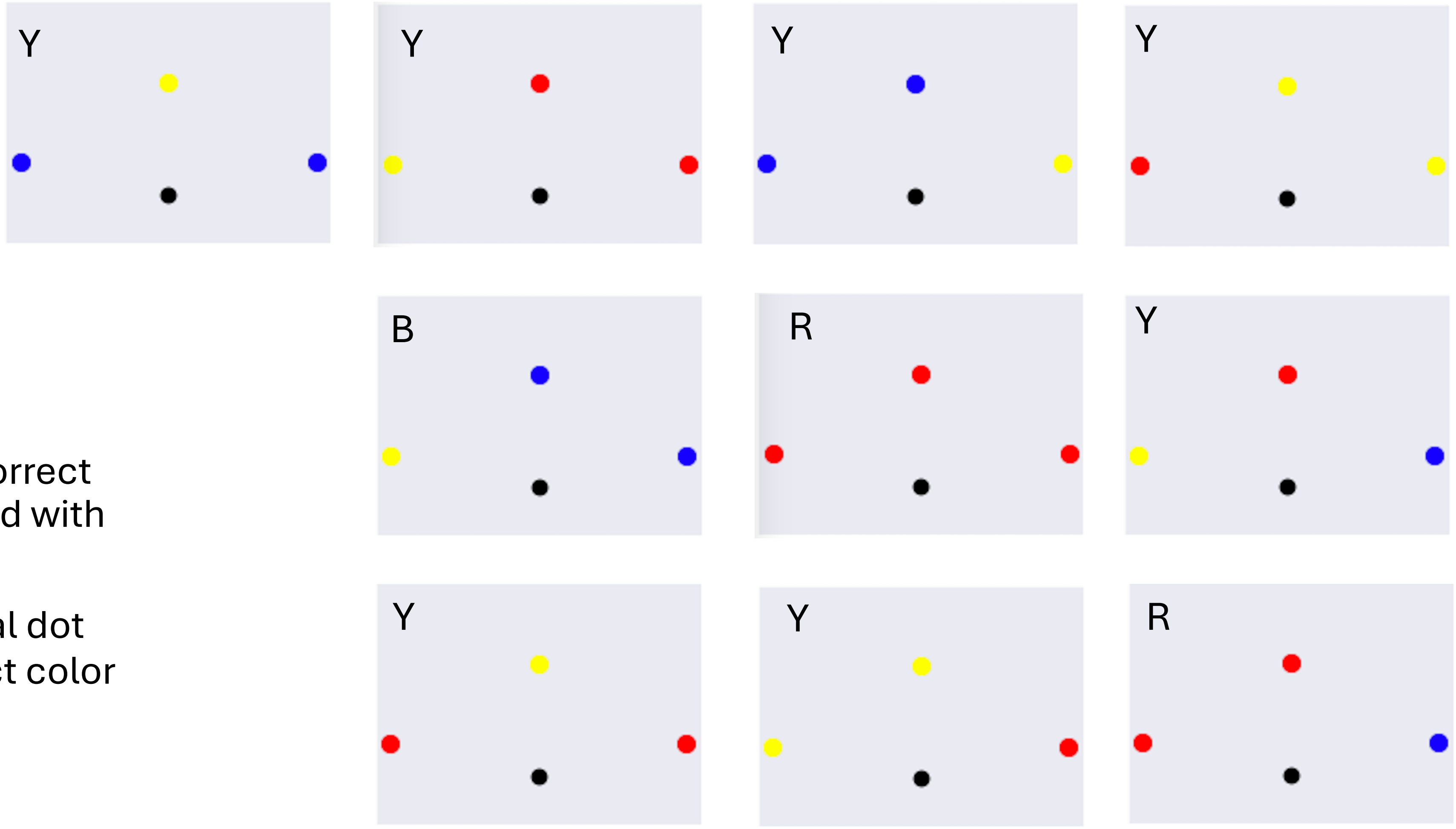
Active Inference, Insight and Curiosity

Based on the work of K. Friston, M. Lin, C. Frith, G. Peluzzo, Neural Computation, 2017

A little yet insightful game



Prior:

- You have to find the correct color that are indicated with the three color dots
- The color of the central dot determines the correct color

Outline

I – Our interior world model

II – A simple example: the Two-step Maze

III – Coming back to the three dots problem

How do we behave optimally ?

You are in a unknown town and you are hungry, what are you doing ?

➢ Most likely, you start looking for a restaurant: reduce your uncertainty about the world: **belief**

Classical setting

$$u_t^* = argmax_{u_t} \quad V(s_{t+1})|u_t)$$
$$= \pi(s_t)$$

Bellman optimality principle:

State action policy

Optimal Control Theory

Reinforcement Learning

Active Inference: optimal action depends on belief about states

$$u_t^* = argmin_{u_t} \quad F(Q(s_{t+1})|u_t)$$

But also subsequent sequence of action

$$\pi^* = argmin \sum_{\tau} \quad F(Q(s_\tau)|\pi)$$
$$u_\tau = \pi(\tau)$$

Hamilton's principle of least action:

Active Inference

Artificial Curiosity

Bayesian Sequence Optimization

Which functional ?

