Reframing the Expected Free Energy:

Four Formulations and a Unification.

Théophile Champion TMAC3@KENT.AC.UK

University of Birmingham, School of Psychology and School of Computer Science,

Birmingham B15 2TT, United Kingdom

Howard Bowman H.BOWMAN@KENT.AC.UK

University of Birmingham, School of Psychology and School of Computer Science,

Birmingham B15 2TT, United Kingdom

University College London, Wellcome Centre for Human Neuroimaging (honorary)

London WC1N 3AR, United Kingdom

Dimitrije Marković DIMITRIJE.MARKOVIC@TU-DRESDEN.DE

Technische Universität Dresden, Department of Psychology

Dresden 01069, Germany

Marek Grześ M.GRZES@KENT.AC.UK

University of Kent, School of Computing

Canterbury CT2 7NZ, United Kingdom

Editor: TO BE FILLED

Abstract

Active inference is a process theory of perception, learning and decision making, that is applied to a range of research fields, such as, neuroscience, robotics, psychology, and machine learning. Active inference rests upon an objective function called the expected free energy, which can be justified by the intuitive plausibility of its formulations, e.g., the risk plus ambiguity and information gain / pragmatic value formulations. This paper seeks to formalize the problem of deriving these formulations from a single root expected free energy definition, i.e., the unification problem. Then, we analyze two approaches to defining expected free energy. More precisely, the expected free energy is either defined as: (1) the risk over observations plus ambiguity, or (2) the risk over states plus ambiguity. In the first setting, no rigorous mathematical justification for the expected free energy has been proposed to date, but all the formulations can be recovered from it, by assuming that the likelihood of target distribution $T(\bar{o}|\bar{s})$ is the likelihood of the generative model $P(\bar{o}|\bar{s})$. Importantly, under this likelihood constraint, if the likelihood is lossless[1] then prior preferences over observations can be defined arbitrarily. However, in the more general case of Partially Observable Markov Decision Processes (POMDPs), we demonstrate that the likelihood constraint effectively restricts the set of valid prior preferences over observations. Indeed, only a limited class of prior preferences over observations is compatible with the likelihood mapping of the generative model. In the second setting, a justification of the root expected free energy definition exists, but this setting only accounts for two formulations, i.e., the risk over states plus ambiguity and entropy plus expected energy formulations. We conclude with a discussion of the conditions under which a unification of expected free energy formulations has been proposed in the literature, by appeal to the free energy principle in the specific context of systems without random fluctuations.

Keywords: Active Inference, Expected Free Energy, Unification Problem

1. Introduction

Active inference (Çatal, Verbelen, Nauta, Boom, & Dhoedt, 2020; Cullen, Davey, Friston, & Moran, 2018; FitzGerald, Dolan, & Friston, 2015; Fountas, Sajid, Mediano, & Friston, 2020; K. Friston et al., 2016; Itti & Baldi, 2009; Millidge, 2019; Sancaktar, van Gerven, & Lanillos, 2020; Schwartenbeck et al., 2018) is a framework for decision-making under uncertainty, in which the agent is equipped with a (generative) model that encodes the environment dynamics, and a variational posterior approximating the true posterior over latent variables. The variational posterior is computed by minimizing a function called the variational free energy (VFE), also known as the negative evidence lower bound in machine learning (Fox & Roberts, 2012; Winn & Bishop, 2005). While the variational posterior defines the most likely state of the environment, it does not indicate which action should be selected.

Instead, the agent aims to reach a set of preferred states or observations, by minimizing the expected free energy (EFE). While the variational free energy has a clear root definition that all other formulations are derived from, the active inference literature does not clearly identify a unique definition for expected free energy. This leaves open the question of antecedents among its formulations. The EFE is an objective function defining the cost of performing a particular policy as a trade-off between explo-

[1] A lossless likelihood implies that $\boldsymbol{A}$ is a permutation matrix, which corresponds to the special case of Markov Decision Processes.

ration and exploitation, e.g., the goal is to maximize pragmatic value (reward), while also maximizing information gain. The pragmatic value relies on the prior preferences of the agent, which specify the preferred states or observations, and provides the agent with its goal-directed behaviour.

In the active inference literature, the prior preferences over observation are usually denoted $P(\bar{o}; C)$, where C is a vector of parameters. In some of the literature, the dependency on C has been dropped and the prior preferences are written as $P(\bar{o})$. This is unfortunate because this overloaded[2] expression can be confused with the marginal likelihood under the generative model $P(\bar{o}|m)$; especially when the model (m) is also left out. It is important to appreciate that the prior preferences $P(\bar{o}; C)$ are not the same as the marginal likelihood $P(\bar{o}|m)$, although some earlier literature associates them. For example, Parr and Friston (2019) wrote: *"The second term[3] speaks to the notion of a 'crooked scientist', who designs experiments to confirm prior beliefs, i.e. preferred outcomes. This preference is the same as the evidence (a.k.a., marginal likelihood) associated with a given model."*

Importantly, the prior preferences $P(\bar{o}; C)$ are not part of the generative model, because the only marginal distribution over observations in the generative model is the marginal likelihood $P(\bar{o}|m)$, and $P(\bar{o}; C) \neq P(\bar{o}|m)$. Thus, we will distinguish between the marginal likelihood and prior preferences by treating the prior preferences as a target distribution T, in a similar way to Hafner et al. (2022).

However, as discussed in Appendix D, when prior preferences over observations, i.e., $P(\bar{o}; C)$, are used within a joint distribution, i.e., $P(\bar{o}, \bar{s}; C) = P(\bar{o}; C)P(\bar{s}|\bar{o}; C)$, it has the potential to constrain active inference in ways that may not be wanted. We also, in Appendix E, consider difficulties that arise from using Bayes theorem with different types of distributions, e.g., the target distribution and the variational posterior. Fortunately, while the issues discussed in Appendix D and E need to be kept in mind, they seems to be restricted to older presentations of the active inference theory, and do not arise in more recent treatments, such as can be found in Parr, Pezzulo, and Friston (2022).

The more current issues that will be discussed in the main-body focus on a problem that remains even after moving from introducing preferences into the model as a prior to incorporating an explicit target distribution. Indeed, the explicit target distribution has to be equipped with the constraint that the likelihood of the target distribution is the likelihood of the generative model. Under this

[2]A reviewer requested that we use this term.

[3]The second term refers to the extrinsic value of the expected free energy, denoted $\ln P(o_\tau)$ in Parr and Friston (2019).

constraint, a specification of the target distribution in terms of target states uniquely determines the target distribution over outcomes. If the likelihood mapping is lossless, a specification of the target distribution over outcomes also uniquely specifies the target distribution over states. However, in most POMDPs, the likelihood is not lossless, i.e., $\boldsymbol{A}$ is not a permutation matrix. Our paper is therefore asking the question: *what if* $\boldsymbol{A}$ *is not a permutation matrix?* As discussed below, this effectively restricts the class of valid preferences, and leads to a definition of the expected free energy (the risk over observations and ambiguity decomposition) that is not currently justified. In the following sections, we borrow heavily from the derivations in Parr et al. (2022), exploring two possible interpretations and explain their limitations. Appendix B and C provide a description of the properties used throughout this paper.

2. Generative model

In active inference, the agent is equipped with a (generative) model of its environment that spans all time steps until the present time t. This model is composed of (a) hidden states $s_{0:t}$ representing states of the environment that the agent does not directly observe, (b) observations $o_{0:t}$, which represent measurements made by the agent, and (c) actions $a_{0:t-1}$ that the agent performed in the environment. For the sake of conciseness, $s_{0:t}$, $o_{0:t}$, and $a_{0:t-1}$ will be denoted $\underline{s}$, $\underline{o}$, and $\underline{a}$, repectively. Moreover, in this paper, we assume that observations depend on states, and each state depends on the state and action at the previous time step. Formally, this setting is called a Partially Observable Markov Decision Process (POMDP), and the model definition is as follows:

$$P(\underline{o}, \underline{s}|\underline{a}) \triangleq P(s_0) \prod_{\tau=0}^{t} P(o_\tau|s_\tau) \prod_{\tau=1}^{t} P(s_\tau|s_{\tau-1}, a_{\tau-1}).$$

3. Variational distribution

The generative model described in the previous section encodes prior beliefs about the environment dynamics. However, when making measurements of key quantities, e.g., $\underline{o}$, the agent needs to compute posterior beliefs about the states, e.g., $P(\underline{s}|\underline{o}, \underline{a})$. These posterior beliefs encode the new beliefs of the agent when taking into consideration the new observations. Unfortunately, computing the true posterior can either be analytically intractable or simply too computationally expensive. Therefore, the true posterior is generally approximated by a variational distribution $Q(\underline{s}|\underline{a})$:

$$\underbrace{Q(\underline{s}|\underline{a})}_{\text{variational posterior}} \approx \underbrace{P(\underline{s}|\underline{o},\underline{a})}_{\text{true posterior}} \propto \underbrace{P(\underline{o},\underline{s}|\underline{a})}_{\text{generative model}}$$

In active inference, the variational posterior 1) factorises over time steps, i.e., a temporal mean-field approximation, but 2) all states still depend on the policy $\underline{a}$. These two assumptions lead to the following definition of the variational distribution:

$$Q(\underline{s}|\underline{a}) \triangleq \prod_{\tau=0}^{t} Q(s_\tau|\underline{a})$$

4. Variational inference and the variational free energy

To sum up, the agent is provided with a generative model $P(\underline{o},\underline{s}|\underline{a})$ and a variational distribution $Q(\underline{s}|\underline{a})$. Given some measurements $\underline{o}$, the variational distribution needs to approximate the true posterior $P(\underline{s}|\underline{o},\underline{a})$. This can be formally expressed as minimising the Kullback-Leibler divergence between the approximate and true posteriors:

$$Q^*(\underline{s}|\underline{a}) = \arg\min_{Q(\underline{s}|\underline{a})} D_{\mathrm{KL}}\left[\, Q(\underline{s}|\underline{a}) || \, P(\underline{s}|\underline{o},\underline{a})\right]$$

Minimising this KL-divergence and minimising the variational free energy (VFE) is equivalent (see proof below). Intuitively, the VFE trades-off accuracy, i.e., how well are the observations predicted, and complexity, i.e., how far the posterior is from the prior. More formally, the VFE is defined as follows:

$$\mathcal{F}\Big[Q(\underline{s}|\underline{a}),\underline{a},\underline{o}\Big] = \underbrace{D_{\mathrm{KL}}\left[\, Q(\underline{s}|\underline{a}) || \, P(\underline{s}|\underline{a})\right]}_{\text{complexity}} - \underbrace{\mathbb{E}_{Q(\underline{s}|\underline{a})}[\ln P(\underline{o}|\underline{s})]}_{\text{accuracy}}$$

