Programmable iontronic neural networks

openaccessgovernment.org/article/programmable-iontronic-neural-networks/189616

12 March 2025

Guo-Xing Miao, Professor at the University of Waterloo, guides us through programmable iontronic neural networks

Inspired by both the biological neurons and the thriving lithium-ion batteries, we construct iontronic memristors for harvesting neuromorphic functionalities.

The impact of artificial intelligence on life

The advancement of artificial intelligence (AI) has been accelerating and has dramatically impacted our daily lives, such as content generation, supply chain optimization, financial management and forecasting, targeted database search, automated vehicles/robots, embedded systems with unsupervised decision-making, and many more scenarios. It is unsurprising that all major players are heavily investing in AI in the current race towards true artificial general intelligence (AGI).

The model training requires a considerable amount of resources, both in terms of data availability and in terms of hardware and energy cost. The effect is clearly visible: the bigger the model, the better the AI capability. This can be analogized with human brains, which contain about 100 billion neurons and 100 trillion synapses/connections, making us much smarter than, say, a mouse with 100 million neurons.

Language processing tasks and more

While the cerebrum in charge of our logical thinking only accounts for 20% of the human brain in terms of neuron numbers, and only about 6% of the cerebrum is used in our language processing tasks. From here, we can see that a normal person's language processing skill is roughly on the same order of magnitude as GPT3, with 175 billion parameters.

While it is probably not solid to quantify that one bio-synapse connection equals one digital training parameter to the bulk order of magnitude, it does match our perception of these models. We can probably safely say that GPT-4, with 1.8 trillion parameters, has better language skills than a normal human being. And on top of that, it has practically unlimited memory covering everything in the world.

If resources are not an issue, mounting GPU clusters and data centers and power plants can indeed get the AI models closer and closer to AGI. The size of the AI model roughly sets the raw power of the AI, and smarter implementation of the hardware can add some further improvement.

The fact that GPT-5 is being debated whether it is already an AGI, or not, indicates that people have made huge steps towards this goal. However, this approach of scaling up the model size will encounter the ultimate challenge when unsupervised, standalone AI units are necessary.

An analogy with human brains

This is another good place to make an analogy with human brains. With a typical power consumption of no more than 10W, a human brain can perform complex tasks superior to any big, large language models (LLMs). Information processing in a brain is driven by the gated discharge of ionic carriers on a time scale much slower than the electrical processing speed in CMOS circuits.

These point to the possibility of realizing analog iontronic neuromorphic computing, rather than digitally synthesized neurons, for more efficient information processing. In analogy to a human brain, in terms of accuracy, our brain is probably not as good as a simple calculator, but in terms of logical comprehension, even GPT5 needs to stand behind.

Here we show a configurable iontronic platform that directly embeds ionic motion into memristors and is also CMOS compatible. Memristors are a type of resistive element that can memorize its history and, therefore, are suitable as the basic training elements in analog AI architectures.

Borrowing the rich knowledge of ion motion in rechargeable battery systems, we construct our devices in a “battery-like” fashion, in which lithium- ions can be controllably injected into and extracted from the cathode-like memristor switching layers via an actual solid-state electrolyte.

The voltage-controlled ion influx leads to corresponding conductance changes that mimic the gated ionic channels in bio-synapses. The conductance change is volatile after the excitation is removed due to the ions self-diffusing back. This mimics the short-term memories and spike-timing-dependent plasticity in biological neural networks.

After a forming process, the same devices can be programmed into their non-volatile version that can permanently store information analogly or digitally, mimicking a brain’s long-term memories. The devices had been further back-end-of-line (BEOL) integrated with mature CMOS chips, giving them complete freedom to advance on the shoulders of giants.

In an idealized architecture, the ion-enhanced memristors are laid down on the CMOS platforms and cross-link into extensive neural networks. The units can be individually programmed into volatile or non-volatile memory nodes; therefore, the same platform can be customized towards the desirable applications – a field- programmable neural network array. This permits a virgin array to adaptively evolve around a task and become specialized processing units. A fully connected array with nodes as many as a real brain is hard to envision now, but not impossible with modern technologies. AI would also evolve towards more general capabilities.

Here, we have seen an intriguing example where nanotechnology, renewable energy technology, and quantum technology all seamlessly work together to produce something new and interesting. The nature of ion motion indicates that the operation speed would be slower than modern electronics, but that is also the nature of a biological brain.

We win out regarding our information processing efficiency, where data are parallelly processed in “concepts” rather than digitally processed “word-by-word.” The coexistence of centralized AI, characterized by its immense power and capabilities, with distributed AI, known for its limited resources but better efficiency, is pivotal for the new era to come.

Contributor Details

- Article Categories
- Fundamental Research

- Article Tags
- AI
- Biology
- Technological Innovations

- Publication Tags
- OAG 046 – April 2025

- Stakeholder Tags
- Institute for Quantum Computing (IQC) - Department of Electrical and Computer Engineering - University of Waterloo