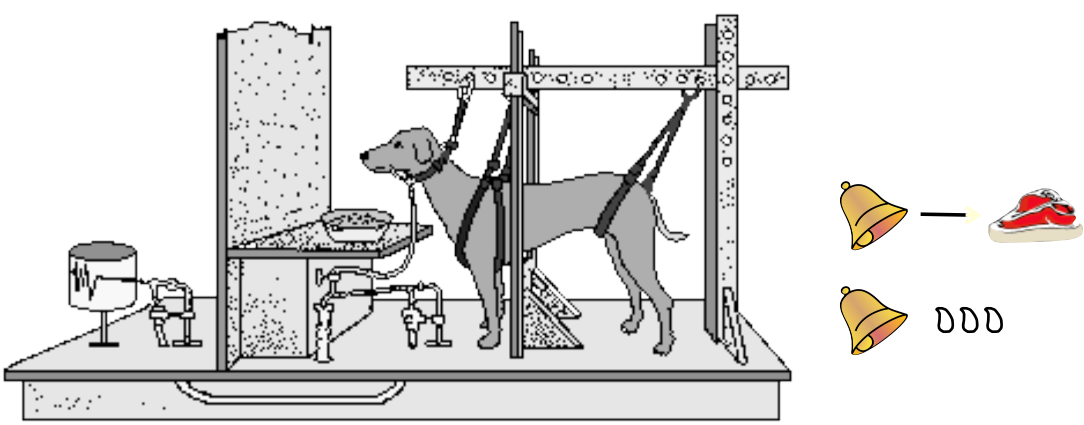
outline

PART I - Basics of classical conditioning

PART II - Some challenging results

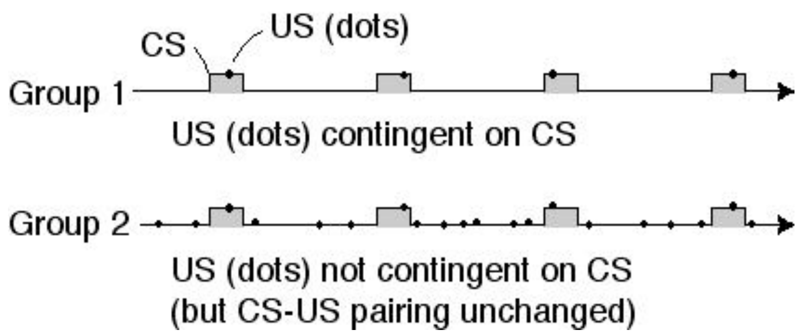
PART III - A theory (model)

predicting the future can help...
...and animals learn predictions!



"pairing a tone (CS) and food (US) is sufficient to induce classical conditioning (prediction learning)"

But... 1) Rescorla's control condition



will Group 2 show a CR to the tone?

temporal contiguity is not enough - need contingency

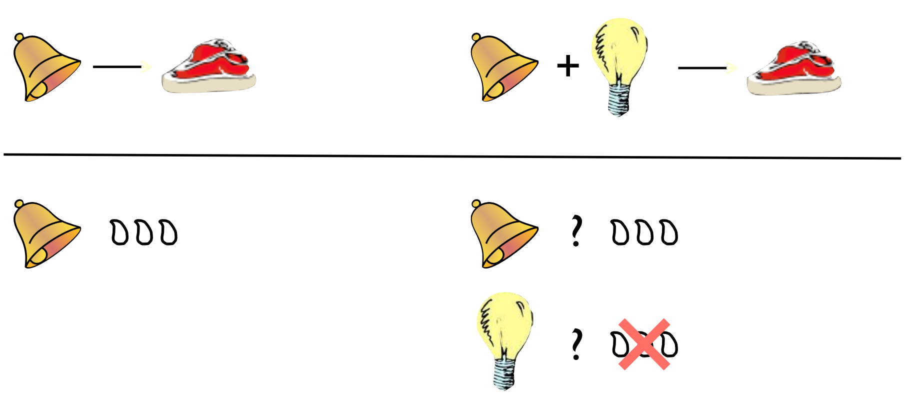
P(food | tone) > P(food | no tone)

Credits: Randy Gallistel

But... 2) Kamin's blocking

Phase I

Phase II



contingency is also not enough.. need <u>surprise</u>

P(food | noise+light) ≠ P(food | noise alone)

And... 3) Reynold's overshadowing

stimuli compete for learning

Summary so far...

