Active Inference

Philipp Schwartenbeck

Wellcome Trust Centre for Human Neuroimaging, UCL

Computational Psychiatry Course 2019, TNU, Zurich

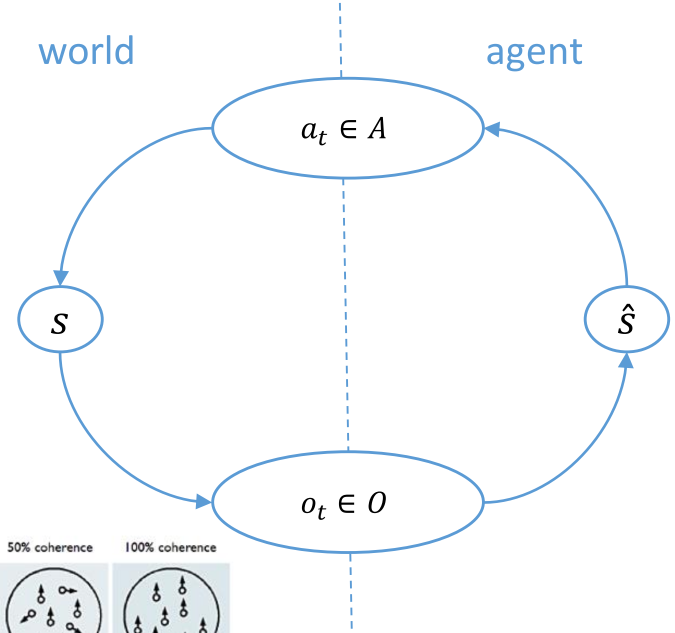
What is active inference?

To make sense of the world, we need to infer its latent structure (hidden causes)

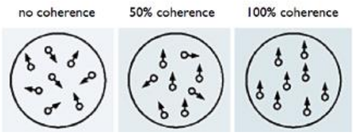
Perception as inference, Bayesian Brain

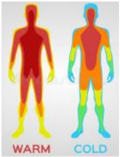
Lee & Mumford, 2003;
Knill & Pouget, 2004;
Doya et al., 2007

v. Helmholtz, 1867



Bayesian Decision Theory, Good Regulator Theorem (Conant & Ashby, 1970), **active inference**





Active inference is concerned with closing the link between perception and action

Inferring the latent structure of the world
- to behave adaptively!

CPC
2019
SCHEDULE

WEDNESDAY – 4th September 2019, location: NO, C60

8:15	8:30	Introduction to Day 3	
8:30	9:25	Active Inference	Philipp Schwartenbeck
9:25	10:20	Drift-Diffusion Model	Ariel Zylberberg
10:20	10:50	Break (catering)	
10:50	11:45	Machine Learning	Andre Marquand
11:45	12:40	Machine Learning Advanced	Andre Marquand
12:20	13:50	Lunch Break (no catering)	
13:50	14:40	DCM for fMRI	Jakob Heinzle
14:40	15:30	DCM for EEG	Dario Schöbi
15:30	16:00	Break (catering)	
16:00	16:50	Advanced models of connectivity	Stefan Frässle, Yu Yao

MODELS OF DECISION MAKING & BIOPHYSICAL MODELS

SCHEDULE

MONDAY, 02.09.
CLINICAL PSYCHIATRY

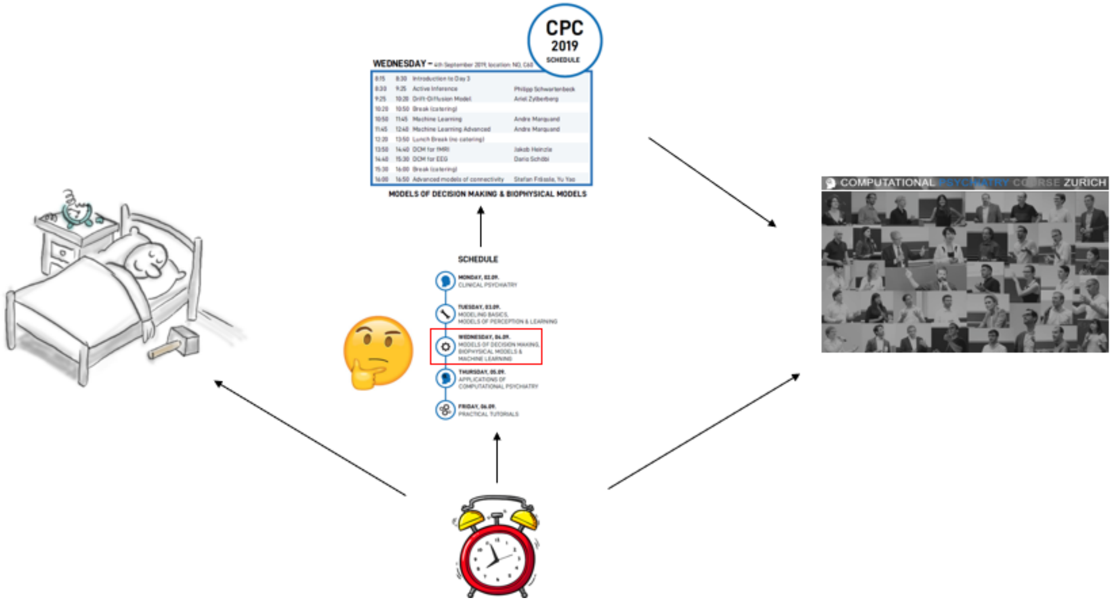
TUESDAY, 03.09.
MODELING BASICS,
MODELS OF PERCEPTION & LEARNING

WEDNESDAY, 04.09.
MODELS OF DECISION MAKING,
BIOPHYSICAL MODELS &
MACHINE LEARNING

THURSDAY, 05.09.
APPLICATIONS OF
COMPUTATIONAL PSYCHIATRY

FRIDAY, 06.09.
PRACTICAL TUTORIALS

Inferring the latent structure of the world - to behave adaptively!



This sounds easy but it is not:

- Necessary to have insight into preferences
- Requires to optimise the balance between information and reward
- Requires knowledge about informative actions -> active learning/inference
- Necessary to get sequence right

Inferring the latent structure of the world
- to behave adaptively!

Central quantity of interest: **(variational) free energy**

- Quantifies the mismatch between **observations** and **beliefs about the world**

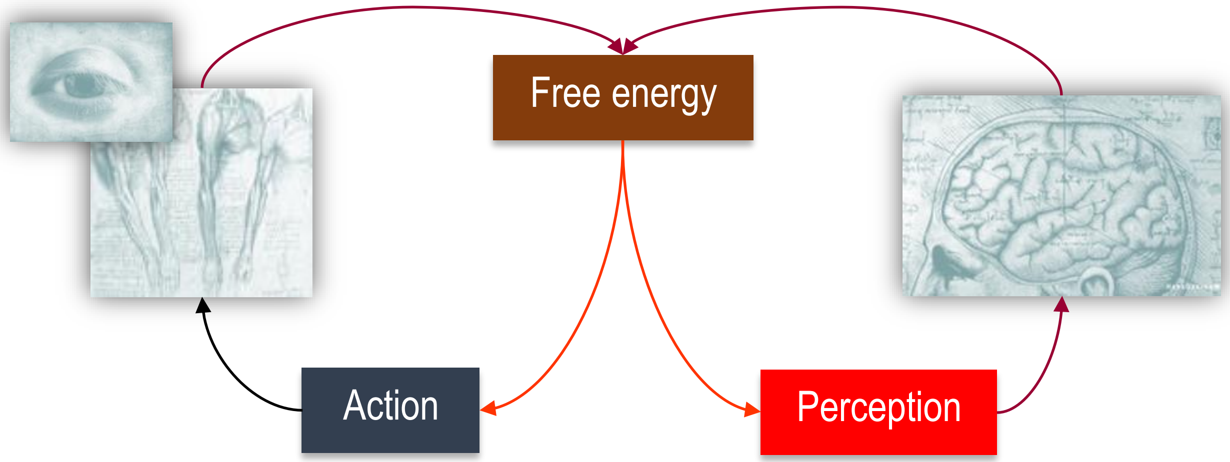
$$F = \ln P(o) - D[Q(s)||P(s|o)]$$

Central idea:

Action fulfils prior expectations

⇔ Maximise *model evidence* (⇔minimise *surprise*)

⇔ minimise *(variational) free energy*



Changing sensations to maximise evidence

$\ln P(o)$

Changing beliefs to minimise divergence

$D[Q(s)||P(s|o)]$

Outline

I. (Variational) inference
  - Generative models and state inference
  - Variational free energy
  - Information theory

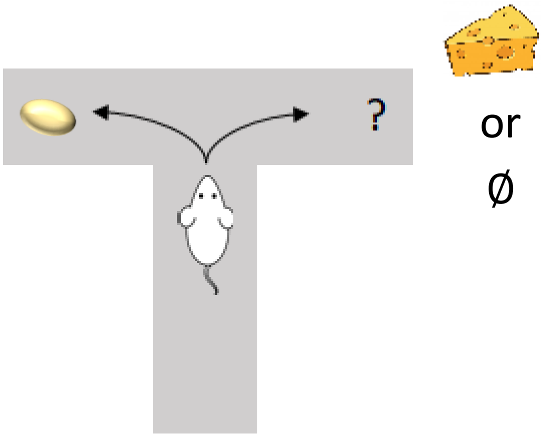
II. Active Inference and active learning
  - Using variational inference for understanding action

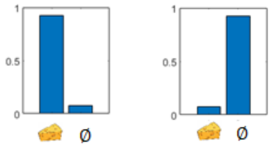
III. Some interesting predictions
  - What makes an action valuable?
  - Different types of information gain

I. (Variational) inference

Inference: generative models and hidden states

Imagine a mouse in a T-maze:





Is there a **high** or **low** probability for obtaining cheese?

Inference is based on a **generative model**

- I.e. a probabilistic mapping based on a **likelihood function** and a **prior density**:

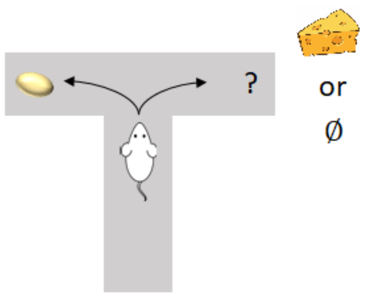
$$p(o, s|m) = p(o|s, m)p(s, m)$$

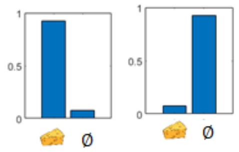
Perform inference on hidden states by applying Bayes rule (‘**model inversion**’):

$$p(s|o, m) = \frac{\overbrace{\boxed{p(o|s, m)}}^{\text{Likelihood}}\overbrace{\boxed{p(s|m)}}^{\text{Prior}}}{\underbrace{\boxed{p(o|m)}}_{\text{Marginal likelihood, model evidence}}}$$

Inference: generative models and hidden states

Imagine a **Bayesian** mouse in a T-maze:





Is there a **high** or **low** probability for obtaining cheese?

