

How do we navigate a world with deep temporal structure?

Navigating our dynamic but structured environment is an essential part of everyday life. For example, when reading this sentence, we pre-empt the context and purpose based on our internal representation of the world and actively sample it to understand what we are reading – and where to look next. This paper suggests how this constructive sampling proceeds in the brain, using active inference based on deep (hierarchical) temporal models. Active inference is a normative theory of brain function that tries to explain how sentient creatures, like us, navigate dynamic environments. It involves creating internal generative models of the world that predict – based upon plausible hypotheses (i.e., hidden states, s^i_τ) – the sort of data (i.e., outcomes, o^i_τ) that will be encountered.

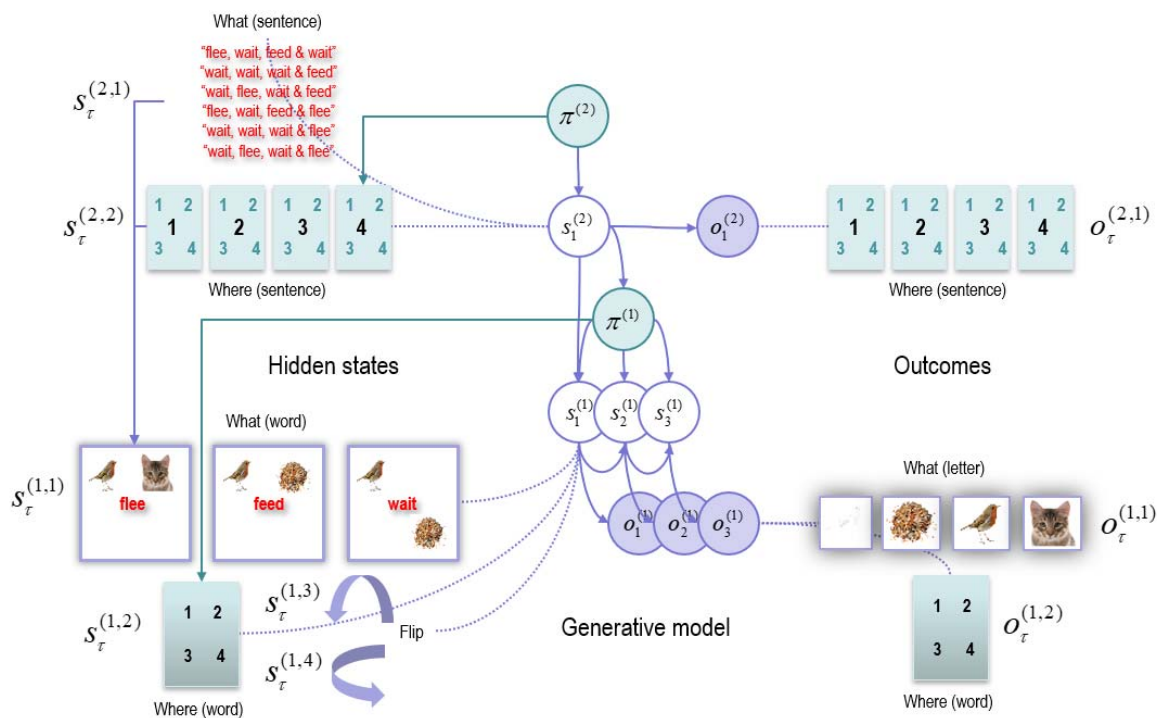


Fig. 1. The generative model used to simulate reading.

A generative model can be specified as a series of discrete states, beliefs about which are updated based on the observed outcomes that are sampled to minimise uncertainty. This allows one to model information-seeking, goal-directed behaviour as ‘planning as inference’. The ‘deep’ aspect of these models rests upon the notion that the outcomes of one (slower) level generate the hidden states at a lower (faster) level – much as a sentence (slow) entails a sequence of (faster) words. This allows the model to accumulate evidence over nested time scales and implicitly make inferences about narratives (i.e., temporal scenes). The outcomes are solicited actively by selecting an appropriate plan of action (i.e. policy, π^i) at each hierarchical level – based on the expected surprise or free energy (i.e., uncertainty).

These policies determine the trajectory of hidden states that predict outcomes (or initial hidden states of the level below). Thus, sensory outcomes generated by the model can be used as evidence to confirm or refute our internal hypotheses; including our beliefs about what we are doing (i.e., policies).

