Neuronal message passing

**Abstract**

Neuronal inferences rely upon local messages, passed between neurons that share synaptic connections. Here we review various (Bayesian) message passing algorithms, and consider their plausibility as descriptions of the computations performed by biological neural networks. Specifically, we discuss variational message passing, belief propagation, and expectation propagation. Each of these may be used to perform inference on probabilistic generative models through the passing of local messages. Each is consistent with active inference, as they can all be shown to be fixed points in approximations to variational free energy. The forms of these messages are subtly different, and imply different sorts of neural signals. We consider the computational optimality of these messages, and the architecture of the wetware required to compute them.

**Keywords:**

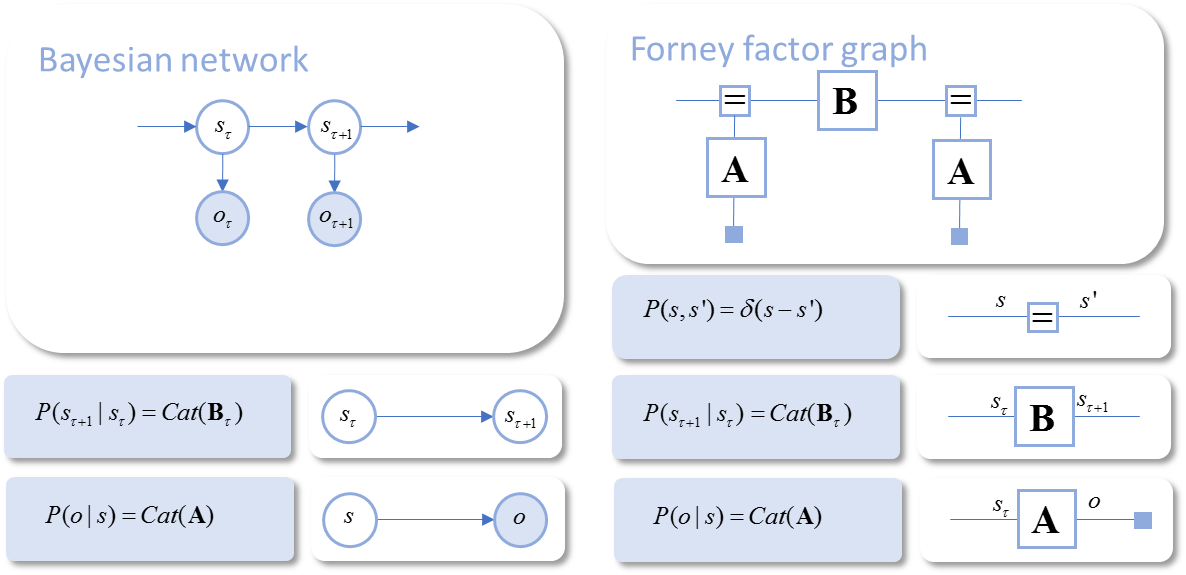
**Introduction**

Under the free energy principle, creatures must engage in Bayesian inference. The reason for this is that, to prevent their own decay, they must ensure that their entropy is bounded…