Free Variational Energy $\ln P(\pi) = \sum_{\tau} F(\pi, \tau)$

$$
\begin{aligned}
-F(\pi,\tau) &= \mathbb{E}_{Q(s_\tau, o_\tau \pi)}(\ln P(o_\tau, s_\tau|\pi)) + H(Q(s_\tau, \pi)) \\
&= \underbrace{\mathbb{E}_{Q(s_\tau, o_\tau \pi)}(\ln P(o_\tau))}_{\text{extrinsic value}} + \underbrace{\mathbb{E}_{Q(o_\tau, \pi)}(KL[Q(s_\tau|o_\tau, \pi)||Q(s_\tau|\pi)])}_{\text{intrinsic value or information gain}} \quad \underbrace{+\mathbb{E}_{\tilde{Q}}[\ln Q(\mathbf{A}) - \ln Q(\mathbf{A}|s_\tau, o_\tau, \pi)]}_{\text{(negative) novelty}}
\end{aligned}
$$

Strong links with known quantities or existing works in neuroscience: **maximizing mutual information and visual salience**

$$KL[Q(s_\tau|o_\tau, \pi)||Q(s_\tau|\pi)Q(o_\tau|\pi)]$$

Generative Model: a Markov decision process

Generative model

$$P\left(o_\tau^m \mid s_\tau^1,\ldots,s_\tau^N\right)=Cat(\mathrm{A}^m)$$

Likelihood

$$P\left(s_{\tau+1}^n \mid s_\tau^n,\pi\right)=Cat(\mathbf{B}_\tau^{n,\pi})$$
$$P(s_1^n)=Cat(\mathbf{D}^n)$$
$$P(o_\tau^m)=\sigma(-\mathbf{C}_\tau^m)$$

priors over hidden states

$$P(\pi)=\sigma(-\mathbf{G})$$

and policies

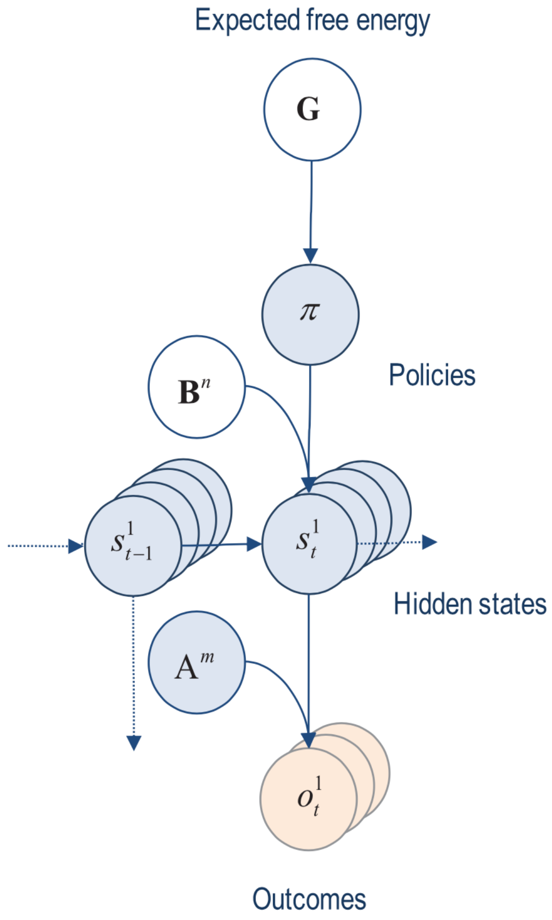
$$P(\mathrm{A}^m)=Dir(a^m)$$

and parameters

$$Q(\tilde{s},\pi,\mathrm{A})=Q(s_1\mid\pi)\ldots Q(s_T\mid\pi)Q(\pi)Q(\mathrm{A})$$
$$Q\left(s_\tau\mid\pi\right)=Q(s_\tau^1\mid\pi)\ldots Q(s_\tau^N\mid\pi)$$
$$Q(\mathrm{A})=Q(\mathrm{A}^1)\ldots Q(\mathrm{A}^M)$$

$$Q\left(s_\tau^n\mid\pi\right)=Cat(\mathbf{s}_\tau^{n,\pi})$$
$$Q(\pi)=Cat(\boldsymbol{\pi})$$
$$Q(\mathrm{A}^m)=Dir(\mathbf{a}^m)$$

