Free energy minimization and information gain: The devil is in the details

Commentary on Friston, K., Rigoli, F., Ognibene, D., Mathys, C., FitzGerald, T., and Pezzulo, G. (2015). Active Inference and epistemic value.

Johan Kwisthout and Iris van Rooij

Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour

Abstract

Contrary to Friston's previous work, this paper describes free energy minimization using categorical probability distributions over discrete states. This alternative mathematical framework exposes a fundamental, yet unnoticed challenge for the free energy principle. When considering discrete state spaces one must specify their granularity, as the amount of information gain is defined over this state space. The more detailed this state space, the lower the precision of the predictions will be, and consequently, the higher the prediction errors. Hence, an optimal trade-off between precision and detail is needed, and we call for incorporating this aspect in the free energy principle.

> *"If you take care of the small things,*
> *the big things take care of themselves.*
> *You can gain more control over your life by*
> *paying closer attention to the little things."*
> Emily Dickinson, 1830–1886

There is much value in Dickinson's advice. In this commentary, we are particularly interested in the epistemic value of *detailed predictions* ('paying closer attention to the little things') for *free energy minimization* ('gaining more control over your life'). We will show that specifying the granularity of state spaces is crucial for minimizing free energy: when the granularity is too low, little information is gained from correct predictions; if it is too high, prediction errors will be needlessly high.

In the target article, Friston and colleagues bring the exploration-exploitation trade-off under the free energy minimization regime, by assuming that the agent's prior beliefs are such that they expect to minimize future free energy and plan their actions accordingly. Formally, they describe their theory using partially observable Markov decision processes (POMDPs) with discrete states and actions; consequently, the generative models are described using categorical probability distributions. This approach overlooks the fact that 'state' and 'action' in these models depend on the granularity (or *level of detail*) of the state space and the actions operating on them. Given that the required granularity cannot be assumed to be fixed, as it may be context dependent, any discrete free energy account will also need to address the question of how the right level of detail is determined.

For example, one may plan to shop for groceries. The action 'shop for groceries' is fairly abstract and may be described more in detail as 'first pick up some croissants at the bakery, then head for the produce market to get vegetables, and don't forget to bring cat food'. Note that the more detailed we make these predictions, the more information they carry; however, they are also more prone to prediction errors. When one expects to buy *this-and-that* flavor of cat food from brand *such-and-so*, then any other flavor or brand would result in a prediction error. If, on the other hand, we expect merely 'to buy cat food', then as a long as we end up buying some brand or flavor of cat food, regardless which one, there would be no prediction error. Hence, increasing the level of detail of predicted and actual outcomes will—everything else being equal—increase average uncertainty, simply because it will increase the entropy of the probability distribution over possible outcomes.

A now classic objection of the free energy principle is that it seems to predict that organisms would seek shelter in a dark cave to defer from any sensory experiences and hence minimize prediction errors (Thornton, 2010). Even though a satisfactory answer may have been given to this objection (Friston, Thornton, & Clark, 2012), we raise a novel problem that seems to generalize that idea and follows naturally from the consideration of the level of detail: Consider that one stands in the middle of Times Square during rush hour, one can minimize prediction errors by simply predicting that 'stuff happens around me' and interpret the sensory inputs accordingly. This very low level of detail of expected and actual outcomes will, by definition, lead to low free energy (prediction errors), simply because there are fewer categories in the probability distribution.

The example illustrates that predicting and interpreting all our sensory experiences as 'stuff happens' is equally ineffective as staying in a dark cave forever. Arguably, an individual that makes more fine-grained predictions, e.g., by discriminating between cars that are parked and cars that are driving, will be more successful in the long run. Making more fine-grained predictions than 'stuff happens' induces potential uncertainty and excessive prediction errors, but it allows us to benefit by making more informative predictions. In contrast, to assess whether we should wait or whether we can walk the street, it is seldom beneficial to make predictions that are too detailed. It is of little use to predict the car type and brand in order to prevent getting run over. The added information from such a prediction is outweighed by the increased prediction error when the prediction turns out to be wrong. These considerations show that somehow a trade-off needs to be made between precision and detail.

As Friston et al. acknowledge, hyperpriors on the (expected) precision are crucial for weighting prediction errors (Friston, 2010; Clark, 2013). As we highlighted with our examples, it is necessary to extend the notion of hyperpriors to govern also level of detail, as precision is a property of predictions at every level of detail. Such an enhanced theory may shed light on why and how we are able to make predictions that trade off information gain and prediction error, and how this fits in with free energy minimization.

References

Clark, A. (2013). Clark, A. (2013b). The many faces of precision (Replies to commentaries on "Whatever next? Neural prediction, situated agents, and the future of cognitive science"). *Frontiers in Theoretical and Philosophical Psychology*, *4*, 270.

Friston, K.L. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138.

Friston, K.L., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, *3*, 130.

Thornton, C. (2010). Some puzzles relating to the free-energy principle: comment of Friston. *Trends in Cognitive Science*, 14, 53–54.