Neural Networks 22 (2009) 1093–1104

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2009 Special Issue

Cortical circuits for perceptual inference

Karl Friston *, Stefan Kiebel

The Wellcome Trust Centre of Neuroimaging, University College London, Queen Square, London WC1N 3BG, United Kingdom

a r t i c l e i n f o

Article history:
Received 26 January 2009
Received in revised form 14 May 2009
Accepted 14 July 2009

Keywords:
Generative models
Predictive coding
Hierarchical
Dynamic
Nonlinear
Circuits
Variational
Birdsong
Free-energy

a b s t r a c t

This paper assumes that cortical circuits have evolved to enable inference about the causes of sensory input received by the brain. This provides a principled specification of what neural circuits have to achieve. Here, we attempt to address how the brain makes inferences by casting inference as an optimisation problem. We look at how the ensuing recognition dynamics could be supported by directed connections and message-passing among neuronal populations, given our knowledge of intrinsic and extrinsic neuronal connections. We assume that the brain models the world as a dynamic system, which imposes causal structure on the sensorium. Perception is equated with the optimisation or inversion of this internal model, to explain sensory input. Given a model of how sensory data are generated, we use a generic variational approach to model inversion to furnish equations that prescribe recognition; i.e., the dynamics of neuronal activity that represents the causes of sensory input. Here, we focus on a model whose hierarchical and dynamical structure enables simulated brains to recognise and predict sequences of sensory states. We first review these models and their inversion under a variational free-energy formulation. We then show that the brain has the necessary infrastructure to implement this inversion and present stimulations using synthetic birds that generate and recognise birdsongs.

1. Introduction

This paper looks at the functional architecture of cortical circuits from the point of view of perception; namely, the fitting or inversion of internal models of sensory data by the brain. Critically, the nature of this inversion lends itself to a relatively simple neural network implementation that shares many formal similarities with real cortical hierarchies in the brain. The basic idea that the brain uses hierarchical inference has been described in a series of papers (Friston, 2005; Friston, Kilner, & Harrison, 2006; Mumford, 1992; Rao & Ballard, 1998). These papers suggest that the brain uses empirical Bayes for inference about its sensory input, given the hierarchical organisation of cortical systems. Here, we focus on how neural networks could be configured to invert these models and deconvolve sensory causes from sensory input.

This paper comprises three sections. In the first, we introduce a free-energy formulation of model inversion or perception, which is then applied to a specific class of models that we assume the brain uses – hierarchical dynamic models. An important aspect of these models is their formulation in generalised coordinates of motion. This lends them a hierarchical form in both structure and dynamics, which can be exploited during inversion. In the second section, we show how inversion can be formulated as a simple gradient descent using neuronal networks and relate these to cortical circuits in the brain. In the final section, we consider how evoked brain responses might be understood in terms of perceptual inference and categorisation, using the schemes of the preceding section.

2. The free-energy formulation

This section considers the problem of inverting generative models of sensory data and provides a summary of the material in Friston (2008). This problem is addressed using ensemble learning or Variational Bayes. These are generic approaches to model inversion that provide an approximation to the conditional density $p(\vartheta|\tilde{y})$ on some causes ϑ of generalised sensory input, $\tilde{y} = [y, y', y'', \ldots]^T$. Generalised input (e.g., the intensity of photoreceptor stimulation) includes the input, its velocity, acceleration, jerk, etc. Causes are quantities in the environment that generate sensory input (e.g., the orientation of an object in the visual field). The approximation of the conditional density (i.e., the probability of a particular set of causes given sensory input) is achieved by optimising a recognition density $q(\vartheta)$ with respect to a bound on the surprise or negative log-evidence $-\ln p(\tilde{y})$ of the sensory input, as we will see next (Feynman, 1972; Friston, 2005; Friston et al., 2006; Hinton & von Cramp, 1993; MacKay, 1995; Neal & Hinton, 1998). This bound is called free-energy

* Corresponding address: The Wellcome Trust Centre for Neuroimaging, Institute of Neurology, Queen Square, London, WC1N 3BG, United Kingdom. Tel.: +44 207 833 7454; fax: +44 207 813 1445.

E-mail address: k.friston@fil.ion.ucl.ac.uk (K. Friston).

0893-6080/$ – see front matter
doi:10.1016/j.neunet.2009.07.023