RECORDING FROM MANY NEURONS SIMULTANEOUSLY

FROM MEASUREMENT TO MEANING

Krishna V. Shenoy

The human brain is comprised of approximately one hundred billion neurons, yet most of what is known comes from measuring the activity of one neuron at a time. Or, at the other extreme, studies rely on measuring the aggregate activity of thousands to millions of neurons at a time. This profound measurement limitation is changing rapidly. It is now possible to measure activity from many hundreds to thousands of individual neurons all at the same time, and it is widely believed that it will soon be possible to measure from many hundreds of thousands, or even millions, of neurons. As game changing as these breakthroughs are, several barriers to converting raw biological measurements into fundamental scientific meaning remain. Two of these challenges—making sense out of activity from large numbers of neurons and the importance of "levels of abstraction"—are discussed below.

Measuring Activity from Large Numbers of Neurons in the Brain

Neuroscientists seek to understand the function and dysfunction of the nervous system, including, ultimately, the human brain. The reasons for this pursuit are simple: to advance scientific knowledge about one of, if not the, most complicated systems in the universe as well as to help alleviate the burden of neurological disease and injury. In order to understand how a system like the brain operates one must measure its internal workings, much like understanding a computer requires measuring voltages and currents throughout its circuitry. In the case of the brain, this means measuring electrical activity (for example, action potentials, field potentials), chemical activity (for example, neurotransmitters, ion concentrations), and likely both throughout its neural circuitry. Pioneers in neuroscience have relied on various measurements in order to take accurate readings of electrical and chemical activity, and these measurements have tended to focus either on individual neurons (for example, intracellular electrode, extracellular electrode) or on aggregate activity from numerous neurons (for example, EEG, MEG, fMRI). Similarly, powerful stimulation technologies have been used to causally perturb neural activity and observe the consequences (for example, electrical microstimulation, TMS, optogenetics).

While many seminal discoveries, insights, and Nobel Prizes have resulted from these measurement (and stimulation) technologies, a renewed appreciation for the complexity of the overall nervous system and the associated need for measuring simultaneously from many individual neurons have arisen in recent years. Fortunately, technological innovation has risen to meet this need, making it now possible to measure from hundreds to thousands of individual neurons at the same time. For example, genetically encoded calcium indicators (for example, GCaMP, see chapter by Ahrens, this volume) allow calcium concentration changes associated with action potentials from thousands to tens of thousands of individual neurons to all be optically imaged simultaneously. The full potential of this class of measurement is still being realized, with animal models ranging from immobilized worms, walking transgenic mice, and freely moving rats already in use, to possibilities on the horizon including monkeys performing a variety of cognitive tasks.

More traditional electrode-based technologies have also scaled up in recent years. One example—a one-hundred-electrode array—is shown in figure 1a. One or more of these arrays can be implanted permanently in the brains of rats, monkeys, and humans (as part of FDA pilot clinical trials focused on neural prostheses to help people with paralysis). These electrodes can measure electrical activity (extracellular action potentials, field potentials) from tens to hundreds of individual neurons while animals perform a variety of cognitive tasks including sensory, decision making, and motor behaviors as shown in figure 1b.