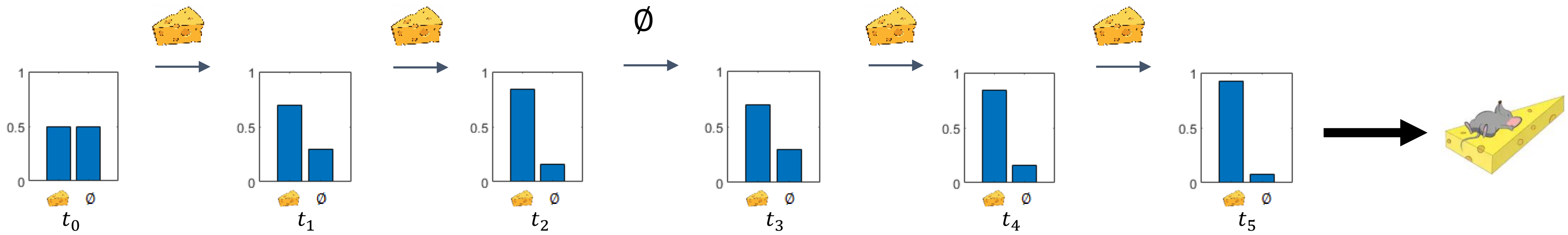
$$p(o_t = 1 | s_t = 1, m) = p(o_t = 2 | s_t = 2, m) = 0.7$$

$$p(o_t = 1 | s_t = 2, m) = p(o_t = 2 | s_t = 1, m) = 0.3$$

$t = 0$: $\quad p(s_t = 1, m) = p(s_t = 2, m) = 0.5$

$t \neq 0$: $\quad p(s_t, m) = p(s_{t-1}, m)$

$$p(s_t | o_t, m) = \frac{p(o_t | s_t, m) p(s_t | m)}{p(o_t | m)} = \frac{p(o_t | s_t, m) p(s_t | m)}{\sum_i p(o_t | s = i, m) p(s = i | m)}$$



Variational Bayes

This usually doesn't work

- Exact inference is generally intractable

$$p(s|o,m) = \frac{p(o|s,m)p(s|m)}{p(o|m)}$$

Variational Bayes allows to cast an inference problem (*difficult*) as a bound optimisation problem (*easier*) *(Beal, 2003)*

- Idea: approximate model evidence in Bayesian inference

Allows to derive specific variational update equations for a given problem

- Maths is not trivial – but only has to be done once!
- Then implement resulting update equations

Provides hypotheses about neuronal implementation of Bayesian inference

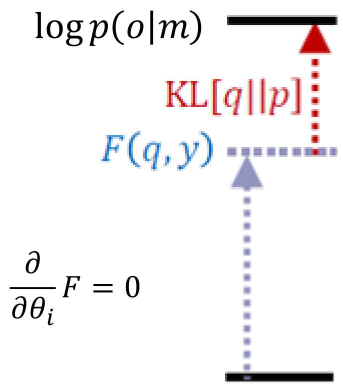
Variational Bayes: free energy

Negative variational free energy provides a **lower bound on model evidence:**

$$F = \boxed{\log p(o|m)} - \boxed{D_{KL}[q(s), p(s|o,m)]}$$

Marginal likelihood,
model evidence

'Distance' between arbitrary distribution
$q(s)$ and true posterior $p(s|o,m)$

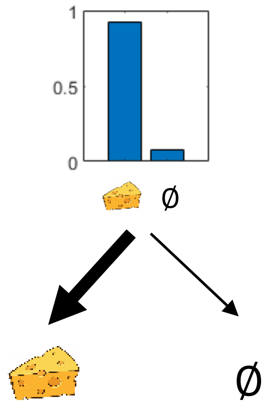


Variational free energy → 0 implies **maximising model evidence** and obtaining a **good approximation of the true posterior**

Note: when $q(s) = p(s|o,m)$, free energy is identical to model evidence

(and inference becomes exact)

Information theory



Information theory quantifies the information content of a signal

- Unlikely events are more informative than likely events

This can be quantified as the **self-information** or **surprise** of a signal *(Shannon, 1948)*:

$$I(o) = \boxed{-logP(o|m)}$$

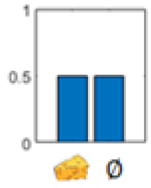
This is negative log-model evidence!

$$F = \underbrace{\boxed{\log p(o|m)}}_{\text{Marginal likelihood, model evidence}} - D_{KL}[q(s), p(s|o,m)]$$

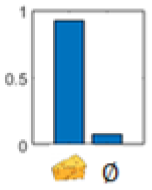
Other important quantities that we will use:

- Expected value of surprise is called (**Shannon**) **entropy**:

$$H(o) = \mathbb{E}_o[I(o)] = -\sum_i p(o_i|m) \cdot \log p(o_i|m)$$



vs.



- **Mutual information**: how much information about a variable can be gained by observing another variable?

$$I(s;o) = H(s) - H(s|o) = \mathbb{E}_o[D_{KL}(p(s|o)||p(s))]$$

(Variational) inference summary

III. Surprise is the negative of model evidence

$$I(o) = -logP(o|m)$$

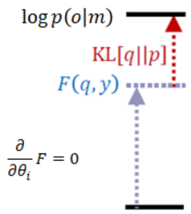
Minimising free energy minimises surprise, and maximises model evidence!

Minimising free energy results in a distribution that approximates the posterior

$$q(s) = \underset{q(s)}{\operatorname{argmin}} F = p(s|o,m)$$

II. Inference is **approximated** based on (**negative**) **variational free energy**

$$F = \log p(o|m) - D_{KL}[q(s), p(s|o,m)]$$



This is hard

I. Inference on hidden states is based on a **generative model** of the world

$$p(s|o,m) = \frac{p(o|s,m)p(s|m)}{p(o|m)}$$

II. Active inference and active learning

Active (variational) inference?

This works well for ‘perception’, but can we use the same approach to understand action?

Yes!

Friston et al. (2013; 2017)

But we need to change the definition of the (variational) free energy slightly, because

1. The free energy should depend on action (policies)
2. The free energy should be about future observations

$$F = \log p(o|m) - D_{KL}[q(s), p(s|o,m)]$$

$$\Leftrightarrow F = \sum_s q(s) \log \frac{q(s)}{p(o,s)}$$

Variational free energy F

$$G = \sum_s q(s_t|\pi) \log \frac{q(s_t|\pi)}{p(o_t, s_t|\pi)}$$

t: some time step in the future

π: a policy (sequence of actions)

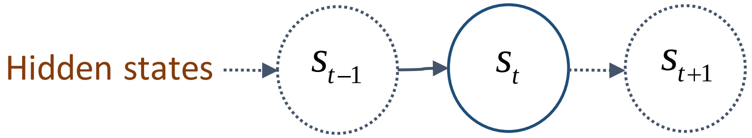
Expected free energy G

Class of problems: Markov Decision Processes

We are dealing with **partially observable Markov decision processes** (**POMDP**)

Key ingredients:

- $1, \dots, \mathrm{T}$ discrete time-steps

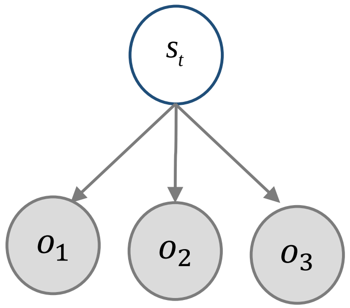


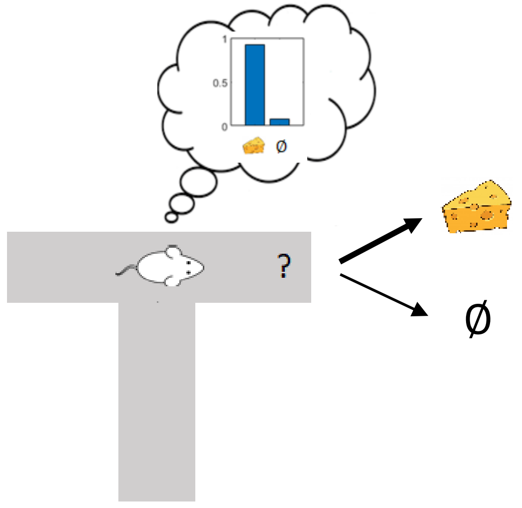
Class of problems: Markov Decision Processes

We are dealing with **partially observable Markov decision processes** (**POMDP**)

Key ingredients:

- $1, \ldots, T$ discrete time-steps
- $P(o_t | s_t)$ not trivial



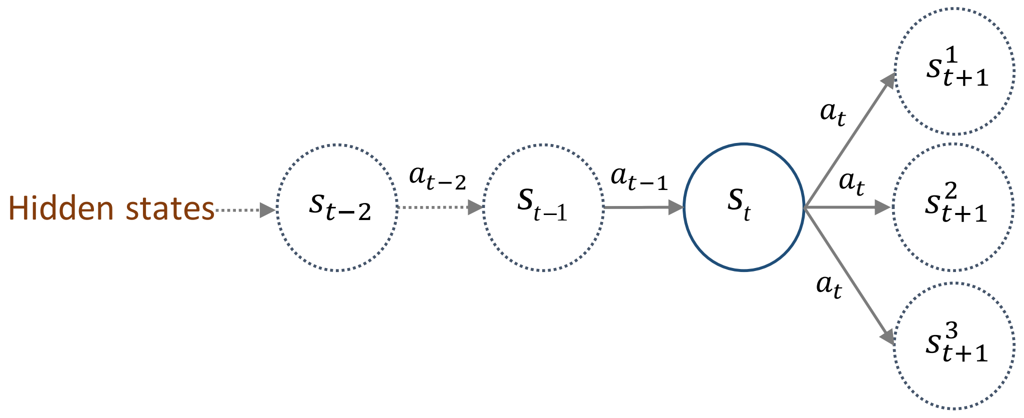


Class of problems: Markov Decision Processes

We are dealing with **partially observable Markov decision processes** (**POMDP**)

Key ingredients:

