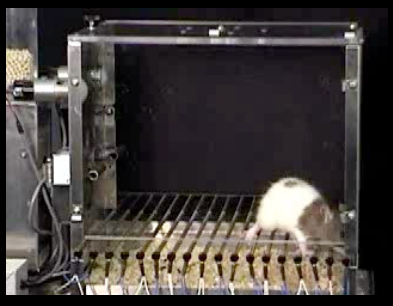
Instrumental Conditioning IV: modeling free operant behavior

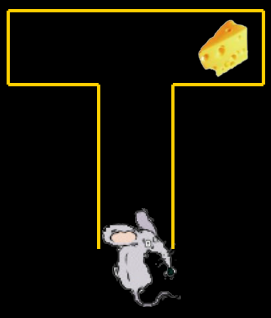
PSY/NEU338: Animal learning and decision making:
Psychological, computational and neural perspectives

what we know, *and what we don't know...*

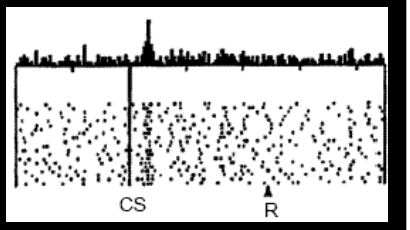
behavior



motivation



dopamine



- Parkinson's disease
- pharmacological manipulations

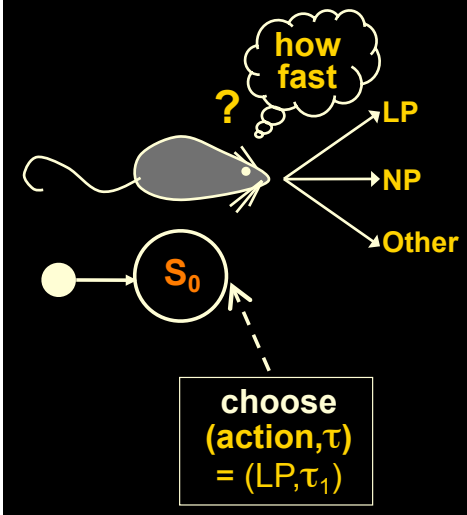
what are the models lacking?

decision making

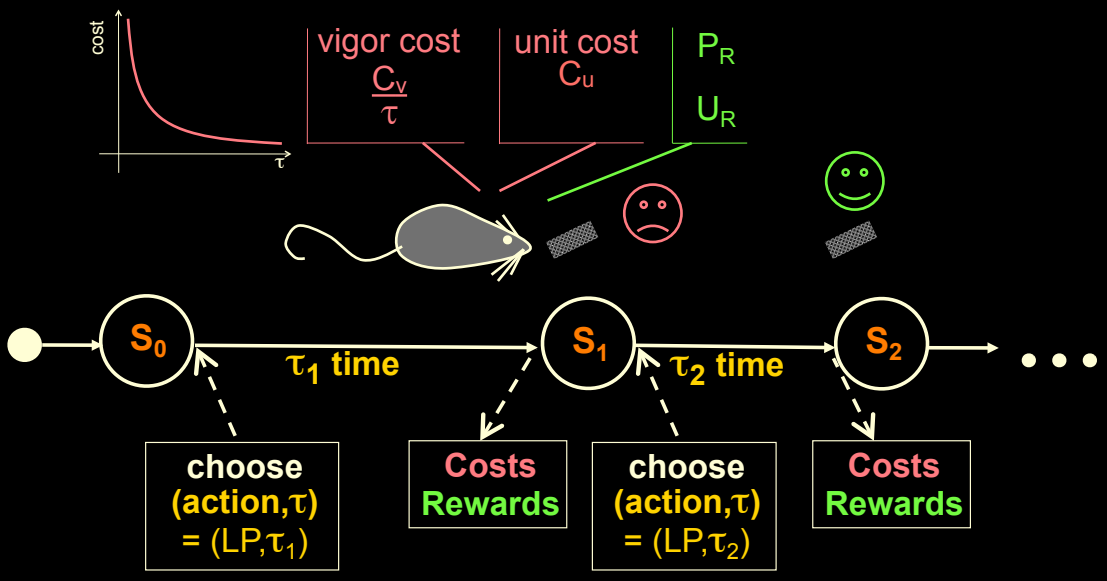
what do to?
(action selection)

how vigorously?
(rate selection)

