【Grant-in-Aid for Transformative Research Areas（A）】

Development and validation of a unified theory of prediction and action（Unified theory）

Head Investigator	RIKEN Center for Brain Science, Unit Leader ISOMURA Takuya	Researcher Number:50802972
Research Area Information	Number of Research Area : 23A402　Project Period (FY) : 2023-2027 Keywords : Neuroscience, AI, Generative model, Free-energy principle	

Purpose and Background of the Research

●Outline of the Research

Elucidating the computational principle of the brain and implementing it in artificial intelligence (AI) is the greatest frontier of natural and computational sciences. Although AI has achieved great success by gaining inspiration from neuroscience (e.g., feature extraction and reinforcement learning), a significant gap still exists between human intelligence and AI.

The brain constructs a 'generative model' that expresses the dynamics of external states. The Bayesian brain hypothesis and the free-energy principle have been proposed to account for the perception, learning, and action of biological organisms. However, the neuronal bases underlying these theories are yet to be elucidated. Recent developments in experimental techniques have enabled the acquisition of high-precision, large-scale data covering multiple layers and regions. Furthermore, reverse engineering of generative models has enabled the mapping of neural circuit quantities to quantities in generative models in a one-to-one manner. These developments have made it practical to identify generative models from experimental data, which will facilitate an understanding of the brain and mind.

This project aims to use state-of-the-art techniques to measure highly accurate, large-scale neuronal activity data from the brains of various animals and reverse engineer generative models from these data, to develop a unified theory of the brain (Fig. 1). We will empirically test its validity by asking whether obtained generative models can predict brain activity, behaviour, and learning in animals. Our theory will be able to explain perceptual prediction, planning, and behavioural control in a unified manner, paving the way for the development of biomimetic AI and early diagnostic methods for psychiatric disorders.

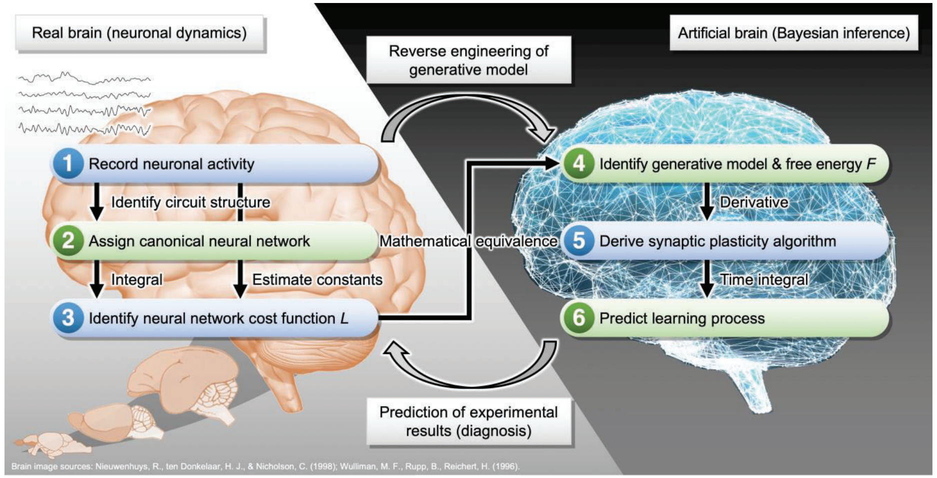

Figure 1. Validation of a unified theory using reverse engineering of generative models

Expected Research Achievements

●Integration of theory and experiments for developing a unified theory

The theory (A00) and experimental (B00) groups will work together to contribute to the goal of developing and validating a unified theory of the brain (Fig. 2, left). State-of-the-art measurements are available for experiments with fish (Okamoto), mice (Doya), marmosets (Ebina, Komatsu), and humans (healthy subjects and patients with mental disorders) (Takahashi). The application of reverse engineering to obtained neural activity data enables to estimate generative models and construct neuromorphic AI. Using diverse animal species and measurement methods, we will demonstrate that reverse engineering can predict a wide range of animal brain activity and behaviour.

Three common tasks are set (Fig. 2, right): Task 1 involves the validation of deep predictive coding for visual and auditory inputs, by comparing human and marmoset brain activity with deep generative models (Suzuki). Task 2 investigates the brain activity of fish and marmosets when solving tasks such as mazes, to elucidate the decision-making and planning mechanisms. Task 3 uses lever-pressing tasks in mice and marmosets to elucidate the neural mechanisms underlying body control. The validity of the free-energy principle and Bayesian brain hypothesis can be examined by these tasks.

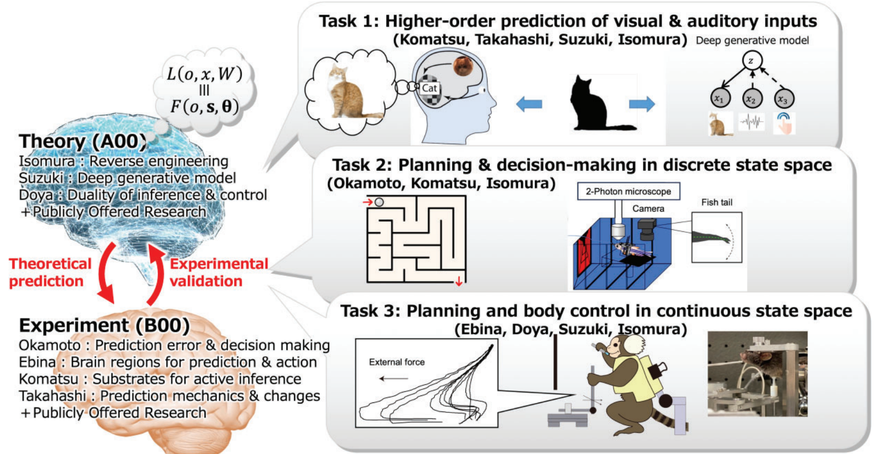

Figure 2. Organisation structure and common tasks

●Expected contributions

Our outcomes will deepen our understanding of the computational mechanisms of the brain, and enable to predict the brain. These can potentially be applied to elucidating the mechanisms of mental disorders, developing early diagnosis and treatment methods, creating new neuro-technologies, and developing biomimetic AI that implements efficient learning algorithms in the brain (Fig. 3).

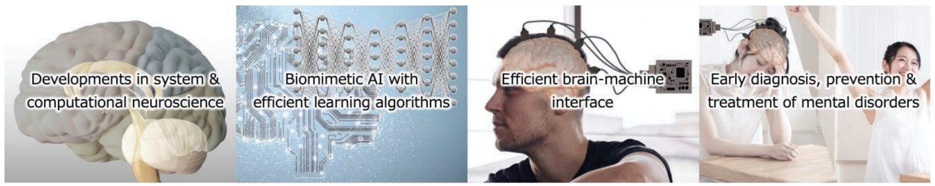

Figure 3. Expected contributions

Homepage Address, etc.	Web page	https://unifiedtheory.jp/en
	YouTube	https://www.youtube.com/@unifiedtheoryjp
	X(Twitter)	https://twitter.com/unifiedtheoryjp
	Contact	https://unifiedtheory.jp/en/contact