ASHBi SEMINAR

Interpreting neural dynamics by modeling beliefs

Lecturer: Xaq Pitkow

Associate Professor,
Department of Neuroscience, Baylor College of Medicine

[NOTE] The event information has been changed.

Date Thursday, 15 June 2023

Time 14:00 – 15:00 [JST]

Venue Seminar Room (B1F, Faculty of Medicine Bldg. B)

*Register via the right QR code

Abstract

Complex behaviors are often driven by an internal model, which integrates sensory information over time and facilitates long-term planning to reach subjective goals. We interpret behavioral data by assuming an agent behaves rationally --- that is, they take actions that optimize their subjective reward according to their understanding of the task and its relevant causal variables. We apply a new method, Inverse Rational Control (IRC), to learn an agent's internal model and reward function by maximizing the likelihood of its measured sensory observations and actions. Technically, we define an animal's strategy as solving a Partially Observable Markov Decision Process (POMDP), and we invert this model to find the task and subjective costs that have maximum likelihood. This is a generalization of both Inverse Reinforcement Learning and Inverse Optimal Control. Our mathematical formulation thereby extracts rational and interpretable thoughts of the agent from its behavior. We apply this method to behavioral data from primates catching fireflies in virtual reality, and use it to understand properties of the mental model monkeys use to navigate by optic flow.

The thoughts imputed to the animal can then serve as latent targets for neural analyses. Using these targets, we provide a framework for interpreting the linked processes of encoding, recoding, and decoding of neural data in light of the rational model for behavior. We first demonstrate the merits of this approach on synthetic neural data during a foraging task. We then analyze real neural activity in primate prefrontal cortex (PFC) and posterior parietal cortex (PPC) to discover computations relevant to foraging tasks. In PFC, we find that reward dynamics are represented in a subspace of the high-dimensional population activity, and predict animal’s subsequent choice better than either the true experimental variables or the raw neural responses. In PPC, we find representations of latent navigation-relevant variables, and find that task manipulations alter the coupling between neurons, suggesting that these interactions reflect the mental model used to perform task-relevant computations. Overall, our approach may identify explainable structure in complex neural activity patterns. This framework lays a foundation for discovering how the brain chooses to act using dynamic beliefs about the uncertain world.

Organizer：Graduate School of Medicine
Institute for the Advanced Study of Human Biology (WPI-ASHBi)
Contact: Dr. Ryo Sasaki
[E-mail] sasaki.ryo.3r@kyoto-u.ac.jp