a model of vigor selection



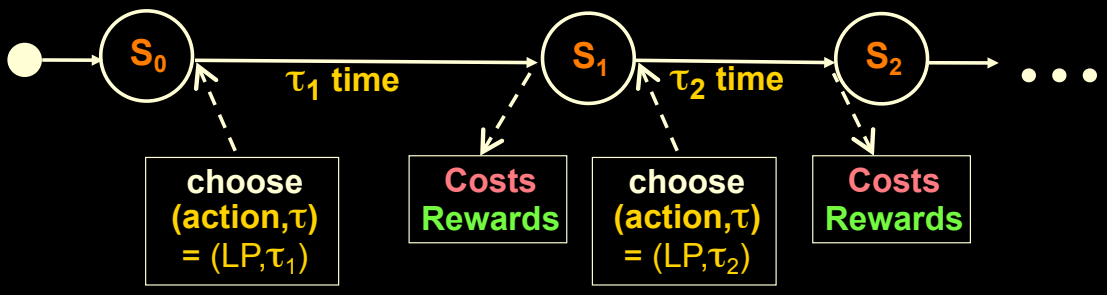
a model of vigor selection



a model of vigor selection

Goal: Choose actions and latencies to maximize the *average rate of return* (rewards minus costs per time)

$$Q(a,\tau|S_t) = (\text{Rewards} - \text{Costs}) + V(S_{t+1}) - \tau\bar{R}$$



intuition: cost/benefit tradeoffs

Choice of action:

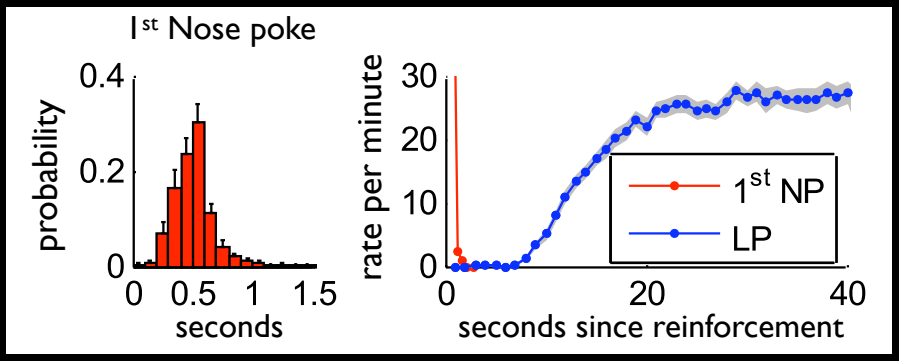
- want to maximize rewards
- and minimize costs

Choice of latency:

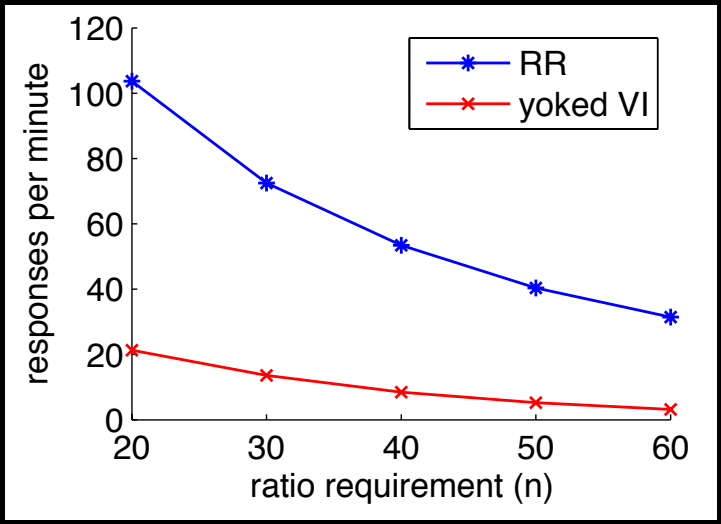
- slow → less costly (vigor cost)
- slow → delays (all) rewards (wastes time)
- what is the cost of time?

$$Q(a,\tau|S_t) = (\text{Rewards} - \text{Costs}) + V(S_{t+1}) - \tau\bar{R}$$

results: optimal policy

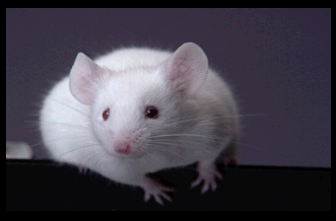


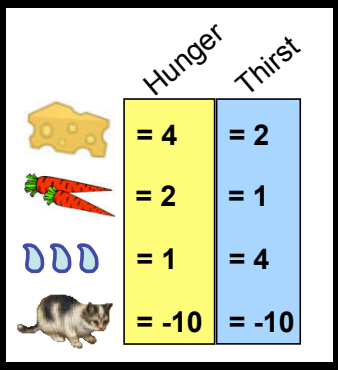
results: optimal policy



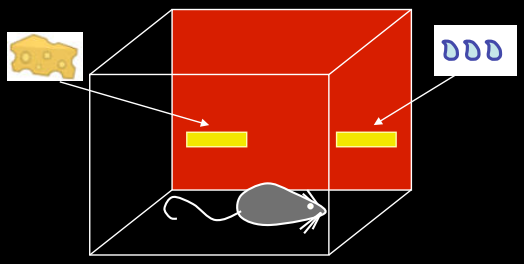
Motivation = ?

