Active Inference for Dummies
(translating Friston)

We observe a phenomenon in our environment, and we take action onto it. The explanation how we make a decision to take that exact action, is what active inference, among other behavioural, neuroscience theories, tries to explain.

Principle of Active Inference.

We have an implicit model of the world; the more experiences or observations **o** we had in our past, the more our model **m** is rich. In that way we can recognize the new situation (observation) and apply the actions that worked out best in the past.

These possible actions (linked to a certain observation) that we can have in our "memory", are called **control states**, **U,** and represent all the possible future actions we can choose to conduct in reality. So, the actions we applied to the world in the past for a certain observation enter our "memory" with a degree of success. We often tend to use the same action or behavior which had success in the past. This success assigns a high value to that action (arranged in a **vector c**, explained later). So using this high-value action also maximizes gain or utility.

Also, we wish to minimize the difference of what we **expect** to happen as a result of our possible action (control state) and the **true** result of our real action. We wish to minimize the surprise of an event and maximize our gain. We are beings who want to control our environment, to control the effect of our own behavior (influence) on the world, but also successfully reach our goals. In order to avoid surprise we must develop a way to predict the consequences of our actions.

However if it is a completely new situation (observation), it becomes harder to act (almost random) but we try to infer the possible best actions from our model of the world, and predict the outcome (observations) -- the richer the given model m, the smaller the proba of being surprised. This minimization of surprise uses (also referred to as) a free-energy minimization approach.

This is a closed-loop system, where the action we just took becomes a cause of the following observation, which in turn causes us to make another action, while storing the previous action in our "memory", becoming part of the **empirical priors**.

Model of the world

What is interesting is the model **m** of the world we have, which is a generative model (from our empirical priors, we generate an action, by iterating through all possible ones). I have already mentioned the **control states, U** (a subset of all possible actions, linked or limited to a certain observation). It is not all, there is something called the **hidden states S** that also take part in the generative model. Why are they hidden? Because they represent our interpretation of the true state of the world, which is in fact hidden to us.

So, the world is in fact in a state, a true state, and we try to infer it.

What we know for sure is what we can observe, so from there we can infer the possible true state of the world, representing it as the hidden states and applying possible control states we have in our “ memory storage”. Both control and hidden states are called the empirical priors.

We not only have the **empirical priors** in our model of the world, but also the **full priors** which define our goal and its importance (how far are we willing to go to get it). They assist in choosing a specific control state or action when in a current hidden state. In formal words, these parameters define the **precision of beliefs about control**. One of them is the prior distribution over last state $\mathbf{s}_{t-1}$, **(a vector c).** It represents the softmax function of control states (revealing the **expected utility**!). This simply means, that we will favorize to some degree one action more over the others, in order to get the expected outcome (gain). So, the vector c is a full prior over the desired outcomes. Examples of argmax will favour (maximize) one possible outcome while penalizing all the others. Our gain depends on these **full priors**, also on the ones which will control the exploration-exploitation trade-off, arranged in ***gamma*** (e.g. favouring precise over rapid actions), it is called gamma because it uses a gamma distribution.

The probability that from a state $\mathbf{s}_1$ we will use a control state $\mathbf{u}_i$, and pass to another state $\mathbf{s}_2$, is regulated by the **transition matrix B**, {state1, possible_action1, state2}, which considers all the priors, in order to maximize gain. This is a typical Markov Decision Process example (state, action, state) following a policy (a cost of transition).

Bayesian model

The Bayesian theory only refers to the posterior probability prediction, i.e. inferring the next future observation by taking knowledge from the past (the empirical priors) or the full priors.

A very simplistic eg. of the classical Bayes is: if you toss two coins, of which one is fake (always showing heads) and the other is true. If you toss them 3 times what is the probability to get heads? The prior probability, given we look for heads, for the fake one is always heads so p = 1, and for the other coin is p = 1/2. After we have tossed them 3 times and got p= 1 for fake and p=⅛ for the true coin, the posterior question would be the inverse from the first question (the likelihood) that is: after tossing 3 times and each time getting heads, what is the probability that the coin is the fake one?

So, by drawing an analogy, the more times we toss, the more certain we will be, and the future observation (that it will be heads or not) is more predictable. And we infer the next observation by using the knowledge from the past (likelihood and priors).

Summary, brief:

1. There is an observation **o**
2. We interpret **o** with our model of the world, **m**
   With the model we infer the true states of the world:
   (**A)** representing them with the hidden states **s** (empirical priors) which
   (B) limit the set of possible actions, control states (empirical priors) linked to this observation or hidden state. Simply put, the hidden states assist in choosing a subset of actions (otherwise it is an enormous amount of possible actions). So, from the current state s_t we can choose from u_t…. u_T possible future actions and lead to s_t+1,
   - We can also choose from true actions we just did (in the past, a_1,a_2...a_t) which enter the set of control states as well.

***(B) control states have values of success from the past, which gives them a degree of probability of being used again or not (likelihood, influenced by the full priors)

3. We wish to choose such action (to pass from one state to another while predicting posterior outcome):
    a. so that the surprise is minimal (free-energy minimization).
    b. so that the utility is maximized (according to full prior parameters).
    c. In which the likelihood is highest -- $P(o|s_t)$ - likelihood is tied to the past observations, and with them it infers the probability that observation o would occur given the state s.

        Inversely, we try to predict the posterior → the state s for such observation o (typical Bayesian method).

