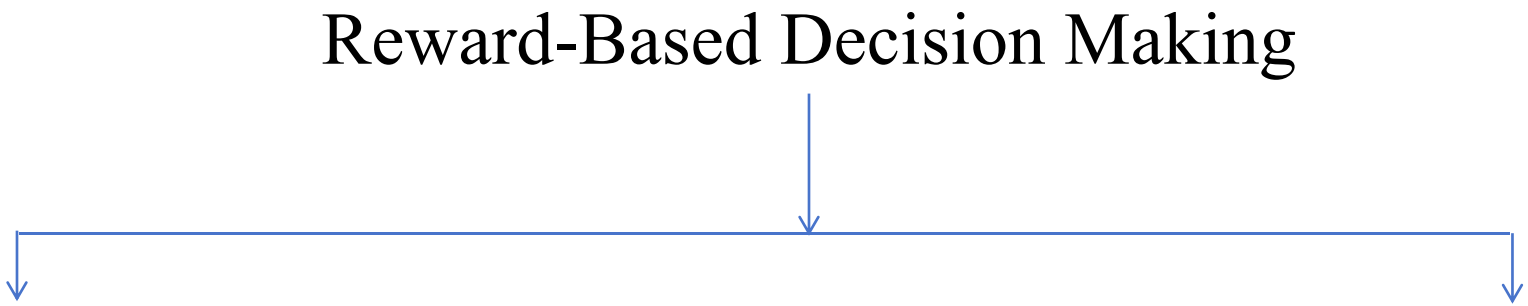
Planning as Inference

Matthew Botvinick, Marc Toussaint

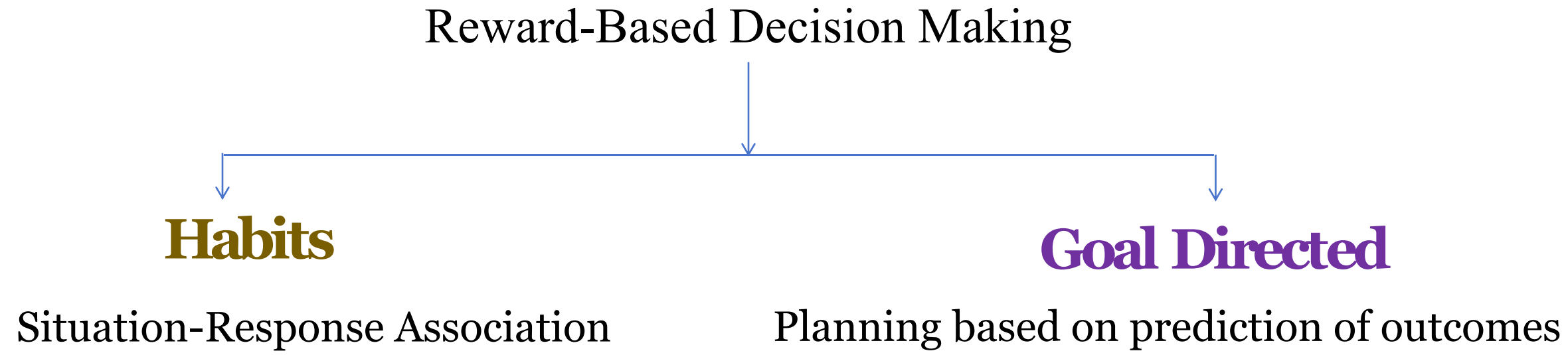
“Planning is accomplished through probabilistic inference.”

- Background: Why do we need a new perspective?
- What is PAI(planning as inference)?
- PAI applications
- Implications for Neuroscience
- Opportunities and Challenges

