



Artificial Intelligence: Midterm Assignment

(2022-2)

Course / Class:	Artificial Intelligence Class 01
Professor:	Dae-Won Kim
Submission Date:	2022. 10. 29
Student ID:	20184757
Name:	주영석

Objective #2: AI Philosophy에 대한 자신만의 고찰 – Strong AI vs. Weak AI

Strong AI, 강인공지능은 인간의 의식, 사고, 생각, 마음을 온전하게 계산하고 구현할 수 있다는 주장 또는 그것이 구현된 인공지능을 뜻하고, Weak AI, 약인공지능은 인공지능이 인간의 마음, 생각을 구현할 수 없거나 가질 필요가 없다는 주장, 또는 한 가지 일에 특화된 인공지능을 뜻한다.

나는 인류의 과학이 더욱 발전하면 인간의 의식, 사고, 생각, 마음을 구현할 수 있다고 생각한다.

현재 인공지능은 인공신경망과 딥러닝의 개발로 다양한 분야에서 인간을 초월하는 성능을 보이고 있으며 많은 주목을 받고 있다. 인공신경망은 생물의 신경 세포가 하는 일을 본따서 만들었다고 알려져 있다. 그러나 이러한 인공신경망을 이용한 현재의 인공지능은 모두 약인공지능에 속하며, 한 가지 일 밖에 수행할 수 있는 능력을 갖추고 있지 않다. 그러면 인공신경망을 이용해서 Strong AI를 만들 수 있을까?

최근 구글의 DeepMind에서 ‘A Generalist Agent’ 논문을 통해 ‘Gato’라는 인공지능을 발표했다. Gato는 다양한 게임을 하고, 이미지 캡션 지정, 인간과의 채팅, 로봇 팔을 제어 하는 등 수백 가지의 일을 할 수 있도록 설계되었다. 해당 논문의 저자 중 Nando de Freitas는 이러한 모델을 규모를 키우고 효율적이고 빠르게 만들면 AGI (Artificial General Intelligence), 즉 Strong AI를 만들 수 있을 것이라고 주장했다. 그러나, 이처럼 다양한 일을 할 수 있음에도 많은 학자들은 Gato의 개발이 AGI를 구축하는 방법을 보여주지 못하며, ‘지능적인’ 것과 거리가 멀다고 주장한다. [1]

인공신경망은, 많은 학습 데이터를 요구하고 주어진 학습 데이터 범위 내에서만 작동하며, 학습이 매우 느리고, 어떠한 결론을 내렸을 때 그 이유와 과정을 설명하는 것이 어렵다. [2] 이는 인간의 사고 과정과는 매우 다르다.

또한 근본적인 문제는, 현재 인공신경망에 사용되는 신경이라고 불리는 것은, 생명 현상을 거의 반영하고 있지 않다는 점이다. 인공신경망에서 신경이라고 불리는 것은 단순히 숫자를 연산하고, 활성화 함수를 거쳐 다음 신경에 가중치를 가지는 시냅스를 통해 전달하는 역할 정도밖에 하지 않는다. 생명현상을 모방했다고는 하지만, 이는 우리의 뇌에서 신경이 작동하는 방식과 다르다.

우리 뇌에서 신경 세포가 활성화 될 때의 스파이크 전위차가 발생하는 빈도, 즉 주파수를 통해서만 정보를 전달한다고 생각한다면, 인공신경망은 그럴 듯 해 보인다. 그러나, 신경 세포는 그렇게 단순한 작업만을 수행하진 않는다. 아직 우리는 하나의 신경 세포가 어떤 일을 하는지, 정보를 어떻게 전달하는지 명확히 알지 못한다. 우리의 신체에는 다양한 종류의 신경 세포가 존재하며, 각기 하는 일도 다르다. 하나의 신경 세포를 모델링한 모델만 해도 단순한 Integrate-and-Fire 모델부터

Izhikevich 모델 [3], Hodgkin-Huxley 모델과 [4] 같은 복잡한 모델들이 존재한다. 이러한 스파이크 모델을 이용해서 인공신경망 모델을 구축한 것을 Spiking Neural Network (SNN)이라고 하며, 해당 인공신경망 모델을 하드웨어로 구현한 뉴로모픽 칩과 함께 에너지 효율적인 차세대 인공신경망으로 관심을 받고 있다. 학자들은 이러한 생물학적 지식을 더 이용해서 AGI, Strong AI에 다가갈 수 있다고 주장한다 [5].

