

# A neurocentric approach to Bayesian inference

Christopher D. Fiorillo

A primary function of the brain is to infer the state of the world in order to determine which motor behaviours will best promote adaptive fitness. Bayesian probability theory formally describes how rational inferences ought to be made, and it has been used with great success in recent years to explain a range of perceptual and sensorimotor phenomena<sup>1–5</sup>. In a recent Review (The free-energy principle: a unified brain theory? *Nature Rev. Neurosci.* **11**, 127–138 (2010))<sup>1</sup>, Friston advocates a ‘free-energy principle’ that seeks to explain how the nervous system performs inference and also to provide a unified theory of the brain. A unified theory is very much needed, and efforts towards that end should be encouraged. However, the relevance of the free-energy principle to the nervous system is questionable. Like most other work on the neural basis of inference, the ‘free energy’ approach is divorced from the biophysical reality of the nervous system. Attention is drawn here to an alternative, ‘neurocentric’ approach that is thoroughly grounded in the physical structure of the brain.

The notion of free energy in inference derives from a mathematical method of inverting a probability distribution. If the brain knows the probability of its sensations (given a “generative model of their causes”), then how can it ‘invert’ this probability to find the probability of their ‘causes’ (the states of the external world) given those sensations? Inversion of probabilities can be difficult mathematically, but it is entirely inconsequential with respect to information. Inversion simply corresponds to two ways of describing the same information, analogous to the conversion between ‘2 + 3’ on the one hand and ‘5’ on the other except of greater mathematical difficulty. The brain does not need to convert 2 + 3 to 5 because 2 + 3 is 5. The alternative view advocated here is that the brain does not need to invert

probabilities and, furthermore, it does not even need to ‘compute’ or ‘identify’ probabilities (unless asked to do so). Probabilities and inference are inherent properties or descriptions of information, and the brain does not need to perform any processing step to go from information to probabilities and inference.

The Bayesian approach to brain function relies on probabilities that are entirely conditional on information that is possessed by the brain. However, Friston departs from a strictly Bayesian view of probabilities when he suggests that survival entails minimizing free energy by avoiding ‘surprising’ states. He defines ‘surprise’ as the negative logarithm of the probability of an event. However, this probability is essentially just the frequency of an event within an imaginary ensemble of states that could unfold over a long period of time. This probability seems not to be conditional on any information and, therefore, it is not a Bayesian probability. Indeed, Friston writes: “A system cannot know whether its sensations are surprising.”<sup>1</sup> There are many reasons to believe that a ‘frequentist’ view of probabilities is incorrect and that it should be completely replaced by a strictly Bayesian view in which all probabilities are conditional on information<sup>6</sup>. Furthermore, Friston’s use of both frequentist and Bayesian probabilities is confusing and may detract from the goal of providing a unified account of brain function.

Friston’s formulation also encounters some difficulties in accounting for well established behavioural phenomena. He proposes that in order to avoid surprising states “the agent will selectively sample the sensory inputs that it expects”<sup>1</sup>. However, in apparent contradiction to his hypothesis animals tend to explore the least predictable sensory inputs while avoiding predictable inputs. For example,

gaze is directed not to the static and predictable parts of a visual scene but to the dynamic and uncertain parts. Exploration provides the brain with new information from its environment.

Most work on inference, by Friston and others, has started with abstract mathematical models and has then tried to find neural correlates of those models. A promising alternative is to take a ‘neurocentric’ approach by starting with known biophysical substrates and then using probabilities to describe what an ion channel, or a neuron, knows about its world. Probability distributions that are conditional only on this biophysical information can be derived using Boltzmann’s equation from statistical mechanics, as previously described<sup>7</sup>. Although the Bayesian approach has so far been mostly restricted to the level of human perception and behaviour, a neurocentric approach gives Bayesian analysis a solid biophysical basis and extends its reach to all levels of the nervous system, starting with molecules at the ‘bottom’ and working up to neurons and systems. A truly unified brain theory will need to bridge the gap between Bayesian principles and biophysical reality, and the neurocentric approach suggests how this can be done.

Christopher D. Fiorillo is at the Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, 335 Gwahangno, Yuseong-gu, Daejeon 305-701, Korea.  
e-mail: [fiorillo@kaist.ac.kr](mailto:fiorillo@kaist.ac.kr)

doi:10.1038/nrn2787-c1

Published online 14 Jul 2010

1. Friston, K. The free-energy principle: a unified brain theory? *Nature Rev. Neurosci.* **11**, 127–138 (2010).
2. Weiss, Y., Simoncelli, E. P. & Adelson, E. H. Motion illusions as optimal percepts. *Nature Neurosci.* **5**, 598–604 (2002).
3. Yang, Z. & Purves, D. A statistical explanation of visual space. *Nature Neurosci.* **6**, 632–640 (2003).
4. Niemeier, M., Crawford, J. D. & Tweed, D. Optimal transaccadic integration explains distorted spatial perception. *Nature* **422**, 76–80 (2003).
5. Kording, K. P. & Wolpert, D. M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244–247 (2004).
6. Jaynes, E. T. *Probability Theory: The Logic of Science* (Cambridge Univ. Press, 2003).
7. Fiorillo, C. D. Towards a general theory of neural computation based on prediction by single neurons. *PLoS ONE* **3**, e3298 (2008).

## Acknowledgements

The author was supported by the World Class University program through the National Research Foundation of Korea, funded by the Ministry of Education, Science and Technology (grant number R32-2008-000-10218-0).