- $1, \ldots, T$ discrete time-steps
- $P(o_t|s_t)$ not trivial



- Markov-property: $P(s_{t+1}|s_t, a_t) = P(s_{t+1}|s_t, a_t, s_{t-1}, a_{t-1}, s_{t-2}, a_{t-2}, \ldots)$

A Markovian generative model

likelihood (empirical and full) priors

$$p(o,s|m) = \boxed{p(o|s,m)}\boxed{p(s,m)}$$

$$P(\tilde{o},\tilde{s},\pi,\gamma,A) = \boxed{\prod_{t=1}^{\tau} P(o_t|s_t)}\boxed{P(s_t|s_{t-1},\pi)P(\pi|\gamma)P(\gamma)P(A)}$$

$$= \boxed{P(\tilde{o}|\tilde{s})}\boxed{P(\tilde{s}|\pi)P(\pi|\gamma)P(\gamma)P(A)}$$

Generative model

$P(o_t|s_t) =: \mathbf{A}$ Observation model: mapping from hidden states to observations

$P(s_{t+1}|s_t,\pi) =: \boldsymbol{B}(\pi(t))$ Transition probabilities: mapping from current to next states contingent on policies

$P(s_0) =: \boldsymbol{d}$ Beliefs about initial states in a task

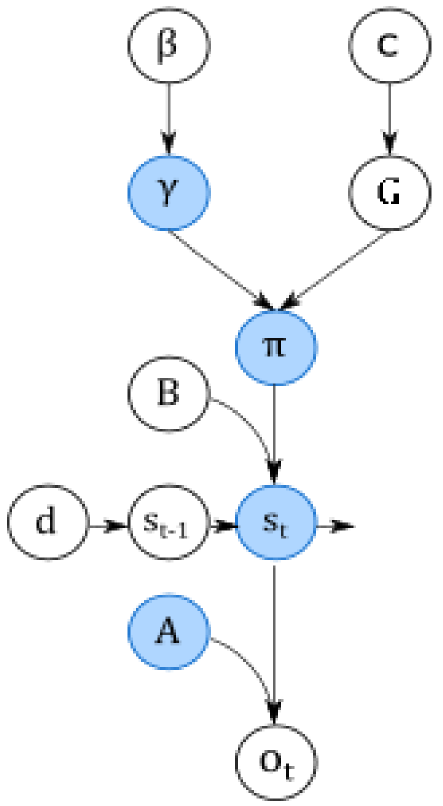
$P(o_t) =: \boldsymbol{c}_t$ Beliefs about outcomes in a task (= preferences)

$P(\pi|\gamma) = \sigma(-\gamma \cdot G(\pi))$ Beliefs about policies

$P(\boldsymbol{A}) = Dir(a)$ Parameters of observation model

$P(\gamma) = \Gamma(1,\beta)$ Precision (inverse stochasticity in behaviour)

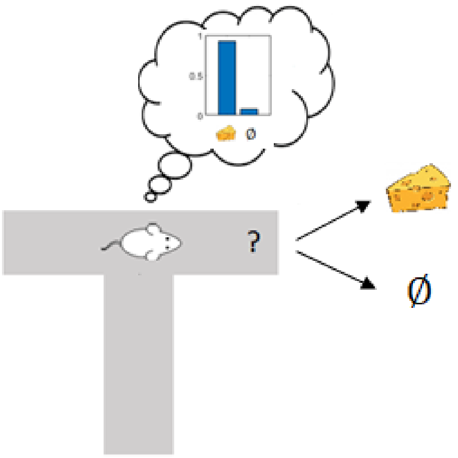
$G(\pi) = \sum_s q(s_t|\pi) \log \frac{q(s_t|\pi)}{p(o_t,s_t|\pi)}$ Expected free energy

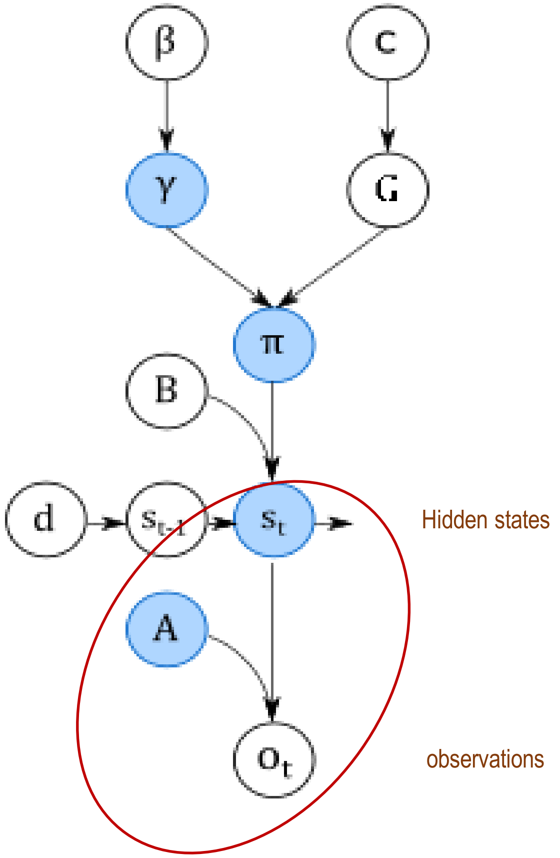


Key idea: use (variational) inference to solve MDPs

A Markovian generative model: observation model

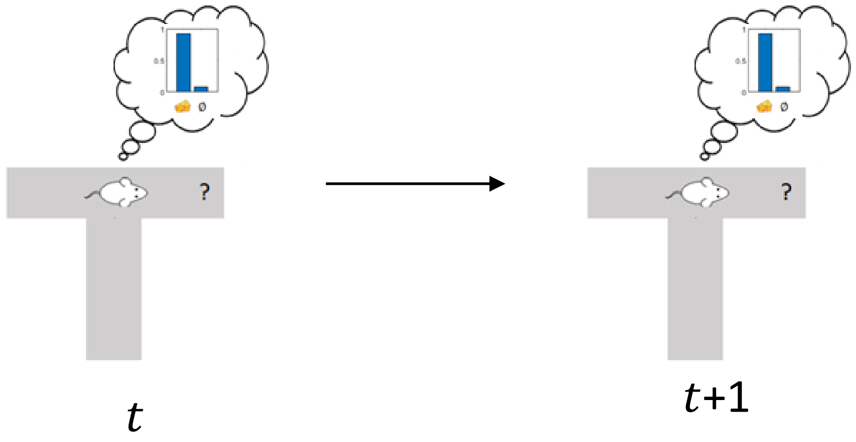
$P(o_t \mid s_t) =: \mathbf{A}$	Observation model: mapping from hidden states to observations
$P(s_{t+1} \mid s_t, \pi) =: \boldsymbol{B}(\pi(t))$	Transition probabilities: mapping from current to next states contingent on policies
$P(s_0) =: \boldsymbol{d}$	Beliefs about initial states in a task
$P(o_t) =: \boldsymbol{c}_t$	Beliefs about outcomes in a task (= preferences)
$P(\pi \mid \gamma) = \sigma(-\gamma \cdot G(\pi))$	Beliefs about policies
$P(\boldsymbol{A}) = Dir(a)$	Parameters of observation model
$P(\gamma) = \Gamma(1, \beta)$	Precision (inverse stochasticity in behaviour)

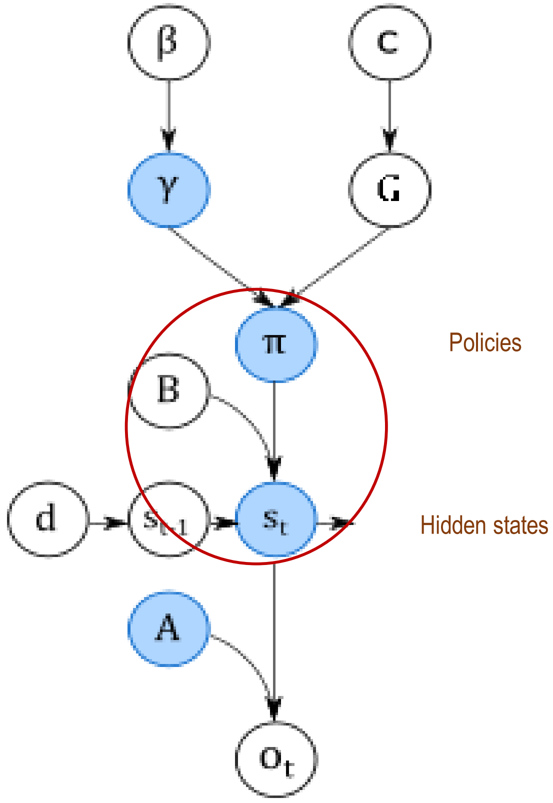




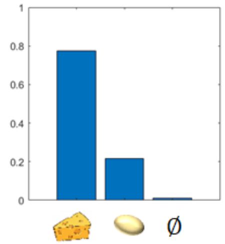
A Markovian generative model : transition probabilities

$P(o_t \mid s_t) =: \mathbf{A}$	Observation model: mapping from hidden states to observations
$P(s_{t+1} \mid s_t, \pi) =: \boldsymbol{B}(\pi(t))$	Transition probabilities: mapping from current to next states contingent on policies
$P(s_0) =: \boldsymbol{d}$	Beliefs about initial states in a task
$P(o_t) =: \boldsymbol{c}_t$	Beliefs about outcomes in a task (= preferences)
$P(\pi \mid \gamma) = \sigma(-\gamma \cdot G(\pi))$	Beliefs about policies
$P(\boldsymbol{A}) = Dir(a)$	Parameters of observation model
$P(\gamma) = \Gamma(1, \beta)$	Precision (inverse stochasticity in behaviour)