P

The derivation of the variational free energy proceeds as follows:

$$
\begin{aligned}
Q^*(\underline{s}|\underline{a}) &= \arg\min_{Q(\underline{s}|\underline{a})} D_{\text{KL}}\left[\, Q(\underline{s}|\underline{a}) || \, P(\underline{s}|\underline{o}, \underline{a})\right] \\
&= \arg\min_{Q(\underline{s}|\underline{a})} \mathbb{E}_{Q(\underline{s}|\underline{a})}\left[\ln Q(\underline{s}|\underline{a}) - \ln \frac{P(\underline{o}|\underline{s})P(\underline{s}|\underline{a})}{P(\underline{o}|\underline{a})} \right] \\
&= \arg\min_{Q(\underline{s}|\underline{a})} \mathbb{E}_{Q(\underline{s}|\underline{a})}[\ln Q(\underline{s}|\underline{a}) - \ln P(\underline{o}|\underline{s})P(\underline{s}|\underline{a})] + \underbrace{\ln P(\underline{o}|\underline{a})}_{\text{constant}} \\
&= \arg\min_{Q(\underline{s}|\underline{a})} \underbrace{D_{\text{KL}}\left[\, Q(\underline{s}|\underline{a}) || \, P(\underline{s}|\underline{a})\right]}_{\text{complexity}} - \underbrace{\mathbb{E}_{Q(\underline{s}|\underline{a})}[\ln P(\underline{o}|\underline{s})]}_{\text{accuracy}} \\
&= \arg\min_{Q(\underline{s}|\underline{a})} \underbrace{\mathcal{F}\Big[Q(\underline{s}|\underline{a}), \underline{a}, \underline{o}\Big]}_{\text{variational free energy}},
\end{aligned}
$$

where the d-separation criterion (Koller & Friedman, 2009) (which simplifies $P(\underline{o}|\underline{s}, \underline{a})$ to $P(\underline{o}|\underline{s})$ in the second line, given the independence of $\underline{o}$ and $\underline{a}$ given $\underline{s}$) and Bayes theorem have been used, along with the linearity of expectation and the log-property.

C

In this paper, we focus on planning, therefore we do not explain how the optimal variational distribution $Q^*(\underline{s}|\underline{a})$ is computed. The interested reader is referred to (Champion, Da Costa, Bowman, & Grześ, 2022; K. J. Friston, Parr, & de Vries, 2017; Parr, Markovic, Kiebel, & Friston, 2019; Winn & Bishop, 2005) for an approximate scheme based on variational message passing, and (Champion, Grześ, & Bowman, 2022b; Kschischang, Frey, & Loeliger, 2001) for an exact scheme based on the sum-product algorithm.

5. Planning and the expected free energy

After performing inference, the agent has at its disposal posterior beliefs encoded by the optimal variational distribution $Q^*(\underline{s}|\underline{a})$. At this point, the agent needs to choose the next action to perform in the environment. In active inference, the cost of a policy is given by the expected free energy (EFE), thus the goal of planning is to identify the policy with the smallest EFE.

Given a time horizon of planning h, the number of policies is $|\mathcal{A}|^{h-t}$, where $\mathcal{A}$ is the set of all actions available to the agent, and $|\mathcal{A}|$ is the cardinality of this set. As the number of policies grows exponentially with the time horizon, computing the EFE of all policies requires exponential time. Therefore, it is important to search the space of policies efficiently. For example, one can use Monte-Carlo tree search (Champion, Bowman, & Grześ, 2022; Champion, Grześ, & Bowman, 2022a), which maintains a balance between exploiting policies with low EFE and exploring rarely visited action sequences. Another solution would be to use sophisticated inference (K. Friston, Da Costa, Hafner, Hesp, & Parr, 2021), which implements a form of backward induction based on a recursive form of expected free energy.

In this paper, we postulate that the expected free energy is based on two distributions: the forecast $F(\overline{o}, \overline{s}|\overline{a})$ and target $T(\overline{o}, \overline{s}|\overline{a})$ distributions, where the future observations, states, and actions are denoted $\overline{o} = o_{t+1:h}$, $\overline{s} = s_{t+1:h}$ and $\overline{a} = a_{t:h-1}$, respectively. The forecast distribution predicts the future according to the agent's best beliefs about the current states of the environment, and its generative model. In contrast, the target distribution encodes the states and observations the agent wants to reach. A general formulation of the target is given here, allowing it to change with policies. Although, in most cases, it will not change.

Q Why is the forecast distribution introduced?

A In the active inference literature, it is frequent to see Bayes theorem being used between factors of the generative model $P(\underline{o}, \underline{s}|\underline{a})$ and the variational distribution $Q(\underline{s}|\underline{a})$, or even to see factors from the generative model $P(x|y)$ being replaced by their variational counterpart $Q(x|y)$. However, Bayes theorem is a corollary of the product rule of probability, i.e.,

$$P(x, y) = P(x|y)P(y) = P(y|x)P(x) \Leftrightarrow P(y|x) = \frac{P(x|y)P(y)}{P(x)}.$$

Thus, technically, Bayes theorem cannot be used between factors of two different distributions. To be really explicit, equality between $P(y|x)$ and the right-hand-side would not hold if, for example, $P(y)$ were replaced by any distribution different to $P(y)$, and similarly for $P(x|y)P(y)$ and $P(x)$. Importantly, it is straightforward to see that the generative model and the variational distribution are two different distributions. The easiest way to see this, is to realise that these two distributions do not even share the same domain, i.e., the generative model is a distribution over states $\underline{s}$ and observations $\underline{o}$, while the variational distribution is a distribution over states $\underline{s}$ only. One might consider observations to be implicitly present, but in this respect, they are a ground term, with a

A

particular value, rather than random variables with a distribution over possible values. Another intuitive argument is that the generative model encodes the prior beliefs of the agent, then upon receiving new data, the agent computes its (approximate) posterior beliefs. If the prior was equal to the posterior, there will be no point in performing inference in the first place. Since the generative model and the variational distribution are two different distributions, one should be careful when replacing factors from one distribution by factors from the other. It is for this reason that the forecast distribution is introduced. Put simply, the forecast distribution provides a bridge between factors of the variational posterior and those of the generative model.

5.1 The unification problem

In this section, we formalise the derivation of four EFE formulations, which can be found in the literature (see below), from the EFE definition, i.e., *the unification problem*. Specifically, the unification problem is a 4-tuple $\mathcal{P} = \langle F, T, \mathcal{G}_{rt}, \mathcal{C} \rangle$, where F is the forecast distribution, T is the target distribution, $\mathcal{G}_{rt}$ is the definition of the expected free energy, and $\mathcal{C} = \{\mathcal{C}_{RSA}, \mathcal{C}_{ROA}, \mathcal{C}_{IGPV}, \mathcal{C}_{3E}\}$ is a set containing the four formulations of the expected free energy. Solving $\mathcal{P}$ consists of finding a definition of F, T, and $\mathcal{G}_{rt}$ such that it is possible to derive $\mathcal{C}_X$ from $\mathcal{G}_{rt}$ for all $\mathcal{C}_X \in \mathcal{C}$.

We now define the four formulations of the expected free energy, which are based upon the definitions in Parr et al. (2022). The formulation for the risk over states and ambiguity is as follows:

$$\mathcal{C}_{RSA}(\overline{a}) \triangleq \underbrace{D_{\text{KL}}\left[\, F(\overline{s}|\overline{a}) \,||\, T(\overline{s}|\overline{a}) \right]}_{\text{risk over states}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})] \big]}_{\text{ambiguity}}.$$

Importantly, the risk over states is the KL-divergence between the predictive posterior over states $F(\overline{s}|\overline{a})$ and the prior preferences over states $T(\overline{s}|\overline{a})$, and the ambiguity is the expected entropy of the likelihood mapping according to the generative model. The risk over states pushes the predictive posterior towards the prior preferences, while the ambiguity encourages the agent to visit states producing a low entropy distribution over observations, i.e., if we arrive at a state, we know which observation(s) to expect. The formulation for the risk over observations and ambiguity is as follows:

$$\mathcal{C}_{ROA}(\overline{a}) \triangleq \underbrace{D_{\text{KL}}\left[\, F(\overline{o}|\overline{a}) \,||\, T(\overline{o}|\overline{a}) \right]}_{\text{risk over observations}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})] \big]}_{\text{ambiguity}}.$$

The ambiguity term is identical, and the risk over observations is a KL-divergence, which pushes the predictive posterior over observations $F(\bar{o}|\bar{a})$ to be as close as possible to the prior preferences over observations $T(\bar{o}|\bar{a})$. The formulation for the information gain and pragmatic value is as follows:

$$\mathcal{C}_{IGPV}(\bar{a}) \triangleq -\underbrace{\mathbb{E}_{F(\bar{o}|\bar{a})}\big[D_{\text{KL}}\big[\,F(\bar{s}|\bar{o},\bar{a})||\,F(\bar{s}|\bar{a})\big]\big]}_{\text{information gain}} - \underbrace{\mathbb{E}_{F(\bar{o}|\bar{a})}[\ln T(\bar{o}|\bar{a})]}_{\text{pragmatic value}}.$$

Importantly, the information gain is a KL-divergence that relies only on factors from the forecast distribution. This prevents degenerate behaviours were the agent stops exploring its environment, i.e., information loss (Champion, Grześ, Bonheme, & Bowman, 2023). In addition, the pragmatic value is based on the preferred observations $T(\bar{o}|\bar{a})$, which provides the agent with its goal directed behaviour. Finally, the expected energy and entropy formulation is as follows:

$$\mathcal{C}_{3E}(\bar{a}) \triangleq -\underbrace{H[F(\bar{s}|\bar{a})]}_{\text{entropy}} - \underbrace{\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln T(\bar{o},\bar{s}|\bar{a})]}_{\text{expected energy}}$$

The entropy term ensures that a good policy is one which keeps our options open by allowing us to reach a wide variety of states, as implied by Jaynes' theory of maximum entropy (Jaynes, 1957a, 1957b). Additionally, as shown in the proof below, the expected energy encourages the agent to reach its preferred states, while also pushing the agent to select states for which the associated distribution over observations has low entropy, i.e., given a state, we know which observation(s) to expect. Additionally, Appendix F shows how the derivations made in the following proof have correspondences to derivations in Parr et al. (2022).

P

Starting with the negative expected energy, one can use the product rule, the log-property and the linearity of expectation to get:

$$-\underbrace{\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln T(\bar{o},\bar{s}|\bar{a})]}_{\text{expected energy}} = -\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln T(\bar{o}|\bar{s},\bar{a})] - \mathbb{E}_{F(\bar{s}|\bar{a})}[\ln T(\bar{s}|\bar{a})].$$

Then, assuming that the forecast distribution is a partially observable Markov decision process, and that the likelihood of the forecast and target distributions are the same, we obtain:

$$\begin{aligned} -\underbrace{\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln T(\bar{o},\bar{s}|\bar{a})]}_{\text{expected energy}} &= -\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln F(\bar{o}|\bar{s},\bar{a})] - \mathbb{E}_{F(\bar{s}|\bar{a})}[\ln T(\bar{s}|\bar{a})] \\ &= -\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln F(\bar{o}|\bar{s})] - \mathbb{E}_{F(\bar{s}|\bar{a})}[\ln T(\bar{s}|\bar{a})] \end{aligned}$$

P

where, $F(\bar{o}|\bar{s},\bar{a}) = F(\bar{o}|\bar{s})$, due to d-separation, observing that $\bar{o}$ and $\bar{a}$ are conditionally independent given $\bar{s}$. Lastly, using this same property again after applying the product rule and then using the definition of entropy gives the final result:

$$-\underbrace{\mathbb{E}_{F(\bar{o},\bar{s}|\bar{a})}[\ln T(\bar{o},\bar{s}|\bar{a})]}_{\text{expected energy}} = \underbrace{\mathbb{E}_{F(\bar{s}|\bar{a})}[H[F(\bar{o}|\bar{s})]]}_{\text{ambiguity}} - \underbrace{\mathbb{E}_{F(\bar{s}|\bar{a})}[\ln T(\bar{s}|\bar{a})]}_{\text{pragmatic value}},$$

This equation shows that maximising expected energy means selecting states that minimise the ambiguity of the likelihood while maximising the pragmatic value of states.

