CellPress

Opinion

The Functional Anatomy of Time: What and When in the Brain

Karl Friston[1,*] and Gyorgy Buzsáki[2,3]

This Opinion article considers the implications for functional anatomy of how we represent temporal structure in our exchanges with the world. It offers a theoretical treatment that tries to make sense of the architectural principles seen in mammalian brains. Specifically, it considers a factorisation between representations of temporal succession and representations of content or, heuristically, a segregation into when and what. This segregation may explain the central role of the hippocampus in neuronal hierarchies while providing a tentative explanation for recent observations of how ordinal sequences are encoded. The implications for neuroanatomy and physiology may have something important to say about how self-organised cell assembly sequences enable the brain to exhibit purposeful behaviour that transcends the here and now.

Trends

Recent studies of hippocampal responses suggest that they have an intrinsic dynamics that may complement (or nuance) spatiotemporal encoding, particularly the encoding of trajectories through space and time and inherent place-cell activity.

Predictive coding and the Bayesian brain now predominate as explanations for much of cognitive neuroscience and functional anatomy in the brain and have clear relevance for the encoding of trajectories through various state spaces.

Recent attempts to understand the form of ordinal or sequential processing in the brain (e.g., navigation, language) emphasise prediction and may be fundamentally informed by recent empirical findings from the study of hippocampal (and neocortical) responses.

The Principles of Functional Anatomy

There are certain architectural principles of neuroanatomy that seem amenable to explanation from a purely theoretical perspective. These range from the existence of axonal processes that form neuronal connections to macroscopic organisational principles such as functional segregation [1]. A key example is the segregation of dorsal and ventral streams into what and where streams [2]. How might these architectural features be explained from a theoretical perspective? In what follows, we appeal to active inference and the Bayesian brain hypothesis [3,4] to suggest that functional segregation emerges from statistical structure in the environment. We then consider the implications of this argument for a fundamental aspect of this structure; namely, the trajectories or ordered sequences of states that we encounter [5]. Our conclusion is that there should be a functional segregation between what and when–a conclusion that seems to explain numerous anatomical and physiological observations, particularly in the hippocampal system.

Good Enough Brains and Good Enough Models

A key theoretical development in neurobiology is the appreciation of the brain as a predictive organ generating predictions of its actions and sensations [4,6–9]. These predictions rest on an internal or **generative model** (see Glossary) of how sensory input unfolds. One can understand much of neuronal dynamics and synaptic plasticity as an optimisation of (Bayesian) **model evidence** as scored by proxies like free energy and prediction errors [9–11]. If one subscribes to this normative theory, the brain must be a good (enough) model of its environment, where recurring sequences of events are the rule. This idea dates back to notions of good regulators in self-organisation and cybernetics [12,13]. In brief, the good regulator theorem states that any system that can control its environment must be a good model of that environment. So what constitutes a good enough model?

Mathematically, a good enough model is simply a model that has sufficient evidence in light of the (sensory) data it has to explain. Evidence is the probability of sensory samples under a model of

[1]Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, UK
[2]NYU Neuroscience Institute, School of Medicine, New York University, New York, NY 10016, USA
[3]Institute of Experimental Medicine, Hungarian Academy of Sciences, Budapest, Hungary

*Correspondence: k.friston@ucl.ac.uk (K. Friston).

 http://dx.doi.org/10.1016/j.tics.2016.05.001

Box 1. Approximate Bayesian Inference

Bayesian inference refers to optimising beliefs about a model or its hidden states (s) in the light of outcomes (o) or evidence. Formally, this can be expressed as minimising a variational free energy bound on Bayesian model evidence [91] with respect to beliefs about hidden states encoded by a probability density $Q(s)$ (with expectation $E[Q(s)] = \mathbf{s}$).

$$F(o,\mathbf{s}) = \underbrace{D[Q(s)||P(s|o)]}_{\text{relative entropy}} - \underbrace{\ln P(o)}_{\text{log evidence}} \geq \underbrace{\ln P(o)}_{\text{log evidence}}$$
$$= \underbrace{D[Q(s)||P(s)]}_{\text{complexity}} - \underbrace{E_Q[\ln P(o|s)]}_{\text{accuracy}}$$

Here, the model is specified by a joint distribution over outcomes and their causes or hidden states: $P(o, s) = P(o|s)P(s)$. The first expression for free energy shows that when free energy is minimised, the relative entropy or Kullback–Leibler (KL) divergence attains its minimum (zero) and free energy becomes the negative logarithm of model evidence. In other words, when free energy is minimised, the approximate posterior beliefs become the true posterior beliefs (i.e., the distribution of hidden states given outcomes) and free energy becomes negative log evidence.