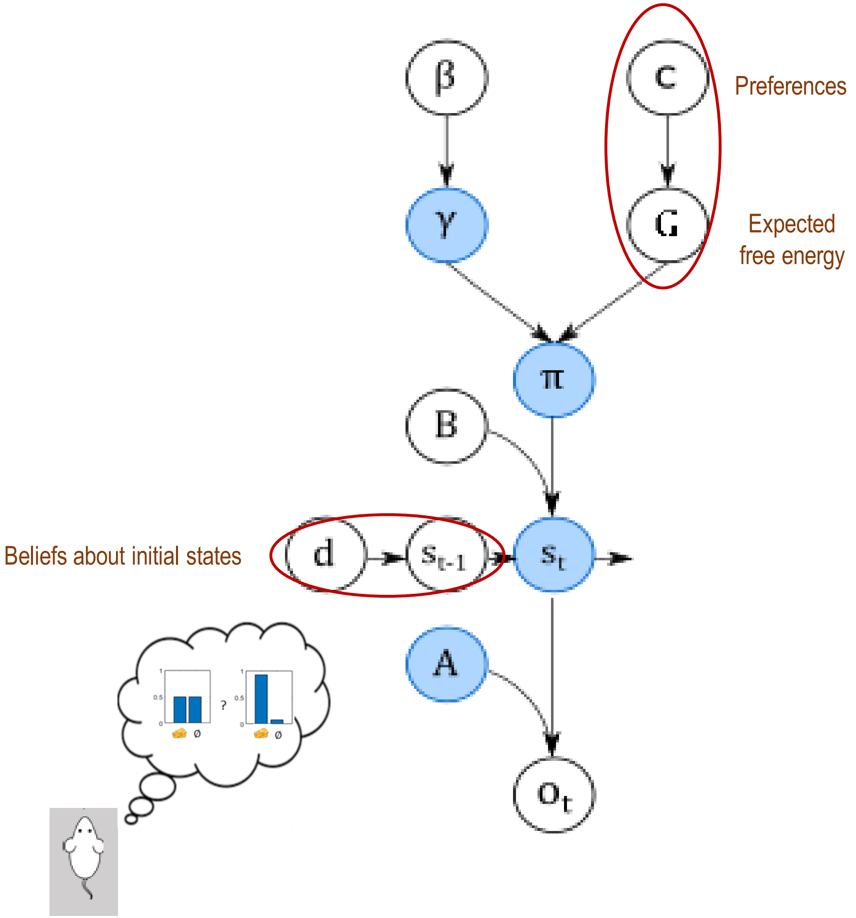




A Markovian generative model: priors about outcomes and states



$P(o_t \mid s_t) =: \mathbf{A}$	Observation model: mapping from hidden states to observations
$P(s_{t+1} \mid s_t, \pi) =: \boldsymbol{B}(\pi(t))$	Transition probabilities: mapping from current to next states contingent on policies
$P(s_0) =: \boldsymbol{d}$	Beliefs about initial states in a task
$P(o_t) =: \boldsymbol{c}_t$	Beliefs about outcomes in a task (= preferences)
$P(\pi \mid \gamma) = \sigma(-\gamma \cdot G(\pi))$	Beliefs about policies
$P(\boldsymbol{A}) = Dir(a)$	Parameters of observation model
$P(\gamma) = \Gamma(1, \beta)$	Precision (inverse stochasticity in behaviour)



A Markovian generative model: priors on observation model

$P(o_t|s_t) =: \mathbf{A}$ Observation model: mapping from hidden states to observations

$P(s_{t+1}|s_t, \pi) =: \boldsymbol{B}(\pi(t))$ Transition probabilities: mapping from current to next states contingent on policies

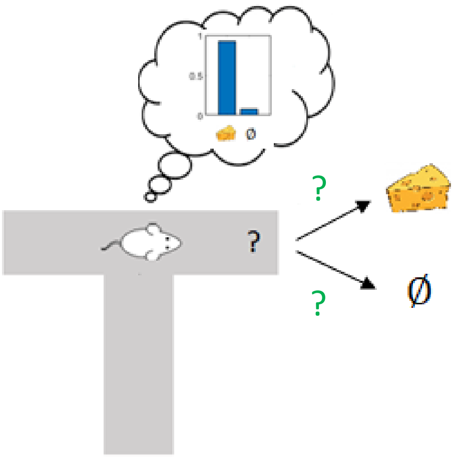
$P(s_0) =: \boldsymbol{d}$ Beliefs about initial states in a task

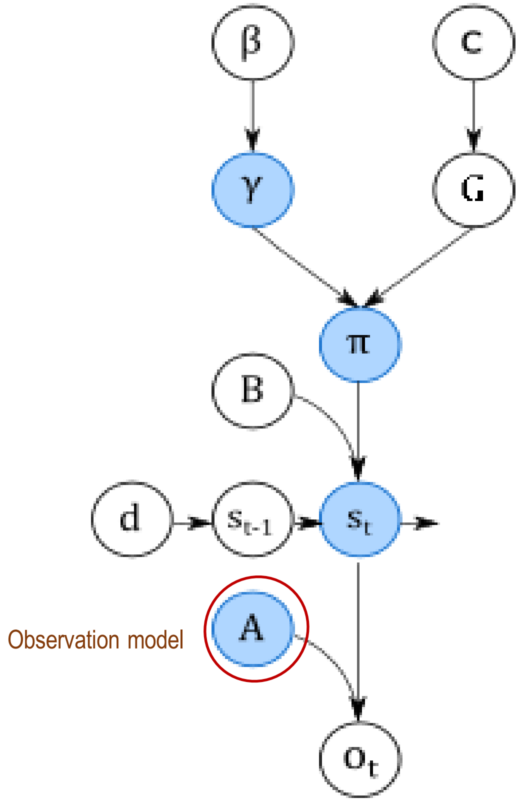
$P(o_t) =: \boldsymbol{c}_t$ Beliefs about outcomes in a task (= preferences)

$P(\pi|\gamma) = \sigma(-\gamma \cdot G(\pi))$ Beliefs about policies

$P(\boldsymbol{A}) = Dir(a)$ Parameters of observation model

$P(\gamma) = \Gamma(1, \beta)$ Precision (inverse stochasticity in behaviour)





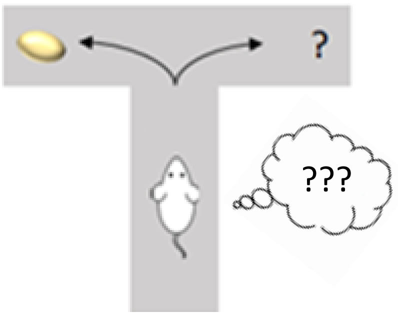
A Markovian generative model: precision

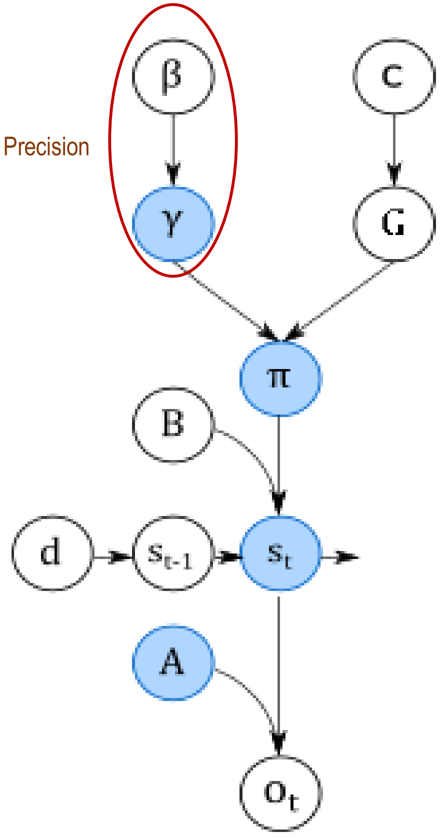
$P(o_t|s_t) =: \mathbf{A}$ — Observation model: mapping from hidden states to observations

$P(s_{t+1}|s_t, \pi) =: \boldsymbol{B}(\pi(t))$ — Transition probabilities: mapping from current to next states contingent on policies
$P(s_0) =: \boldsymbol{d}$ — Beliefs about initial states in a task
$P(o_t) =: \boldsymbol{c}_t$ — Beliefs about outcomes in a task (= preferences)

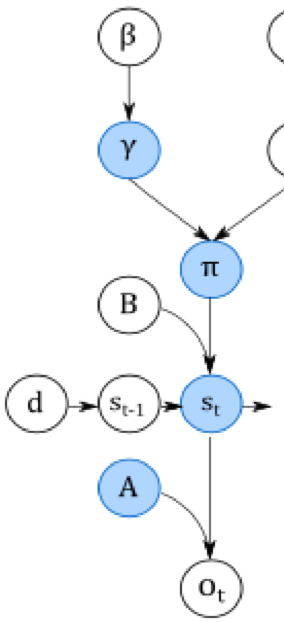
$P(\pi|\gamma) = \sigma(-\gamma \cdot G(\pi))$ — Beliefs about policies

$P(\boldsymbol{A}) = Dir(a)$ — Parameters of observation model
$P(\gamma) = \Gamma(1, \beta)$ — Precision (inverse stochasticity in behaviour)





(Variational) belief updating



$$F = \log p(o|m) - D_{KL}[q(s), p(s|o,m)]$$
$$\Leftrightarrow F = \sum_s q(s) \log \frac{q(s)}{p(o,s)}$$

Variational free energy F

$$Q(\hat{s}_t, \pi, A, \gamma) = \underset{Q(\hat{s}_t,\pi,A,\gamma)}{\operatorname{argmin}} F = P(\hat{s}_t, \pi, A, \gamma | o, m)$$

$$F = -E_Q[\ln P(o, s_t, \pi, A, \gamma | m)] - E_Q[\ln Q(s_t, \pi, A, \gamma)]$$

$$= \log A \cdot o_t + \log B^{\pi}_{t-1} \cdot \hat{s}_{t-1} + \log B^{\pi}_t \cdot \hat{s}_{t+1} + \cdots$$

Belief updating

Perception	$\hat{s}_t = \sigma(\log A \cdot o_t + \log B^{\pi}_{t-1} \cdot \hat{s}_{t-1} + \log B^{\pi}_t \cdot \hat{s}_{t+1})$	$\frac{\partial}{\partial_s} F = 0$
Policy selection	$\hat{\pi} = \sigma(-\gamma \cdot G)$	$\frac{\partial}{\partial_\pi} F = 0$
Precision	$\hat{\beta} = \beta - \hat{\pi} \cdot G$	$\frac{\partial}{\partial_\gamma} F = 0$

Approximate posterior

$Q(\tilde{s}, \pi, A, \gamma) = Q(s_1|\pi) \ldots Q(s_T|\pi) Q(\pi) Q(A) Q(\gamma)$

$Q(s_t|\pi) = Cat(s_t|\pi)$

$Q(\pi) = Cat(\boldsymbol{\pi})$

$Q(\boldsymbol{A}) = Dir(\boldsymbol{a})$

$Q(\gamma) = \Gamma(1, \boldsymbol{\beta})$

Friston et al., 2015, *Cognitive Neuroscience*;
Friston et al., 2017, *Neural Computation*;
Bogacz, 2017, *Journal of Mathematical Psychology*