5.2 Forecast distribution

As previously mentioned, the forecast distribution predicts the future according to the agent's best beliefs about the current states of the environment, and its generative model. More formally, the forecast distribution factorizes as follows:

$$F(\bar{o},\bar{s}|\bar{a}) \triangleq F(s_{t+1}|a_t) \prod_{\tau=t+1}^{h} F(o_\tau|s_\tau) \prod_{\tau=t+2}^{h} F(s_\tau|s_{\tau-1},a_{\tau-1}).$$

Additionally, we make three assumptions to define the factors of the forecast distribution. These assumptions define the forecast distribution in terms of factors from the generative model and the variational distribution. But importantly, these definitions are made explicit here, leaving no uncertainty about these relationships amongst probability distributions. First, we assume that the likelihood in the future $F(o_\tau|s_\tau)$ is the same as in the past $P(o_\tau|s_\tau)$, i.e., the likelihood of the forecast distribution is the same as the likelihood of the generative model. Second, the temporal transition in the future $F(s_\tau|s_{\tau-1},a_{\tau-1})$ is the same as in the past $P(s_\tau|s_{\tau-1},a_{\tau-1})$, i.e., the temporal transition of the forecast distribution is the same as the temporal transition of the generative model. Third, the agent's best prior over the current state $F(s_t)$ is given by the optimal variational posterior $Q^*(s_t|\underline{a})$. More formally:

$$F(o_\tau|s_\tau) = P(o_\tau|s_\tau)$$
$$F(s_\tau|s_{\tau-1}, a_{\tau-1}) = P(s_\tau|s_{\tau-1}, a_{\tau-1})$$
$$F(s_t) = Q^*(s_t|\underline{a})$$

Using the above assumptions, we obtain $F(s_{t+1}|a_t)$ by taking the expectation of the temporal transition w.r.t. the prior over states at time t. For generality, we define this as an integral, but this could be specified to a summation in the discrete case:

$$F(s_{t+1}|a_t) = \int_{s_t} F(s_{t+1}|s_t, a_t)F(s_t)\ ds_t = \int_{s_t} P(s_{t+1}|s_t, a_t)Q^*(s_t|\underline{a})\ ds_t$$

5.3 Target distribution

The second distribution of interest is the target distribution, which encodes the states and observations that the agent wants to reach. In the following section, we define the target distribution as follows:

$$T(\overline{o}, \overline{s}|\overline{a}) \triangleq \prod_{\tau=t+1}^{h} T(o_\tau|s_\tau)T(s_\tau|\overline{a}),$$

where $T(o_\tau|s_\tau) = P(o_\tau|s_\tau)$ and $T(s_\tau|\overline{a}) = \text{Cat}(s_\tau; \boldsymbol{C}_s)$. Note, $\boldsymbol{C}_s$ is the vector known as the $\boldsymbol{C}$ vector in the active inference literature, where the prior preferences are defined w.r.t. states. Also, the target's likelihood is equal to $P(o_\tau|s_\tau)$, which means that given any state, the agent wants to reach the observations that will arise naturally according to the likelihood of the generative model, i.e., $P(o_\tau|s_\tau)$. If the target changes with policy $\overline{a}$, this would have to be built into the $\boldsymbol{C}_s$ vector, but in most cases, the target distribution would be fixed.

5.4 Solving the unification problem

With the forecast and target distributions laid out, we now focus on the unification problem. We will explore whether any of the EFE formulations could serve as a root definition from which all other formulations could be derived. First, we define the root expected free energy as the risk over observations plus ambiguity:

$$\mathcal{G}_{rt}(\overline{a}) \triangleq \underbrace{D_{\text{KL}}\left[\, F(\overline{o}|\overline{a}) || \, T(\overline{o}|\overline{a})\right]}_{\text{risk over observations}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big]}_{\text{ambiguity}} = \mathcal{C}_{ROA}(\overline{a}). \tag{1}$$

5.4.1 The information gain / pragmatic value formulation

In this section, we demonstrate that the information gain / pragmatic value formulation can be recovered from the $\mathcal{C}_{ROA}(\overline{a})$-as-root expected free energy definition. The derivation relies on the following equality:

$$\frac{F(\overline{s}|\overline{a})}{F(\overline{s}|\overline{o}, \overline{a})} = \frac{F(\overline{o}|\overline{a})}{F(\overline{o}|\overline{s})}, \tag{2}$$

which holds because the forecast distribution is a partially observable Markov decision process.

P

We start by re-arranging Bayes theorem as follows:

$$F(\overline{s}|\overline{o}, \overline{a}) = \frac{F(\overline{o}|\overline{s}, \overline{a}) F(\overline{s}|\overline{a})}{F(\overline{o}|\overline{a})} \Leftrightarrow \frac{F(\overline{s}|\overline{a})}{F(\overline{s}|\overline{o}, \overline{a})} = \frac{F(\overline{o}|\overline{a})}{F(\overline{o}|\overline{s}, \overline{a})}.$$

Then, in a partially observable Markov decision process, $\overline{o} \perp\!\!\!\perp \overline{a} \mid \overline{s}$, (i.e., the Markov property ensures that observation sequences are conditionally independent of policies, if the sequence of states is known), thus $F(\overline{o}|\overline{s}, \overline{a}) = F(\overline{o}|\overline{s})$ and:

$$\frac{F(\overline{s}|\overline{a})}{F(\overline{s}|\overline{o}, \overline{a})} = \frac{F(\overline{o}|\overline{a})}{F(\overline{o}|\overline{s})}.$$

Importantly, by starting with the definition of $\mathcal{C}_{IGPV}(\overline{a})$ and using (2), one can show that:

$$\mathcal{G}_{rt}(\overline{a}) = -\underbrace{\mathbb{E}_{F(\overline{o}|\overline{a})}[D_{\text{KL}}\left[\, F(\overline{s}|\overline{o}, \overline{a}) || \, F(\overline{s}|\overline{a})\right]]}_{\text{information gain}} - \underbrace{\mathbb{E}_{F(\overline{o}|\overline{a})}[\ln T(\overline{o}|\overline{a})]}_{\text{pragmatic value}} = \mathcal{C}_{IGPV}(\overline{a}). \tag{3}$$

P

Starting with the definition of $\mathcal{C}_{IGPV}(\overline{a})$:

$$\mathcal{C}_{IGPV}(\overline{a}) = -\underbrace{\mathbb{E}_{F(\overline{o}|\overline{a})}[D_{\text{KL}}\left[\, F(\overline{s}|\overline{o}, \overline{a}) || \, F(\overline{s}|\overline{a})\right]]}_{\text{information gain}} - \underbrace{\mathbb{E}_{F(\overline{o}|\overline{a})}[\ln T(\overline{o}|\overline{a})]}_{\text{pragmatic value}},$$

and using the KL-divergence definition, and that $F(\overline{o}, \overline{s}|\overline{a}) = F(\overline{s}|\overline{o}, \overline{a})F(\overline{o}|\overline{a})$, by the product rule, we obtain:

$$\mathcal{C}_{IGPV}(\overline{a}) = -\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{s}|\overline{o}, \overline{a}) - \ln F(\overline{s}|\overline{a})] - \mathbb{E}_{F(\overline{o}|\overline{a})}[\ln T(\overline{o}|\overline{a})]$$

Then, by using the log-properties and (2) to replace $\frac{F(\overline{s}|\overline{o},\overline{a})}{F(\overline{s}|\overline{a})}$ by $\frac{F(\overline{o}|\overline{s})}{F(\overline{o}|\overline{a})}$, we get:

$$\mathcal{C}_{IGPV}(\overline{a}) = -\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{o}|\overline{s}) - \ln F(\overline{o}|\overline{a})] - \mathbb{E}_{F(\overline{o}|\overline{a})}[\ln T(\overline{o}|\overline{a})]$$

P

Next, the linearity of expectation can be applied to re-arrange the expression as follows:

$$\mathcal{C}_{IGPV}(\overline{a}) = -\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{o}|\overline{s})] + \mathbb{E}_{F(\overline{o}|\overline{a})}[\ln F(\overline{o}|\overline{a}) - \ln T(\overline{o}|\overline{a})]$$

Lastly, recognizing the entropy and KL-divergence definitions leads to the final results:

$$\mathcal{C}_{IGPV}(\overline{a}) = \underbrace{D_{\mathrm{KL}}\left[\, F(\overline{o}|\overline{a}) || \, T(\overline{o}|\overline{a})\right]}_{\text{risk over observations}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big]}_{\text{ambiguity}} = \mathcal{G}_{rt}(\overline{a}).$$

5.4.2 The risk over states / ambiguity formulation

In this section, we demonstrate that the risk over states plus ambiguity is an upper bound of the expected free energy. Restarting from the EFE definition, one can show that:

$$\mathcal{G}_{rt}(\overline{a}) \leq D_{\mathrm{KL}}\left[\, F(\overline{o}, \overline{s}|\overline{a}) || \, T(\overline{o}, \overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big].$$

P

We follow the proof provided in Appendix B of Parr et al. (2022), but using our notation and going from the end of the proof to the beginning. Restarting from the EFE definition:

$$\mathcal{G}_{rt}(\overline{a}) = \underbrace{D_{\mathrm{KL}}\left[\, F(\overline{o}|\overline{a}) || \, T(\overline{o}|\overline{a})\right]}_{\text{risk over observations}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big]}_{\text{ambiguity}},$$

P

we obtain an upper bound on the EFE by adding the following bound, which is the expectation of a KL-divergence and cannot be negative:

$$\mathcal{G}_{rt}(\overline{a}) \leq \underbrace{D_{\mathrm{KL}}\left[\, F(\overline{o}|\overline{a}) || \, T(\overline{o}|\overline{a})\right]}_{\text{risk over observations}} + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big] + \underbrace{\mathbb{E}_{F(\overline{o}|\overline{a})}\big[D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{o},\overline{a}) || \, T(\overline{s}|\overline{o},\overline{a})\right]\big]}_{\text{bound}}.$$

Next, using the linearity of expectation and the log-property, the bound can be merged to the risk over observations:

$$\mathcal{G}_{rt}(\overline{a}) \leq D_{\mathrm{KL}}\left[\, F(\overline{o},\overline{s}|\overline{a}) || \, T(\overline{o},\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big]. \tag{4}$$