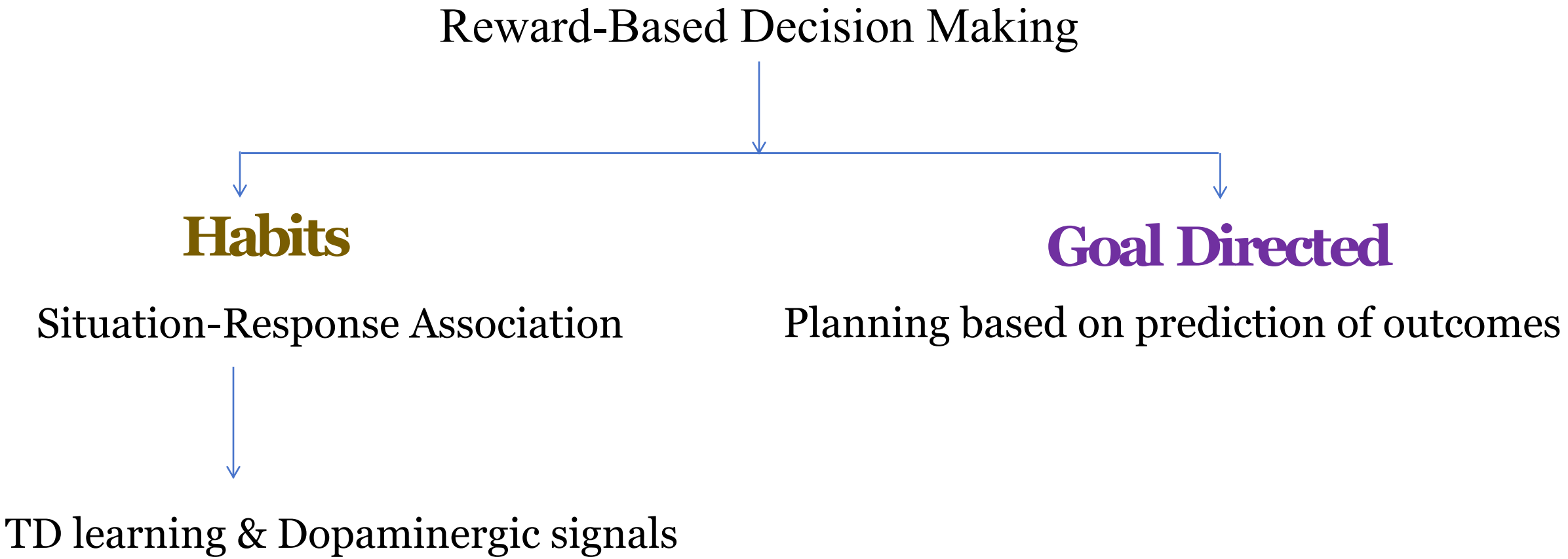
Background



Background

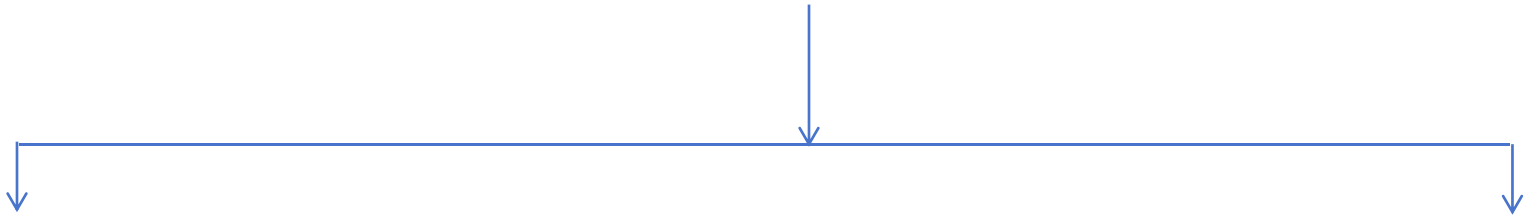

Reward-Based Decision Making
Habits
Goal Directed
Situation-Response Association
Planning based on prediction of outcomes

Background


Reward-Based Decision Making
Habits
Goal Directed
Situation-Response Association
Planning based on prediction of outcomes
TD learning & Dopaminergic signals

Background

Reward-Based Decision Making



Habits

Situation-Response Association

TD learning & Dopaminergic signals

Goal Directed

Planning based on prediction of outcomes

Not Clear:
What's the information processing operations?