(Variational) inference

Belief updating

Perception

$$\hat{s}_t = \sigma(\log A \cdot o_t + \log B^{\pi}_{t-1} \cdot \hat{s}_{t-1} + \log B^{\pi}_t \cdot \hat{s}_{t+1})$$

"Infer the **current state** based on your **observations** and beliefs about **transitions between states**"

Policy selection

$$\hat{\pi} = \sigma(-\gamma \cdot G)$$

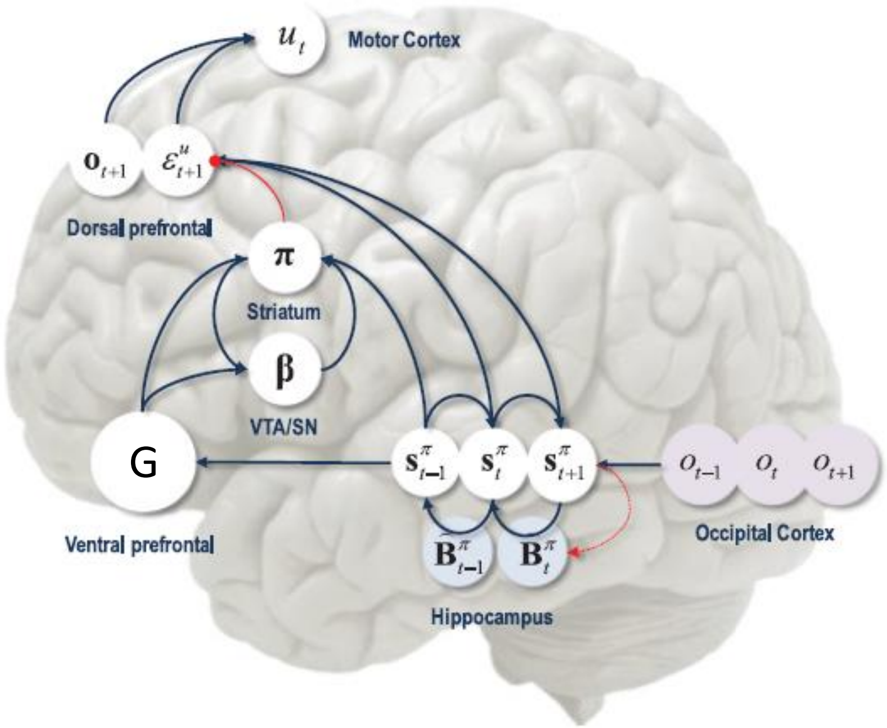
"Infer the **best policy** based on the **value of policies** and your **goal-directedness**"

Precision

$$\hat{\beta} = \beta + (\hat{\pi} - \pi_0) \cdot G$$

"Infer the **right level of goal-directedness** based on a **prediction error** between prior and posterior **expected free energies**"

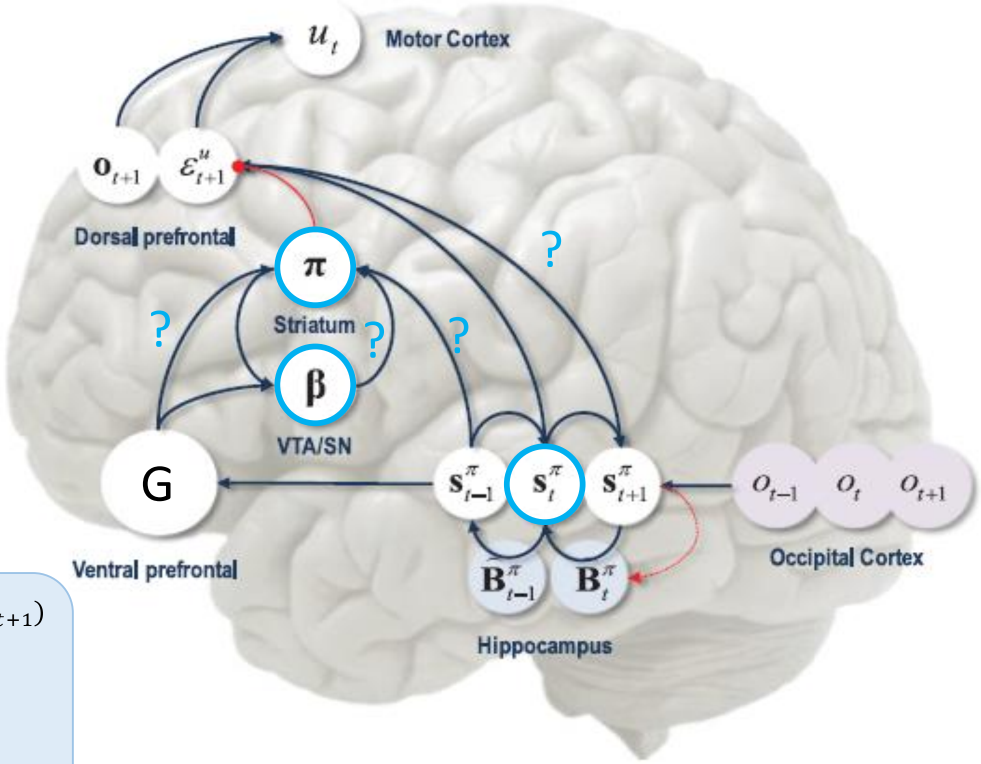
Friston et al., 2017, *Neural Computation*



Makes predictions for neuronal implementation, e.g. different message passing schemes (Parr, Markovic, Kiebel & Friston, 2018)

(Active) learning

We use generative models to perform inference on hidden states.



But we also learn and update our models.

Belief updating

Perception	$\hat{s}_t = \sigma(\log A \cdot o_t + \log B^\pi_{t-1} \cdot \hat{s}_{t-1} + \log B^\pi_t \cdot \hat{s}_{t+1})$
Policy selection	$\hat{\pi} = \sigma(-\gamma \cdot G)$
Precision	$\hat{\beta} = \beta - \hat{\pi} \cdot G$

Learning

Observation model $P(A) = Dir(a)$

where

$$a_t = a_{t-1} + \eta \cdot \sum_t o_t \otimes s_t$$

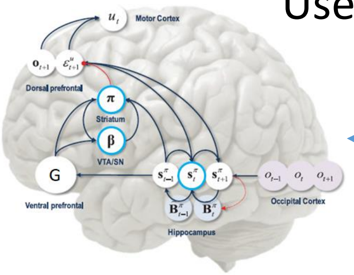
Counting number of transitions from one particular hidden state to a particular outcome

- Modulated by learning rate η

Friston et al., 2017, *Neural Computation*

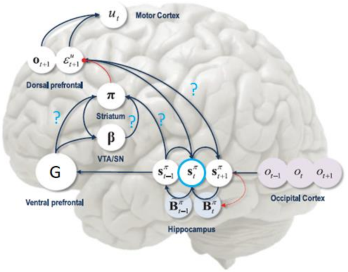
Variational inference applied to POMDPs predicts specific types of belief updating and learning

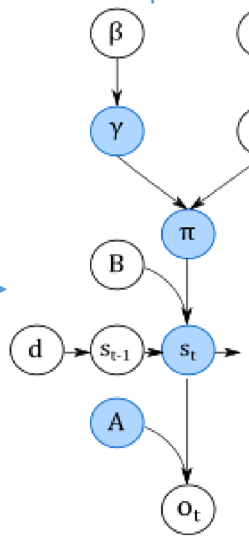
Active inference and active learning summary



Use this to understand **inference**…

…and **learning**





Need to define **variational free energy** wrt

- Future states and observations
- Contingent on policies

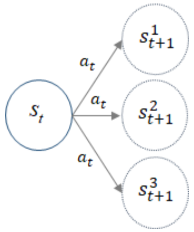
$$G = \sum_s q(s_t|\pi) \log \frac{q(s_t|\pi)}{p(o_t, s_t|\pi)}$$

Expected free energy G

Let's use **variational inference** during **action** and **planning**

$$F = \log p(o|m) - D_{KL}[q(s), p(s|o,m)]$$

Apply to a **POMDP**



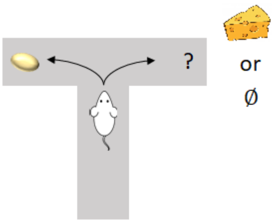
III. Some interesting predictions

What makes an action valuable?

Actions become valuable if they

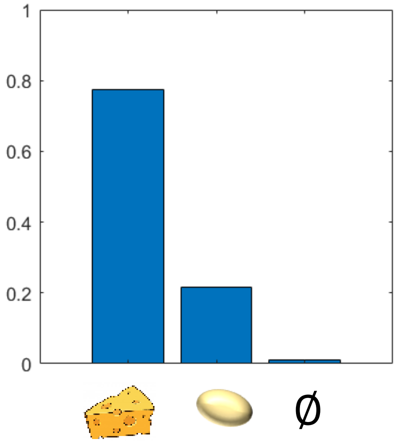
- Maximise **reward/utility**
- Allow us to minimise **uncertainty** about the world
  - Uncertainty about **states** (active inference)
  - Uncertainty about **models** of the world (active learning)

Utilities (preferences) are defined as log-expectations over outcomes:



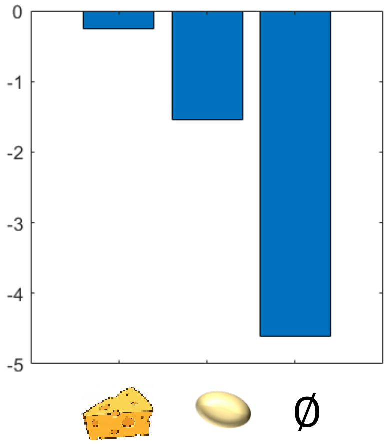
$P(o_t|m)$

Preferences are probabilities



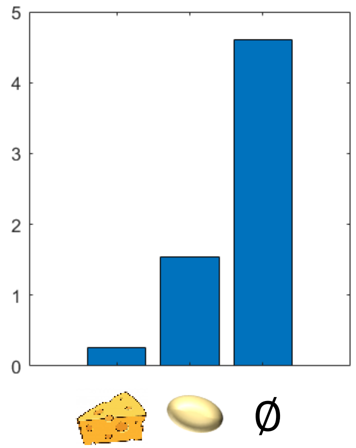
$\log P(o_t|m)$ $\longrightarrow$

Turn this into model evidence



$-\log P(o_t|m)$ $\longrightarrow$

Turn this into surprise



Fulfilling these preferences minimises surprise $-\log P(o_t|m)$!

- This can be approximated with variational free energy!

What makes an action valuable?