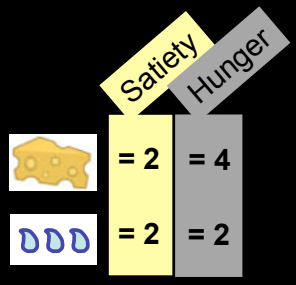
a mapping from outcomes to subjective utilities



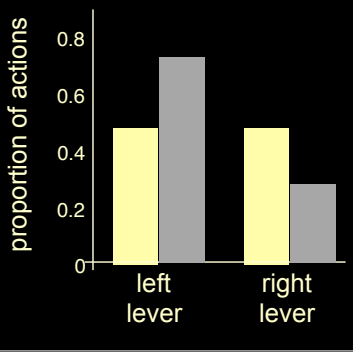


two effects of motivation

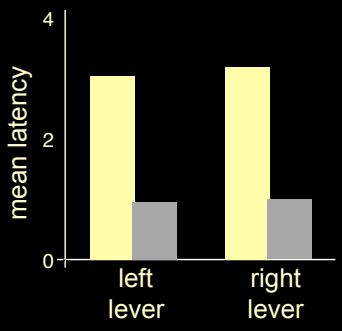




1. directing, through U_r

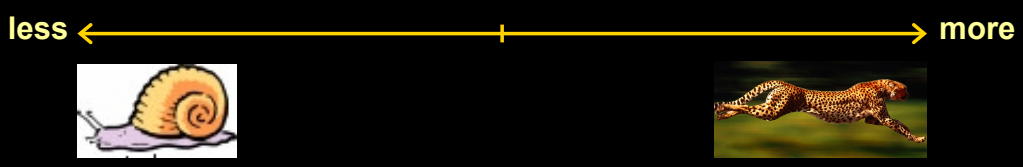


2. energizing, through $\bar{R}$



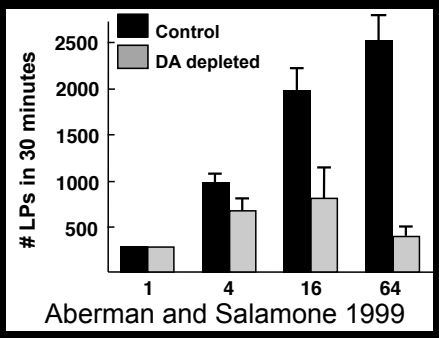
finally: what about dopamine?

- phasic dopamine: reward prediction error
- but also another mode of dopamine signaling: *tonic dopamine*

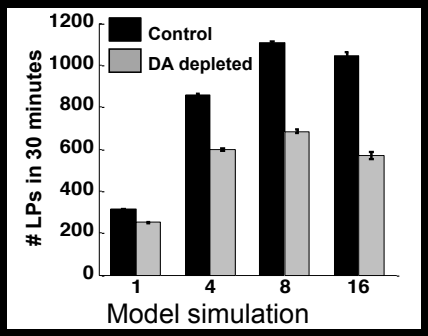


tonic dopamine hypothesis

tonic level of dopamine = average reward rate

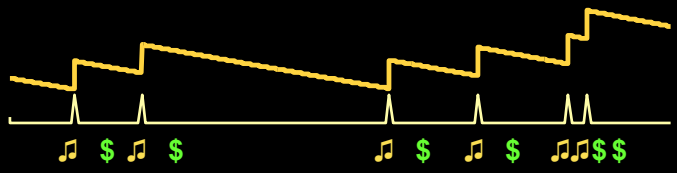


Aberman and Salamone 1999

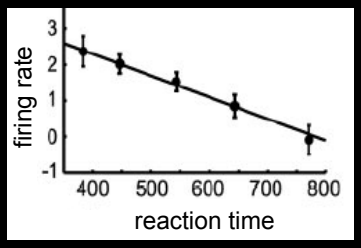


Model simulation

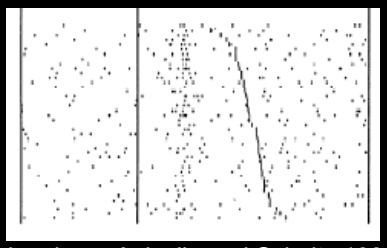
tonic dopamine hypothesis



NB. also explains effect of (phasic) dopamine on reaction times:



Satoh and Kimura 2003


Ljungberg, Apicella and Schultz 1992

summary so far...

- Pavlovian conditioning: all about prediction learning, modeled using error-correcting rules, related to dopamine
- Instrumental conditioning: using predictions to improve actions, modeled using reinforcement learning in MDPs, related to dopamine-dependent learning in the basal ganglia
- Powerful modeling framework (can also model free operant etc.)
- connections to the brain are still under investigation

practice exam questions

- animals that are given amphetamine (a dopamine agonist) seem to behave much faster than they normally do (hence the name 'speed'). Explain why this might be normative (that is, why more dopamine activation *should* make you behave faster)
- on which of the two free operant schedules, a random ratio and a random interval schedule, will animals typically show faster responding and why?
- bonus: suggest an experiment that can test whether motivation has a general energizing effect on behavior.