Approximate posterior



Belief updates

Variational updates

Perception or state-estimation

$$\mathbf{s}_\tau^{n,\pi} = \sigma(\hat{\mathbf{s}}_\tau^{n,\pi})$$

$$\dot{\hat{\mathbf{s}}}_\tau^{n,\pi} = \hat{\mathbf{B}}_{\tau-1}^{n,\pi}\mathbf{s}_{\tau-1}^{n,\pi} + \hat{\mathbf{B}}_\tau^{n,\pi} \cdot \mathbf{s}_{\tau+1}^{n,\pi} + \sum_m o_\tau^m \cdot \hat{\mathbf{A}}^m \circ \mathbf{s}_\tau^{/n,\pi} - \hat{\mathbf{s}}_\tau^{n,\pi}$$

$$\mathbf{o}_\tau^{m,\pi} = \mathbf{A}^m \circ \mathbf{s}_\tau^\pi$$

Evaluation and policy selection

$$\boldsymbol{\pi} = \sigma(-\mathbf{G})$$

$$\mathbf{G}_\pi = \sum_{m,\tau} \mathbf{o}_\tau^{m,\pi} \cdot (\hat{\mathbf{o}}_\tau^{m,\pi} + \mathbf{W}^m \circ \mathbf{s}_\tau^\pi + \mathbf{C}_\tau^m) + \mathbf{H}^m \circ \mathbf{s}_\tau^\pi$$

$$u_t^n = \max_u \boldsymbol{\pi} \cdot [U_t^{n,\cdot} = u]$$

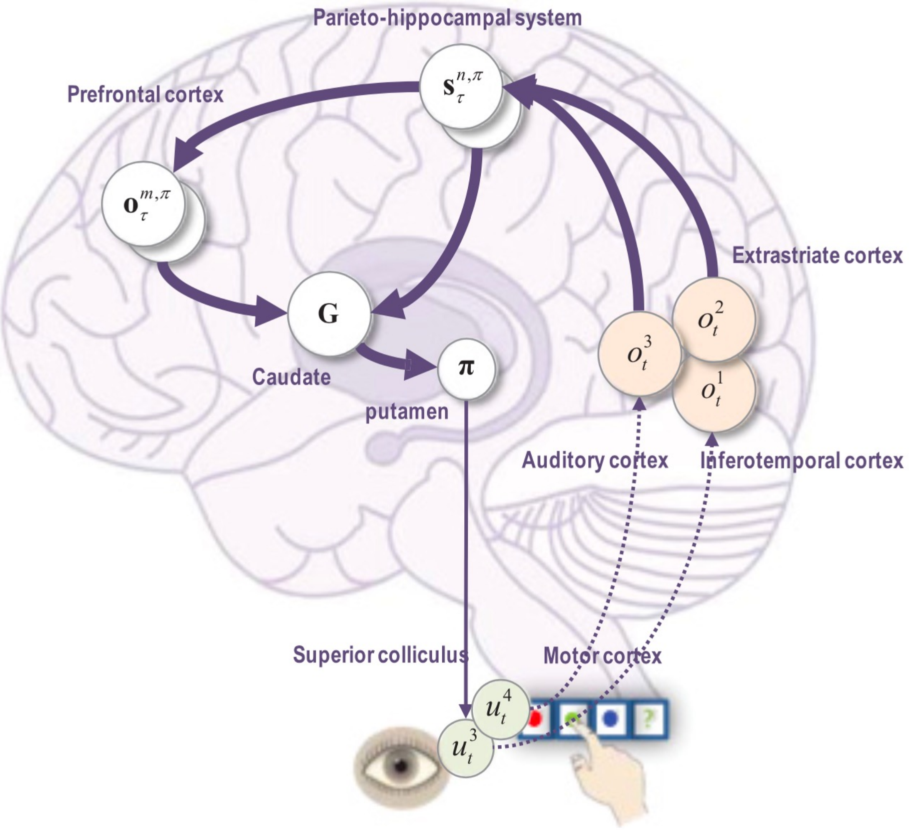
Bayesian model averaging and learning

$$\mathbf{s}_\tau^n = \sum_\pi \boldsymbol{\pi}_\pi \cdot \mathbf{s}_\tau^{n,\pi}$$