Additionally, if one assumes that $T(\overline{o}|\overline{s}) = F(\overline{o}|\overline{s})$, then restarting from Equation 4, one can show that the risk over states plus ambiguity is an upper bound of the expected free energy, i.e.,

$$\mathcal{G}_{rt}(\overline{a}) \leq D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{a}) || \, T(\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big] = \mathcal{C}_{RSA}(\overline{a}). \tag{5}$$

P

Once again, we keep following the proof presented in Appendix B of Parr et al. (2022) backward. Let us restart from Equation 4:

$$\mathcal{G}_{rt}(\overline{a}) \leq D_{\mathrm{KL}}\left[\, F(\overline{o},\overline{s}|\overline{a}) || \, T(\overline{o},\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big].$$

Then, using the definition of the KL-divergence, the linearity of expectation and the log-property, we can split the KL-divergence as follows:

$$\begin{aligned}\mathcal{G}_{rt}(\overline{a}) \leq\ & D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{a}) || \, T(\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big] \\ & + \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{o}|\overline{s})] - \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln T(\overline{o}|\overline{s})].\end{aligned}$$

Next, using our assumption that $T(\overline{o}|\overline{s}) = F(\overline{o}|\overline{s})$, we can get:

$$\begin{aligned}\mathcal{G}_{rt}(\overline{a}) \leq\ & D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{a}) || \, T(\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\big[H[F(\overline{o}|\overline{s})]\big] \\ & + \underbrace{\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{o}|\overline{s})] - \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln F(\overline{o}|\overline{s})]}_{=0},\end{aligned}$$

P

which simplifies to:

$$\mathcal{G}_{rt}(\overline{a}) \leq \underbrace{D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{a}) || \, T(\overline{s}|\overline{a})\right]}_{\text{risk over states}} + \underbrace{\mathbb{E}_{F(\overline{s}|\overline{a})}\left[H[F(\overline{o}|\overline{s})]\right]}_{\text{ambiguity}} = \mathcal{C}_{RSA}(\overline{a}).$$

Importantly, since the risk over states plus ambiguity is an upper bound of the EFE, minimising the upper bound will also minimise the EFE.

5.4.3 Expected energy vs entropy formulation

Finally, restarting from the risk over states plus ambiguity in Equation 5, one can demonstrate that:

$$\mathcal{G}_{rt}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = -\underbrace{H[F(\overline{s}|\overline{a})]}_{\text{entropy}} - \underbrace{\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln T(\overline{o},\overline{s}|\overline{a})]}_{\text{expected energy}} = \mathcal{C}_{3E}(\overline{a})$$

Thus, the expected energy vs entropy formulation can be recovered in this setup.

P

Restarting from Equation 5:

$$\mathcal{G}_{rt}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = D_{\mathrm{KL}}\left[\, F(\overline{s}|\overline{a}) || \, T(\overline{s}|\overline{a})\right] + \mathbb{E}_{F(\overline{s}|\overline{a})}\left[H[F(\overline{o}|\overline{s})]\right].$$

and using the KL-divergence and entropy definitions, we obtain:

$$\mathcal{G}_{rt}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = \mathbb{E}_{F(\overline{s}|\overline{a})}\left[\ln F(\overline{s}|\overline{a}) - \ln T(\overline{s}|\overline{a})\right] - \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}\left[\ln F(\overline{o}|\overline{s})\right].$$

Then, using our assumption that $F(\overline{o}|\overline{s}) = T(\overline{o}|\overline{s})$, as well as the linearity of expectation and the log-property, we get:

$$\begin{aligned}\mathcal{G}_{rt}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) &= \mathbb{E}_{F(\overline{s}|\overline{a})}\left[\ln F(\overline{s}|\overline{a}) - \ln T(\overline{s}|\overline{a})\right] - \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}\left[\ln T(\overline{o}|\overline{s})\right] \\ &= \mathbb{E}_{F(\overline{s}|\overline{a})}[\ln F(\overline{s}|\overline{a})] - \mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln T(\overline{o},\overline{s}|\overline{a})].\end{aligned}$$

Finally, recognizing the entropy definition, we obtain the desired result:

$$\mathcal{G}_{rt}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = -\underbrace{H[F(\overline{s}|\overline{a})]}_{\text{entropy}} - \underbrace{\mathbb{E}_{F(\overline{o},\overline{s}|\overline{a})}[\ln T(\overline{o},\overline{s}|\overline{a})]}_{\text{expected energy}} = \mathcal{C}_{3E}(\overline{a}).$$

6. Limitations

In the previous section, we defined the expected free energy as the risk over observations plus ambiguity, i.e., $\mathcal{C}_{ROA}(\overline{a})$, and showed that it is equal to the information gain and pragmatic value formulation, i.e., $\mathcal{C}_{IGPV}(\overline{a})$. Then, following the proof in Appendix B of Parr et al. (2022), the risk over states plus ambiguity $\mathcal{C}_{RSA}(\overline{a})$ was shown to be an upper bound of $\mathcal{C}_{ROA}(\overline{a})$. Finally, the entropy plus expected energy formulation $\mathcal{C}_{3E}(\overline{a})$ was derived from $\mathcal{C}_{RSA}(\overline{a})$. In summary:

$$\mathcal{G}_{rt}(\overline{a}) \triangleq \mathcal{C}_{ROA}(\overline{a}) = \mathcal{C}_{IGPV}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = \mathcal{C}_{3E}(\overline{a}). \tag{6}$$

Importantly, the proofs leading to equation (6) rely on the assumption that the likelihood of the forecast and target distributions are equal, i.e., $F(\overline{o}|\overline{s}) = T(\overline{o}|\overline{s})$. In the following subsections, we study the limitations of the formalism presented in Section 5.

6.1 Prior preferences over observations

In this section, we study the assumptions made in Section 5.4 and their consequences. For simplicity, we only consider the case where the time horizon h is equal to $t+1$, and where o_{t+1}, s_{t+1}, as well as a_t are discrete random variables. In this case, $\overline{o} = o_{t+1:h} = o_{t+1}$, $\overline{s} = s_{t+1:h} = s_{t+1}$, and similarly $\overline{a} = a_{t:h-1} = a_t$. Note, in Section 5.4, the EFE is defined as the risk over observations plus ambiguity, i.e.,

$$\mathcal{G}_{rt}(a_t) \triangleq \underbrace{D_{\text{KL}}\left[\, F(o_{t+1}|a_t) \,||\, T(o_{t+1}|a_t)\right]}_{\text{risk over observations}} + \underbrace{\mathbb{E}_{F(s_{t+1}|a_t)}\big[H[F(o_{t+1}|s_{t+1})]\big]}_{\text{ambiguity}} = \mathcal{C}_{ROA}(a_t).$$

6.1.1 The assumptions seemingly lead to an equation with no valid solution

In this section, we show that the assumptions made in Section 5, seemingly lead to an equation with no valid solution. First, note that in the active inference literature, the prior preferences over observations are defined as follows: $T(o_{t+1}|a_t) = \text{Cat}(o_{t+1}; \boldsymbol{C}_o)$. Additionally, recall that the proof in Section 5.4.2 relies on the assumption that $T(o_{t+1}|s_{t+1}) = F(o_{t+1}|s_{t+1})$, where the likelihood of the forecast distribution is also the likelihood of the generative model, i.e., $T(o_{t+1}|s_{t+1}) = F(o_{t+1}|s_{t+1}) = P(o_{t+1}|s_{t+1})$. In the active inference literature, the likelihood of the generative model is defined as: $P(o_{t+1}|s_{t+1}) = \text{Cat}(o_{t+1}|s_{t+1}; \boldsymbol{A})$. To sum up, we have $T(o_{t+1}|a_t) = \text{Cat}(o_{t+1}; \boldsymbol{C}_o)$ and $T(o_{t+1}|s_{t+1}) = \text{Cat}(o_{t+1}|s_{t+1}; \boldsymbol{A})$.

Using the sum and product rules of probability, we have:

$$T(o_{t+1}|a_t) = \sum_{s_{t+1}} T(o_{t+1}|s_{t+1}, a_t)T(s_{t+1}|a_t) \Leftrightarrow \boldsymbol{C}_o = \boldsymbol{A}\boldsymbol{C}_s, \tag{7}$$

where without loss of generality, we let $T(s_{t+1}|a_t) = \text{Cat}(s_{t+1}; \boldsymbol{C}_s)$. Importantly, the above assumes that:

$$T(o_{t+1}|s_{t+1}, a_t) = T(o_{t+1}|s_{t+1}),$$

which holds because in the target distribution, the observations are independent of the policy given the states, i.e., using the d-separation criteria one can show that: $o_{t+1} \perp\!\!\!\perp a_t \mid s_{t+1}$. Re-starting from (7), one can solve for $\boldsymbol{C}_s$ and get:

$$\boldsymbol{C}_o = \boldsymbol{A}\boldsymbol{C}_s \Leftrightarrow \boldsymbol{C}_s = \boldsymbol{A}^{-1}\boldsymbol{C}_o.$$

However, the above equation may not have any valid solution. For example, if:

$$\boldsymbol{A} = \begin{bmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C}_o = \begin{bmatrix} 0.8 \\ 0.2 \end{bmatrix}, \tag{8}$$

then one can show that:

$$\boldsymbol{A}^{-1} = \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{C}_s = \boldsymbol{A}^{-1}\boldsymbol{C}_o = \begin{bmatrix} 2 \\ -1 \end{bmatrix},$$

which is not a valid solution as $\boldsymbol{C}_s$ are the parameters of a categorical distribution. Thus, the elements of $\boldsymbol{C}_s$ should add up to one, and be between zero and one. Importantly, finding a renormalization $\bar{\boldsymbol{C}}_s$ of $\boldsymbol{C}_s$ is impossible. Indeed, if $\boldsymbol{A}$ is invertible, then the inverse is unique. Thus, $\boldsymbol{C}_s$ is the only matrix that satisfies: $\boldsymbol{C}_s = \boldsymbol{A}^{-1}\boldsymbol{C}_o$.

6.1.2 The class of valid prior preferences over observations

The problem described in the previous section occurs because we are defining two distributions over the random variable o_{t+1}, and these distributions are not compatible with each other. This indicates that in the setting of Section 5, one cannot define the prior preferences over observations arbitrarily. Instead, given the likelihood mapping $T(o_{t+1}|s_{t+1}) = \text{Cat}(o_{t+1}|s_{t+1}; \boldsymbol{A})$, one can only define prior preferences over states, i.e., $T(s_{t+1}|a_t) = \text{Cat}(s_{t+1}; \boldsymbol{C}_s)$, Then, the prior preferences over observations can be computed as follows:

$$T(o_{t+1}|a_t) = \sum_{s_{t+1}} T(o_{t+1}|s_{t+1}, a_t) T(s_{t+1}|a_t) \Leftrightarrow \boldsymbol{C}_o = \boldsymbol{A}\boldsymbol{C}_s. \tag{9}$$

While this tells us how the prior over observations can be computed, it does not characterise the class of all valid prior preferences over observations. We can characterise this class, since $n \times n$ matrices can be understood as linear transformations of the n-dimensional Euclidean space. We elaborate on this link to linear transformations in Appendix G. Since $\boldsymbol{A}$ is a linear transformation, we can better understand $\boldsymbol{C}_o = \boldsymbol{A}\boldsymbol{C}_s$, and which prior preferences over observations $\boldsymbol{C}_o$ are compatible with the likelihood mapping $\boldsymbol{A}$.