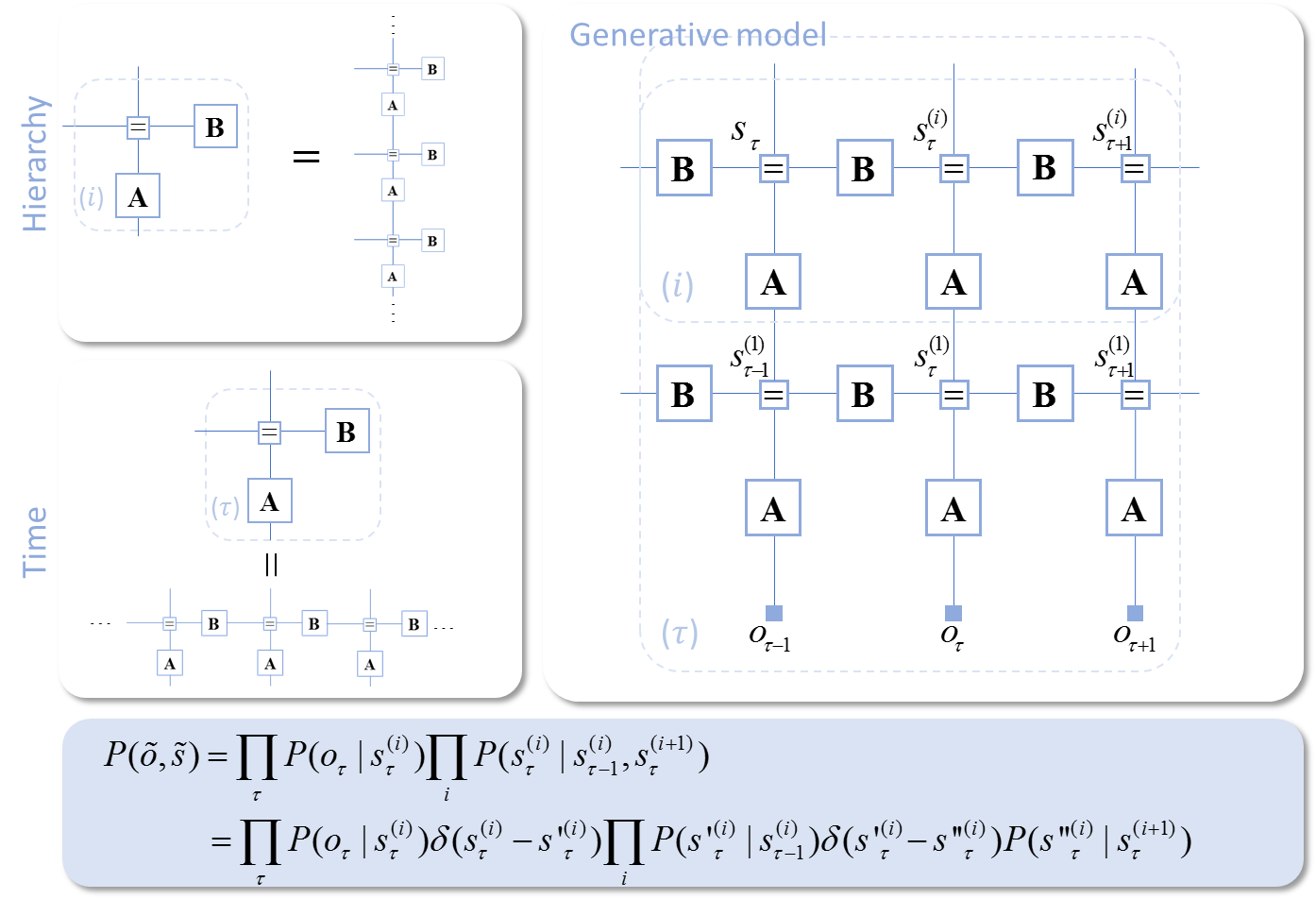
In the following, we first introduce the kinds of generative models that biological agents must contend with. For the purposes of this paper, we assume that the agent exerts no influence over the causes of its sensations. This is a drastic assumption, but simplifies the form of the models we need to provide a minimal illustration of the issues surrounding neuronal message passing. Two features that are necessary for this purpose are hierarchy and temporal progression. Having introduced a model that incorporates these features, we consider three different methods to perform (approximate) Bayesian inference, minimising free energy. Each of these rests upon local message passing across the nodes in the generative model. We then consider the neuroanatomy of within-region and between region connections needed to perform each of these forms of inference.

**Generative models**

The graphical representations of probabilistic models used here are based on normal (Forney) factor graphs (Forney 2001, Friston *et al.* 2017). The advantage of using these graphs is that all messages passed between variables can be displayed in a simple way.

****

**Figure 1 –** The graph on the left depicts a simple probabilistic model as a Bayesian network. Here, latent variables are shown in unfilled circles, with data in filled circles. Arrows indicate conditional dependencies, as shown in the lower left. The same model is shown on the right as a Forney factor graph. Here, variables are represented on the edges (lines) of the graph, with the square nodes representing the factors of the distribution over all of the variables. To ensure that each edge connects exactly two factor nodes, we duplicate variables that participate in more than two factors and use equality constraint nodes (small squares with ‘=’ signs) to connect these duplicates. Data is represented by a small filled square.