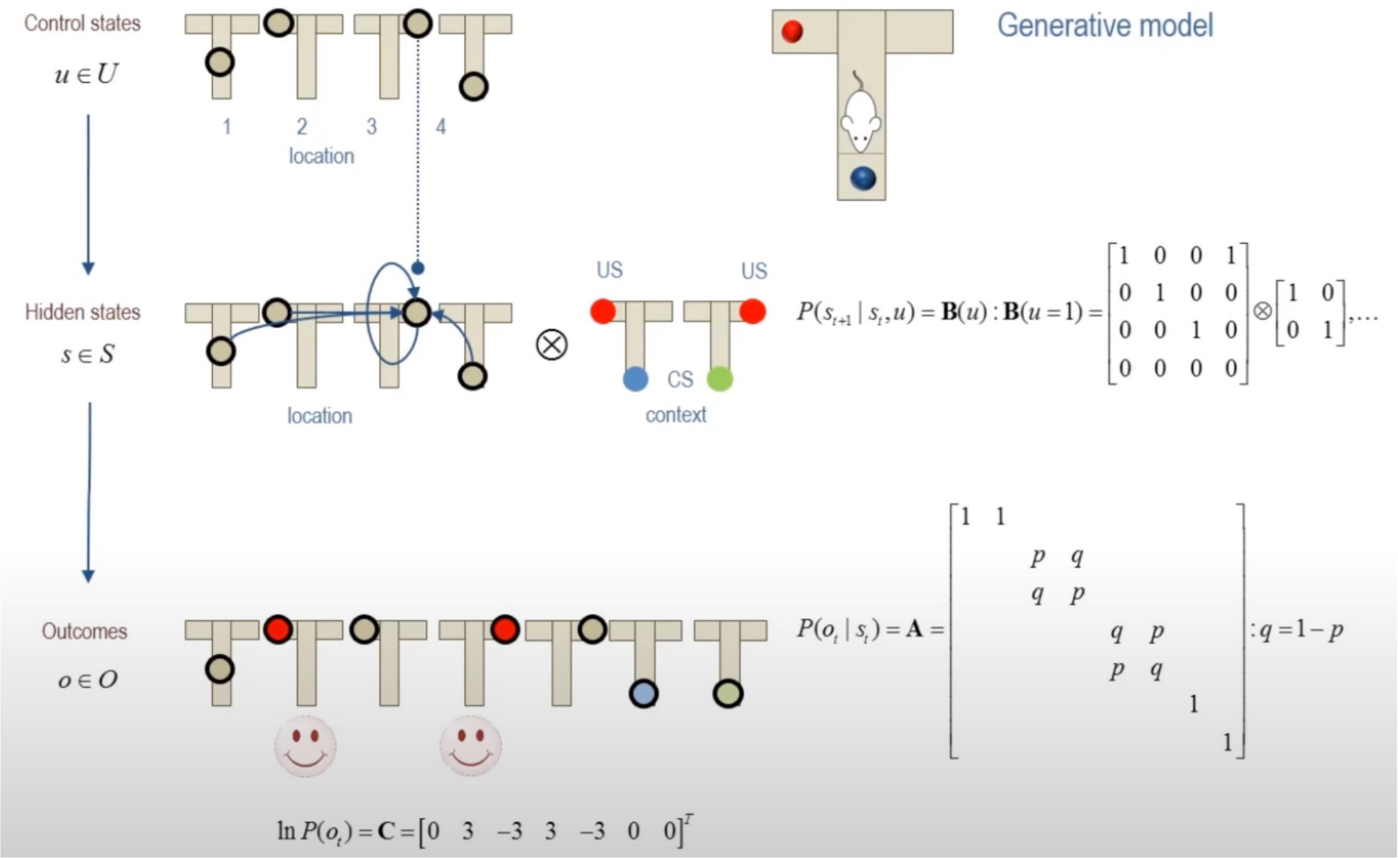
$$\hat{\mathbf{A}}^m = \psi(\mathbf{a}^m) - \psi(\mathbf{a}_0^m)$$

$$\mathbf{a}^m = a^m + \sum_\tau o_\tau^m \otimes \mathbf{s}_\tau^1 \otimes \mathbf{s}_\tau^2 \ldots$$

