Modelling **Meta-awareness** & **Attentional control** using **deep parametric active inference** framework[1]

–

Modeling in Neuroscience & elsewhere
Supervised by Jean-Pierre Nadal

Romy BEAUTÉ

ENS Paris-Saclay & Université de Paris, France

Master MVA • 2021-2022

école normale supérieure paris—saclay | université PARIS-SACLAY

[1] Sandved-Smith et al., 2021

Table of content

Idea : Propose a formal model of **awareness** & **attentional control** through **mental action**

Model particularities :

1. *relies on* **active inference** framework
2. *casts* mental action [a] as **policy selection** over higher-level cognitive states
3. *adds* hierarchical level : **model meta-awareness**

   $\Longrightarrow$ *modulates precision*[b] *mapping between observations & cognitive processes*

[a] goal-directed internal behavior

[b] expected confidence

Simulation of cycle of attention & its regulation during an attentional task

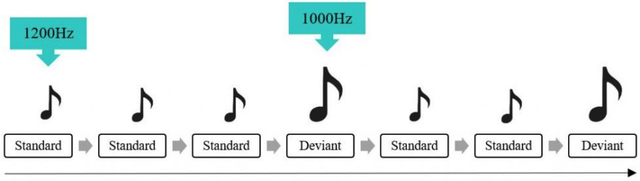


Figure 1: Example of auditory oddball paradigm task

Task : Detection deviant stimuli

⇒ involves **sustained selective attention** on a perceptual object

Table of content

Ability to '**explicitly**' notice current content of conscious episode.

Enable **effective self-regulation** *(of attention)* by providing :

- Knowledge about objects & processes of experience **(ACCESS)**
- Revision of beliefs **(EVALUATION)**
- **Increased control over cognitive processes (CONTROL)**

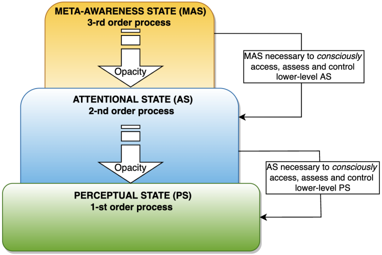


Figure 2: The 3 states as hierarchical levels. Higher order levels enable process to become aware of lower levels, i.e **'opaque'** to the system.

State accessibility

The capacity of a system to access some subset of its own states has been theorized under the rubric of 'opacity' versus 'transparency' [2].

Mental states can be decomposed into :

- **Opaque states** : Cognitive processes constructing these states **perceptible** as such, available to awareness.
- **Transparent states** : Cognitive processes constructing these states **imperceptible** as such. Only aware of the content they make available.

[2] *Metzinger, T. (2003). Phenomenal transparency and cognitive self-reference*

1. Conscious mental action is **grounded in higher-level access to cognitive states**
2. Distinction between opaque *vs* transparent states & processes can be formalized under **deep active inference** framework
3. **Understanding inferential architecture** $\Rightarrow$ advanced for formal, computational neurophenomenological account of **cognitive control**

1. Explicitly **model meta-awareness** (MA) states (*as a 3rd order level*)
2. Account for the computational & phenomenal consequences of MA and attentional states on cognitive control
3. Provide a computational account of mental (*covert*) action
4. Extend the work from Sandved-Smith et al., (2021) to simulate ADHD condition

Table of content

Computational architecture

Agent seeks to **control/regulate their mental states** by becoming **aware of the state as a cognitive process (opacity [3])**

⇒ **need notion of confidence/reliability** in effective self regulation :

Precision

Quantifies **confidence/reliability** of *implicit (posterior Bayesian) beliefs :*

- knowledge about states of the world
- their relation to sensory observations
- how they change over time

[3] gives access new source of data, allowing for **further processing** & **action over** state

Goals

1. Implementation self-reflective, hierarchically structured inferences
2. Simulate attentional control dynamics linked with :
   - cognitive opacity
   - mental action selection

Implementation : **parametrically deep active inference framework**

1. **Active inference**
   - neurocognitive & behavioral modeling framework
   - knowledge driven action & perception
   - adaptation in dynamical & embodied environment

   $\Rightarrow$ **gather evidences for embodied and enactive generative model**

2. **Parametric depth**
   - cognitive architecture property
   - nested beliefs/knowledge structures
   - inferential results as data for further processings

   $\Rightarrow$ **access to self-states & meta awareness**

Offers a unified mathematical framework for modeling cognitive processes (perception, learning, decision making)

⇒ treats each of these psychological processes, and their interactions, as **interdependent forms of inference** :

Idea : cognitive agents are assumed to infer the probability of different external states and events in the environment

The inferences are :