**Figure 2 –** The panel on the right shows the form of the hierarchical generative model we use in this paper. It is the graphical (factor graph) expression of the factorisation shown in the lower panel. The dotted lines indicate elements of the graph that are repeated an arbitrary number of times, as illustrated for both hierarchical levels, and for time, in the panels on the left. Note we have assumed that . This ensures pairwise factors – i.e. only two edges per factor (with the exception of the equality constraint nodes).

**Variational free energy**



Approximate probability.

Variational energy



Variational entropy



Variational free energy



Mean-field approximation



Bethe approximation



Mean-field free energy



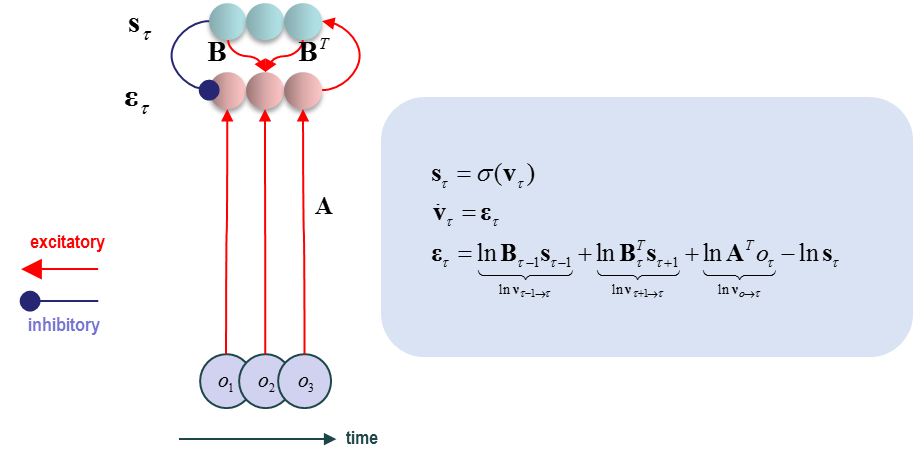
Bethe free energy



**Variational message passing**



The form of these messages suggests that we should sum all of the messages arriving at an equality constraint node. Each message to a given edge is constructed from a single factor, and the posterior belief about the variable represented on the opposite edge…(Winn 2004, Dauwels 2007)



**Belief propagation**

(Yedidia *et al.* 2005)

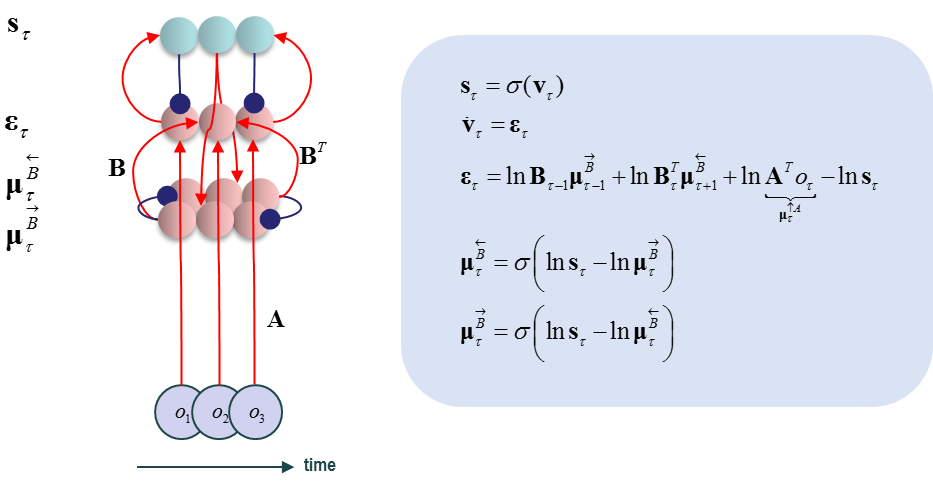
****

These messages are a little more complicated than they first appear. Unlike those for variational message passing, for which the expectation is with respect to the approximate posterior, these messages are averaged with respect to the incoming messages from other edges. In other words, messages are computed recursively from other messages, not directly from marginal beliefs. This forces us to consider two possibilities. Firstly, can we re-express these messages in terms of their marginal beliefs? Secondly, should we reject the notion that neurons signal marginal probabilities, and instead assume that they represent the messages themselves?