Note that $\boldsymbol{C}_o$ is a linear transformation of $\boldsymbol{C}_s$, where the transformation is defined by the elements of the matrix $\boldsymbol{A}$. Moreover, $\boldsymbol{C}_s$ are the parameters of a categorical distribution, which means that its elements are positive and sum to one. Geometrically, this means that $\boldsymbol{C}_s$ is in the 1-dimensional simplex of the Euclidean space. Additionally, since $\boldsymbol{A}$ defines the probability of each observation given each state, all the elements of $\boldsymbol{A}$ are positive and the columns of $\boldsymbol{A}$ sum up to one. In other words, the columns of $\boldsymbol{A}$ correspond to points on the 1-dimensional simplex of the Euclidean space.

Figure 1 illustrates the linear transformation corresponding to the $\boldsymbol{A}$ matrix of the previous section, c.f., Equation 8. Recall that the columns of $\boldsymbol{A}$ (i.e., $\vec{a}_1$ and $\vec{a}_2$) correspond to points on the 1-dimensional simplex (represented in blue on the left of Figure 1). Importantly, the standard basis vector $\vec{\imath}$ and $\vec{\jmath}$ are mapped by $\boldsymbol{A}$ to $\vec{a}_1$ and $\vec{a}_2$, respectively. While this is happening, the gray grid (Figure 1 left) will be squeezed into the red grid (Figure 1 right), and the 1-dimensional simplex represented by a long blue segment will be squeezed into a shorter blue segment. This short blue segment is the class of valid prior preferences over observations $\boldsymbol{C}_o$.

Indeed, the matrix $\boldsymbol{A}^{-1}$ performs the inverse linear transformation, i.e., $\boldsymbol{A}^{-1}$ transforms the red grid into the gray grid. Therefore, if a point is on the short blue segment in the right-hand-side of Figure 1, it will be mapped back to the original 1-dimensional simplex (the long blue segment). However, if a point starts on the blue dotted line (outside the short blue segment), it will be mapped outside of the original 1-dimensional simplex.

To conclude, the class of valid prior preferences over observations is the class of all vectors $\boldsymbol{C}_o$ that can be obtained as a linear combination of the columns of $\boldsymbol{A}$, where the weights of the linear combination are positive numbers between zero and one that sum up to one, i.e., all the vectors that satisfies the equation $\boldsymbol{C}_o = \boldsymbol{A}\boldsymbol{C}_s$. Geometrically, the class of valid prior preferences over observations corresponds to the short blue segment obtained by applying $\boldsymbol{A}$ to the 1-dimensional simplex.

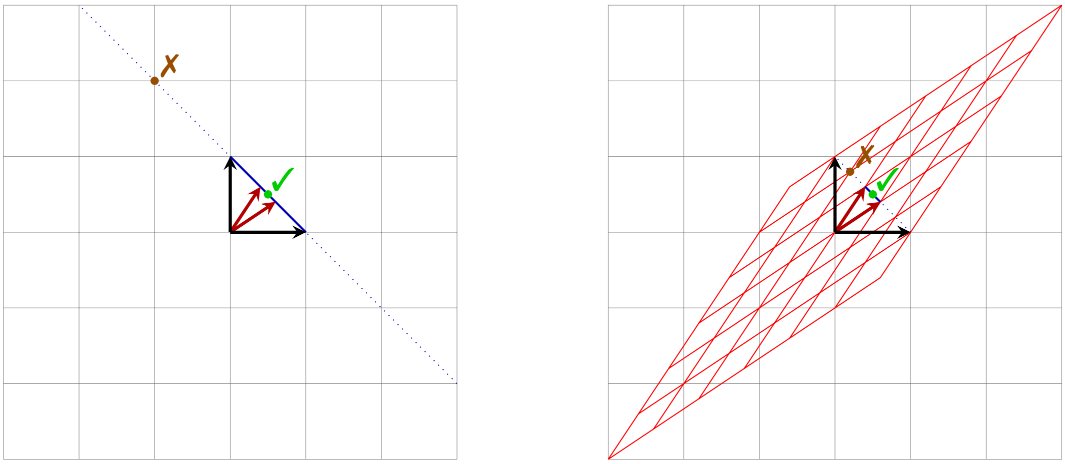

Figure 1: This figure illustrates the linear transformation corresponding to the $\boldsymbol{A}$ matrix in Equation 8. The gray grid represents a set of points in the Euclidean space on which the linear transformation corresponding to the $\boldsymbol{A}$ matrix will be applied. The red grid represents where the gray grid lands when applying this linear transformation. The long (blue) segment on the left-hand-side of the figure corresponds to the 1-dimensional simplex in which the prior preferences over states $\boldsymbol{C}_s$ lives. The short (blue) segment on the right-hand-side of the figure corresponds to where the 1-dimensional simplex lands when applying the linear transformation, i.e., the class of valid prior preferences over observations $\boldsymbol{C}_o$. The green point (✓ on the right-hand-side) corresponds to a point that lives on the projected simplex (i.e., the short blue segment). This point will be projected back to the original 1-dimensional simplex (i.e., the long blue segment) by $\boldsymbol{A}^{-1}$. The brown point (✗ on the right-hand-side) corresponds to a point that lives outside the short blue segment. This point will be projected outside of the long blue segment by $\boldsymbol{A}^{-1}$.

6.2 Justification of the expected free energy

In this section, we discuss the expected free energy justification presented in Appendix B of Parr et al. (2022). The appendix derives $\mathcal{C}_{RSA}(\overline{a})$ from the assumption that an agent aims to reach its prior preferences, i.e., $F(s_\tau) = T(s_\tau|a_{\tau-1})$ for some time step τ in the future. Importantly, this derivation provides a justification for $\mathcal{C}_{RSA}(\overline{a})$, i.e., an upper bound of $\mathcal{G}_{rt}(\overline{a})$ as defined in Section 5, but not for $\mathcal{G}_{rt}(\overline{a})$ itself. Thus, the expected free energy of Section 5, i.e., $\mathcal{C}_{ROA}(\overline{a})$, remains justified only by its intuitive definition. To resolve this issue, we could try to define the expected free energy as the risk over states plus ambiguity:

$$\mathcal{C}_{ROA}(\overline{a}) = \mathcal{C}_{IGPV}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a}) = \mathcal{C}_{3E}(\overline{a}) \triangleq \mathcal{G}_{rt}(\overline{a}).$$

In this case, we have a justification for the expected free energy, i.e., $\mathcal{C}_{RSA}(\overline{a})$, but $\mathcal{C}_{ROA}(\overline{a})$ and $\mathcal{C}_{IGPV}(\overline{a})$ are now lower bounds of $\mathcal{G}_{rt}(\overline{a})$. As minimizing a lower bound of the expected free energy does not imply minimizing the expected free energy, $\mathcal{C}_{ROA}(\overline{a})$ and $\mathcal{C}_{IGPV}(\overline{a})$ cannot be recovered in this setting. The results of the derivations presented in this paper are summarized in Table 1. Out of two potential definitions for the expected free energy, one is justified but can only recover two EFE formulations, and the other is not currently justified but can recover the four formulations.

	$\mathcal{C}_{IGPG}$	$\mathcal{C}_{RSA}$	$\mathcal{C}_{ROA}$	$\mathcal{C}_{3E}$	justified
risk over states vs ambiguity as definition	✗	✓	✗	✓	✓
risk over observations vs ambiguity as definition	✓	✓	✓	✓	✗

Table 1: This table summarises the results of the derivations of the expected free energy formulations, and whether a justification of the EFE has been provided in the active inference literature. The first row corresponds to the definition of Parr et al. (2022) where the expected free energy is defined as the risk over states vs ambiguity formulation, while the second row corresponds to another interpretation of Parr et al. (2022) where the expected free energy is defined as the risk over observations vs ambiguity formulation. Cells containing ✗ mean that the formulation cannot be recovered (or no justification is given), while cells containing ✓ correspond to the case where the formulation can be recovered (or a justification is given).

7. Discussion

In this paper, we presented a theoretical framework in which all four formulations of the expected free energy can be derived. However, when preferences over both observations and states are required, this framework restricts the class of valid prior preferences over observations. More precisely, only the prior preferences over observations that can be expressed in terms of prior preferences over states are valid in this framework. Therefore, in this situation, while defining prior preferences over observations, practitioners should be careful that these preferences are compatible with the likelihood mapping. Importantly, when preferences over states and observations are required, the prior preferences over observations can only be arbitrarily defined when the likelihood mapping is a permutation matrix. Another important special case arises when the likelihood mapping is an identity matrix. In this case, the bound is zero, and the risk over states is equal to the risk over observations, which means that the inequality in Equation (6) becomes an equality, i.e.,

$$\mathcal{G}_{rt}(\overline{a}) \triangleq \mathcal{C}_{ROA}(\overline{a}) = \mathcal{C}_{IGPV}(\overline{a}) = \mathcal{C}_{RSA}(\overline{a}) = \mathcal{C}_{3E}(\overline{a}). \tag{10}$$

Importantly, the issue of moving outside the simplex, is a problem of mathematical derivations, but does not imply that coherent implementations of active inference do not exist, they clearly do. Standard tabular implementations of active inference can employ preferences over observations or preferences over states, with the corresponding expected free energy formulation, and obtain a tenable planning system. However, because of the issues we highlight in this paper, the relationship between those different implementations is not always straightforward, i.e. our findings suggest that one cannot directly map between EFE definitions in these implementations. Additionally, difficulties would arise if one sought to develop more complex realisations of active inference in which both preferences / target distributions over observations and over states were utilised in the same framework.

The above discussion prompts an interesting question: *should the prior preferences be defined in terms of states or observations?* As previously mentioned, prior preferences over states are always valid. However, when preferences over observations and over states are being required, preferences over observations are not always valid. Thus, in this respect, it is safer to define preferences over states. Unfortunately, this may not be possible in some contexts, for example, in deep active inference, the latent space is learned by optimizing the parameters of deep neural networks. The resulting latent representation may be entangled and lack interpretability, making the prior preference over states difficult to specify.

In many circumstances, it may be more practical to define preferences over observations. This is because a full generative model is required before states and particularly preferred states are available, and this may not initially be available, perhaps requiring previous structure learning.