- Created by combining **prior beliefs** with sensory (internal or external) inputs
- **Active** : the agent infers the actions most likely to generate preferred sensory input

⇒ Leads to preference towards actions optimizing trade-off between reward and information gain

Based on 2 concepts

- Agent actively engage with their environments to gather information, seek out 'preferred' observations, and avoid non-desired observations
- **Bayesian inference** : statistical procedure describing the optimal way to update one's beliefs (i.e probability distributions) when making new observations (i.e new sensory input) based on the rules of probability

$\Rightarrow$ Beliefs updated in light of observations using Baye's theorem : Variational inference : find the posterior distribution $p(s|o)$ to infer how the states s of the world have changed based on new observations o.This corresponds to finding the posterior distribution $p(s|o)$. However, the calculation of the marginal likelihood $p(o)$ is

Active inference agents : must deploy inference and action to guesstimate their generative processes and structure

Generative model : probabilistic model of the process that generated this (sensory) data [4]

$\Rightarrow$ *capture dynamics of how the environment generates observations.*

Variational free energy

Quantifies :

- how much evidence for the model is provided by the data
- discrepancy between **observed outcomes** (i.e. the sensory states of organism) & **outcomes predicted** under statistical (generative) model

[4] Friston et al. 2016

Parameters :

- D : **initial state vector** $p(S_1)$
  $\Rightarrow$ Encodes **prior beliefs about initial hidden states** S_t
- S : **Hidden state**
  $\Rightarrow$ causing O
- A : **likelihood mapping** $p(O_t|S_t)$
  $\Rightarrow$ Encodes beliefs about the **relationship hidden states** (S_t) **and observable outcomes** (O_t)
- O : **Observable outcomes**
  $\Rightarrow$ caused by S
- $\bar{S}$: **Approximated most probable state** $p(S|\bar{o})$.

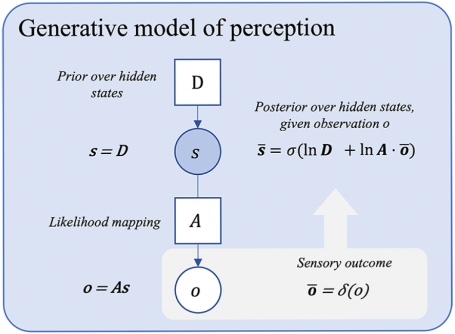


Figure 3: Basic generative model of perception.

Implements beliefs about **state transitions**, noted B
$\Rightarrow$ controls observations through selection of actions (policy selection).

Additional parameters :

- C : **prior preference mapping** $p(S_1)$
  $\Rightarrow$ prior beliefs about sensory outcomes $\hat{o}$
- E : **prior over policies**
- B **(matrices): serries of beliefs about state transitions**
- π : **policy selection**
  $\Rightarrow$ selecting the sequence of B associated with least expected free energy $min(G)$ *(i.e minimizes 'ambiguity' and 'risk')*
- G : **expected free energy.**

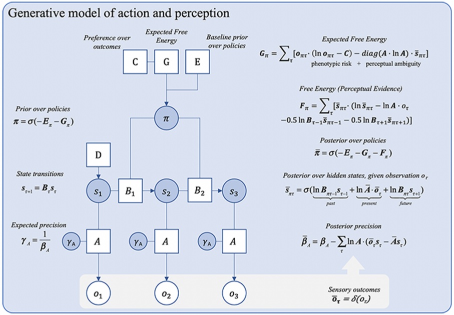


Figure 4: Generative model of action & perception.

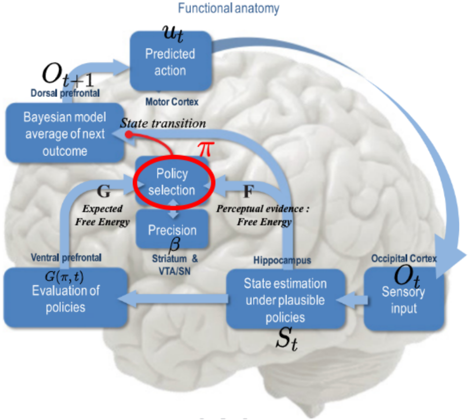


Figure 5: Overview of belief updates for discrete Markovian models and processes assignment to functional anatomy.[5] Sensory evidence (O_t) is accumulated to optimise expectations about the current state (S_t), which are constrained by expectations of past (S_{t-1}) & future (S_{t+1}) states. Corresponds to state estimation under each policy the agent entertainments. Once an action has been selected, it generates a new observation (O_{t+1}) and the cycle begins again.