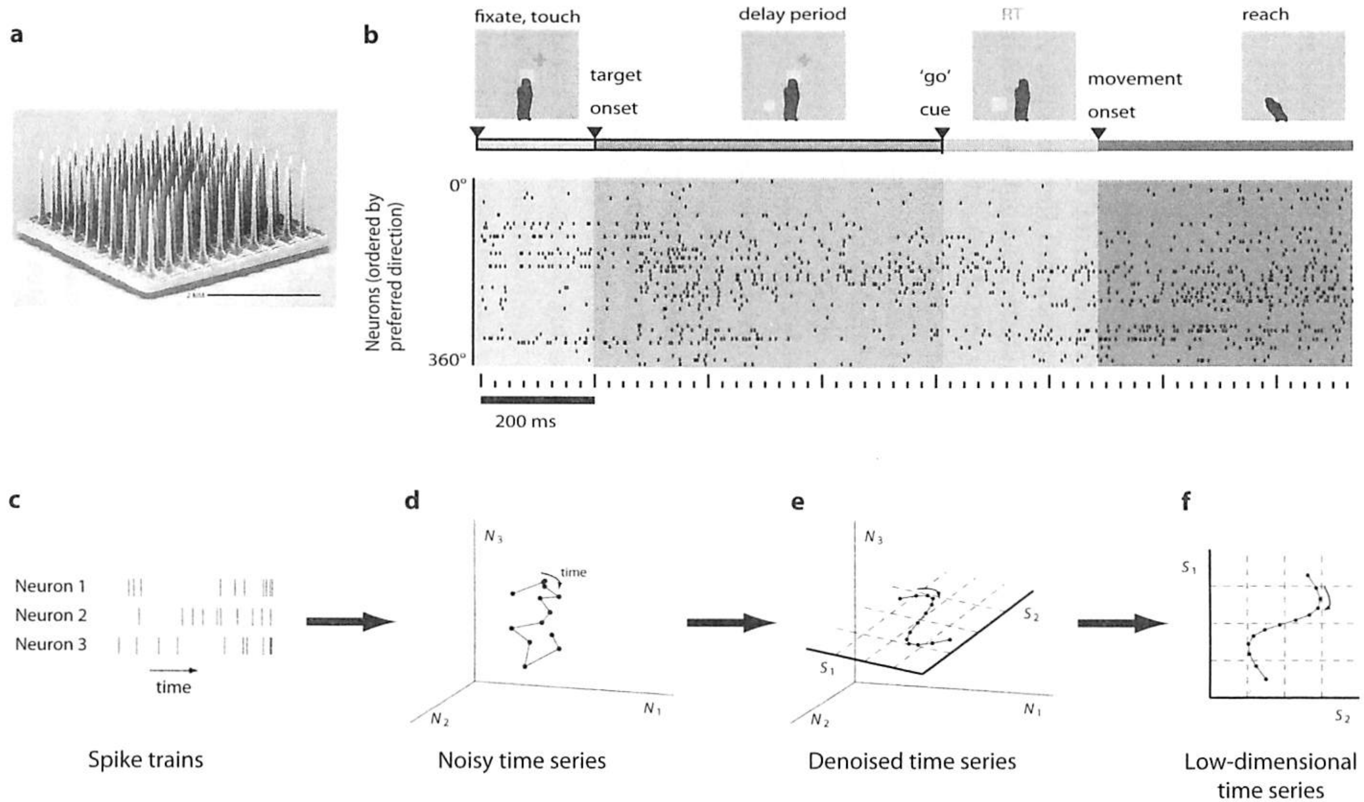
a
b
fixate, touch
target
onset
delay period
'go'
cue
RT
movement
onset
reach
0°
360°
Neurons (ordered by preferred direction)
200 ms
c
Neuron 1
Neuron 2
Neuron 3
time
Spike trains
d
N_3
time
N_2
N_1
Noisy time series
e
N_3
S_2
S_1
N_2
N_1
Denoised time series
f
S_1
S_2
Low-dimensional
time series

Figure 1. Key steps in extracting low-dimensional, single-trial, neural population state-space trajectories from multichannel neural data. a. A silicon-based, 4 x 4 mm, 100 electrode array with 1 mm long electrodes made by Blackrock Microsystems Inc. The analysis begins with multineuron data from such a device. b. The instructed-delay center-out reaching task, with the neural response from an electrode array surgically implanted in premotor cortex. Central targets are fixated and touched with the eye and arm. A spot of light then appears on the screen indicating the target that should eventually be touched. Following a delay period a "go" cue is given, leading to an arm movement and touching of the target. "Brain states" associated with each part of this type of task can be measured from the activity of a population of neurons in the appropriate brain region(s). Each row of dots represents the times of action potentials (spikes) from one of 44 neurons. c–f. Computational steps for extracting a population neural trajectory from multiple spike trains on a single trial. For clarity, c illustrates spike trains recorded simultaneously from 3 (of the 44) neurons shown in panel b. d. Depicts the time evolution of the recorded neural activity plotted in a 3-dimensional space, where each axis measures the instantaneous firing rate of a neuron (e.g., N1 refers to neuron 1). Firing rates may be estimated in brief time bins (e.g., 10 ms). e. Depicts the population neural trajectory (a denoised version of the trajectory in d), which is shown to lie within a 2-dimensional space with coordinates S1 and S2. Finally, f. shows the population neural trajectory visualized directly in a low-dimensional state space and can be referred to using its low-dimensional coordinates (S1, S2). This final panel illustrates a low-dimensional, single-trial, neural population state-space trajectory computed from neural data measured simultaneously from many neurons. For further reading see Yu et al. (2009).

Even more revolutionary measurement technologies are also being developed, but the two technologies described above serve as examples that it is now possible to record simultaneously from hundreds to thousands of individual neurons.