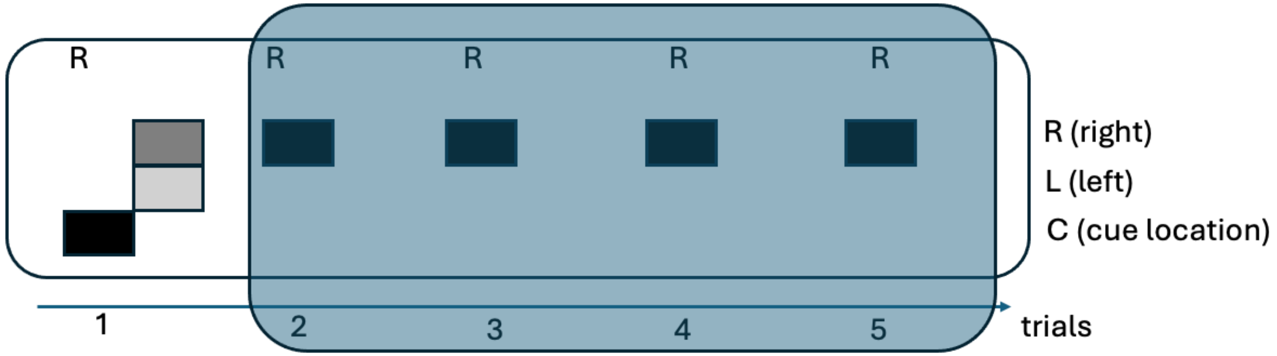
Functional anatomy

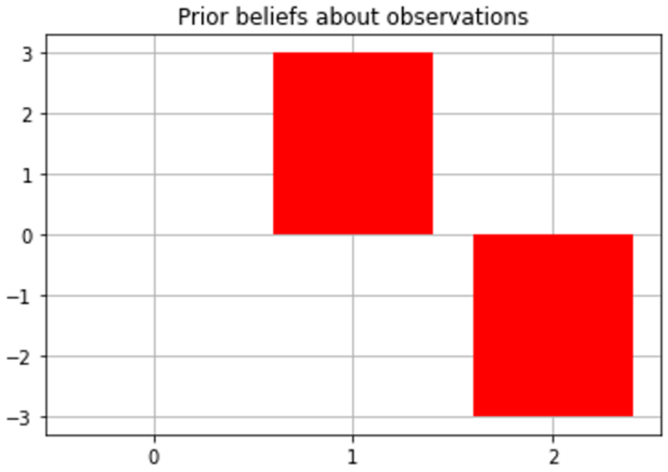


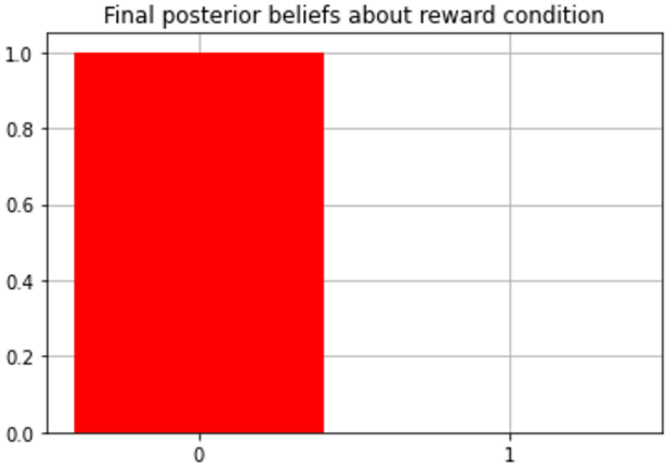
The two step maze



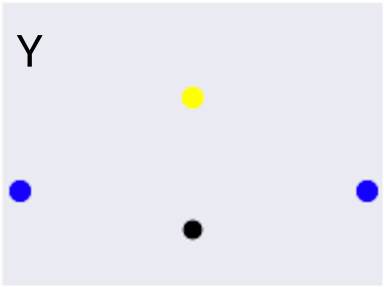
Results

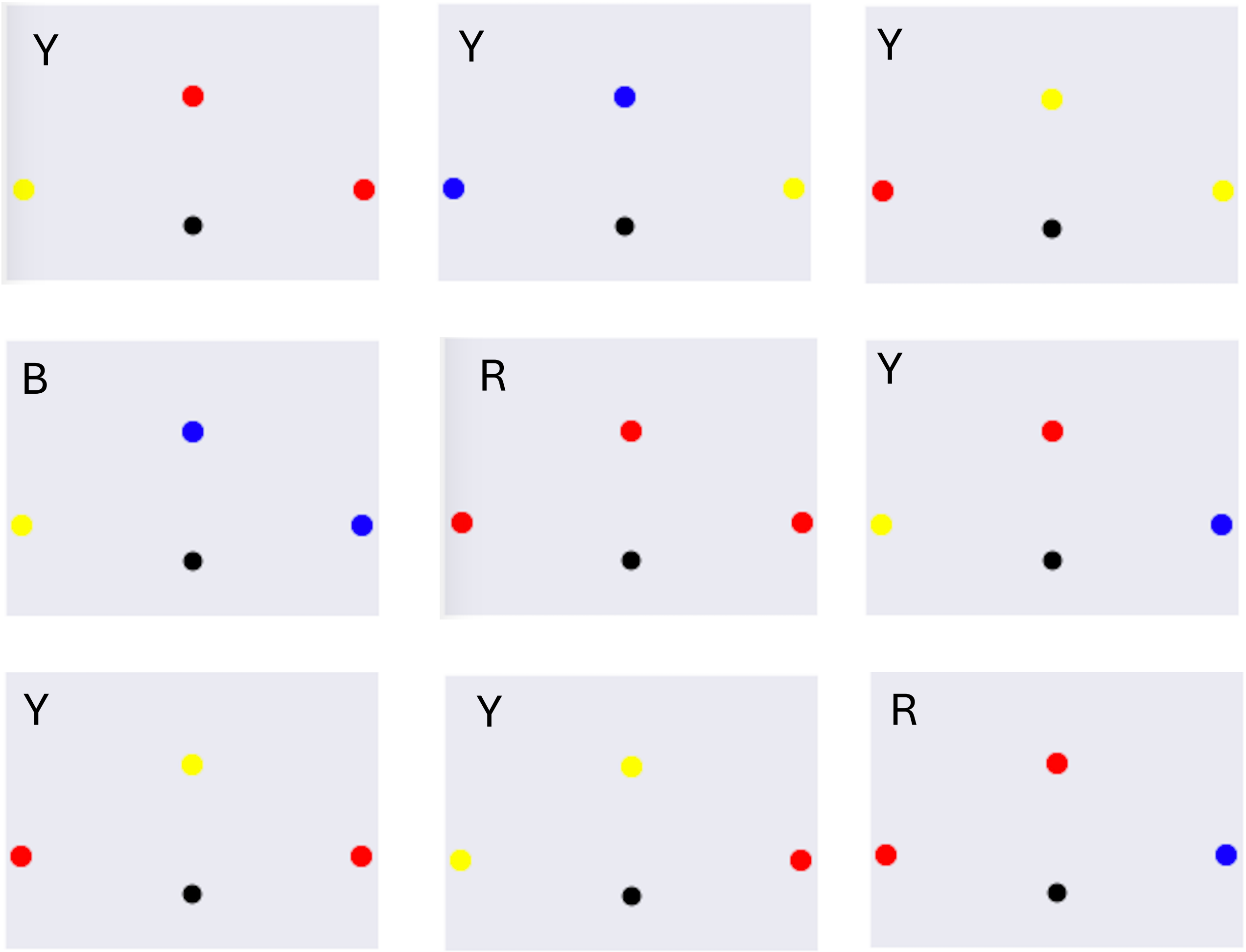

R
R
R
R
R
R (right)