[5] adapted from (Friston et al., 2016)

Variational free energy quantifies the discrepancy between observed outcomes (i.e. the sensory states of an organism) and the outcomes that would be predicted under a statistical (generative) model of how their sensory data were produced

Idea I : AI assumes that perception and learning can be understood as minimizing a quantity known as **Variational free energy (FE)**

Idea II : The role of action selection (policy) is to minimize the **expected free energy (EFE)**

$\Rightarrow$ EFE quantifies the VFE of various actions based on expected future outcomes

$\Rightarrow$ Use Bayesian inference framework to use and derive these quantities

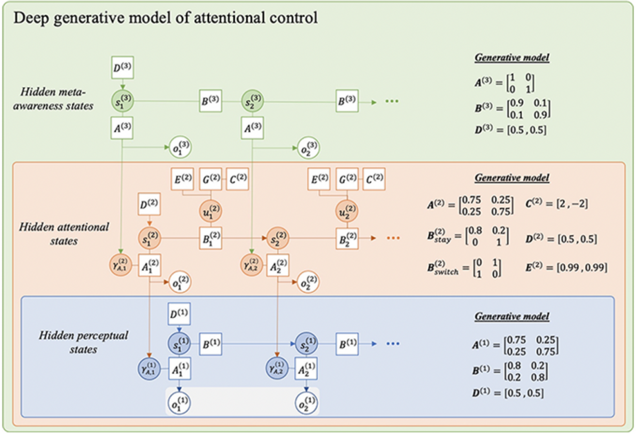
Table of content

Limitation active inference :
Do not model agent's access, evaluation & control of its own attentional states (i.e no meta-awareness)

Extension

- deep, higher-level hierarchical architecture
- higher level policy selection over attentional states
- parametric depth : precision-dependant state inference

Probabilistic Graphical Model of mental action : Active Inference architecture



- **First-order level** : perception of the external environment $S^{(1)}$
- **Second-order level** : perception of internal attentional states $S^{(2)}$
- **Third-order level** : perception of meta-awareness states $S^{(3)}$

Influence of attentional states on lower-order (perceptual) states - non-ADHD condition

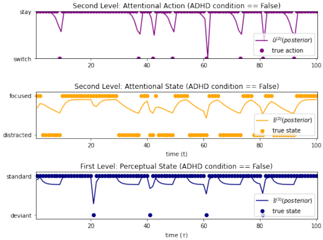


Figure 6: Replication in non-ADHD state. Attentional states during the first and second part of the experiment have been respectively set to "focus" and "distracted".

Influence of attentional states on lower-order (perceptual) states - non-ADHD condition

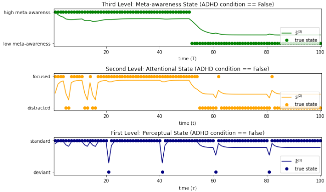


Figure 7: Replication in non-ADHD condition. Implementation of the 3-rd state level (meta-awareness level)

Influence of attentional states on lower-order (perceptual) states - ADHD condition

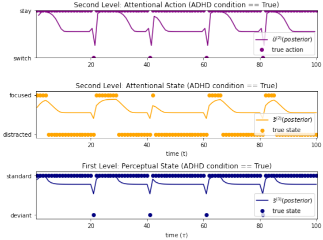


Figure 8: Replication in ADHD state. Attentional states during the first and second part of the experiment have been respectively set to "focus" and "distracted".

Influence of attentional states on lower-order (perceptual) states - ADHD condition

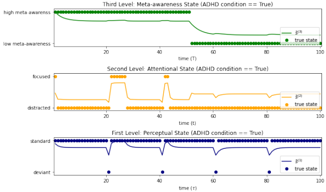


Figure 9: Replication in ADHD condition. Implementation of the 3-rd state level (meta-awareness level)

Table of content

Control of cognitive (attentional) state through mental action involved in psychiatric conditions.

- ADHD
- Schizophrenia (Brown et al., 2013)
- Autism spectrum (Kiverstein et al., 2020)

Perceptual, mood, thinking dysfunctions :

$\Rightarrow$ **Atypical inferences** about higher-order states[a] control

$\Rightarrow$ **Erroneous control** associated processes

[a] Higher-order : govern attribution of lower-level (eg perception) confidence

Computational implementation of **inferential architecture** enabling emergence of ability to access & control cognitive states.

- Interesting model to **formalize specific phenomenal properties** (*here transparency and opacity of cognitive states, with respect to attention, and meta-cognition states*)
- Framework presents as an interesting avenue to shed lights on the **influence of priors in behavioral outcome**.
- Particularly interesting for areas of research that are interested in studying **variational properties** of universal and individual features (e.g metacognitive skills) across states, cultures, conditions

Thank you for your attention!

romy.beaute@ens.psl.eu

Github of this project available here