Background

Traditional Perspectives on Planning:

- a specific a priori goal ⟷ the generic goal of reward maximization

Background

Traditional Perspectives on Planning:

- a specific a priori goal ⟷ the generic goal of reward maximization
- deterministic action outcomes: unrealisitic ⟷ MDP

Background

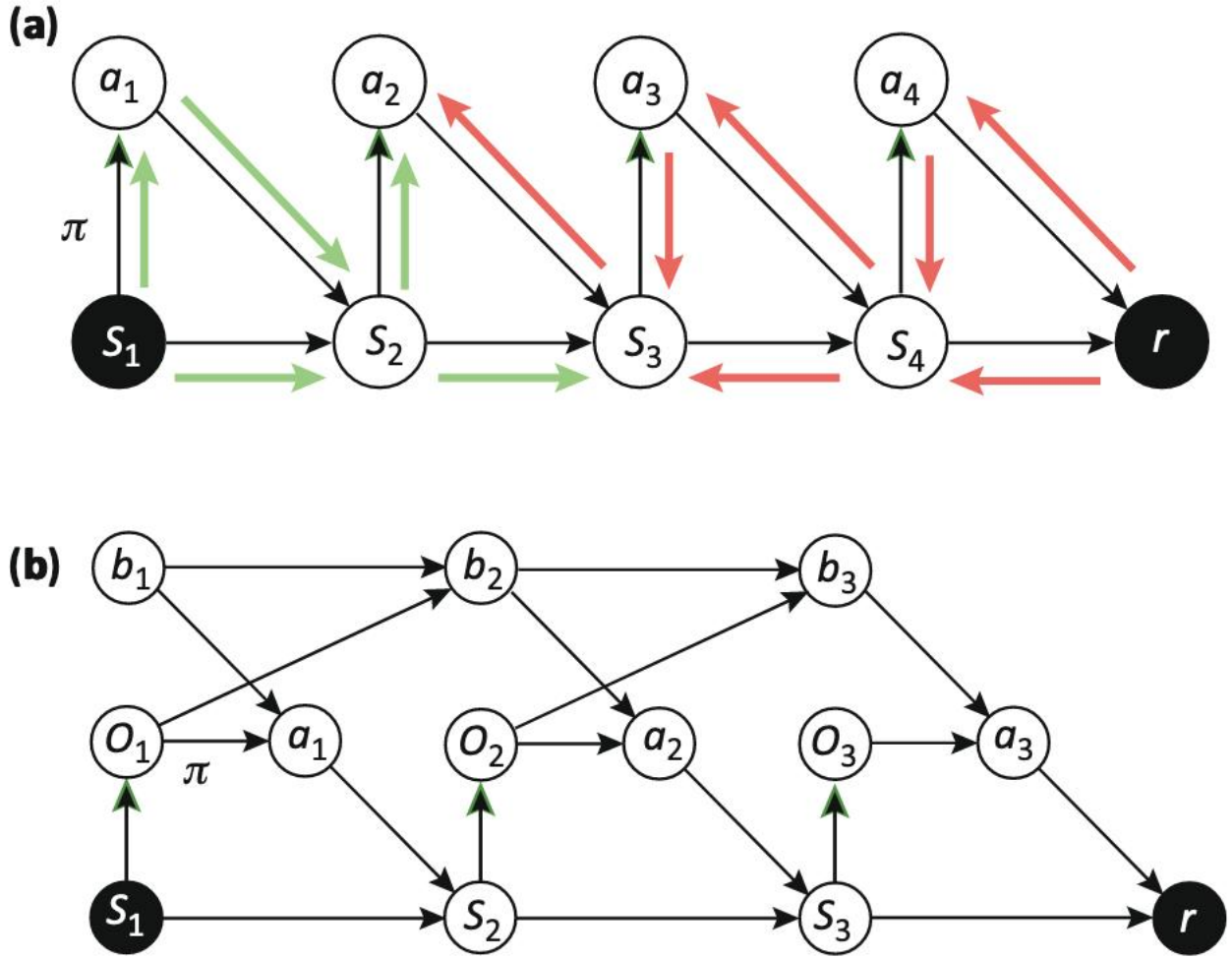
Traditional Perspectives on Planning:

- a specific a priori goal ⟷ the generic goal of reward maximization
- deterministic action outcomes: unrealisitic ⟷ MDP
- Dynamic Programming, Model-Based RL