- Naïvely it seemed that pairing a CS and a US is enough for conditioning (prediction learning)...
- But now we see that we also need contingency and surprise
- And that different stimuli compete for learning such that even with contingency and surprise learning is not guaranteed
- What is going on here?

outline

PART I - Basics of classical conditioning

PART II - Some challenging results

PART III - A theory (model)

desiderata/goals of a learning theory

- Explain why the CS comes to elicit a response
- When (under what conditions) does this happen?
- Basic phenomena: gradual learning and extinction curves
- More elaborate behavioral phenomena (blocking, overshadowing)
- (explain neural data)

Rescorla & Wagner (1972)

<u>The idea</u>: error-driven learning

Change in value is proportional to the difference between actual and predicted outcome

$$\Delta V(CS_i) = \eta[R_{US} - \sum_{j \in \text{trial}} V(CS_j)]$$

Two assumptions/hypotheses:

(1) learning is driven by error (formalize notion of surprise)

(2) summations of predictors is linear

Rescorla & Wagner (1972)

Lets assume a simpler scenario: one CS and one US only

$$\Delta V = \eta[R - V]$$

$$\Delta V(CS_i) = \eta[R_{US} - \sum_{j \in \text{trial}} V(CS_j)]$$

Rescorla & Wagner (1972)

Lets assume a simpler scenario: one CS and one US only

$$\Delta V = \eta[R - V]$$

or: $$V_{T+1} = V_T + \eta[R_T - V_T]$$ (note: T counts trials)

imagine a scenario with 50% reinforcement: 11010001101

1) what would V be on average after learning?

a. V = 1
b. V = 0.5
c. V = 0
d. it is impossible to know

Rescorla & Wagner (1972)

Lets assume a simpler scenario: one CS and one US only

$$\Delta V = \eta[R - V]$$

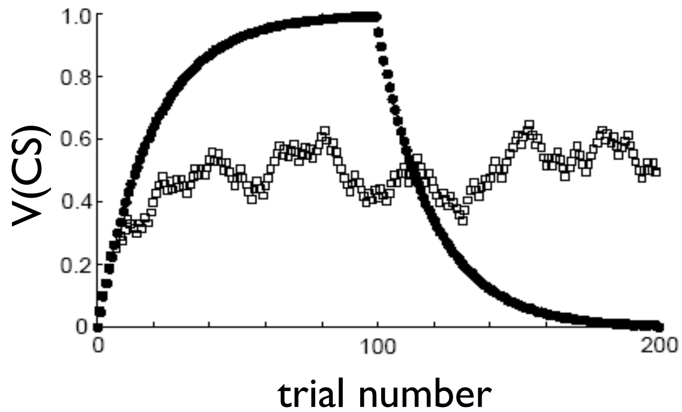
or: $$V_{T+1} = V_T + \eta[R_T - V_T]$$ (note: T counts trials)

imagine a scenario with 50% reinforcement: 11010001101

2) what would the prediction error be after learning?

a. PE = 1
b. PE = 0.5
c. PE = 0
d. it depends on the trial

Rescorla & Wagner (1972)



Can you estimate what learning rate (or step size) η was used in this simulation?

(try to think how you could do the same from behavioral data)

Rescorla & Wagner (1972)

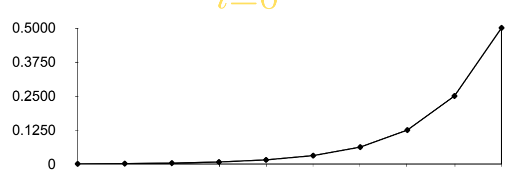
$$V_{T+1} = V_T + \eta[R_T - V_T]$$

how is the value (=prediction) in trial (*T*+*1*) influenced by reinforcements on trials (*T*), (*T*-*1*),(*T*-2) ...?

$$V_{T+1} = V_T(1-\eta) + \eta R_T$$

$$V_{T+1} = \eta \sum_{i=0}^{T} (1-\eta)^{T-i} R_i$$

the R-W rule predicts reinforcement using a weighted average of past reinforcements



Recent reinforcements are weighted more heavily.
Why is this sensible?
Learning rate = forgetting rate!

what does the theory explain?

	R-W	
acquisition	√	
extinction	√X	
blocking	√	
overshadowing	√	
temporal relationships	X	
overexpectation	√	

Summary so far...

- Animals learn predictions (classical conditioning can be seen as a "pure" instance of prediction learning)
- Prediction learning can be explained by an error-correcting learning rule: predictions are learned from experiencing the world and comparing predictions to reality (ie, learning from prediction errors)
- R-W: A simple model - but very powerful!

optional reading

Rescorla & Wagner (1972) - *A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement* - the original chapter that is so well cited (and well written!)

self-practice questions

1. What is an error-correcting learning rule?

____

____

____

2. How does the Rescorla-Wagner learning rule explain the phenomenon of blocking?

____

____

____