Additionally, since $\mathcal{C}_{RSA}(\overline{a})$ is an upper bound of the expected free energy, i.e., $\mathcal{G}_{rt}(\overline{a}) = \mathcal{C}_{ROA}(\overline{a}) \leq \mathcal{C}_{RSA}(\overline{a})$, another interesting (empirical) question goes as follows: *is there any behavioural differences that occur when minimizing $\mathcal{C}_{RSA}(\overline{a})$ instead of $\mathcal{C}_{ROA}(\overline{a})$?* Also, when the prior preferences are learned using a Dirichlet prior (Sajid, Ball, Parr, & Friston, 2021): *are the learned prior preferences compatible with the likelihood mapping?* Although these questions hold interest, they fall beyond the scope of this paper.

In contrast, our framework provides an answer to the following question: *what is the relationship between the variational and expected free energies?* Put simply, the variational free energy is concerned with inferring the latent states, i.e., computing posterior beliefs, while the expected free energy is concerned with planning, i.e., estimating the loss associated with performing a particular policy. Additionally, the variational free energy is defined in terms of the generative model and variational distribution, while the expected free energy is based on the forecast and target distribution.

Although a correspondence was posited in the older literature (e.g. Parr and Friston (2019)), it is now recognised that the variational and expected free energy are very different. Millidge, Tschantz, and Buckley (2021) was an early paper to highlight this difference, and it is also discussed in Champion et al. (2023), (four sentences following eqn 5, paragraphs following figure 12 and eqn 15); Champion, Grześ, Bonheme, and Bowman (2024); and K. Friston, Da Costa, Sajid, et al. (2023) (paragraph following equation 39). Although, of course, inferring the forecast distribution, upon which the expected free energy is calculated, does involve minimisation of the variational free energy.

7.1 Foundations of the EFE[4]

This paper has considered how to relate key formulations of the Expected Free Energy (EFE) that can be found in the literature and provided a partial unification of those definitions. This was done

[4]An anonymous reviewer suggested this treatment of K. Friston, Da Costa, Sajid, et al. (2023); K. Friston, Da Costa, Sakthivadivel, et al. (2023).

through direct manipulation of the EFE definitions. K. Friston, Da Costa, Sajid, et al. (2023); K. Friston, Da Costa, Sakthivadivel, et al. (2023) has explored an alternative strategy for providing a foundation for the EFE. This starts with the free energy principle (FEP) viewed from a (physics) statistical mechanics perspective and investigates in what way the EFE can be seen to arise from the probability density over paths of a specific class of random dynamical systems. Working from the free energy, (K. Friston, Da Costa, Sajid, et al., 2023) identify conditions under which Equation 10 holds. The key condition that (K. Friston, Da Costa, Sajid, et al., 2023) propose is that active inference is conservative (i.e. non-dissipative), implying that random fluctuations are infinitesimally small and uncertainty in sensing the environment is minimal. Focussing on the risk over observations plus ambiguity formulation of the EFE, the near absence of random fluctuations means that the ambiguity term is redundant and the risk over states and observations are the same. This is because, the lack of random fluctuations, mean there is little if any loss of information when moving between states and observations (K. Friston, Da Costa, Sajid, et al., 2023). For a discrete state space, the corresponding condition requires the likelihood mapping (A) to be a permutation matrix (K. Friston, Da Costa, Sajid, et al., 2023). As an example, if A is an identity matrix, $F(o|a) = F(s|a)$ and $T(o) = T(s)$.

Of course, though, the risk over observations plus ambiguity formulation of the EFE does contain an ambiguity term. K. Friston, Da Costa, Sajid, et al. (2023) argue that "Expected free energy supplements risk with ambiguity; namely, the expected inaccuracy or negative sensory likelihood given an autonomous path. Clearly, this is redundant as a description of conservative particles that — by definition — have minimal ambiguity (because there are no random sensory fluctuations). However, in applications of the FEP, the ambiguity term is retained to ensure particle or agent seeks out unambiguous regimes of state space; thereby, evincing the behaviour of conservative particles".

The findings in (K. Friston, Da Costa, Sajid, et al., 2023) provide a link between the expected free energy and the variational free energy in a well-behaved context (of conservative systems), enabling a simple unification of expected free energy definitions. However, this still leaves open the relationship between EFE definitions, as well as between free energy and EFE, in more general contexts, e.g. in the full context of Partially Observable Markov Decision Process, in which the prediction of observations from states contains uncertainty, reflecting inherent uncertainty in the world. It is to this broader context that our, if only partial, unifications of EFE formulations have contributed.

K. Friston, Da Costa, Sajid, et al. (2023) may raise the possibility that the marginal likelihood $P(o|m)$ (where m is a model) and the target distribution $T(o)$ could become equal in a specific context. This is because in the K. Friston, Da Costa, Sajid, et al. (2023) framework, the target over observations is defined to be a non-equilibrium steady state solution of a dynamical system. The marginal likelihood could be thought to describe the (marginal) probability that a state is in this steady state solution. Thus, in this particular situation, the marginal likelihood may be the target. However, an active inference system is not always at a non-equilibrium steady state, thus, equality between these two distributions is not guaranteed, and they have accordingly been distinguished in this paper[5].

8. Conclusion

This paper aimed to revisit the expected free energy definition, as well as the problem of deriving its four formulations found in the literature, i.e., the unification problem. When the expected free energy is defined as the risk over observations plus ambiguity, all formulations can be recovered, and can therefore be used in practice. However, an important contribution of this paper was to show that some prior preferences over observations are incompatible with the likelihood mapping. Thus, when preferences over both states and observations are required, we are left with a dilemma, either the modeller has to carefully pick the prior preferences of the agent to avoid any conflict, or we have to let go of the theoretical connection between the four formulations.

Another issue is the absence of a justification for the risk over observations plus ambiguity formulation. While there exists a justification for the risk over states plus ambiguity formulation, justifying a lower bound is not enough to justify the expected free energy. Therefore, future research should focus on finding a derivation of the risk over observations plus ambiguity from first principles. Importantly, while the risk over states plus ambiguity is justified, this definition of the expected free energy does not allow us to recover the four formulations. Thus, it does not constitute a valid solution to the unification problem.

Note, we only studied two possible definitions of the expected free energy. An alternative set of proofs and/or another factorization of the forecast and target definitions may allow us to recover all four decompositions, while also removing the conflict between prior preferences and likelihood. However, testing all possible factorizations and proofs is outside the scope of this paper.

Finally, this paper provides a solid foundation for future research, especially in the context of deep active inference. Indeed, this paper clarifies the expected free energy definition, but unfortunately, it

[5]this line of argument was indicated by a reviewer.

does explain how to compute it using deep neural networks. Thus, additional research is required to implement and empirically evaluate the proposed expected free energy definition.

References

Çatal, O., Verbelen, T., Nauta, J., Boom, C. D., & Dhoedt, B. (2020). Learning perception and planning with deep active inference. In *2020 IEEE international conference on acoustics, speech and signal processing, ICASSP 2020, barcelona, spain, may 4-8, 2020* (pp. 3952–3956). IEEE. Retrieved from https://doi.org/10.1109/ICASSP40776.2020.9054364 doi: 10.1109/ICASSP40776.2020.9054364

Champion, T., Bowman, H., & Grześ, M. (2022). Branching time active inference: Empirical study and complexity class analysis. *Neural Networks*, *152*, 450-466. Retrieved from https://www.sciencedirect.com/science/article/pii/S0893608022001824 doi: https://doi.org/10.1016/j.neunet.2022.05.010

Champion, T., Da Costa, L., Bowman, H., & Grześ, M. (2022). Branching time active inference: The theory and its generality. *Neural Networks*, *151*, 295-316. Retrieved from https://www.sciencedirect.com/science/article/pii/S0893608022001149 doi: https://doi.org/10.1016/j.neunet.2022.03.036

Champion, T., Grześ, M., Bonheme, L., & Bowman, H. (2023). *Deconstructing deep active inference.* Retrieved from https://arxiv.org/abs/2303.01618

Champion, T., Grześ, M., Bonheme, L., & Bowman, H. (2024). Deconstructing deep active inference: a contrarian information gatherer. *Neural Computation*.

Champion, T., Grześ, M., & Bowman, H. (2022a, 09). Branching Time Active Inference with Bayesian Filtering. *Neural Computation*, *34*(10), 2132-2144. Retrieved from https://doi.org/10.1162/neco_a_01529 doi: 10.1162/neco_a_01529

Champion, T., Grześ, M., & Bowman, H. (2022b). *Multi-modal and multi-factor branching time active inference.* arXiv. Retrieved from https://arxiv.org/abs/2206.12503 doi: 10.48550/ARXIV.2206.12503

Cullen, M., Davey, B., Friston, K. J., & Moran, R. J. (2018). Active inference in openai gym: A paradigm for computational investigations into psychiatric illness. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(9), 809 - 818. Retrieved from http://www.sciencedirect.com/science/article/pii/S2451902218301617 (Computational Methods and Modeling in

Psychiatry) doi: https://doi.org/10.1016/j.bpsc.2018.06.010

FitzGerald, T. H. B., Dolan, R. J., & Friston, K. (2015). Dopamine, reward learning, and active inference. *Frontiers in Computational Neuroscience*, *9*, 136. Retrieved from https://www.frontiersin.org/article/10.3389/fncom.2015.00136 doi: 10.3389/fncom.2015.00136

Fountas, Z., Sajid, N., Mediano, P. A. M., & Friston, K. J. (2020). Deep active inference agents using Monte-Carlo methods. In H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, & H. Lin (Eds.), *Advances in neural information processing systems 33: Annual conference on neural information processing systems 2020, neurips 2020, december 6-12, 2020, virtual.* Retrieved from https://proceedings.neurips.cc/paper/2020/hash/865dfbde8a344b44095495f3591f7407-Abstract.html

Fox, C. W., & Roberts, S. J. (2012, Aug 01). A tutorial on variational Bayesian inference. *Artificial Intelligence Review*, *38*(2), 85-95. Retrieved from https://doi.org/10.1007/s10462-011-9236-8 doi: 10.1007/s10462-011-9236-8

Friston, K., Da Costa, L., Hafner, D., Hesp, C., & Parr, T. (2021, 03). Sophisticated Inference. *Neural Computation*, *33*(3), 713-763. Retrieved from https://doi.org/10.1162/neco_a_01351 doi: 10.1162/neco_a_01351

Friston, K., Da Costa, L., Sajid, N., Heins, C., Ueltzhöffer, K., Pavliotis, G. A., & Parr, T. (2023). The free energy principle made simpler but not too simple. *Physics Reports*, *1024*, 1–29.

Friston, K., Da Costa, L., Sakthivadivel, D. A., Heins, C., Pavliotis, G. A., Ramstead, M., & Parr, T. (2023). Path integrals, particular kinds, and strange things. *Physics of Life Reviews*.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., Doherty, J. O., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience & Biobehavioral Reviews*, *68*, 862 - 879. doi: https://doi.org/10.1016/j.neubiorev.2016.06.022

Friston, K. J., Parr, T., & de Vries, B. (2017, 12). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, *1*(4), 381-414. Retrieved from https://doi.org/10.1162/NETN_a_00018 doi: 10.1162/NETN_a_00018

Hafner, D., Ortega, P. A., Ba, J., Parr, T., Friston, K., & Heess, N. (2022). *Action and perception as divergence minimization.*

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295 - 1306. Retrieved from http://www.sciencedirect.com/science/article/pii/S0042698908004380 (Visual Attention: Psychophysics, electrophysiology and neuroimaging) doi:

https://doi.org/10.1016/j.visres.2008.09.007

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *Physical review*, *106*(4), 620.