Making Sense out of Activity from Large Numbers of Neurons

The first challenge to converting this newfound torrent of neural measurements into fundamental scientific meaning is to ask how to "make sense of the data." This is a deceptively simple-sounding question, as it would appear that we could just keep analyzing the measured data as we always have but now do so with a lot more presumably beneficial data. However, this would be overlooking many likely benefits of having massively parallel neural measurements where each neuron is measured with high temporal precision. Moreover, there may also be additional new information available such as cell type, axonal and dendritic projection pattern, and synaptic connection strengths. By analogy again to a computer, if presented with the opportunity to measure from one thousand transistors *simultaneously* it would save time relative to measuring one thousand transistors one at a time—but there are other far more important advantages as well. The reasons for this are developed more fully below.

Are there different ways forward? There are undoubtedly many potential ways forward, and at least one has been pursued in recent years and is termed the "dynamical systems approach" since it is borrowed and adapted from physical science and engineering where dynamical systems design and analysis is a staple. Three central elements to the dynamical systems approach are as follows. First, measured neural data constitute a time series, where there is correlation structure between measurements nearby in time. As such some form of temporal smoothing may be appropriate, and may help combat noise inherent in neural measurement. This is depicted in figure 1c–e. Second, the simultaneously measured neural data constitute a high-dimensional dataset but putatively actually occupies fewer dimensions. Dimensionality reduction, a major topic in machine learning and statistics, can be used to infer a lower-dimensional manifold on which the data reside. This is depicted in figure 1f. Taken together it is possible to visualize the nominally important dimensions that vary in the data and to then see how these population neural trajectories correspond to cognitive variables, such as the time it takes the arm to start moving following a "go" cue (reaction time, RT) or the direction in which the arm will move (see again figure 1b). It is important to note that very-low-dimensional visualizations, such as two- or three-dimensional figures drawn on paper, almost certainly miss some information. Thus such visualizations are useful for building intuition, but answering scientific questions must be done with higher-dimensional data where little if any information is lost. Finally, the dynamical systems approach seeks to estimate, quantitatively, the rules governing the evolution of the population neural state. This is akin to ascertaining Newton's laws from observations of a ball rolling on an uneven surface such that momentum, friction, and elasticity can be characterized. Together, visualizing lower-dimensional population neural trajectories, so as to generate hypotheses about how the neural circuit is working as a whole and relates to (single-trial) behavior, and identifying the equations of motion (for example, using a family of techniques known as systems identification) are a framework for leveraging massively parallel neural measurements into nominally meaningful scientific insights.

The Importance of "Levels of Abstraction"

The second challenge to fruitfully converting unprecedented volumes of neural data into scientific discoveries and insights—as opposed to potentially "drowning in data"—is to know what to pay attention to. This is certainly easier said than done when it comes to the brain, which is still poorly understood and it is unclear what details matter at a given level of investigation. Does the detailed connection pattern and synaptic strengths for each neuron matter when attempting to relate population neural activity to an arm movement? Does the exact pattern of action potential emission times matter when neurons must constantly contend with (probabilistic) synaptic failure? These questions, and countlessly many more, are open questions in neuroscience. Nevertheless, we can likely benefit by at least being aware that other fields in physical science and engineering contend with similar problems by adopting a well-proven philosophy for the design and analysis of physical systems.

This ubiquitous and essential concept to understanding and designing physical systems is termed "levels of abstraction." We anticipate that levels of abstraction will be of growing importance when investigating biological systems, including the brain. We describe here an analogy between a well-understood electronic system and the nervous system in order to highlight the potential merits of increasingly employing levels of abstraction in brain science.

Modern computer systems are comprised of several integrated circuits ("chips") connected together, and connected with peripherals such as displays, keyboards, and networked devices. Consider just one of these chips, the central processing unit (CPU), and how we can understand how it works. At the smallest level are atoms arranged precisely to bestow transistors with the desired electrical properties. Transistors come in a multitude of sizes and types, number in the billions, and form the next level of the CPU. The third level is the wiring between the transistors, which can be quite complex and have hundreds of millions of individual wires, due to wires bridging over and tunneling under each other similar to a metropolitan highway system. The fourth and final level, again broadly speaking, is the software. Software ranges from the detailed control of specific hardware (machine code) through the more global coordination of resources and data (operating system, algorithms). The software level is distinct from the other three because it resides in the pattern of electrical states (1s and 0s), as opposed to being physically manifest, and because it can grow to essentially arbitrary complexity by expanding well beyond the not uncommon millions of lines of code.