L (left)
C (cue location)
1
2
3
4
5
trials


Prior beliefs about observations
3
2
1
0
−1
−2
−3
0
1
2


Final posterior beliefs about reward condition
1.0
0.8
0.6
0.4
0.2
0.0
0
1

Coming back to our game

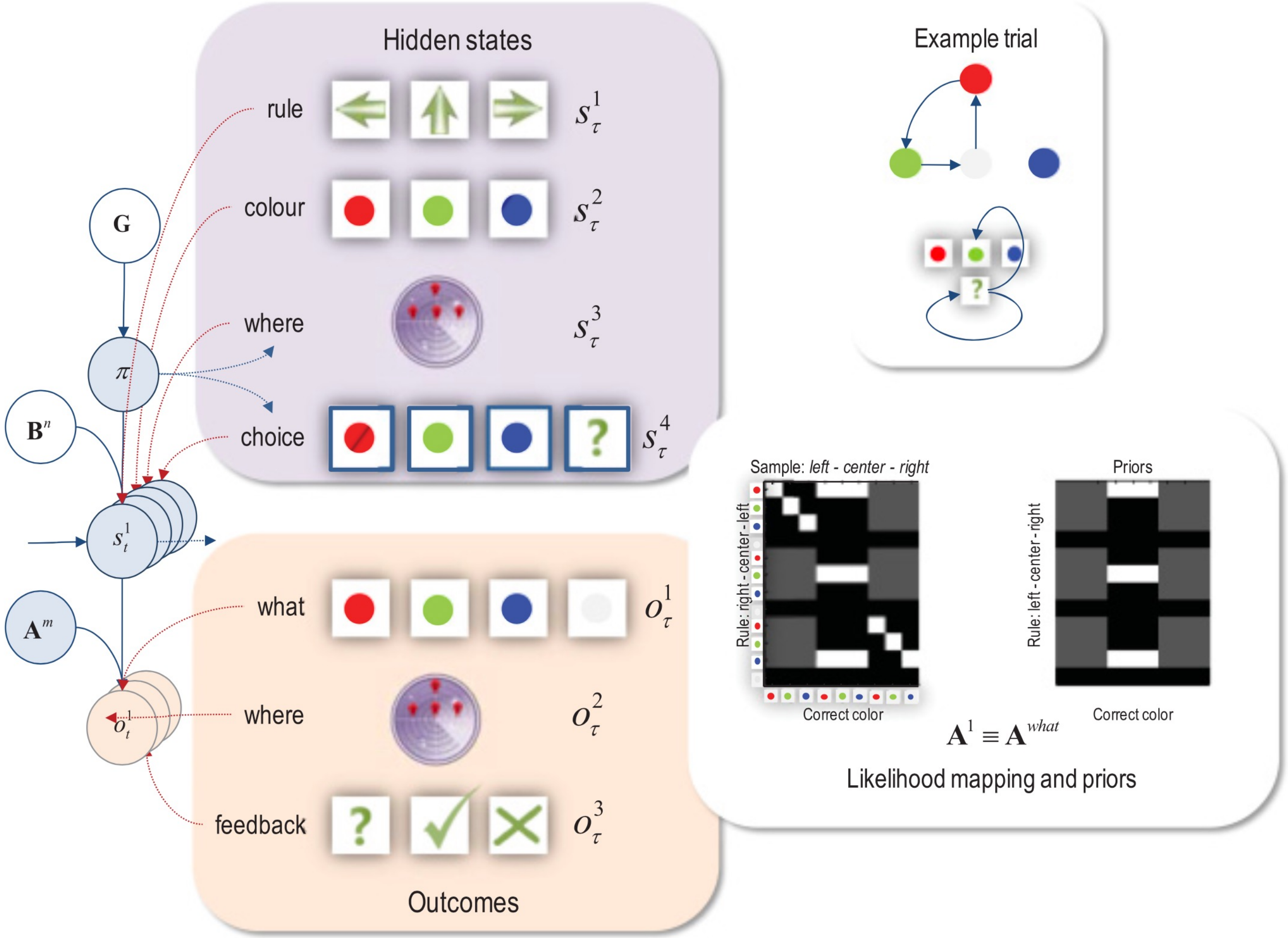




Prior:

