日本沖繩科學技術大學院大學 Kenji Doya 教授演講：What Can We Further Learn From the Brain for AI and Robotics

文／洪偉 資科工博士生、 解佳穎 資科工博士生

Prof. Kenji Doya 於東京大學取得博士學位，目前於沖繩科學技術大學院大學 (OIST) 擔任教授，Prof. Doya 的研究領域涵蓋人工智慧、深度學習、機器學習、神經科學和機器人學，帶領的神經計算團隊致力於透過神經生物學和計算工程等研究方法的結合，嘗試了解大腦或人工智慧作為 agent，如何對未知的環境進行強化學習，並進一步提出在學習方面穩健且靈活的演算法。

Prof. Kenji Doya 於 2023 年 2 月 13 日受資訊學院和神經科學研究所邀請，在大腦與人工智慧學術講座進行專題演講，主題為” What Can We Further Learn From the Brain for AI and Robotics”。在演講中，Prof. Doya 分享了神經科學與人工智慧過去的研究和歷史發展。許多機器學習架構能和大腦的生物特性能進行類比，如神經元細胞和 perceptron 結構，或是對應不同功能的腦區劃分和神經網路中不同目標的模組切割等等。經由這些觀察，能發現腦科學與深度學習的高度關聯性，因此以腦科學的角度去發想研究，進一步提升人工智慧和機器人學的發展，成為一門研究課題。其中，機器學習領域中的強化學習和神經科學中大腦的學習機制存在許多相似的例子；除了大腦可以被視為 multi agent 加上不同學習方法的複雜系統，強化學習的時序差分學習和多巴胺的制約機制，大腦基底核的獎勵預測和 deep Q network 原理等也顯示結構之外的相似性。

上述的例子再次顯示出神經科學與人工智慧發展之間密不可分的關係。因此，基於腦部神經科學的進展，講者與我們分享了人工智慧的潛在發展方向，以及現階段進行中的研究主題。討論的方向主要涵蓋三大主題，包括能源效率、數據效率和自主性與社交性。基於 biophysical computing 的概念，研究者利用分布式的記憶體發展出了具有極佳能源效率的神經形態晶片。除此之外，大腦科學提出了 "mental simulation"，通過重組預先訓練好的預測模型和行動策略來幫助學習，這能有效地利用現有的數據。最後，就如人類的自我學習，能夠自行找到新理論、算法或目標的人工智能就會是現階段最有潛力的研究方向，然而背後還有許多問題我們需要審慎評估，舉例來說，人工智能的自我發展是否會違反現階段的道德標準，並且是否會被有心人士作為犯罪的用途。

除了介紹腦科學、神經科學和強化學習的關聯，Prof. Doya 也展示了他帶領的團隊在前述具自主性或社交性的人工智慧領域近年的一些研究成果。例如經由特殊設計的 value function，透過調整其中的某些 factor，可以讓機器人對於 reward 的反應像是不同個性的人類對於風險的評估態度，或是藉由強化學習，讓機器人主動調整位置達成行走或站立等指定任務等等。這些和機器人行為或決策有關的研究能作為參考，奠定開發對環境適應更有彈性或更自動化的人工智慧或機器人的基礎。

Professor Kenji Doya's speech on What Can We Further Learn from the Brain for AI and Robotics

Professor Kenji Doya received his Ph.D. degree from the University of Tokyo and is currently a professor at the Okinawa Institute of Science and Technology Graduate University (OIST).
His research interests include artificial intelligence, deep learning, machine learning, neuroscience, and robotics. He leads a neural computation team dedicated to understanding how the brain or artificial intelligence, as agents, can perform reinforcement learning in unknown environments and further develop robust and flexible algorithms for learning.

On February 13, 2023, Professor Doya was invited by the College of Computer Science and the Institute of Neuroscience to give a keynote speech on "What Can We Further Learn From the Brain for AI and Robotics". In his speech, Dr. Doya shared the history and development of neuroscience and artificial intelligence research. Many machine learning architectures can be analogized with the biological properties of the brain, such as the structure of neurons and perceptrons, or the division of different brain regions for different functions and module segmentation for different goals in neural networks. Through these observations, the high correlation between neuroscience and deep learning can be discovered. Therefore, by approaching research from the perspective of neuroscience, further advancements can be made in the development of artificial intelligence and robotics has become a research topic.Among them, there are many similar examples between the learning mechanism of the brain in neuroscience and the reinforcement learning in machine learning. For instance, the brain can be regarded as a complex system of multi-agents with different learning methods, the dopamine constraint mechanism, the reward prediction of the basal ganglia in the brain, and the principle of the deep Q network. These examples demonstrate similarities that go beyond their structural aspects."

The above examples once again demonstrate the close relationship between neuroscience and the development of artificial intelligence. Therefore, based on the advancements in brain neuroscience, the speaker shared with us the potential directions of AI development and the current research topics. The discussion primarily covers three main themes: energy efficiency, data efficiency, and autonomy and sociality. Building on the concept of biophysical computing, researchers have developed neural morphology chips that exhibit excellent energy efficiency through distributed memory. Additionally, mental simulation was proposed to utilize reconfigured pre-trained predictive models and action strategies to aid learning, as well as to leverage existing data. Lastly, similar to human's self-learning, the most promising research direction at this stage would be artificial intelligence that can independently discover new theories, algorithms, or objectives. However, there are still many issues that need to be carefully evaluated. For example, whether the self-development of artificial intelligence might violate current ethical standards and whether it could be exploited by malicious individuals for criminal purposes.

In addition to introducing the connection between neuroscience, brain science, and reinforcement learning, Prof. Doya also showcased some recent research achievements in the field of autonomous or social artificial intelligence led by his team. For example, robots can exhibit responses to rewards that resemble the risk assessment attitudes of different personalities in humans. Furthermore, through reinforcement learning, robots can actively adjust their positions to accomplish specific tasks such as walking or standing. These studies related to robot behavior and decision-making serve as references and foundation for developing more flexible and automated artificial intelligence or robots that can adapt to different environments.