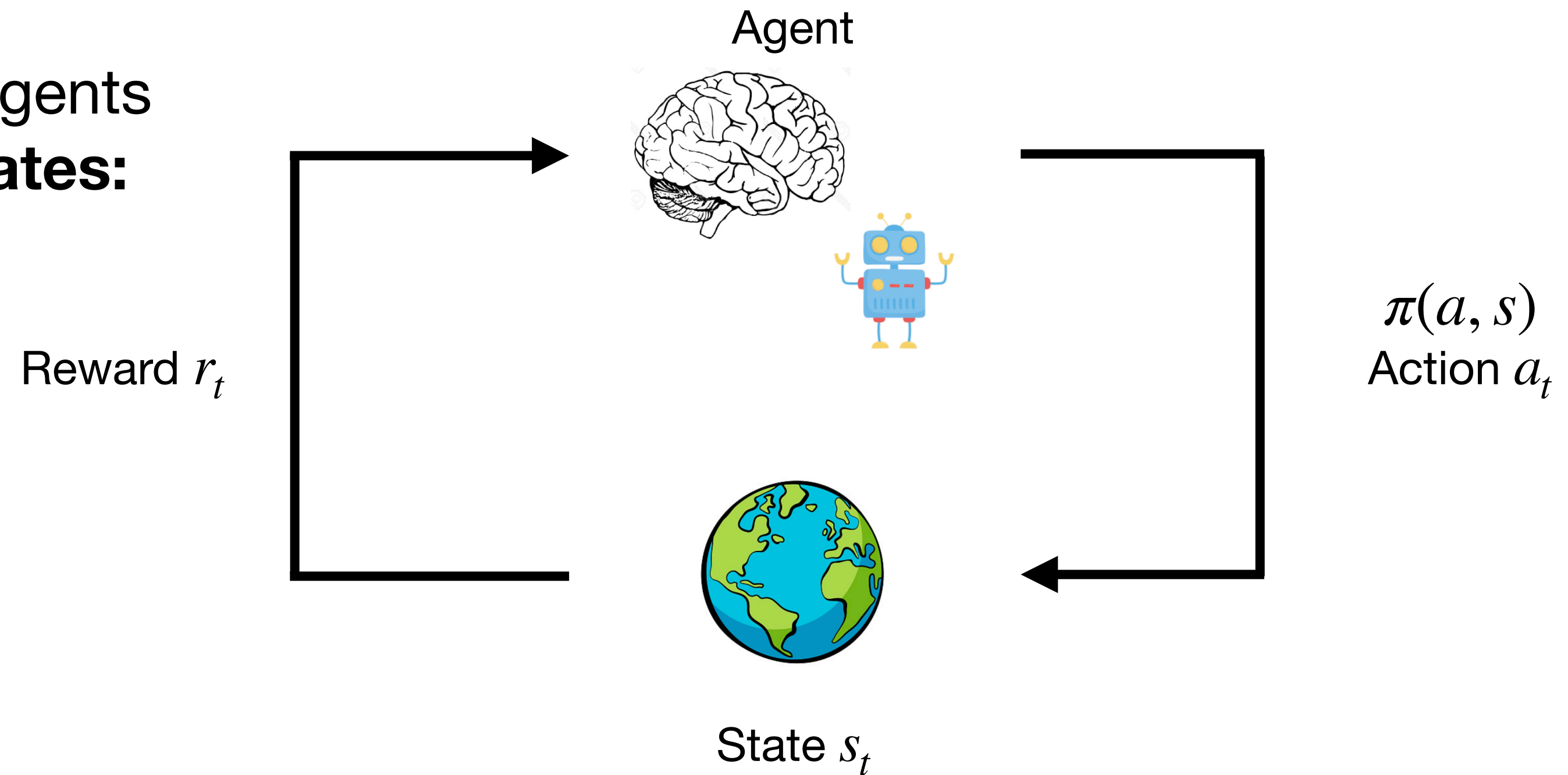# An introduction to Reinforcement Learning

31st of May 2022

Recap: Temporal Difference Learning

Based on a reward signal, agents learn **values of actions/states:**

$$V_\pi(s) = \mathbb{E}_\pi[R \,|\, s_0 = s]$$

Agent

Reward r_t

$\pi(a, s)$
Action a_t

State s_t

TD Learning:

Prediction error

$$V(s_t) \leftarrow V(s_t) + \alpha \cdot (r + \gamma \cdot V(s_{t+1}) - V(s_t))$$

Learning rate   Discount rate

Rescorla Wagner Learning:

