



Special Features of Deep Learning and Symbol Emergence

Yutaka Matsuo¹

Published online: 16 March 2020

© Ohmsha, Ltd. and Springer Japan KK, part of Springer Nature 2020

Keywords Deep learning · Symbol manipulation · World model

This special issue entitled “Deep Learning and Symbol Emergence” presents discussion of technology beyond deep learning toward human intelligence. In recent years, deep learning has achieved phenomenal success in various fields such as image recognition, speech recognition, game-playing, and natural language processing. Nevertheless, discussion persists among researchers on the limitations of deep learning. The most popular criticism is that what deep learning has accomplished to date is mere pattern recognition. To realize more advanced intelligence toward achieving human intelligence requires greater integration with symbol manipulation such as logical reasoning and causal reasoning [1]. Bengio recently delivered an invited talk at the NuerIPS 2019 conference entitled “From system 1 deep learning to system 2 deep learning,” which emphasized the necessity of integrating causal reasoning with current deep learning methods. A dichotomy of system 1 and system 2 was presented by Karneman, a Nobel laureate in the area of economy and psychology [2]: System 1 refers to “fast thinking” as pattern recognition by current deep learning. System 2 refers to “slow thinking,” which is typically done as conscious and logical thinking.

Integration of current deep learning with symbol manipulation is not an easy path. Several hurdles remain. First, we must produce some means of simulating the environment of an agent, a so-called world model. Several studies have examined world models [3, 4], which attempt to model the environment around an agent using deep learning. To date, most studies have specifically examined prediction based on sensory data, but integration of sensor and motor data are important. Second, it is necessary to integrate the world model with symbols. Some concepts or features learned in the world model can be associated with linguistic patterns such as a word or a phrase. Prediction of the utterances from other people might serve as a strong prior to promote learning.

This special issue includes three papers addressing the direction laid out above. The first paper, by Ito et al. [5], targets modeling of the sensory-motor signals of a robot. This task is fundamentally important to produce a world model so that a

✉ Yutaka Matsuo
matsuo@weblab.t.u-tokyo.ac.jp

¹ The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

robot can obtain important concepts and features while finding a low-dimensional manifold that describes the agent itself and the surrounding environment. To have such world models, one usually uses deep generative models, which is the topic of the second paper by Taniguchi et al. [6]. The paper presents a new framework called Neuro-SERKET, which can accommodate advanced versions of deep generative models. It integrates variational auto-encoder (VAE), which is among the most popular generative models, Gaussian mixture model (GMM), and latent Dirichlet allocations (LDA), which are also used polarly to model data of various kinds, and automatic speech recognition (ASR). It can accommodate from sensori-motor data to speech signals, which can be regarded as linking system 1 and system 2. The third paper by Yamakawa [7] addresses symbolic aspects, which specifically examine how one can predict the next utterance. It describes current deep learning methods such as attentions and transformers, and their relations to brain activities.

In summary, we attempted to show the future direction of deep learning research toward human intelligence. We hope the three papers can provide some insights to readers.

References

1. Ford, M.: Architects of Intelligence: The Truth About AI from the people building it. Packt Publishing, Brimingham (2018)
2. Kahneman, D.: Thinking Fast and Slow. Farrar, Straus and Giroux, New York (2013)
3. D. Ha, J. Schmidhuber: World Models (2018). (arXiv:1803.10122)
4. Eslami, A., et al.: Neural scene representation and rendering. *Science* **360**(6394), 1204–1210 (2018)
5. Ito, H., et al.: Evaluation of generalization performance of visuo-motor learning by analyzing internal state structured from robot motion. *New Gener Comput* (2020). <https://doi.org/10.1007/s00354-019-00083-x>
6. Taniguchi, T., et al.: Neuro-SERKET: development of integrative cognitive system through the composition of deep probabilistic generative models. *New Gener Comput* (2020). <https://doi.org/10.1007/s00354-019-00084-w>
7. Yamakawa, H.: Attentional reinforcement learning in the brain. *New Gener Comput* (2020). <https://doi.org/10.1007/s00354-019-00081-z>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.