Values of policies $\boldsymbol{G(\pi)}$ defined as expected free energy:

$$P(\pi) = \sigma\big(-\gamma \cdot G(\pi)\big)$$

$$G(\pi) = \sum\nolimits_{\tau} G(\pi, \tau)$$

$$\begin{aligned} G(\pi, \tau) &= \mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau, A|\pi) - \ln P(o_\tau, s_\tau, A|\pi)] \\ &= \mathbb{E}_{\tilde{Q}}[\ln Q(A) + \ln Q(s_\tau|\pi) - \ln P(A|s_\tau, o_\tau, \pi) - \ln P(s_\tau|o_\tau, \pi) - \ln P(o_\tau)] \\ &\approx \mathbb{E}_{\tilde{Q}}[\ln Q(A) + \ln Q(s_\tau|\pi) - \ln Q(A|s_\tau, o_\tau, \pi) - \ln Q(s_\tau|o_\tau, \pi) - \ln P(o_\tau)] \end{aligned}$$

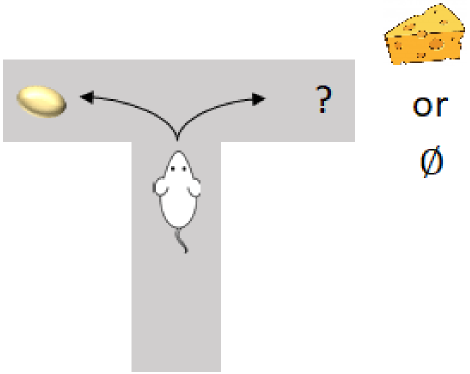
$$G(\pi, \tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A) - \ln Q(A|s_\tau, o_\tau, \pi)]}_{\substack{\text{‘Model exploration’} \\ \text{(Active learning)}}} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau, \pi)]}_{\substack{\text{‘Hidden state exploration’} \\ \text{(Active Inference)}}} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{\substack{\text{Realising preferences} \\ \text{(Exploitation)}}}$$

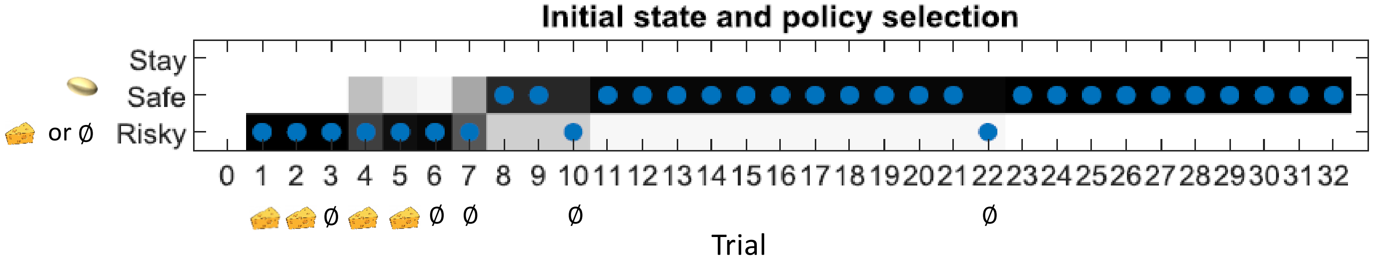
Mutual information

Expectations over outcomes

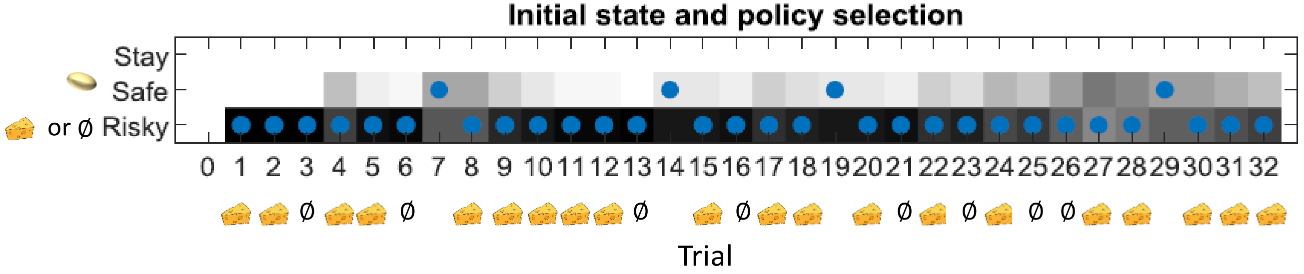
see Friston et al., 2017 *Neural Computation;* Parr & Friston, 2017 *Cerebral Cortex;* Schwartenbeck et al. 2019, *eLife;* Solopchuck 2019 *Tutorial on Active Inference* and *Free Energy, Action Value and Curiosity*

Active learning: ‘model exploration’





Prob high reward = 0.5

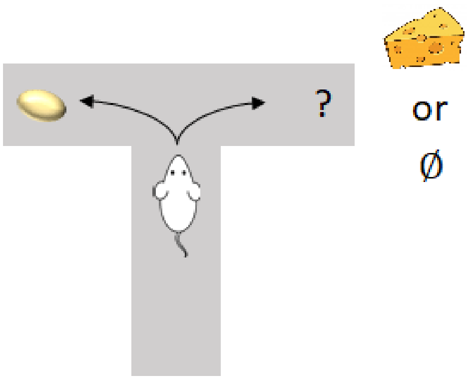


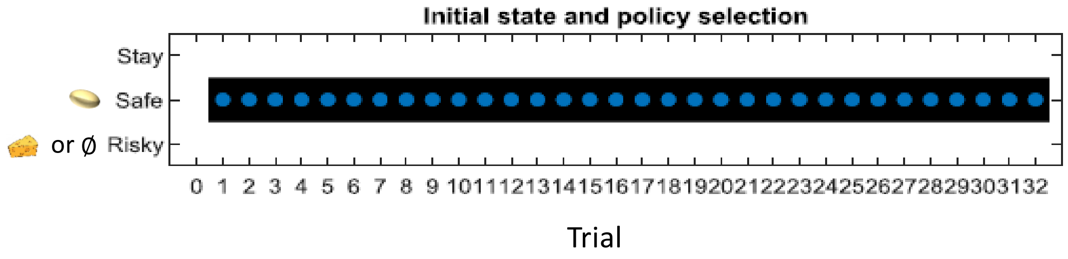
Prob high reward = 0.75

$$G(\pi,\tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A) - \ln Q(A|s_\tau, o_\tau, \pi)]}_{\substack{\textbf{‘Model exploration’} \\ \textbf{(Active learning)}}} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau,\pi)]}_{\substack{\text{‘Hidden state exploration’} \\ \text{(Active Inference)}}} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{\substack{\text{Realising preferences} \\ \text{(Exploitation)}}}$$

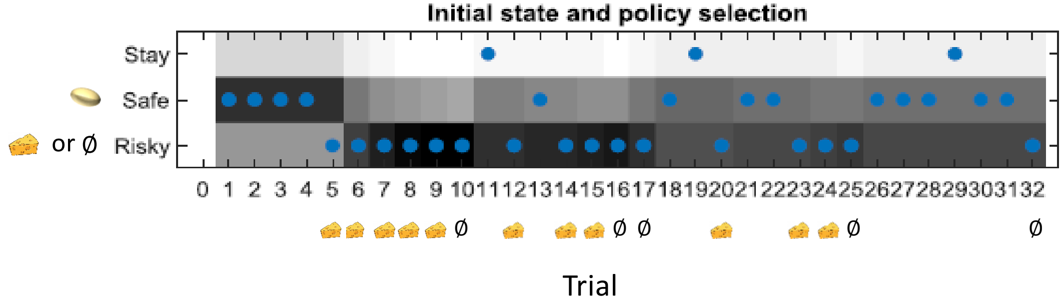
Schwartenbeck et al., 2019, eLife

‘Broken’ active learning





Prob high reward = 0.75



Prob high reward = 0.75

$$G(\pi,\tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A \quad |s_\tau, o_\tau, \pi)]}_{} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau,\pi)]}_{} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{}$$

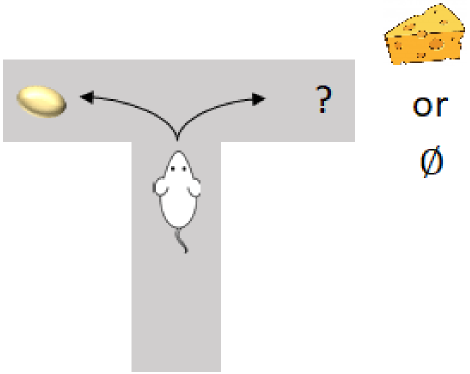
‘Mo ion’
(Active learning)

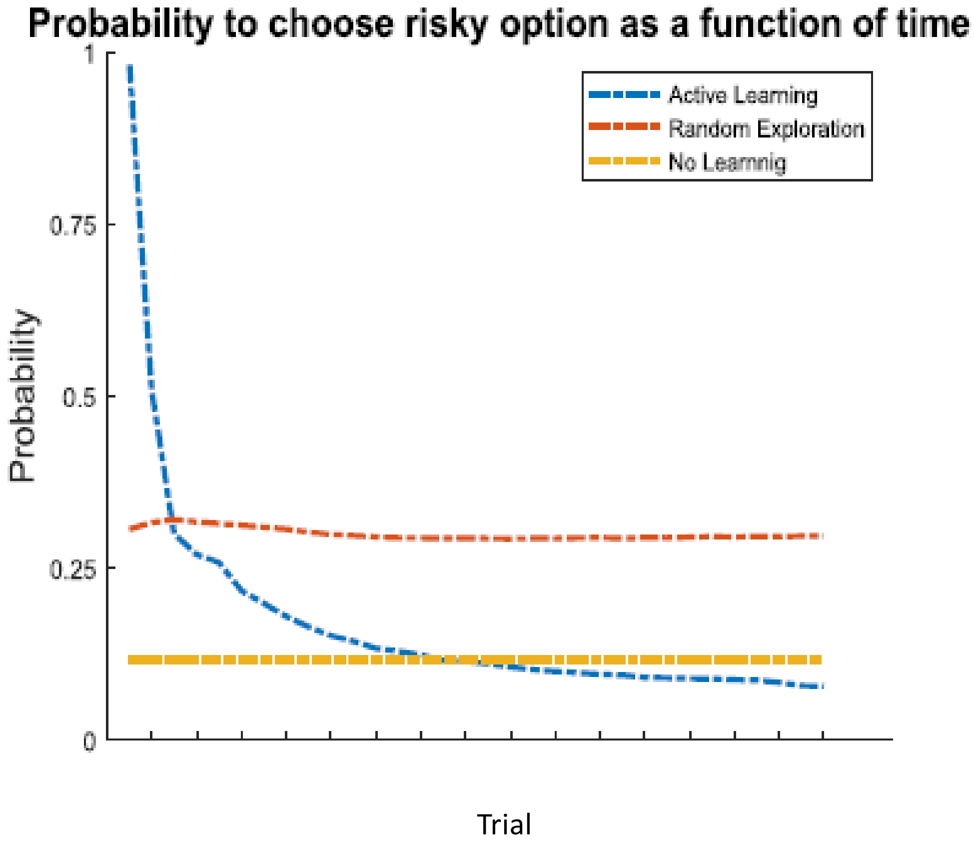
‘Hidden state exploration’
(Active Inference)