Planning as Inference

Planning as Inference

The brain/agent has a generative model of the world

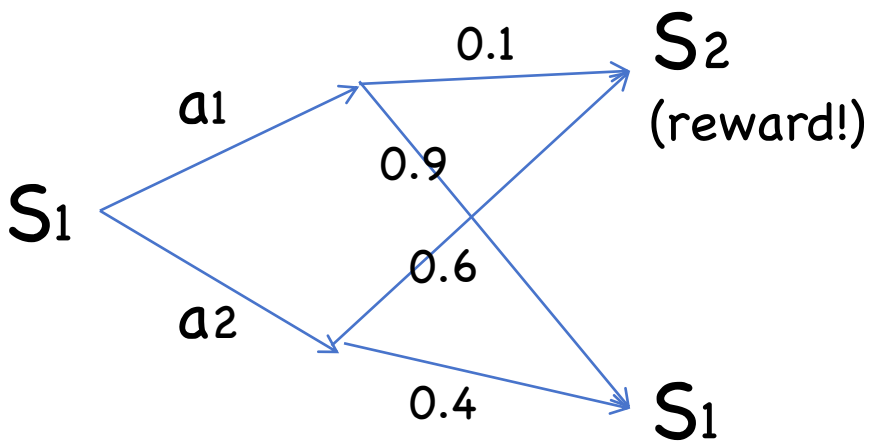


a joint probability distribution over actions, outcome states, and rewards

Planning as Inference

The brain/agent has a generative model of the world

Condition on reward: assume reward , what's the most probable action/policy that had been taken ?



Planning as Inference

The brain/agent has a generative model of the world

Condition on reward: assume reward , what's the most probable action that had been taken ?

Probabilistic Inference/ Inverse Inference

Dynamic Programming, Model-Based RL:

Forward-looking, Value optimization

Planning as Inference(PAI):

Backward-looking, Probabilistic Inference

PAI applications

- Success in Machine Learning and Robotics
- Discover optimal solutions more quickly
- Tackle complex problems(otherwise intractable)

PAI applications

Two Critical Developments in PAI theory:

- Rigorous Mathematical Formulations:

$$\min_{\pi} D_{KL}\left(P_{\pi}(\tau) \,||\, P(\tau|R=1)\right)$$

- Efficient Algorithms(e.g. EM)

PAI applications

E-step(expectation/inference)

Estimate the posterior distribution of optimal trajectories, q(τ), given the current policy π_k.

inverse inference

the improved policy

$$\pi_{k+1} = \arg\min_{\pi} D_{\mathrm{KL}}(q_k(\tau) \,\|\, \pi(\tau))$$

$$q_k(\tau) = P(\tau \mid R = 1, \pi_k)$$

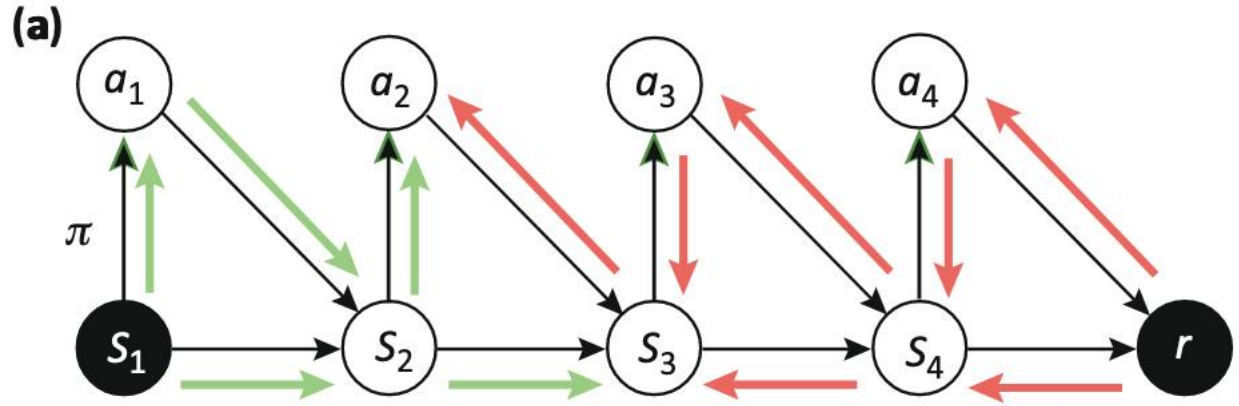
the ideal distribution

M-step(maximization/planning)