Jaynes, E. T. (1957b). Information theory and statistical mechanics. ii. *Physical review*, *108*(2), 171.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques.* MIT press.

Kschischang, F. R., Frey, B. J., & Loeliger, H.-A. (2001). Factor graphs and the sum-product algorithm. *IEEE Transactions on information theory*, *47*(2), 498–519.

Millidge, B. (2019). *Combining active inference and hierarchical predictive coding: A tutorial introduction and case study.* Retrieved from `https://doi.org/10.31234/osf.io/kf6wc`

Millidge, B., Tschantz, A., & Buckley, C. L. (2021). Whence the expected free energy? *Neural Computation*, *33*(2), 447–482.

Parr, T., & Friston, K. J. (2019). Generalised free energy and active inference. *Biological cybernetics*, *113*(5), 495–513.

Parr, T., Markovic, D., Kiebel, S. J., & Friston, K. J. (2019). Neuronal message passing using mean-field, bethe, and marginal approximations. *Scientific reports*, *9*(1), 1889.

Parr, T., Pezzulo, G., & Friston, K. J. (2022). *Active inference: the free energy principle in mind, brain, and behavior.* MIT Press.

Sajid, N., Ball, P. J., Parr, T., & Friston, K. J. (2021, 03). Active Inference: Demystified and Compared. *Neural Computation*, *33*(3), 674-712. Retrieved from `https://doi.org/10.1162/neco_a_01357` doi: 10.1162/neco_a_01357

Sancaktar, C., van Gerven, M. A. J., & Lanillos, P. (2020). End-to-end pixel-based deep active inference for body perception and action. In *Joint IEEE 10th international conference on development and learning and epigenetic robotics, icdl-epirob 2020, valparaiso, chile, october 26-30, 2020* (pp. 1–8). IEEE. Retrieved from `https://doi.org/10.1109/ICDL-EpiRob48136.2020.9278105` doi: 10.1109/ICDL-EpiRob48136.2020.9278105

Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H. B., Kronbichler, M., & Friston, K. (2018). Computational mechanisms of curiosity and goal-directed exploration. *bioRxiv*. Retrieved from `https://www.biorxiv.org/content/early/2018/09/07/411272` doi: 10.1101/411272

Winn, J., & Bishop, C. (2005). Variational message passing. *Journal of Machine Learning Research*, *6*, 661-694.

Appendix A: comments, questions, answers, and proofs

In this paper, we used the following blocks:

C A comment.

Q A question.

A An answer.

P A proof.

Appendix B: important properties

In this paper, we made extensive use of four basic properties. The first is the sum-rule of probability. Given two sets of random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$, the sum-rule of probability states that:

$$P(\boldsymbol{Y}) = \int_{\boldsymbol{X}} P(\boldsymbol{X}, \boldsymbol{Y}) \, d\boldsymbol{X}. \tag{11}$$

The sum-rule can then be used to sum out random variables from a joint distribution. The second property is called the product-rule of probability, and can be used to split a joint distribution into conditional distributions. Given two sets of random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$, the product-rule of probability states that:

$$P(\boldsymbol{X}, \boldsymbol{Y}) = P(\boldsymbol{X}|\boldsymbol{Y})P(\boldsymbol{Y}). \tag{12}$$

Next, if $\boldsymbol{a}$ and $\boldsymbol{b}$ are two real numbers, then a relevant property of logarithm is the following:

$$\ln(\boldsymbol{a} \times \boldsymbol{b}) = \ln(\boldsymbol{a}) + \ln(\boldsymbol{b}). \tag{13}$$

Put simply, this allows us to turn the logarithm of a product into a sum of logarithms, and we will refer to the above equation as the "log-property". Finally, the last property is called the linearity of expectation. Given a random variable $\boldsymbol{X}$, and two real numbers $\boldsymbol{a}$ and $\boldsymbol{b}$, the linearity of expectation states that:

$$\mathbb{E}[\boldsymbol{a}\boldsymbol{X} + \boldsymbol{b}] = \boldsymbol{a}\mathbb{E}[\boldsymbol{X}] + \boldsymbol{b}, \tag{14}$$

where the expectation is w.r.t. the marginal distribution over $\boldsymbol{X}$, i.e., $P(\boldsymbol{X})$.

Appendix C: d-separation criterion

Given a Bayesian network, the d-separation criterion provides a bridge between the graph topology and the independence assumptions holding within the Bayesian network. Let $\mathcal{G}_{rt} = (\mathcal{V}, \mathcal{E})$ be a directed graph corresponding to a Bayesian network, where $\mathcal{V}$ and $\mathcal{E}$ are the graph's vertices and edges, respectively. A trail is a sequence of vertices $(V_1, V_2, ..., V_k)$ such that there is an edge $V_i \rightarrow V_{i+1}$ or $V_{i+1} \rightarrow V_i$ for all $i \in \{1, ..., k-1\}$. Intuitively, trails connect two variables V_1 and V_k; conceptually if a trail is blocked, then V_1 does not provide any new information about V_k through this trail. The notion of blocked trail is based on colliders:

Definition 1 (Collider) *Within a trail $(V_1, V_2, ..., V_k)$, a collider is a vertex V_j s.t. $V_{j-1} \rightarrow V_j \leftarrow V_{j+1}$.*

Definition 2 (Blocked trail) *Given a set of vertices $\boldsymbol{S} \subseteq \mathcal{V}$, and two vertices $V_1, V_2 \in \mathcal{V}$, we say that a trail between V_1 and V_2 is blocked by $\boldsymbol{S}$ if at least one node of the trail is a collider not in $\boldsymbol{S}$ and with no descendants in $\boldsymbol{S}$, or at least one vertex of the trail that is not a collider is in $\boldsymbol{S}$.*

Importantly, two vertices in the graph can be connected through multiple trails; if all trails between these two vertices are blocked, then we say that these vertices are d-separated. This can be generalized to sets of vertices as shown below:

Definition 3 (D-separated) *Given a set of vertices $\boldsymbol{S} \subseteq \mathcal{V}$, and two vertices $V_1, V_2 \in \mathcal{V}$, we say that V_1 and V_2 are d-separated by $\boldsymbol{S}$ if all trails between V_1 and V_2 are blocked.*

C

Given three sets of vertices $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{S} \subseteq \mathcal{V}$. We say that $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are d-separated by $\boldsymbol{S}$, if each node in $\boldsymbol{V}_1$ is d-separated from all nodes in $\boldsymbol{V}_2$ given the nodes in $\boldsymbol{S}$.

Finally, the d-separation theorem states that: if two sets of vertices are d-separated in the graph, then the associated random variables are conditionally independent, i.e.,

Theorem 4 (d-separation and independence) *Given three sets $\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{S} \subseteq \mathcal{V}$. If $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are d-separated by $\boldsymbol{S}$, then $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are conditionally independent given $\boldsymbol{S}$, i.e., $\boldsymbol{V}_1 \perp\!\!\!\perp \boldsymbol{V}_2 \mid \boldsymbol{S}$.*

In practice, the d-separation theorem is used on the generative model's graph. The goal is to know whether a conditional assumption holds, i.e., whether $\boldsymbol{V}_1 \perp\!\!\!\perp \boldsymbol{V}_2 \mid \boldsymbol{S}$ holds. If it does, then: $P(\boldsymbol{V}_1 | \boldsymbol{S}, \boldsymbol{V}_2) = P(\boldsymbol{V}_1 | \boldsymbol{S})$.

Appendix D: Preferences in Likelihood and Posterior

The enforcement of preferences in the prior over observations, which is common in the Active Inference literature, e.g., Parr and Friston (2019) and Parr et al. (2022), has implications for key derivations involving those observations. One often wants to take a joint over observations and states, e.g., $P(o_\tau, s_\tau)$, and apply the product rule to obtain a marginal and a conditional distribution. However, this marginal might become the preferences. So, the question is how to do this in a fashion that does not have unwanted impact on the rest of the theory.

As part of these derivations, one might equate the likelihood with preferences to the standard likelihood, i.e., $P(o_\tau|s_\tau, C) = P(o_\tau|s_\tau)$, an equality that we have in our target distribution (see Section 5.3 "Target distribution", which contains the definition). This does not cause problems for the theory. However, if one adds the equating of a posterior with preferences and a standard posterior, i.e., $P(s_\tau|o_\tau, C) = P(s_\tau|o_\tau)$, the theory "folds in on itself". This needs to be kept in mind in future work, but fortunately only seems to arise in earlier publications, e.g. Parr and Friston (2019). For example, in Parr and Friston (2019), there is a derivation from equation (9) to (10), which includes the following sub-derivation:

$$\ln P(o_\tau, s_\tau) = \ln P(s_\tau|o_\tau) + \ln P(o_\tau)$$

To simplify, we analogously take this derivation to be:

$$P(o_\tau, s_\tau) = P(s_\tau|o_\tau)P(o_\tau).$$

In equation (10), $P(o_\tau)$ is referred to as the preferences. So, the derivation is actually:

$$P(o_\tau, s_\tau) = P(s_\tau|o_\tau)P(o_\tau|C).$$

Even allowing for overloading to change the meaning of distributions, we are unclear how this derivation is made without consequences for the Active Inference theory. There are two possibilities for this derivation:

1. Starting with simple joint, i.e. $P(o_\tau, s_\tau)$; or,

2. Starting with preferences joint, i.e. $P(o_\tau, s_\tau|C)$.

Note, from what is written in the paper, we do not know which of these is correct, because of the overloading. So, let's try to investigate the derivation from both of these starting points.

1. Starting with simple joint

This will give us:

$$P(o_\tau, s_\tau) = P(s_\tau|o_\tau)P(o_\tau),$$

where we can only get to the $P(o_\tau|C)$ we need, by requiring that $P(o_\tau|C) = P(o_\tau)$. However, there is no reason why the preferences over observations should be the same as a marginal likelihood over observations for all τ. So, starting from a simple joint, constrains Active Inference in an unwanted manner, i.e., $P(o_\tau|C)$ must be equal to $P(o_\tau)$.

2. Starting with preferences joint

This will give us:

$$P(o_\tau, s_\tau|C) = P(s_\tau|o_\tau, C)P(o_\tau|C),$$

but the problem now is going from $P(s_\tau|o_\tau, C)$ to $P(s_\tau|o_\tau)$. There seems to be no alternative to requiring these to be, by assumption, equal, in a similar way to what is done with the likelihood. That is, we are left with the following two relationships:

$$P(o_\tau|s_\tau, C) = P(o_\tau|s_\tau) \quad \text{and} \quad P(s_\tau|o_\tau, C) = P(s_\tau|o_\tau).$$

Unfortunately, if both of these are assumed, the active inference theory is significantly constrained. To formalise this, we drop the τ for conciseness and one can reason as follows. Assumptions:

a. $P(o|s, C) = P(o|s)$ [the likelihood constraint]

b. $P(s|o, C) = P(s|o)$ [the posterior constraint]

Proposition D.1:

Given $P(s) = \text{Cat}(s; D)$ and $P(s|C) = \text{Cat}(s; C_s)$, for any observation, o, the following holds:

$$P(s|o) = P(s|o, C) \Rightarrow P(s) = P(s|C) \Rightarrow D = C_s,$$

that is:

$$P(s|o) = P(s|o, C) \Rightarrow D = C_s.$$

Thus, if one assumes the posterior constraint, as well as the likelihood constraint, the prior over states and the prior preferences over states must be the same.

