

Expected Free Energy

Théophile Champion

University of Kent, School of Computing
Canterbury CT2 7NZ, United Kingdom

TMAC3@KENT.AC.UK

Editor: TO BE FILLED

Abstract

~~Section ?? discusses the literature surrounding the definition and justification~~ This document discusses the definition and implementation of the expected free energy, ~~while focusing on the discrete time and state space version of in the context of deep~~ active inference. ~~Then, after discussing the definition and justification of the expected free energy (EFE),~~ Section 1 discusses one possible definition of the EFE and explains how the EFE can be estimated in the context of deep active inference. Finally, Section 2 ~~summarises the questions asks the question~~ for which an answer would be valuable.

Keywords: Expected Free Energy, Variational Free Energy

1. Expected Free Energy

0.1 Definitions

~~The expected free energy is defined as:~~

$$G(\pi) = \sum_{\tau=0}^T G(\pi, \tau)$$

~~? defined the expected free energy of a policy at time step τ as:~~

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

~~where $\tilde{o} = (o_1, \dots, t)$ is the sequence of observations until the current time step t . ? defined the expected free energy of a policy at time step τ as:~~

$$G(\pi, \tau) = \mathbb{E}_{Q(o_\tau, s_\tau | \pi)} [\ln Q(s_\tau | \pi) - \ln P(s_\tau, o_\tau)].$$

~~Those two definitions of the expected free energy seems to be different. Indeed, in , P is a conditional distribution, while in , it is a joint distribution, i.e., $P(s_\tau, o_\tau | \tilde{o}, \pi) \neq P(s_\tau, o_\tau)$.~~

0.1 Justifications

0.0.1 ~~THE EXPECTED FREE ENERGY IS THE PATH INTEGRAL OF EXPECTED (VARIATIONAL) FREE ENERGY~~

~~? claimed the following:~~

~~“However, beliefs about policies rest on outcomes in the future, because these beliefs determine action and action determines subsequent outcomes. This means that policies should, a priori, minimise the free energy of beliefs about the future. Eq. (1.e) expresses this formally by making the log probability of a policy proportional to the free energy expected under that policy. The expected free energy of a policy follows from Eq. (4).”~~ where Eq. (1.e) refers to the prior over policy $P(\pi)$, and Eq. (4) refers to the variational free energy. **According to this view the expected free energy is obtained by taking the expectation of the variational free energy. Importantly, in this view, an active inference agent behaves according to the free energy principle because the expected**

~~free energy is the path integral of expected (variational) free energy, i.e., by minimising its expected free energy, the agent also minimises its variational free energy.~~

~~The variational free energy is the KL-divergence between the generative model and the variational distribution. However, in the definition of the expected free energy given by , the generative model is conditioned on \tilde{o} and π . Why is that? Put simply, I do not understand the algorithm or proof behind the sentence: “The expected free energy of a policy follows from Eq. (4).”. Why does it follows from Eq. (4)?~~

~~The relationship between the expected and variational free energy is central to active inference. Indeed, if no relationship can be established between the two, then one cannot claim that: “an active inference agent behaves according to the free energy principle because ...by minimising its expected free energy, the agent also minimises its variational free energy”.~~

0.0.1 THE PRECISE AGENT ASSUMPTION

~~? claimed that the expected free energy can be derived from the precise agent assumption, i.e., **Definition 5.1** (Precise agent). An agent is precise when it evolves deterministically in a (possibly) stochastic environment, i.e., when $P(\pi|s)$ is a Dirac measure for any s . A Partially Observable Markov Decision Process (POMDP) is not a precise agent. Since, active inference is based on the POMDP formalism, the proof does not seems to apply to active inference.~~

1. Deep Active Inference

1.1 Definition of the expected free energy

~~Re-starting from ,~~ Starting from the expected free energy ~~is defined as~~ (?) defined as follows:

$$G(\pi) = \sum_{\tau=t+1}^T G_{\tau}(\pi) = \sum_{\tau=t+1}^T \mathbb{E}_{P(o_{\tau}|s_{\tau})Q(s_{\tau}|\pi)} [\ln Q(s_{\tau}|\pi) - \ln P(o_{\tau}, s_{\tau})], \quad (1)$$

where $P(o_{\tau}|s_{\tau})$ is the likelihood mapping, $Q(s_{\tau}|\pi)$ is the variational distribution, and in the literature $P(o_{\tau}, s_{\tau})$ is the target distribution encoding the prior preferences of the agent. Re-arranging the above equation leads to:

$$\begin{aligned} G_{\tau}(\pi) &= \mathbb{E}_{P(o_{\tau}|s_{\tau})Q(s_{\tau}|\pi)} [\ln Q(s_{\tau}|\pi) - \ln P(o_{\tau}, s_{\tau})] \\ &= \mathbb{E}_{P(o_{\tau}|s_{\tau})Q(s_{\tau}|\pi)} [\ln Q(s_{\tau}|\pi) - \ln P(s_{\tau}|o_{\tau}) - \ln P(o_{\tau})] \\ &\approx \mathbb{E}_{P(o_{\tau}|s_{\tau})Q(s_{\tau}|\pi)} [\ln Q(s_{\tau}|\pi) - \ln Q(s_{\tau}) - \ln P(o_{\tau})] \\ &= \underbrace{D_{\text{KL}} [Q(s_{\tau}|\pi) || Q(s_{\tau})]}_{\text{epistemic value}} - \underbrace{\mathbb{E}_{P(o_{\tau}|s_{\tau})Q(s_{\tau}|\pi)} [\ln P(o_{\tau})]}_{\text{extrinsic value}}, \end{aligned} \quad (2)$$