Update the policy to π_k+1 to minimize the KL divergence with the estimated posterior q(τ)

PAI applications

"message passing" in the generative model



Decentralized,
Localized

forward messages

backward messages

update policy

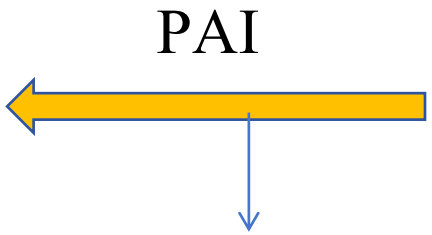
Iterative Procedure

Implications

"all cognitive and neural computation can be understood in terms of probabilistic inference"

Implications

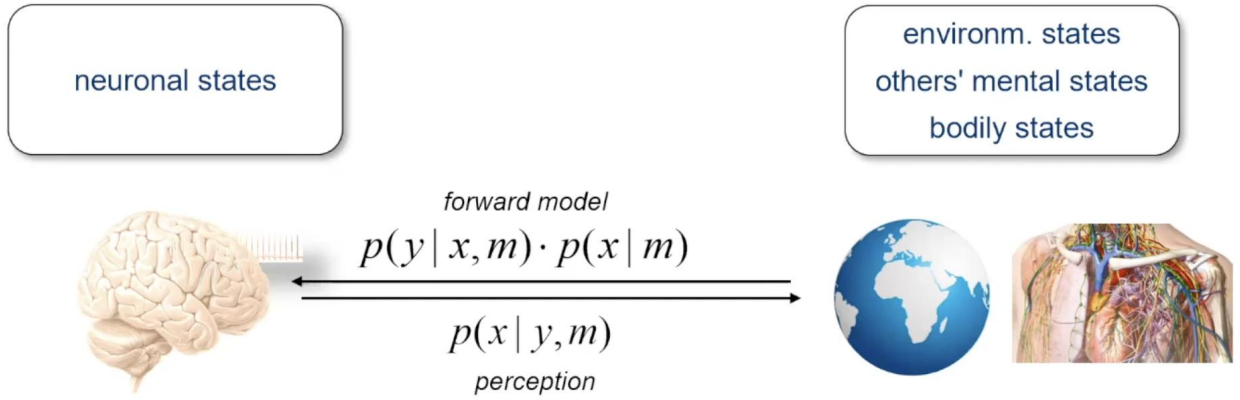
"all cognitive and neural computation can be understood in terms of probabilistic inference"