- You have to find the correct color that are indicated with the three color dots
- The color of the central dot determines the correct color

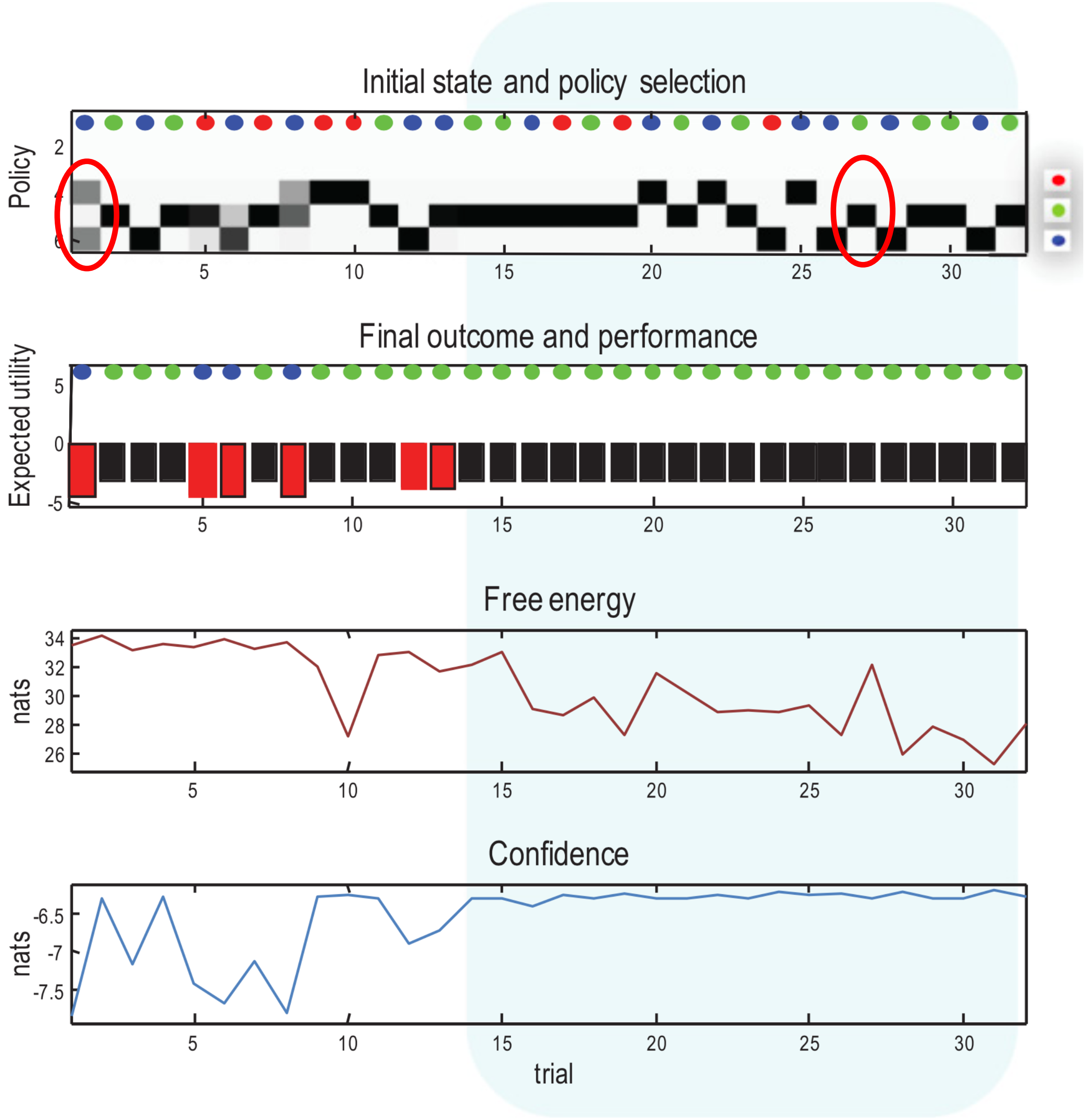
The model and the A matrix



Results

Confidence is $-\pi.\hat{\pi}$

The agent can retry if fail !



