World Wide NeuRise

Talks by rising stars of neuroscience

Learning static and dynamic mappings with local self-supervised plasticity

Pantelis Vafeidis
(Caltech)

Animals exhibit remarkable learning capabilities with little direct supervision. Likewise, self-supervised learning is an emergent paradigm in artificial intelligence, closing the performance gap to supervised learning. In the context of biology, self-supervised learning corresponds to a setting where one sense or specific stimulus may serve as a supervisory signal for another. After learning, the latter can be used to predict the former. On the implementation level, it has been demonstrated that such predictive learning can occur at the single neuron level, in compartmentalized neurons that separate and associate information from different streams. We demonstrate the power such self-supervised learning over unsupervised (Hebb-like) learning rules, which depend heavily on stimulus statistics, in two examples: First, in the context of animal navigation where predictive learning can associate internal self-motion information always available to the animal with external visual landmark information, leading to accurate path-integration in the dark. We focus on the well-characterized fly head direction system and show that our setting learns a connectivity strikingly similar to the one reported in experiments. The mature network is a quasi-continuous attractor and reproduces key experiments in which optogenetic stimulation controls the internal representation of heading, and where the network remaps to integrate with different gains. Second, we show that incorporating global gating by reward prediction errors allows the same setting to learn conditioning at the neuronal level with mixed selectivity. At its core, conditioning entails associating a neural activity pattern induced by an unconditioned stimulus (US) with the pattern arising in response to a conditioned stimulus (CS). Solving the generic problem of pattern-to-pattern associations naturally leads to emergent cognitive phenomena like blocking, overshadowing, saliency effects, extinction, interstimulus interval effects etc. Surprisingly, we find that the same network offers a reductionist mechanism for causal inference by resolving the post hoc, ergo propter hoc fallacy.

Event link:
https://www.crowdcast.io/e/wwneurise/

https://neurise.github.io

sponsored by eLife