Title of Project : Correspondence and Fusion of Artificial Intelligence and Brain Science

Kenji Doya
(Okinawa Institute of Science and Technology Graduate University, Neural Computation Unit, Professor)

Research Project Number : 16H06561 Researcher Number : 80188846

【Purpose of the Research Project】

The purpose of this research project is to bring together artificial intelligence research and brain science research, which separated apart while their own sophistication, and to promote developments of novel learning algorithms and deeper understanding of brain mechanisms. We aim to develop efficient algorithms and clarify brain's realization of **supervised learning of internal models, reinforcement learning by exploration and evaluation**, and representation learning to facilitate them. We further try to realize flexible artificial intelligent systems based on the understanding of the whole-brain architecture for flexibly linking those learning modules.

【Content of the Research Project】

Artificial intelligence and brain science have had a swinging relationship of convergence and divergence. In the early days of pattern recognition, multi-layer neural networks based on the anatomy and physiology of the visual cortex played a key role, but subsequent sophistication of machine learning promoted methods that are little related to the brain. Recently, however, the remarkable success of deep neural networks in learning from big data has re-evoked the interests in brain-like artificial intelligence.

Evidence suggests that the cerebellum, the basal ganglia, and the cerebral cortex are respectively specialized in supervised learning, reinforcement learning, and unsupervised representation learning, (Fig. 1). On the other hand, the recent success of artificial intelligence in beating a Go world champion has demonstrated that exquisite combination of supervised learning, reinforcement learning, and representation learning by deep neural networks can achieve human level or higher intelligence.

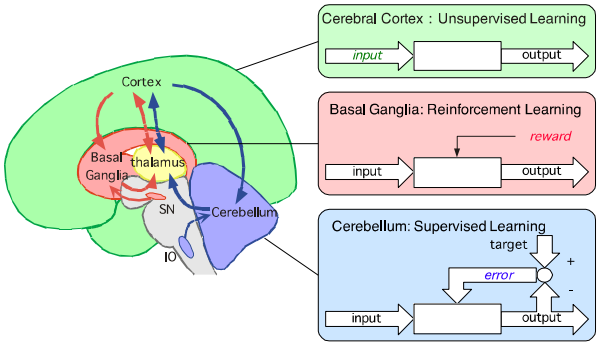

Fig. 1: Learning algorithms of cerebellum, basal ganglia, and cerebral cortex (Doya, 1999).

This project brings together leading researchers in artificial intelligence and brain science, let them learn the latest developments in each other's fields, and promote identifying the seeds of novel discovery and developments. In addition to developing novel algorithms for supervised learning, reinforcement learning, and deep representation learning and clarifying their implementation in the brain, we try to elucidate how those learning modules are flexibly integrated in the brain depending on the behavioral needs, and utilize the knowledge for developing artificial intelligence that allows human-like flexible actions and communication.

【Expected Research Achievements and Scientific Significance】

We expect to achieve better understanding of the mechanisms of deep neural network learning, data-efficient learning algorithms for humanoid robots, and hierarchical models to understand human intentions, for example. In a long run, we try to understand the self-organizing mechanism of learning modules in the brain and to derive the design principles for realizing artificial general intelligence.

The project will also aim to produce young scientists who can lead the crossing edges of artificial intelligence and brain science, through training programs such as summer schools, hackathons and international exchange programs.

【Key Words】

Deep neural network: a multi-layer network for discovering statistical features hidden in the data, from simple ones to gradually complex ones. It is widely used for image and speech recognition.

【Term of Project】 FY2016-2020

【Budget Allocation】 1,119,100 Thousand Yen

【Homepage Address and Other Contact Information】

http://www.brain-ai.jp
ncus@oist.jp