Realising preferences
(Exploitation)

Schwartenbeck et al., 2019, eLife

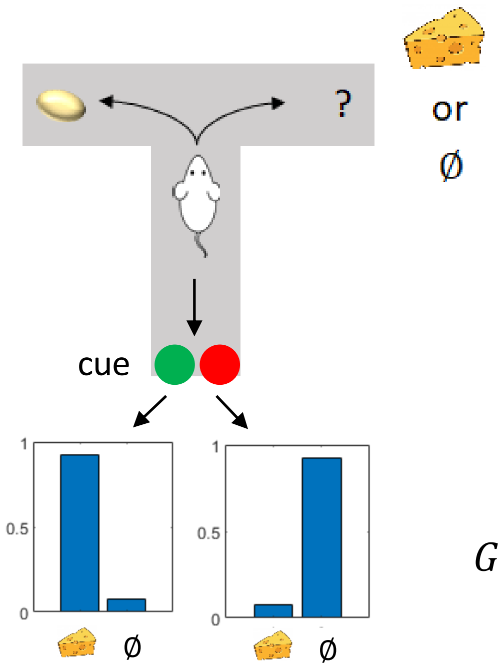
Active learning: 'model exploration'

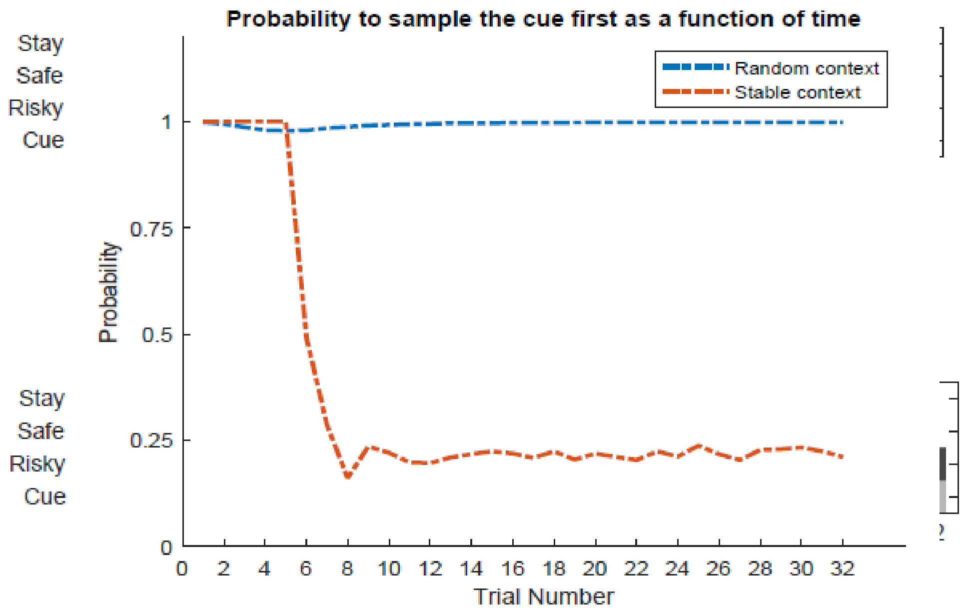




Schwartenbeck et al., 2019, eLife

Active Inference: ‘hidden sate exploration’





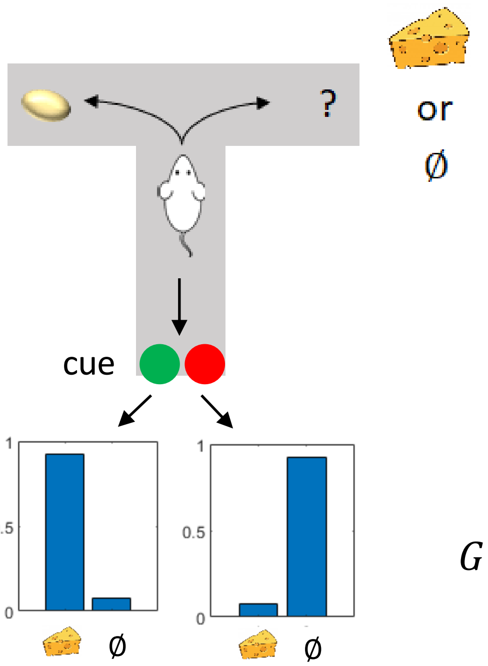
Prob high reward changes trial by trial

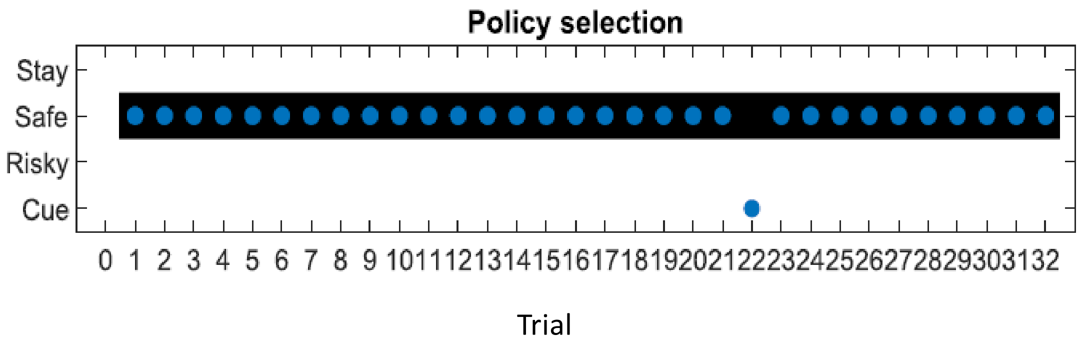
Prob high reward = 0.75 (stable)

$$G(\pi,\tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A) - \ln Q(A|s_\tau,o_\tau,\pi)]}_{\substack{\text{'Model exploration'} \\ \text{(Active learning)}}} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau,\pi)]}_{\substack{\textbf{'Hidden state exploration'} \\ \textbf{(Active Inference)}}} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{\substack{\text{Realising preferences} \\ \text{(Exploitation)}}}$$

Schwartenbeck et al., 2019, eLife

‘Broken’ active Inference

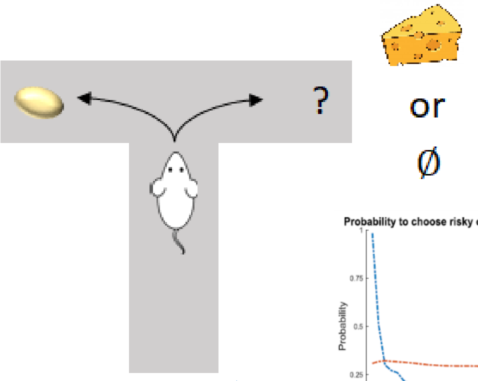


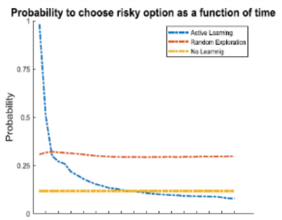


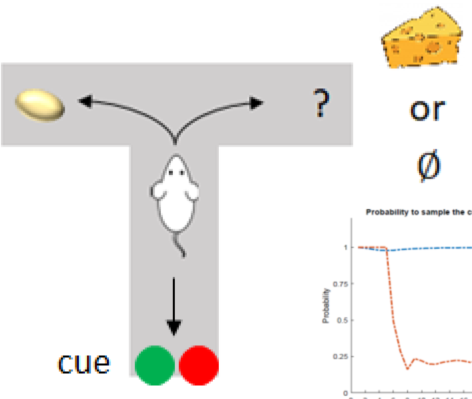
$$G(\pi,\tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A) - \ln Q(A|s_\tau,o_\tau,\pi)]}_{\text{‘Model exploration’ (Active learning)}} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau \ldots (s_\tau|o_\tau,\pi)]}_{\textbf{‘Hidde … ration’ (Active Inference)}} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{\text{Realising preferences (Exploitation)}}$$

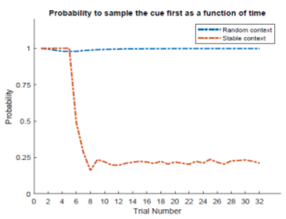
Schwartenbeck et al., 2019, eLife

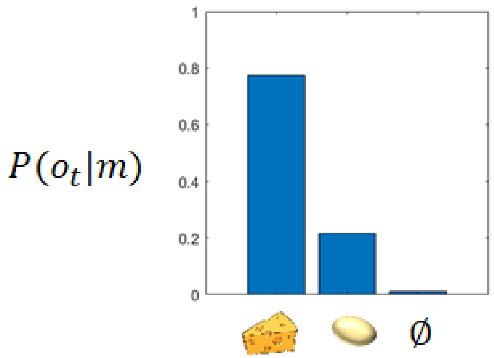
Value of actions summary











About goal-directed sampling of states from which we can *infer* (**active inference**)

About goal-directed sampling of observations from which we can *learn* (**active learning**)

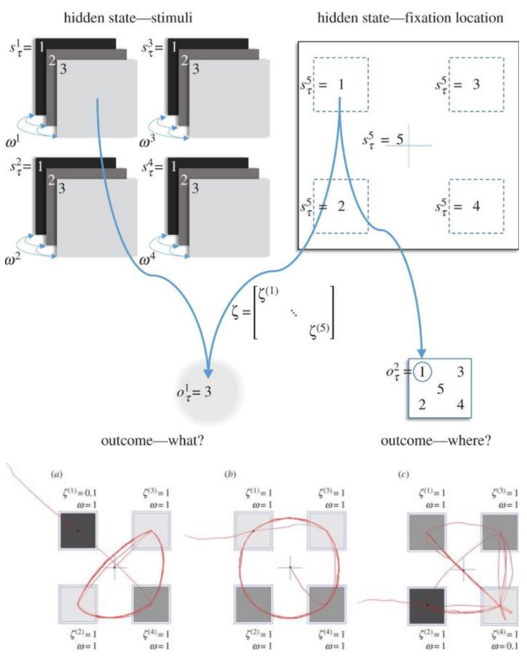
Whilst ensuring that preferences (expectations over outcomes) are fulfilled (**exploitation**)