The key aspect of these hierarchical models is that hidden states at higher levels contextualise trajectories of hidden states at lower levels; generating a deep dynamic narrative. We illustrate this in terms of the epistemic foraging implicit in reading. The (simulated) subject is tasked with categorising (pictographic) sentences into specific narratives by sampling letters (images of bird, cat, seeds or nothing) from four quadrants into a word (flee, feed or wait). This generative model (Fig. 1) has two levels. The higher level has three hidden factors – sentence, current word and decision. The current word and sentence specify the hidden state at the lower level. The lower level includes letter location and spatial transformations (e.g. to simulate order invariant understanding of words). The current word and letter location specify the outcome (letter; cat, bird, seed or nothing). At both levels, the hidden locations specify an outcome in terms of movements that can be at higher (e.g., head) and lower (e.g., eye) levels. Finally, a decision state determines feedback with three possibilities; namely, nothing, right or wrong. There are policies at each level. The high-level policy determines which word the agent is currently reading, while the lower level dictates the transitions among the quadrants containing letters.

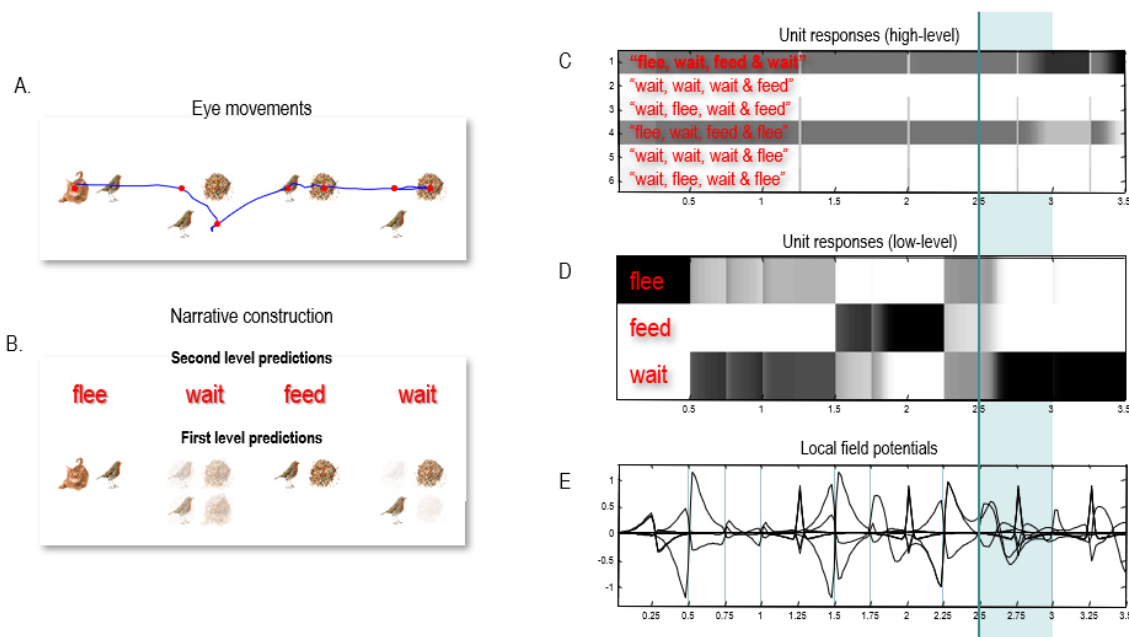


Fig. 2. A & B show simulated behavioural responses during reading and C, D & E show simulated electrophysiological (firing rate activity [C,D] and local field potentials [E]) responses. Vertical cyan line represents saccade onsets.

The simulation highlights some of the key aspects of deep temporal inference. For example, we may form precise (confident) beliefs about letters without seeing them; e.g., the subject believes there is a bird in the second quadrant of the first word, despite never looking there (Fig. 2.A & 2.B). Conversely, we may experience uncertainty

about letters, even though we are confident about the word (e.g., spatial transformation introduces a certain degree of uncertainty when predicting the second word – wait in Figure 2.B). Furthermore, the simulated electrophysiological responses associated with belief updating look remarkably like empirical responses (Purpura, et al, 2003) (Fig. 2.E). We can extend this model (motivated by work in computational psychiatry) to simulate mismatch responses (such as the mismatch negativity) to unexpected stimulus features at multiple hierarchical levels. Thus, the active inference scheme presented takes a potentially important step towards explaining hierarchical behaviour and how it may be orchestrated by the brain.

*Noor Sajid, Karl J. Friston, Cathy J. Price, Howard Bowman
Wellcome Trust Centre for Neuroimaging, Institute of Neurology,
University College London, WC1N 3BG, United Kingdom*

Publication

[Deep temporal models and active inference.](#)

Friston KJ, Rosch R, Parr T, Price C, Bowman H
Neurosci Biobehav Rev. 2017 Jun