Prediction error

$$V(s_t) \leftarrow V(s_t) + \alpha \cdot (r - V(s_t))$$
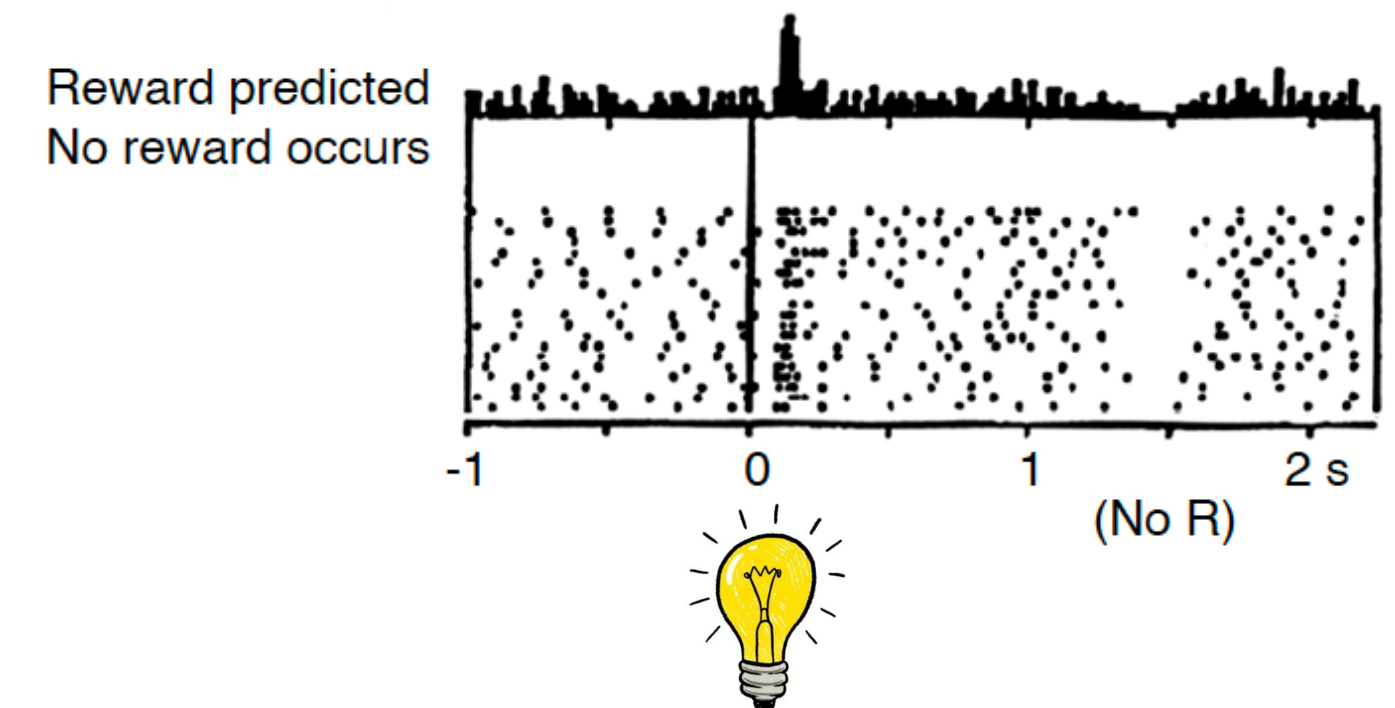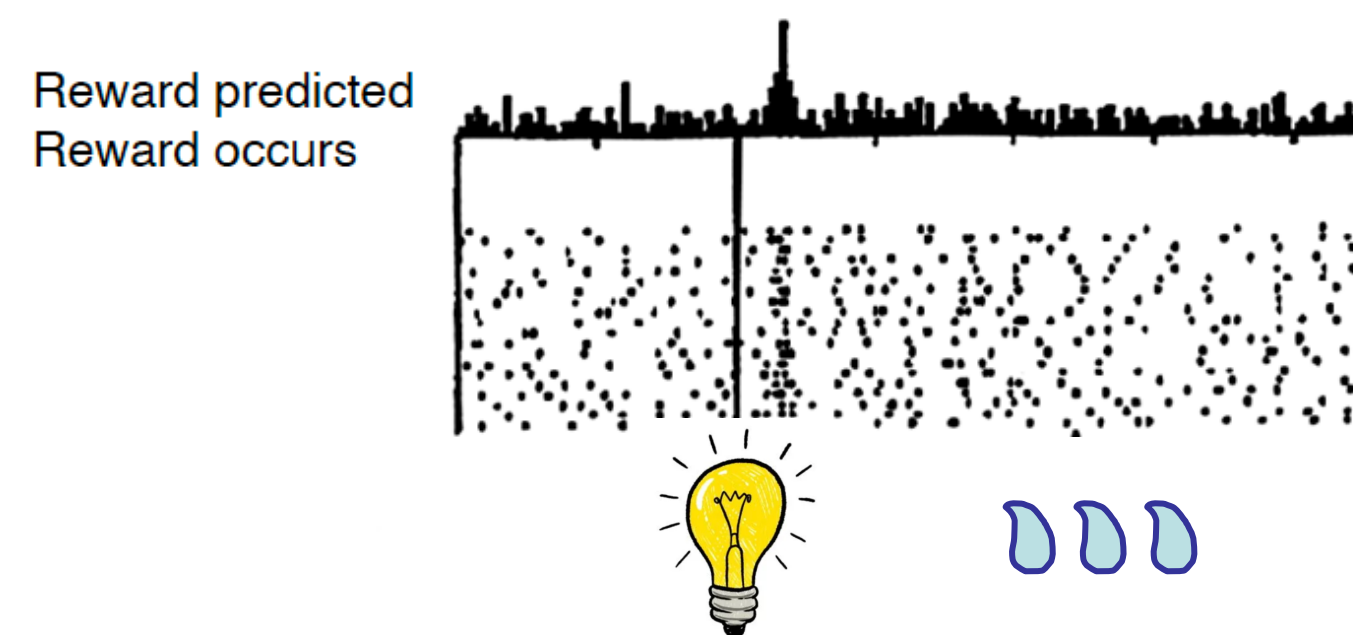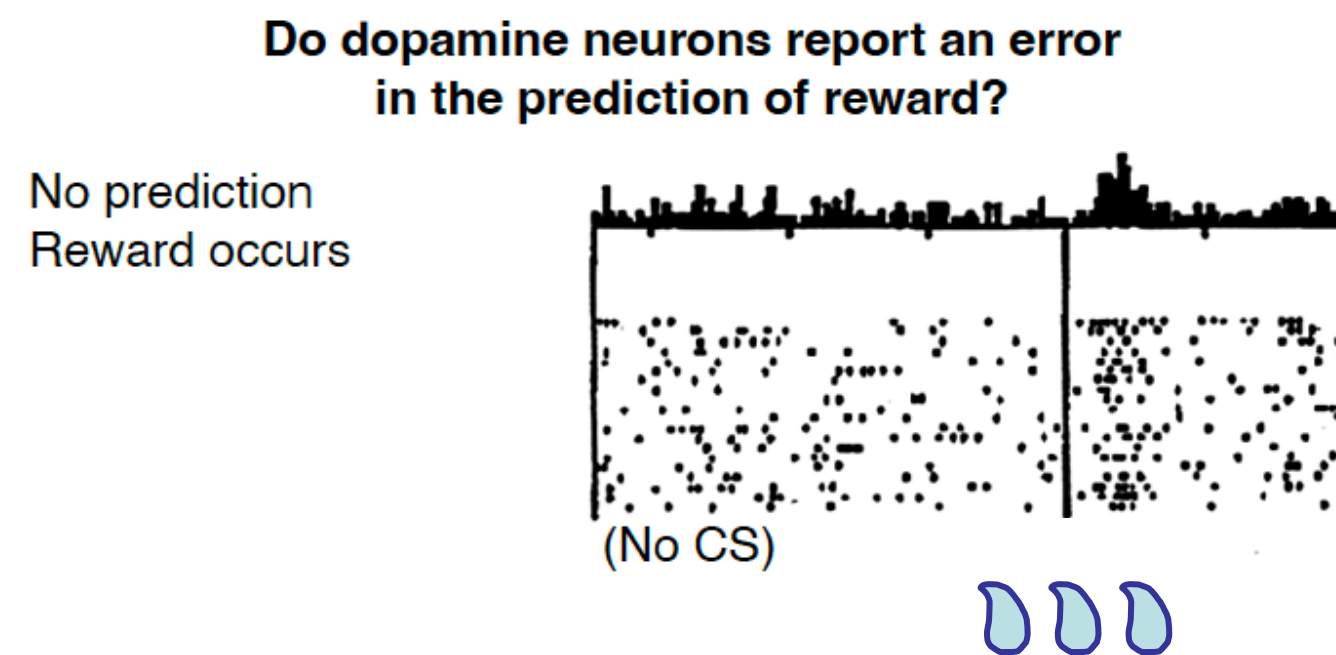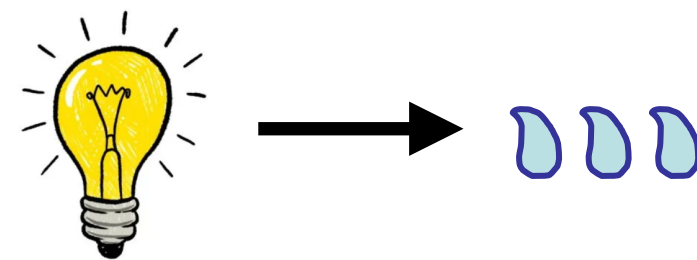
Learning rate

Recap: Can RL tell us anything about the brain?

$$V(s_t) \leftarrow V(s_t) + \alpha \cdot (r + \gamma \cdot V(s_{t+1}) - V(s_t))$$

- It looks like DA signals the reward prediction error in TD learning (Schultz, Dayan & Montague Science, 1997)

Do dopamine neurons report an error in the prediction of reward?

No prediction
Reward occurs

(No CS)

Reward predicted
Reward occurs

Reward predicted
No reward occurs

(No R)

Coding: TD Learning

https://github.com/schwartenbeckph/RL-Course/tree/main/2022_05_24