Hyomin Seo

Title : Paper Analyzation on "Neural Archimetric Logic Unit", " Hybrid Computing"

Reading Links :

https://arxiv.org/pdf/1808.00508.pdf: Neural Archimetric Logic Unit
https://www.nature.com/articles/nature20101: Hybrid Computing Using a Neural Network with Dynamic External Memory

The paper 'Neural Archimetric Logic Unit' addresses some limitations of classical neural networks, several arithmetic operations that are considered challenging to classical neural networks. Even though these neural networks are flexible, if the test data exceeds the training range, they are most likely to fail immediately. Such failure proves that while a multi-layer perceptron is theoretically capable of working with any continuous functions, the choices such as training data, learning schedule, and architecture significantly affect the training of the function, inevitably resulting in a biased result. Neural Arithmetic Logic Unit (NALU) is presented as a new module that can produce a good performance in such conditions. NALU can track time, perform arithmetic over images of numbers, translate numerical language into real-valued scalars, execute computer code, and count objects in images.

The next paper regarding the adoption of Dynamic External Memory discusses the method that adopts external memory blocks to surpass short-term memory issues in deep neural networks. This external differentiable memory block is modeled referring to human memory theory, focusing on theories such as attribute theory, similarity measures, a search of associative memory, temporal context model, and serial recall. Three different attention mechanisms are defined for reading/writing and replacing in external memory: soft attention mechanism, the attention of temporal links for replacing memory and attention for writing in memory matrices. These methods are end to end differentiable extensions of Neural Turing Machines. The paper suggests and adduces that the proposed methods can be used for multi-domain association problems with guaranteed sufficient time for self-correction.

Both methods provide a pin-point solution for the specific limitation of typically used models, and it is shown that when these methods are adopted, the limitation is resolved. On the other hand, both models exchange their ability to excel in that limitation with the general advantage of the typical network; while external memory provides the multi-purpose capacity to neural networks, it is not able to generalize learning. The absence of generalizing ability is as critical as the advantage that can be achieved by adopting external memory blocks, which might be the reason that the mainstream is not yet fully applying such models into their bulk system.