A conversation with Professor Markus Meister, September 24, 2019

Participants

- Professor Markus Meister - Anne P. and Benjamin F. Biaggini Professor of Biological Sciences at the California Institute of 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. Meister.

Summary

Open Philanthropy spoke with Prof. Markus Meister of the California Institute of Technology 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 evidence that the retina provides in this regard.

Current understanding of the retina

The computations performed in the retina are fairly well-understood. There is more to learn, of course, but the core framework is in place. We have a standard model of the retina that can account for a lot of retinal processing, as well as predict new observations.

Modeling requirements and results

If your goal is to predict the spiking outputs of the retina, you don't need a highly intricate model (for example, you don't have to simulate the details of every neuron using multi-compartmental models). Rather, you can use very compact models known as "point neuron models," which you can connect together with simple synapses. This type of model has worked well in predicting retinal outputs.

Current retinal models can explain something like 80% of the variance in the retina's response to visual stimuli drawn from natural scenes. It has taken more effort to simulate retinal responses to natural scenes than to artificial stimuli used in labs (e.g. spots, flashes, moving bars).

To create a functional model of the whole retina, in the extreme case you'd need a point-neuron model for every cell. However, you can probably get away with less than that, because there are a lot of regularities that can be simplified computationally. Thus, rather than simulating every cell, you can capture the activity of a population of cells more efficiently.

Adaptation

Experience can alter the state/function of the retina. For example, the retina adapts to changes in lighting conditions, and to the statistics of its environment. This adaptation occurs over timescales of seconds to many minutes, as opposed to the ten ms timescales of the immediate visual response. Most light adaptation occurs in the photoreceptors, but some occurs in bipolar cells as well.

It's probably reasonable to assume that the retina is less plastic than other circuits in the brain. People have looked for more sophisticated forms of learning in the retina, like image storage, but haven't found them.

The biochemistry involved in retinal light adaptation is well-understood, and it can be captured using a simplified computational model. Specifically, you can write down a three-variable dynamical model that gets it about 80% correct. The compute required to run a functional model of the retina would probably be dominated by the feedforward processing in the circuit, rather than by capturing adaptation.

Scientific advantages of peripheral systems

The retina is probably the best understood part of the brain. In general, the peripheral sensory and motor systems are much better understood than the inner parts of the nervous system. This is for a number of reasons.

Experimental access

In a sensory system, you have full control over at least one of the variables that matters -- e.g., the sensory stimulus. And in the motor system, you can measure the output (e.g., muscle fibre activity) very accurately.

Functional understanding

You also know that what you're measuring (e.g., the transduction of sensory signals, or the muscle fiber output) is the relevant quantity. In the middle of the brain, by contrast, there is more uncertainty about what matters. E.g., if you're measuring spike trains in the hippocampus, you don't know which aspects of those spike trains are fundamental to what the system is trying to do, and which are epiphenomenal/accidents of implementation.

Functional specialization

There is a long history, in neuroscience, of attempting to assign understandable computational roles to little chunks of brain matter (e.g., "the anterior cingulate cortex is for X"). Prof. Meister believes that this program is not going to be very successful, because these regions are massively interconnected, and we now know that if you inject signals into one part of the brain, you find them in many other parts of the brain.

In the periphery, though, this is less of a problem. For example, we can specify the function of the retina fairly unambiguously: it is the source of all visual input to the brain, it has to transduce light into neural signals, and it performs some initial processing on those signals. Information in the retina also flows in an almost exclusively feedforward direction (though there are some feedback signals, and it is an interesting question what those fibers do).

We can state the function of other neural circuits involved in motor outputs -- for example, central pattern generators that produce rhythmic movements involved in locomotion -- with similar clarity.

Retinal prostheses

Despite 30 years of effort, attempts to create functional artificial retinas have met with very little success. Recent performance tests show that people implanted with the devices are functionally blind -- e.g., they cannot read, and they cannot distinguish between letters unless the letters occupy the entire visual field.

However, this lack of success is not about computation. People in the field generally agree that if you could make the right kind of one-to-one connection to the optic nerve fibers, you could compute spike trains that would allow the brain to see. The obstacle is actually making the interface between an electrical device and the retina. Electrodes on top of the retina stimulate many nerve fibers at once; you don't know ahead of time which fiber you'll be stimulating or what type of retinal ganglion cell you're connected to, and you can't get data into the eye at the right rate.

Generalizing to the rest of the brain

There is nothing particularly simplistic about the retina, relative to other neural circuits. It probably has a hundred different cell types, it probably uses almost every neurotransmitter we know of, and it has very intricate microcircuitry. Prof. Meister would be sympathetic to scaling up from the retina as a way of putting an upper limit on the difficulty of simulating the brain as a whole. Prof. Meister has not actually done this back-of-the-envelope calculation, but budgeting based on the rate at which action potentials arrive at synapses, multiplied by the number of synapses, seems like roughly the right approach.

It is theoretically possible that the brain's task-performance draws on complex chemical computations, implemented by protein circuits, that would require models much more complicated than those that have been successful in the retina. But Prof. Meister's approach is to ask: is there any evidence that forces us to think in this more complicated way? That is, he starts with the simplest possible explanation of the phenomena, and then adds to this explanation when necessary. Some neuroscientists take a different approach. That is, they ask "what is the most complicated way that this thing could work?", and then assume that nature is doing that.

There is evidence that single point neuron models are not sufficient to explain all neural phenomena. For example, in cortical pyramidal cells, the basal dendrites and soma operate with different dynamics than the apical tuft. Using two point-neuron models (one for the soma, and another for the apical tuft), you can capture this fairly well. These are more powerful models, but they are not dramatically more computationally complex: e.g., it's basically a factor of two. We don't know exactly what phenomena this additional complexity allows you to capture, because the range of phenomena that matter in the cortex is less clear than it is in the retina.

Complexity of human cognition

Simulating the brain may not be the right way to reproduce human-level cognitive performance. For example, the brain's solutions may be terribly inefficient. Indeed, we know this is true in some cases. Synapses are noisy, and silicon isn't; and the brain uses huge numbers of neurons to represent the same variable, probably because a single neuron can't do it robustly. Prof. Meister expects that human-level AI systems will use methods more naturally suited to silicon devices. This would suggest compute estimates lower than what scaling up from the retina would suggest.

Prof. Meister thinks that people often overestimate the sophistication of the tasks that humans perform, which tend to involve low-bandwidth outputs. People have measured the bits per second involved in different types of motor outputs (e.g., typing, playing piano, athletics, speaking speed, etc), and the numbers are in the range of 10-40 bits per second. Similarly, people have tried to measure the information rate of human thought (for example, by seeing how much information humans can retain per second in reading), and it's in the same ballpark.

Bits are a reasonable common currency for comparing humans with artificial systems, and we make systems that operate with much higher data-rates than humans all the time. There are some tasks, like prime factorization, that require large amounts of compute despite low-bandwidth inputs and outputs. But Prof. Meister does not see positive evidence that the tasks that humans perform are like this.

What's interesting about humans is that they can do many different things, and change what they are doing very fast. However, Prof. Meister does not think that this is all that conceptually mysterious.