또한 현재 인공지능의 가장 큰 문제점은, 인간의 사고체계와는 완전히 다르다는 것이다. 인간의 사고 체계 및 기억의 중요한 요소 중 하나는 추상화와 일반화인데, 현재 인공신경망에는 이러한 메커니즘이 전혀 담겨있지 않다. 인간의 뇌에서 해당 기능이 어떻게 작동하는지는 알려진 바가 없으나, 신경과학 분야에서 ‘Grandmother Cell’의 개념이 있다. ‘Jennifer Aniston Cell’, ‘Halle Berry Neuron’로도 불리는 이 Grandmother Cell은 특정 개념 혹은 대상에 선택적으로 반응하는 뉴런을 뜻한다 [6]. 현재 아직 가설로만 존재하며, 학자들 사이에서 논쟁이 있는 개념이지만, 많은 실험 자료가 Grandmother Cell이 inferotemporal cortex, superior temporal sulcus, temporal lobe 등 뇌의 특정 부분에 존재함을 뒷받침한다 [7, 8, 9]. Grandmother Cell은 기억의 기능에도 연관되어 있는 것으로 생각되며, 해당 개념을 이용하여 추상화 및 일반화 기능과 기억에 관한 기능을 인공지능, 인공신경망에 구현할 수 있을 것으로 생각되고, Strong AI에 더욱 근접할 수 있다고 생각된다.

이러한 생물학적, 특히 뇌과학적 지식을 기반으로 인공지능을 개발하면 언젠가는 Artificial General Intelligence, Strong AI에 도달할 수 있을 것으로 생각된다. 설령 Strong AI를 개발하는 것에 실패하더라도, 뇌과학의 연구와 이를 이용한 인공지능 연구가 진행됨에 따라 인류는 존재에 대한 의미와 생각이라는 과정에 대해 얻는 통찰이 더욱 많아질 것으로 생각된다.

Reference

- [1] M. Heikkilä, The hype around DeepMind's new AI model misses what's actually cool about it, MIT Technology Review, <https://www.technologyreview.com/2022/05/23/1052627/deepmind-gato-ai-model-hype/amp/>
- [2] M. Świechowski, Deep Learning and Artificial General Intelligence: Still a Long Way to Go, <https://doi.org/10.48550/arXiv.2203.14963>
- [3] E.M. Izhikevich, Simple model of spiking neurons, IEEE Transactions on Neural Networks, <https://doi.org/10.1109/TNN.2003.820440>
- [4] A. L. Hodgkin, A. F. Huxley, A quantitative description of membrane current and its application to conduction and excitation in nerve, <https://doi.org/10.1113/jphysiol.1952.sp004764>
- [5] L. Deng, H. Tang, K. Roy, Editorial: Understanding and Bridging the Gap Between Neuromorphic Computing and Machine Learning, Frontiers in Computational Neuroscience, <https://doi.org/10.3389/fncom.2021.665662>
- [6] J. S. Bowers, Grandmother cells and localist representations: a review of current thinking, Language, Cognition and Neuroscience, <https://www.tandfonline.com/doi/full/10.1080/23273798.2016.1267782>
- [7] C. G. Gross et al., Visual properties of neurons in inferotemporal cortex of the Macaque, 10.1152/jn.1972.35.1.96
- [8] C. Bruce et al., Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque, 10.1152/jn.1981.46.2.369
- [9] S. M. Landi et al., A fast link between face perception and memory in the temporal lobe, 10.1126/science.abi6671