Let’s take a step back

We programmed our agent optimally with Bayesian framework and variational updates (gradient descent on the free energy)

However, the optimal agent takes 14 moves to find out the rule, and us humans: far less ! Why and how this happens ?

We **know** that there is a rule at play, and how rules are generally built:

If A then B

No mixture of outcomes, low entropy: the state in which we search our optimal A is reduced

Bayesian Model Reduction

- Comparing the evidence of different models
- To encode a rule, we switch off connections if there is not enough evidence that there is a connection between hidden states and outcomes

$$\Delta F = \ln \mathrm{B}(\mathbf{a}) + \ln \mathrm{B}(a') - \ln \mathrm{B}(a) - \ln \mathrm{B}(\mathbf{a} + a' - a)$$

In other words, we change the concentration parameters *a* of the likelihood matrix
How different is free enrergy if we had started observing outcomes with simpler prior beliefs

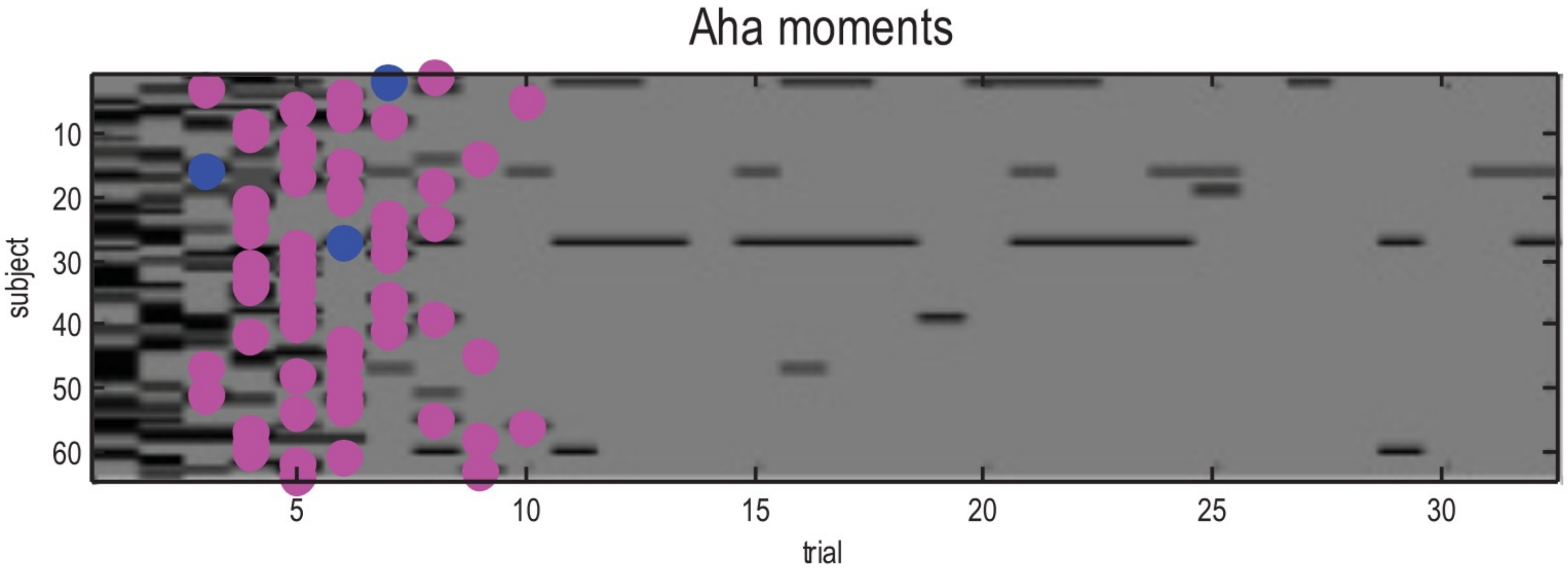
➢ Take a nap, or rehearse the past experiences


Posteriors
Correct color


Reduced
Correct color

Reflection (model reduction) at each time step

- We perform Bayesian Model Reduction after each trial to simplify the *A* matrix, to encode the prior « what is a rule »
- Emergence of insight is comparable to that of humans!



Thank you!

- Karl J. Friston et al. "Active Inference, Curiosity and Insight". en. In: Neural Computation 29.10 (Oct. 2017), pp. 2633–2683.
- Lectures and interviews of K. Friston, especially at University of Edinburgh in 2017. Link: https://www.youtube.com/watch?v=Y1egnoCWgUg