This makes interesting predictions

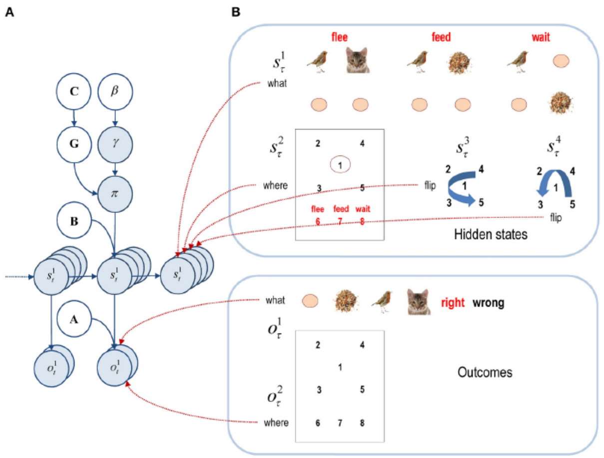
$$G(\pi,\tau) = \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(A) - \ln Q(A|s_\tau, o_\tau, \pi)]}_{\text{'Model exploration' (Active learning)}} + \underbrace{\mathbb{E}_{\tilde{Q}}[\ln Q(s_\tau|\pi) - \ln Q(s_\tau|o_\tau, \pi)]}_{\text{'Hidden state exploration' (Active Inference)}} - \underbrace{\mathbb{E}_{\tilde{Q}}[\ln P(o_\tau)]}_{\text{Realising preferences (Exploitation)}}$$

Let's use **variational free energy** during **action** and **planning**

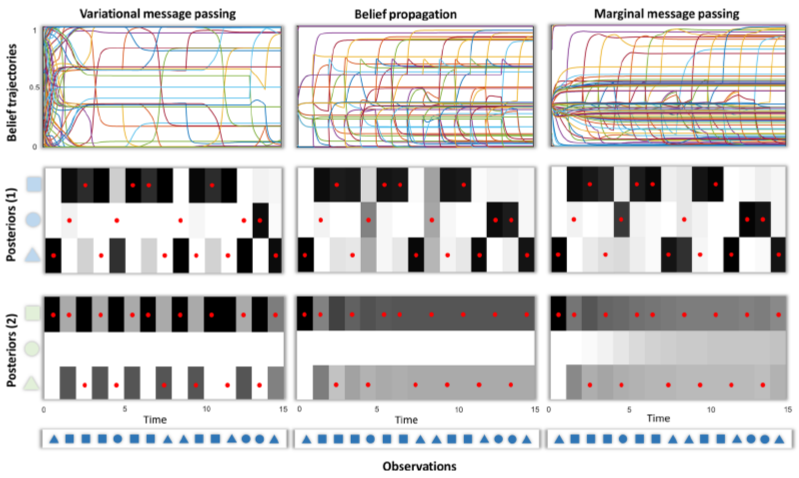
This can be applied in lots of other (more relevant) contexts



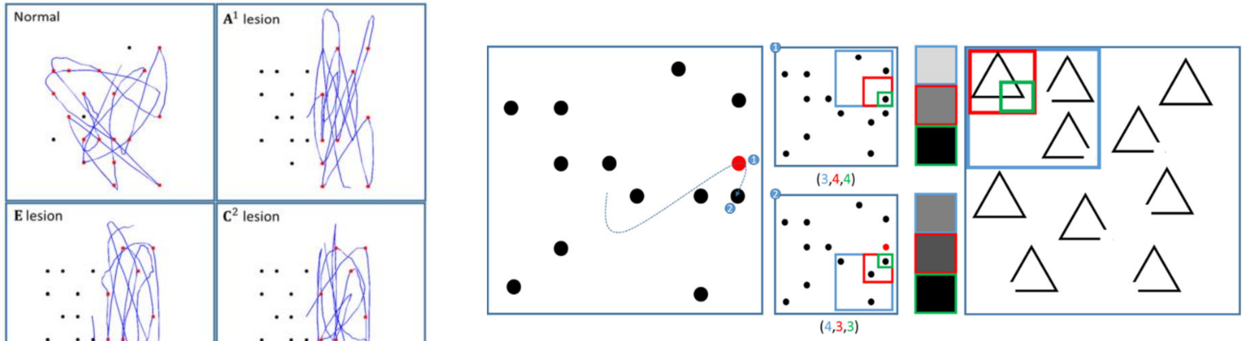
Modelling ‘epistemic’ (foraging) saccades based on a factorised MDP with different levels of precision (Parr & Friston, 2017)



Saccades and scene construction (Mirza & Adams, Mathys & Friston, 2016)



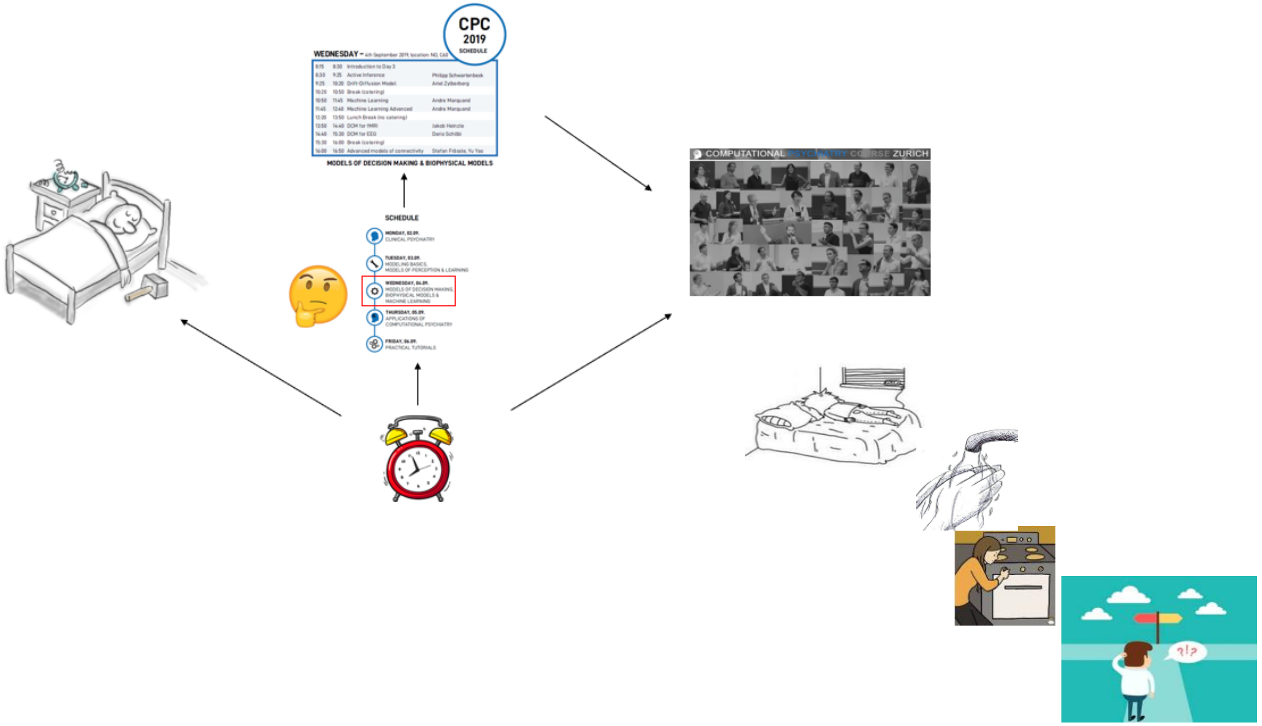
Different message passing schemes underlying active inference (Parr, Markovic, Kiebel & Friston, 2019)



Use this to understand failures in epistemic foraging, such as visual hemineglect (Parr & Friston, 2019)

Computational Phenotyping in active inference

All models are wrong, but some are useful - for understanding how things can break



Active inference/learning provides a tool for understanding

Failures in **active inference**

- Lack of insight into informative states

Failures in **active learning**

- Lack of insight into observations that can reduce uncertainty

Suboptimal **preferences**

Suboptimal **model configurations**, i.e.

- Suboptimal action sequences
- Wrong/noisy observation model, transition probs, …

Failures of **precision**

- Of action selection
- Of different aspects of value

Other introductory resources

Blog by Oleg Solopchuk:

- Tutorial on Active Inference (https://medium.com/@solopchuk/tutorial-on-active-inference-30edcf50f5dc)
- Free Energy, Action Value and Curiosity (https://medium.com/@solopchuk/free-energy-action-value-and-curiosity-514097bccc02)

"What does the free energy principle tell us about the brain?" by Sam Gershman (*arXiv*, 2019)

"A tutorial on the free-energy framework for modelling perception and learning" by Rafal Bogacz (*Journal of Mathematical Psychology*, 2017)

"The free energy principle for action and perception: A mathematical review" by Buckley, Kim, McGregor & Seth (*Journal of Mathematical Psychology*, 2017)

"Combining active inference and hierarchical predictive coding: a tutorial introduction and case study" by Beren Millidge (*PsyArXiv*, 2019)

Take home messages

I. Active inference applies **variational inference** to **Markov Decision Processes**
   - Central idea: actions fulfil expectations ⇔ minimise surprise ⇔ maximise model evidence

II. This predicts
   - **Inference** on the current state, policy and goal-directedness
   - **Learning** (model optimisation) of the observation model, transition probabilities, …

III. Defining the value of policies as **expected free energy** over future observations predicts
   - Exploration of hidden states
   - Exploration of 'model parameters'
   - Exploitation (fulfilling preferences)

IV. Provides a computational framework for *active inference* and *active learning* - and how this might break
   - Failures of inference, learning, preferences, precision, …

Thank you

Collaborators

Tim Behrens
Ray Dolan
Karl Friston

Rick Adams
Alon Baram
Ben Castine
Tom FitzGerald
Daniel Freinhofer
Dorothea Hammerer
Tobias Hauser
Lilla Horvath
Jakob Howy
Martin Kronbichler
Zeb Kurth-Nelson
Yunzhe Liu
Andrei Manoliu
Christoph Mathys
Shirley Mark
Matthew Nour
Thomas Parr
Johannes Passecker
Giovanni Pezzulo
Natalie Rens
Nitzan Shahar
Ryan Smith