To be more precise, the next hidden state is what we expect to happen in the future, and the next observation is what will really happen. Note that the next observation depends solely upon the hidden states because they influence and limit the possible actions to be made in the future, and the observation or outcome can be modified only by acting on it).

Dependencies (closed loop)

Observations → from which we infer true states of the world (**hidden states** to us) → which lead us to select a subset of possible actions (**control states**) → which will modify the next observation → which we again interpret with our model of the world.

Conclusion

The more priors (empirical) we have, the richer our model is. In this way we reduce the possibility of being surprised, and maximize the expected gain.

The world is in a state which is reflected onto our limited interpretation of it -- a hidden state which we infer from what we can observe. With these hidden states we infer the possible actions we can take to maximize our gain, and minimize the surprise of the next future world state → observation → hidden state.

Active inference uses the Bayes theory to explain the choice of action a bounded, rational being makes to infer the future observations, in order to minimize the surprise, or to have control over his environment and maximize the gain. In order to store the empirical priors and assign the values to them (favouring ones over the others) it needs a framework filled with vectors and transition matrices (the ABC).

The ABC (matrices) of Active Inference

A represents **the likelihood** that an observation **o** will occur given a state **s** (mentioned above) -- the likelihood says that observations depend on hidden states.

B the probability to transition from one state to the other (control and hidden states modify each other's probabilities of transition).

C represents an important element of the full priors, which influence the likelihood (transitions between the states), influences the choice of the control states in order to achieve the desired outcome. So in short it is described as the prior over the outcomes.

Active Inference in BCI

Example (P300-speller) and future work

The machine needs to infer the true states of the user, in a specific situation, a P300 speller. We want to expand the rationalization, and think further than a simple implementation, but consider the cause of such implementation (the user and task). Active inference enables such expanded framework, and so what has been done so far (till machine 2), needs to be implemented (expanded) in the Active Inference framework, and add other adaptive features:

Machine 0. Basic (without Active Inference)

For **t (until end of word)** number of trials:

- Flash randomly, **n (fixed)** number of times for each trial **t :**
    - Accumulate evidence to spell item (1 out of 36)
- Spell item

Machine 1. Optimal Stopping

For **t** (until end of word) number of trials:

- For **n (until item estimated)** number of times for each trial **t** :
    - Flash in random order x columns and x rows (x fixed)
        - Accumulate evidence for spelling the item (1 of 36)
- Spell item

Machine 2. Optimal Stopping + Correct Feedback

For **t** (until end of word) number of trials:

- For **n (until item estimated)** number of times for each trial **t** :
    - Flash in random order x columns and x rows (x fixed)
        - Accumulate evidence for spelling the item (1 of 36)
- Spell item
- If error potentials spell the second most probable item

Machine 3. Optimal Stopping + Flashing

For **t** (until end of word) number of trials:

- For **n** (until item estimated) number of times for each trial **t** :
    - Flash those columns and rows which best reveal the target item
    - Accumulate evidence for spelling the item (1 of 36)
- Spell item
- If error potentials spell the second most probable item

Machine 4. (3) + Look Away (idem)

For **t** (until end of word **or look away)** number of trials:

- For **n** (until item estimated) number of times for each trial **t** :
  - Flash those columns and rows which best reveal the target item
  - Accumulate evidence for spelling the item (1 of 36)
- Spell item
- If error potentials spell the second most probable item

Machine 5. (4) + theory of mind (low level)

For **t** (until end of word or look-away) number of trials:

- For **n** (until item estimated) number of times for each trial **t** :
  - Flash those columns and rows which best reveal the target item
  - **Penalize** flashing x consecutive columns/rows
  - Accumulate evidence for spelling the item (1 of 36)
- Spell item
- If error potentials spell the second most probable item

These machines will be validated and compared to verify the effects of active inference.

What the machine can **observe**, is the EEG (in our case), or ECoG, MEG etc.
From this data the machine infers the user's **true states** or intentions using the **user** and **task** models. The **purpose** of the machine, in this case, is to spell a word. So the machine knows a priori that the user wants to spell something, considering it is using the machine for that purpose.
The user model represents the **prior** information the machine has about the user (gender, age, language, handedness etc.), but also contains the current **hidden states** (fatigue, attention, workload, mood swings etc) the machine needs to infer during the task, within multiple trials or runs. However, within shorter time intervals (trial-to-trial, flash-to-flash) the hidden states are, in this example, all the possible symbols the user wants to spell (36 possibilities); or in motor imagery, e.g. go left or right etc.
The task model contains the **prior** task information such as the BCI **purpose** (communicate, control, monitor etc.), or BCI paradigm (evoked or spontaneous: P300, SSVEP, SMR, SCP etc.). However during the task, within runs, **control states** are represented as e.g. the instructions given to the user as strategy for manipulating the system (counting for spelling, imagining playing an instrument for hands imagined movement, or imagining feet movement, or mental calculations etc.). This part is not intuitive, but it represents a set of actions offered by the machine to the user. It is not what the machine does at the instant but what it offers as possible actions. So for a P300 the offered task is to focus on the letter you want to spell, count the number of times the target letter was flashed etc. By knowing the set of possible actions in a longer time period, it assists in the short time instances, to limit the possible actions of the machine and to be more concrete (spell the letter, flash for P300, or move cursor right/left for MI BCI etc.).

For now, we are using the short-time instances adaptation, such as optimal stopping/flashing or adapting feedback for motor imagery BCI (tux ref).