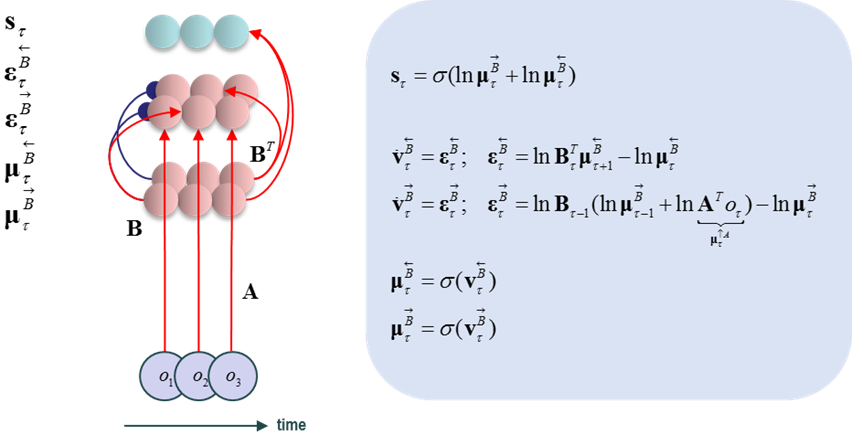
A simple way to compute messages from marginal posteriors is to rearrange the equation above to give



Importantly, this still leaves the message dependent on other messages. For the lowest level of the hierarchy, these equations suggest updates of the form shown in Figure X. An analogous formulation has been used to describe hallucinations in Schizophrenia (Jardri *et al.* 2013). This is because the subtractions in the above equation suggest there must be important local inhibitory neurons in play. These participate in ascending or descending ‘loops’. Failure of this inhibition means that the message passed on is larger than it should be (referred to as an ‘overcounting’ of the message). This is consistent with some of the abnormalities in cortical micro-circuitry described in psychosis…



Alternatively, we could assume that the marginal beliefs have no influence over the messages passed, but are just computed from them (Steimer *et al.* 2009, Steimer *et al.* 2013). While this seems like a simple solution, we run into difficulties when trying to render this biologically plausible. To do so, we require a gradient ascent as in the above formulations, rather than discrete updates. This suggests the following set of equations.



**Expectation propagation?**

A third inference method that can be expressed in terms of local message passing is expectation propagation. Like the methods outlined above, this is derived through minimisation of an approximate free energy functional. However, this involves replacing the KL-Divergence with…

**The anatomy of message passing**

*Intrinsic connections*

*Extrinsic connections*

**Conclusion**

**Notes:** Implementing belief propagation in spiking neuronal networks as a sampling process from Boltzmann distribution (Buesing *et al.* 2011, Pecevski *et al.* 2011).

**Conclusion**

**References**

Buesing, L., J. Bill, B. Nessler and W. Maass (2011). "Neural dynamics as sampling: a model for stochastic computation in recurrent networks of spiking neurons." PLoS computational biology **7**(11): e1002211.

Dauwels, J. (2007). On variational message passing on factor graphs. Information Theory, 2007. ISIT 2007. IEEE International Symposium on, IEEE.

Forney, G. D. (2001). "Codes on graphs: Normal realizations." IEEE Transactions on Information Theory **47**(2): 520-548.

Friston, K. J., T. Parr and B. d. Vries (2017). "The graphical brain: belief propagation and active inference." Network Neuroscience **0**(ja): 1-78.

Jardri, R. and S. Denève (2013). "Circular inferences in schizophrenia." Brain **136**(11): 3227-3241.

Pecevski, D., L. Buesing and W. Maass (2011). "Probabilistic inference in general graphical models through sampling in stochastic networks of spiking neurons." PLoS computational biology **7**(12): e1002294.

Steimer, A. and R. Douglas (2013). "Spike-based probabilistic inference in analog graphical models using interspike-interval coding." Neural computation **25**(9): 2303-2354.

Steimer, A., W. Maass and R. Douglas (2009). "Belief propagation in networks of spiking neurons." Neural Computation **21**(9): 2502-2523.

Winn, J. M. (2004). Variational message passing and its applications, Citeseer.

Yedidia, J. S., W. T. Freeman and Y. Weiss (2005). "Constructing free-energy approximations and generalized belief propagation algorithms." IEEE Transactions on Information Theory **51**(7): 2282-2312.