Planning

Bayesian theories of visual perception

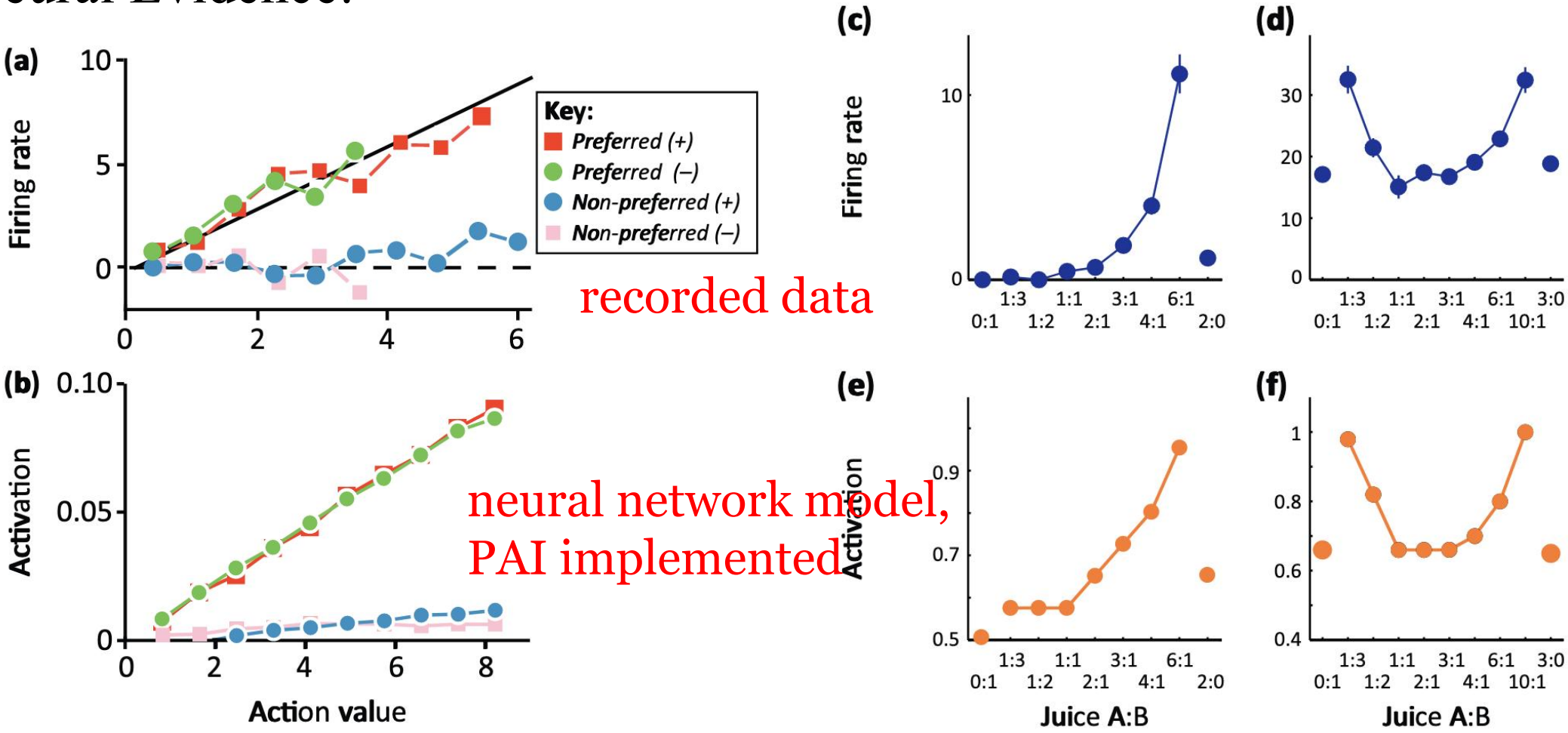
Generative models as a concept for brain function: the "Bayesian brain" hypothesis