1.2 Challenges of estimating the expected free energy

Now, the question is how to estimate (2), and we focus on the case where $\tau = t+1$. Note, because $\tau = t+1$, the policy π contains only one action a_t , i.e., $\pi = a_t$. In the tabular version of active inference, the variational distribution is composed of a factor $Q(s_{\tau}|\pi)$. **However, in the deep active inference literature, the variational distribution does not contain such a factor.** Generally, a Monte-Carlo estimate is used as follows:

$$Q(s_{t+1}|a_t) = \mathbb{E}_{Q_{\phi_s}(s_t)} [P_{\theta_s}(s_{t+1}|s_t, a_t)] \approx \frac{1}{N} \sum_{i=1}^N P_{\theta_s}(s_{t+1}|s_t = \hat{s}_t^i, a_t), \quad (3)$$

where $\hat{s}_t^i \sim Q_{\phi_s}(s_t)$. In what follows, we use $N = 1$ leading to a simplified version of the estimate:

$$Q(s_{t+1}|a_t) \approx P_{\theta_s}(s_{t+1}|s_t = \hat{s}_t, a_t). \quad (4)$$

Note, in the above equation, \hat{s}_t^i is denoted \hat{s}_t because there is only one sample, i.e., $N = 1$. At this point, we have an estimate for $Q(s_{t+1}|a_t)$ and $Q_{\phi_s}(s_t)$ is the variational distribution. The only missing piece is an estimate of the

extrinsic value. In the tabular version of active inference, the preferences of the agent can be related to the rewards from the reinforcement learning literature. In this paper, we follow (?) and define the prior preferences as:

$$P(o_\tau) = \frac{\exp(\psi r_\tau[o_\tau])}{\sum_{o_\tau} \exp(\psi r_\tau[o_\tau])},$$

where ψ is the precision of the prior preferences, and $r_\tau[o_\tau]$ is the reward obtained when making observation o_τ . Taking the logarithm of the above equation leads to:

$$\begin{aligned} \ln P(o_\tau) &= \psi r_\tau[o_\tau] - \ln \sum_{o_\tau} \exp(\psi r_\tau[o_\tau]) \\ &= \psi r_\tau[o_\tau] + C, \end{aligned} \tag{5}$$

where we used the fact that the summation over all o_τ is a normalisation term, i.e., a constant. Using (5), we can now create an estimate of the extrinsic value as follows:

$$\mathbb{E}_{P_{\theta_o}(o_\tau|s_\tau)Q(s_\tau|a_t)}[\ln P(o_\tau)] \approx \frac{1}{M} \sum_{i=1}^M \ln P(o_\tau = \hat{o}_\tau^i) = \frac{1}{M} \sum_{i=1}^M \psi r_\tau[o_\tau^i] + C,$$

where $\hat{o}_\tau^i \sim P_{\theta_o}(o_\tau|s_\tau = \hat{s}_\tau^i)$ and $\hat{s}_\tau^i \sim Q(s_\tau|a_t)$. In what follows, we use $M = 1$ and discard the constant¹, which leads to a simplified version of the estimate:

$$\mathbb{E}_{P_{\theta_o}(o_\tau|s_\tau)Q(s_\tau|a_t)}[\ln P(o_\tau)] \triangleq \psi r_\tau[o_\tau] \triangleq \psi r_\tau,$$

where we simplified the notation by denoting $r_\tau[o_\tau^i]$ as r_τ . To conclude, we have the following estimate for the EFE at time $\tau = t + 1$:

$$\begin{aligned} G_{t+1}(a_t) &\approx D_{\text{KL}}[Q(s_{t+1}|a_t) || Q_{\phi_s}(s_{t+1})] - \mathbb{E}_{P_{\theta_o}(o_{t+1}|s_{t+1})Q(s_{t+1}|a_t)}[\ln P(o_{t+1})] \\ &\approx D_{\text{KL}}[P_{\theta_s}(s_{t+1}|s_t = \hat{s}_t, a_t) || Q_{\phi_s}(s_{t+1})] - \psi r_{t+1}, \end{aligned} \tag{6}$$

where $\hat{s}_t \sim Q_{\phi_s}(s_t)$, $P_{\theta_s}(s_{t+1}|s_t, a_t)$ is known from the generative model, $Q_{\phi_s}(s_{t+1})$ is known from the variational distribution, the KL-divergence can be estimated using an analytical solution, ψ is a hyperparameter modulating the precision of the prior preferences, and r_{t+1} is the reward obtained at time step $t + 1$.

2. ~~Questions~~Question

1. What is the ~~definition of the expected free energy?~~
2. ~~What is the justification of the expected free energy?~~
3. ~~What is the proper definition of the expected free energy in the context of deep active inference?~~
4. ~~Why are equations and incorrect? (last time I presented, I was told that it was not the proper definition)~~ cost of this inability to calculate $Q(s_\tau|\pi)$ directly?

~~A point by point response where the open questions of research are clearly acknowledged would be very valuable.~~

Acknowledgments

TO BE FILLED

1. Removing a constant does not influence which policy is the best. Indeed, $\pi^* = \arg \max_\pi G(\pi) = \arg \max_\pi G(\pi) - C$.