Name: Will Dorrell

Affiliation: Gatsby Unit, University College London

URL: https://www.williamdorrell.co.uk/
E-mail: dorrellwec@gmail.com

Talk Title: Efficient Coding of Cognitive Maps: A Normative Theory of Internals Models

Abstract:

Understanding neural tuning is a fundamental problem in neuroscience. Most theories of neural representations descend from the Efficient Coding Hypothesis. This argues that neural populations are informatively encoding certain variables under biological efficiency constraints. While broadly successful for sensory encodings, it's been difficult to apply the same principles to cognitive representations (e.g. grid cells). We argue this is because many cognitive tasks require an internal model which is not captured by just efficiently encoding variables. To remedy this, we present a normative theory of neuronal computations using an internal model (i.e. answering: "if I take this action how will this variable change?"). We show this theory introduces additional constraints on the representation that alter the optimal representation, and we'll apply this theory in three settings. First, our theory provides a complete normative theory of multi-modular grid cells, and makes predictions we verify in data, such as the relative inter-modular angle. Second, we study the representation of shared structure in families of related tasks and show how it can be used to understand the representation of task structure in prefrontal cortex. Third, time permitting, we will apply the same ideas to families of action sequences, and show it can be used to understand supplementary motor responses. Neural activity is the substrate of the computations that undergird behaviour. To understand these computations we need theories that comprehensibly link (i.e. more than neural networks) computations and neural responses. Our theory hopes to build one such link.

Biographical information:

Will Dorrell is a PhD student at the Gatsby Unit in London working with Peter Latham, James Whittington, and Tim Behrens. He, like most neuroscientists, is interested in the neural basis of behaviour. He previously worked on olfaction with Cengiz Pehlevan at Harvard, and bio-plausible RL with Sergio Verduzco-Flores in Okinawa. In a previous life he nearly became a Condensed Matter Physicist.