perception = inference = inversion of a generative model

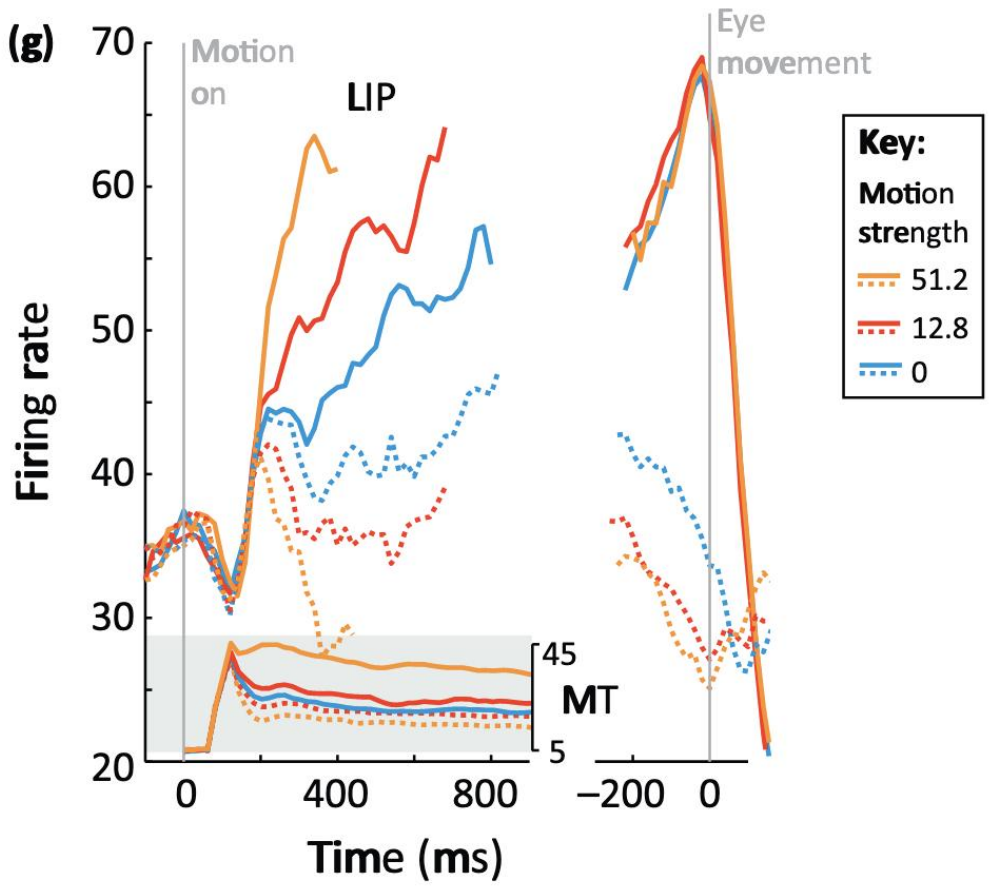
Implications

Neural Evidence:

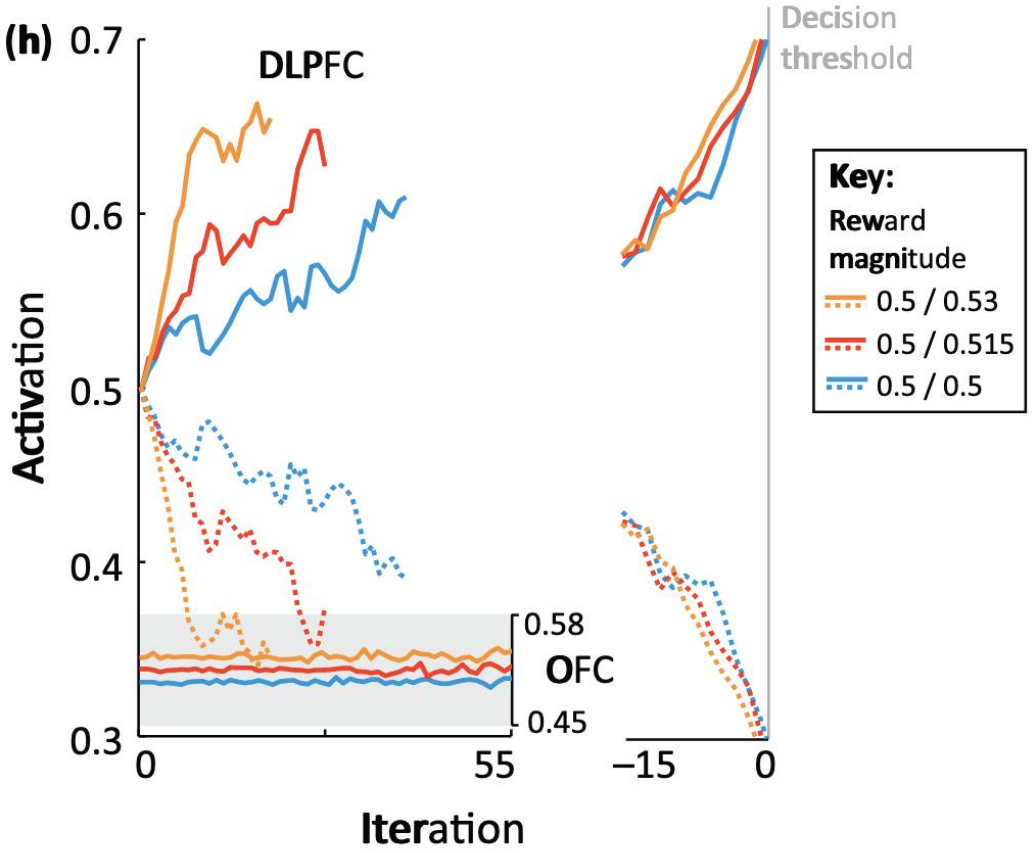


Implications

Neural Evidence:



Iteration
(integration of momentary evidence)



neural network model,
PAI implemented

Opportunities and Challenges

current work on PAI mostly theoretical  empirical evidence?

Opportunities and Challenges

current work on PAI mostly theoretical  empirical evidence?

"planning occurs under strict capacity limitations"(bounded rationality)

In ML: approximation techniques(e.g. sampling based)

approximation applied to PAI might offer new insights