Inferring Neural Activity Before Plasticity: A Foundation for Learning Beyond Backpropagation

Rafał Bogacz[1]

[1] *IMRC Brain Networks Dynamics Unit, University of Oxford, Oxford, United Kingdom*
* rafal.bogacz@bndu.ox.ac.uk

For both humans and machines, the essence of learning is to pinpoint which components in its information processing pipeline are responsible for an error in its output – a challenge that is known as credit assignment. How the brain solves credit assignment is a key question in neuroscience, and also of significant importance for artificial intelligence. It has long been assumed that credit assignment is best solved by backpropagation, which is also the foundation of modern machine learning. However, it has been questioned whether it is possible for the brain to implement backpropagation and learning in the brain may actually be more efficient and effective than backpropagation. This talk will present a fundamentally different principle on credit assignment, called prospective configuration. In prospective configuration, the network first infers the pattern of neural activity that should result from learning, and then the synaptic weights are modified to consolidate the change in neural activity. During the talk it will be demonstrated that this distinct mechanism, in contrast to backpropagation, (1) underlies learning in a well-established family of models of cortical circuits, (2) enables learning that is more efficient and effective in many contexts faced by biological organisms, and (3) reproduces surprising patterns of neural activity and behaviour observed in diverse human and animal learning experiments. These findings establish a new foundation for learning beyond backpropagation, for both understanding biological learning and building artificial intelligence.