

A conversation with Professor Erik De Schutter, September 17, 2019

Participants

- Professor Erik De Schutter - Professor, Computational Neuroscience, Okinawa Institute of Science and Technology
- Joseph Carlsmith - Research Analyst, Open Philanthropy

Note: These notes were compiled by Open Philanthropy and give an overview of the major points made by Prof. De Schutter.

Summary

Open Philanthropy spoke with Prof. Erik De Schutter as part of its investigation of what we can learn from the brain about the computational power (“compute”) sufficient to match human-level task performance. The conversation focused on the complexity of different processes in the brain.

Neuroscientific uncertainty

People often confuse the question of how to build AI systems with the question of which features of the human brain give rise to functional intelligence. In principle, you could probably build AI systems that have very little to do with how the human brain works.

The latter, brain-focused question is very open. Neuroscientific knowledge has improved a lot, but there is a lot we still don’t know. Some scientists ignore this uncertainty, and others worry about it. And with improvements in imaging and cell biology techniques, we discover all sorts of new complexities that we didn’t know were there.

Evidence about adequate modeling detail

Prof. De Schutter thinks that at this point, we simply are not in a position to place any limits on the level of biological detail that might be relevant to replicating the brain’s task-performance. Many common simplifications do not have solid scientific foundations, and are more at the level of “the way we do things.” The best way forward is to try to explore and understand the function of the brain’s underlying mechanisms -- a project that may eventually lead to an understanding of what can be simplified. But to try to simplify things too early, before you understand them, is a dangerous game.

Modeling neural networks at the level of simple spiking neuron models or rate-based models is very popular. Prof. De Schutter thinks the field would benefit from a greater diversity of approaches.

Limitations of data about input-output behavior

At this point, we have no way to reliably measure the input-output transformation of a neuron, where the input is defined as a specific spatio-temporal pattern of synaptic input. You can build models and test their input-output mappings, but you don't really know how accurate these models are.

There is a tradition of integrate and fire modeling that achieves very accurate fits of neuron firings in response to noisy current injection into the soma (more accurate, indeed, than could be achieved by current biophysical models). However, this is a very specific type of experiment, which doesn't tell you anything about what happens to synaptic input in the dendrites.

In live imaging, it's very difficult to see what's happening at synapses. Some people do calcium imaging of pre-synaptic terminals, but this is only for one part of the overall synaptic input (and it may create artefacts). Currently, you cannot get a global picture of all the synaptic inputs to a single neuron. You can't stain all the inputs, and for a big neuron you wouldn't be able to image the whole relevant volume of space.

Using glutamate uncaging, you can reliably activate single dendritic spines *in vitro*, and you can even do this in a sequence of spines, thereby generating patterns of synaptic input. However, even these patterns are limited. For example, you can't actually activate synapses simultaneously, because your laser beam needs to move; there's only so much you can do in a certain timeframe; and because it's glutamate, you can only activate excitatory neurons. What's more, you don't actually know what the physiological pattern of inputs is.

Increased resolution of detailed modeling

Old multi-compartmental models, based on cable theory, described voltage in one dimension, and the typical resolution was on the order of tens of microns per compartment. That is adequate for modeling voltage, but molecular events happen on much smaller scales.

Researchers now have much more computing power available to them, and so can build more ambitious models. For instance, they can now use fully stochastic, three-dimensional "mesh" models with sub-micron resolution (typically on the order of 100 nanometers). These can incorporate molecular reactions, as well as features of cell biology like spatial models of synaptic vesicles.

Examples of biophysical dynamics simpler models leave out

Here are some examples of possibly important biophysical dynamics that simpler models leave out.

- Synaptic transmission in the central nervous system is stochastic. Only about 20% of spikes actually trigger neurotransmitter release. This has never been fully explained, and it is almost never included in neural network models. Some hypothesize that it's about energy efficiency, but there is no proof of this. It's an open question whether you could capture this stochasticity by drawing from a relatively simple distribution, or whether the brain manipulates synaptic stochasticity in more computationally complex ways.
- Networks of neurons can rewire themselves fairly quickly, over timescales of tens of minutes. These changes correlate with improvements in performance on tasks.
- Scientists have recently discovered that dendrites can release neurotransmitters to neighboring cells.
- Astrocytes can release vesicles loaded with mRNA to neighboring cells. This might be a mechanism for expediting axonal repair.
- There is some work on how a learning process analogous to backpropagation might be implemented via the interaction between different signals arriving at the basal and apical dendrites (though how this interaction takes place is currently the weakest part of the model). This idea is interesting (though it is not empirically proven), and it illustrates the possibility that simplified models (e.g., those without dendrites) can leave out important dynamics.
- The distribution of receptors on the postsynaptic side of the synaptic cleft has been optimized for diffusion. You probably don't need to model the diffusion of individual neurotransmitter molecules to capture what matters about synaptic transmission, but understanding the brain at this level of detail may provide insight into the system's limitations.

Engineered vs. evolved systems

The brain was not engineered. Rather, it evolved, and evolution works by adding complexity, rather than by simplification.

There are good reasons for this complexity. In order to evolve, you can't have systems, at any level (proteins, channels, cells, brain regions), with unique functions. If you did, and a single mutation knocked out the function, the whole system would crash. Whereas if you have overlapping functions, performance suffers somewhat, but something else can take over. If you don't allow for this, you can't evolve, since evolution works by random mutations, and most mutations are not positive.

In Prof. De Schutter's view, neuroscience hasn't adequately appreciated this point. You often see papers claiming that thing X has function Y, without exploring the possibility that it has other functions as well.

Indeed, in general, many scientists who approach the brain from an engineering perspective end up on the wrong footing. Engineering is an appropriate paradigm for building AI systems, but if you want to understand the brain, you need to embrace the fact that it works because it is so complicated. Otherwise, it will be impossible to understand the system.

Quantum dynamics

Prof. De Schutter has not delved deeply into the possibility that quantum dynamics play a functional role in the brain, and available theoretical frameworks and empirical evidence might not be strong enough to support rigorous investigation. He is also wary of the history of comparing the brain to the latest engineering technology (e.g., a steam engine, a classical computer, now maybe a quantum computer).

However, if it is possible to implement quantum computing with the brain's biological infrastructure, he thinks it would be a mistake to assume that the brain would not be able to find a way to exploit this possibility.

*All Open Philanthropy conversations are available at
<http://www.openphilanthropy.org/research/conversations>*