Another way of conceptualising free energy is in terms of accuracy and **complexity**, as shown in the second equality. This equality shows that minimising free energy minimises complexity. Here, complexity is the KL divergence between posterior beliefs and **prior beliefs** (prior to any outcomes). In other words, complexity reflects the degrees of freedom–above and beyond prior beliefs–needed to provide an accurate account of observed data. It follows that when one is absolutely certain about the hidden states causing data, the complexity increases with the number of hidden states entertained by the model.

The imperative to minimise complexity is known as Occam's principle and is the basis of approximations to model evidence provided by the Akaike and Bayesian information criteria [92]. The role of complexity will become important below, when we consider models with a large number of states encoding joint distributions over two factors relative to parsimonious models (with greater model evidence) that encode just the factors or marginal densities (Box 2). In terms of the equations above, this distinction can be expressed as the mean field approximation $Q(s) = Q(s^{where})Q(s^{what})$.

how those samples were generated (Box 1). In this sense, any brain can be viewed as (self-) organising itself to maximise model evidence. Here we are implicitly appealing to the Bayesian brain hypothesis [14] while gently sidestepping big questions about its utility and falsifiability (e.g., [15,16]). In what follows, we assume that the imperative to maximise model evidence is a (possibly tautological) truism [17] and consider the implications for functional anatomy. Our focus is on the notion of a **mean field approximation** that is an integral part of **approximate Bayesian inference**.

A key conclusion–that follows from the Bayesian brain–is that the structure of a good brain will recapitulate the (statistical) structure of how sensations are caused; in short, the model resides in the structure of the brain. For example, why does the brain have extensive connections while the liver seems to operate perfectly happily without them? An obvious answer is that the brain has to model sparse dependences induced by regularities in the world. In other words, our sensory inputs are generated by a small number of underlying causes that act on each other (usually at a distance) in a lawful and structured way. This lawful structure requires a relatively sparse dependency among the causes, such as gravity causing things to fall or visual objects causing sensory impressions. In short, the probabilistic structure of our world should, in principle, provide a sufficient explanation for the structure and fabric of connections of any brain that is trying to model that world. For example, our sensations are generated in a way that conforms to logarithmic rules (e.g., Weber's law). These statistical rules may then be transcribed into the lognormal statistics of synaptic physiology (implicit in divisive normalisation [18]) or the connectome that supports this physiology [19,20]. Simply noting that causal regularities in the world are transcribed into neuronal architectures may sound self-evident. However, this conjecture does not get to the heart of principles such as functional segregation. To understand how maximising model evidence leads to functional segregation, we have to consider the constraints under which evidence is optimised. This brings us to the notion of approximate Bayesian inference (Box 1).

Good Enough Brains and Approximate Bayesian Inference

Any system or procedure that optimises (maximises) Bayesian model evidence can be regarded as implementing Bayesian inference. However, exact Bayesian inference is generally impossible

Glossary of Bayesian terms

Approximate Bayesian inference: Bayesian belief updating in which approximate posterior distributions are optimised by minimising variational free energy, ensuring that the approximate posterior convergences to the true posterior.
Bayesian belief updating: the combination of prior beliefs about the causes of an observation and the **likelihood** of that observation producing a posterior belief about its hidden causes. This updating conforms to Bayes' rule.
Bayesian model evidence: this is the probability that some observations were generated by a model. It is also known as the marginal or integrated likelihood because it does not depend upon the hidden causes.
Complexity: the difference or divergence between prior and posterior beliefs. The complexity of a model reflects the change in prior beliefs produced by Bayesian belief updating (also known as Bayesian surprise).
Expectation: the mean or average (the first-order moment of a probability distribution).
Factorisation: decomposition of a quantity into the product of factors such that multiplying the factors reproduces the original quantity.
Generative model: a probabilistic specification of the dependencies among causes and consequences; usually specified in terms of a prior belief and the likelihood of observations, given their causes.
Hidden causes or states: the unobserved (including fictive) causes of observed data. They are hidden because they are random variables that can only be inferred from observations.
Likelihood: the probability of an observation under a generative model, given its causes.
Marginal: a marginal probability distribution of a joint distribution over random variables is obtained by marginalising or averaging over all of the variables apart from the variable of interest.
Mean field approximation: approximating a joint distribution over two or more random variables with the product of their marginal distributions.
Posterior beliefs: a probability distribution over the hidden causes of