P

Starting from the posteriors, by Bayes' theorem, we have:

(a)

$$P(s|o) = \frac{P(o|s)P(s)}{P(o)};$$

(b)

$$P(s|o, C) = \frac{P(o|s, C)P(s|C)}{P(o|C)}.$$

P

Then, we can start with the posterior constraint and reason as follows:

$$P(s|o) = P(s|o, C) \tag{15}$$

$$\Leftrightarrow \{ \text{ from (a) and (b) } \}$$

$$\frac{P(o|s)P(s)}{P(o)} = \frac{P(o|s, C)P(s|C)}{P(o|C)} \tag{16}$$

$$\Leftrightarrow \{ \text{ from the likelihood constraint } \}$$

$$\frac{P(o|s)P(s)}{P(o)} = \frac{P(o|s)P(s|C)}{P(o|C)} \tag{17}$$

$$\Leftrightarrow \{ \text{ cancelling } P(o|s) \}$$

$$\frac{P(s)}{P(o)} = \frac{P(s|C)}{P(o|C)} \tag{18}$$

$$\Leftrightarrow \{ \text{ rearranging } \}$$

$$P(s) = \frac{P(o)}{P(o|C)} P(s|C) \tag{19}$$

$$\Leftrightarrow \{ \text{ for all } o, \frac{P(o)}{P(o|C)} \text{ is a scalar } \}$$

$$P(s) \propto P(s|C) \tag{20}$$

$$\Leftrightarrow \{ \text{ since both } P(s) \text{ and } P(s|C) \text{ are distributions } \}$$

$$P(s) = P(s|C) \tag{21}$$

$$\Leftrightarrow \{ \text{ definitions of priors } \}$$

$$D = C_s \tag{22}$$

Thus, proposition D.1 shows that care is required when preferences over observations are included in the Active Inference theory. Fortunately, though, as previously indicated, more recent treatments of the theory, e.g. in Parr et al. (2022), do not equate the posterior with preferences to the standard posterior.

Appendix E: Bayes Theorem over Multiple Distributions

A further potential issue that can be observed in the literature is application of Bayes theorem across different types of distributions. For example, in equation (11) in Parr and Friston (2019), one is moving liberally between P's and Q's. Thus, Bayes theorem is being applied between factors of different distributions, i.e. the posterior with preferences $P(s_\tau|o_\tau, C)$ and factors from what seems to be the variational distribution, for example: $Q(o_\tau|\pi)$ and $Q(s_\tau|\pi)$. As Bayes' theorem is a corollary of the product rule of probability, it cannot be used between factors of different types of distribution.

Appendix F: Similarity to derivations in Parr et al. (2022)

We show here that the derivations contained in the expected energy proof in Section 5.1 (see P box) correspond to derivations in Parr et al. (2022). As previously indicated, the following derivation can be found in Appendix B, Section B.2.5 Expected Free Energy of Parr et al. (2022):

$$\begin{aligned}
&\mathbb{E}_{Q(s_\tau|\pi)}[H[P(o_\tau|s_\tau)]] + D_{\mathrm{KL}}\left[\, Q(s_\tau|\pi) ||\, P(s_\tau|C)\right] \\
&\quad = \mathbb{E}_{Q(s_\tau|\pi)}[H[P(o_\tau|s_\tau)]] + D_{\mathrm{KL}}\left[\, Q(s_\tau|\pi) ||\, {\color{red}P(s_\tau|C)}\right] \\
&\qquad + \underbrace{\mathbb{E}_{Q(s_\tau|\pi)P(o_\tau|s_\tau)}[\ln P(o_\tau|s_\tau)] - \mathbb{E}_{Q(s_\tau|\pi)P(o_\tau|s_\tau)}[\ln {\color{red}P(o_\tau|s_\tau)}]}_{=0} \\
&\quad = \mathbb{E}_{Q(s_\tau|\pi)}[H[P(o_\tau|s_\tau)]] + D_{\mathrm{KL}}\left[\, Q(o_\tau, s_\tau|\pi) ||\, {\color{red}P(o_\tau, s_\tau|C)}\right] \qquad \text{(B.27)} \\
&\quad = \mathbb{E}_{Q(s_\tau|\pi)}[H[P(o_\tau|s_\tau)]] + D_{\mathrm{KL}}\left[\, Q(o_\tau|\pi) ||\, {\color{red}P(o_\tau|C)}\right] \\
&\qquad + \mathbb{E}_{Q(o_\tau|\pi)}[D_{\mathrm{KL}}\left[\, Q(s_\tau|o_\tau, \pi) ||\, {\color{red}P(s_\tau|o_\tau, C)}\right]] \\
&\quad \geq \mathbb{E}_{Q(s_\tau|\pi)}[H[P(o_\tau|s_\tau)]] + D_{\mathrm{KL}}\left[\, Q(o_\tau|\pi) ||\, P(o_\tau|C)\right] = G(\pi)
\end{aligned}$$

This suggests a derivation of the following kind is part of B.27 (steps going from the 2nd and 3rd lines to the 4th):

$$\ln P(s_\tau|C) + \ln P(o_\tau|s_\tau) = \ln P(s_\tau|C) + \ln P(o_\tau|s_\tau, C) = \ln\left[P(s_\tau|C)P(o_\tau|s_\tau, C)\right] = \ln P(o_\tau, s_\tau|C),$$

where, in particular, $P(o_\tau|s_\tau, C) = P(o_\tau|s_\tau)$, which is the constraint that the likelihood is shared, i.e. the likelihood over preferences is equal to the standard generative model likelihood. And also, a derivation of the following kind (steps going from 4th to 5th and 6th lines):

$$\ln P(o_\tau, s_\tau|C) = \ln\left[P(o_\tau|C)P(s_\tau|o_\tau, C)\right] = \ln P(o_\tau|C) + \ln P(s_\tau|o_\tau, C).$$

If one maps the notation of Parr et al. (2022) to that of our paper, one would have the following, where our notation is on the right:

$$
\begin{aligned}
P(o_\tau, s_\tau|C) &= T(o_\tau, s_\tau)\\
P(s_\tau|C) &= T(s_\tau)\\
P(o_\tau|C) &= T(o_\tau)\\
P(s_\tau|o_\tau, C) &= T(s_\tau|o_\tau)\\
P(o_\tau|s_\tau) &= F(o_\tau|s_\tau)\\
P(o_\tau|s_\tau, C) &= T(o_\tau|s_\tau) \triangleq F(o_\tau|s_\tau) \triangleq P(o_\tau|s_\tau)
\end{aligned}
$$

and where the last line contains the constraint that the target likelihood is equal to the standard (generative model) likelihood, which we highlighted earlier for the likelihood over preferences. If we adapt the above derivations to our notation, one obtains,

$$\ln T(s_\tau) + \ln F(o_\tau|s_\tau) = \ln T(s_\tau) + \ln T(o_\tau|s_\tau) = \ln [T(s_\tau)T(o_\tau|s_\tau)] = \ln T(o_\tau, s_\tau),$$

and

$$\ln T(o_\tau, s_\tau) = \ln [T(o_\tau)T(s_\tau|o_\tau)] = \ln T(o_\tau) + \ln T(s_\tau|o_\tau).$$

In the P box of section 5.1 of the main body of the paper, we use the first of these, but in the opposite direction, i.e.

$$\ln T(o_\tau, s_\tau) = \ln [T(s_\tau)T(o_\tau|s_\tau)] = \ln T(s_\tau) + \ln T(o_\tau|s_\tau) = \ln T(s_\tau) + \ln F(o_\tau|s_\tau),$$

where we use the relationship/constraint $T(o_\tau|s_\tau) = F(o_\tau|s_\tau) = P(o_\tau|s_\tau)$, which, as previously emphasized, is an exact analogue of the constraint $P(o_\tau|s_\tau, C) = P(o_\tau|s_\tau)$ used in the appendix of Parr et al. (2022).

Appendix G: Geometric interpretation of linear transformation

In Section 6.1.2, we discuss how an $n \times n$ matrix can be viewed as linearly transforming an n-dimensional Euclidean space. We illustrate this idea further here. To understand this link to linear transformations, let's take the following 2×2 matrix as an example:

$$\boldsymbol{B} = \begin{bmatrix} | & | \\ \vec{b}_1 & \vec{b}_2 \\ | & | \end{bmatrix},$$

and let:

$$\vec{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and: } \vec{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

be the two standard basis vectors. One can see that: $\vec{b}_1 = \boldsymbol{B}\vec{i}$ and $\vec{b}_2 = \boldsymbol{B}\vec{j}$. In other words, the matrix $\boldsymbol{B}$ transforms the vector $\vec{i}$ into the vector $\vec{b}_1$, and the vector $\vec{j}$ into the vector $\vec{b}_2$. More generally, the matrix $\boldsymbol{B}$ transforms each vector $\vec{x}$ into a vector $\vec{y} = \boldsymbol{B}\vec{x}$. Geometrically, this can be understood as mapping each point of the Euclidean space $\vec{x}$, to another point $\vec{y}$ in the space. For example, the transformation corresponding to the following matrix:

$$\boldsymbol{B} = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \end{bmatrix},$$

is illustrated in Figure 2. Importantly, since the transformation is linear, the lines of the grid in Figure 2 remain parallel and equally spaced. Thus, knowing where $\vec{i}$ and $\vec{j}$ land under the transformation $\boldsymbol{B}$ is enough to know where all the other points on the grid land, i.e., $\vec{i}$ and $\vec{j}$ defines one parallelogram and all the others (parallelograms) are obtained by copy and pasting this parallelogram along the transformed axes.

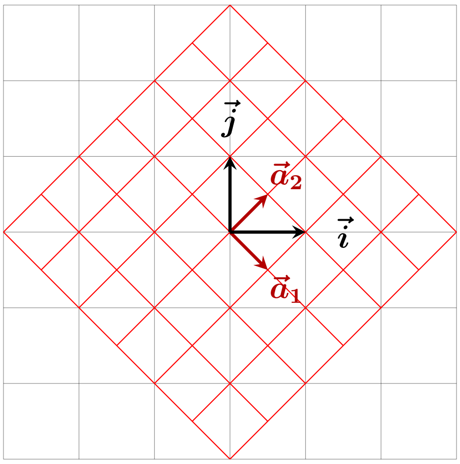

Figure 2: This figure illustrates how matrices can be seen as linear transformations of the Euclidean space. The example taken is a matrix that rotates the Euclidean space by 45 degrees clockwise and scales each axis by a factor of 0.5.