What does this have to do with the brain? Any detailed, literal comparison between the brain and a CPU is doomed. Examples of this type of flawed, detailed comparison that have been put forth in recent years include likening a computationally rich neuron to a computationally impoverished transistor (that is, a simple switch in a digital system), or likening the three-dimensional point-to-point connections between neurons to the essentially two-dimensional and relatively less general connections among transistors. Nevertheless, a broad comparison may help highlight how the levels of abstraction concept is anticipated to help shed insights on how the brain works. Importantly, this concept is related to David Marr's trilevel hypothesis, where in broad terms Marr's computational

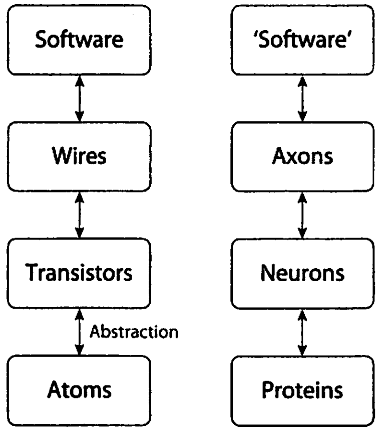

Figure 2. Levels of abstraction for a CPU (left column) and the brain (right column). Arrows indicate how detailed information at one level is abstracted away, so as to pass along only the essential operating principles and characteristics to the next level. Arrows are bi-directional to indicate that abstraction is beneficial both to understanding how physical implementation impacts software capability (bottom-up) as well as how software requirements impact physical design (top-down).

and algorithmic/representational levels are grouped, for brevity, into the software level, and his physical level appears here as the first three levels to reflect the increasingly detailed physical information available.

At the smallest level, there is similarity between material science focused on atomic design of silicon, dopants, and oxygen and molecular neuroscience focused on channel proteins, synapses, and neurotransmitters (see figure 2). While this detailed understanding is critical, some of the detail must be "abstracted away" in order to facilitate understanding (and the ability to design) at the next level, or else complexity will grow rapidly and the fundamental principles will be obscured. For example, aggregate properties and statistical descriptions of the materials must be brought forward, but specific locations of individual atoms must be left behind.

At the next level, there is similarity between device engineering focused on converting materials properties into transistors sizes and types so as to achieve the needed electrical properties and dynamics, and cellular neuroscience focused on neuron geometry, channel conductance, and membrane potential so as to understand electrical and neurotransmitter properties and dynamics. Again, the exquisitely interesting and important transistor designs must be abstracted away, passing to the next level only a few simple current-voltage rules. Without such abstraction, or simplification, understanding and designing the next level would be intractable both analytically and computationally. What to include or exclude when abstracting away detail when it comes to cellular neuroscience, or molecular neuroscience before that, is of course an open question, and we do not propose an answer. Instead we highlight the need for this question to be addressed, since for many physical systems, including the CPU considered here, a comprehensive understanding and the ability to design would simply not be possible without abstraction between levels.

At the third level, there is similarity between circuit design and computer architecture focused on the optimal wiring between transistors and between chips, and neuro-anatomy and connectomics (see chapters by Sporns, Zador, and Hawrylycz, this volume) focused on the detailed wiring and wiring rules between neurons within a brain area and between brain regions. Again abstraction is essential in the CPU case as the overall hardware capabilities and limits are of paramount importance when working at the next (software) level, and, similarly, it is anticipated that the overall neural "hardware" capabilities and limits are of primary importance when working at the next (neural "software") level. How best to abstract away detail in the neural context is again an open question, perhaps especially so as the neural hardware changes through time (that is, development, learning, plasticity), unlike most electronic hardware.

At the fourth and final level, there is similarity between computer architecture and computer science—focused on designing machine codes, operating systems, and algorithms that orchestrate all information processing—and on systems and cognitive neuroscience, including network modeling—focused on the relationship between neural activity and sensation, perception, decisions, actions, and more abstract thought. In broad terms, this is the level of the CPU that faces the greatest challenge if the levels of abstraction discipline is not followed. This is because inheriting the full complement of details from the three prior levels would leave one attempting to understand an existing CPU (that is, reverse engineering) or designing a new CPU hopelessly confused in the morass of information; without any prioritization as to the properties that are of direct relevance and those that, while critical to each prior level, are no longer essential to understanding at the final level one cannot see the forest through the trees.

With the levels of abstraction concept in place, it becomes possible to glean new insights into the fundamental operation of a CPU at this final level and, we anticipate, the same will be possible for the brain. As an example, consider what could be learned about a CPU with a few hundred oscilloscopes. With one oscilloscope it is possible to measure the electronic waveform from one transistor terminal, discover that voltages tend to be either high or low (that is, binary), see that voltages change very fast (for example, ns) and do so according to a master clock (for example, 5 GHz), and one could then conjecture that the transistor is part of an adder, memory register, or data bus. Moreover, if it is possible to place the CPU in exactly the same state again and now measure from a different transistor terminal it should be possible to, across many such measurements, build up a more complete picture.

If instead a few hundred oscilloscopes measure a few hundred transistor terminals at the same time then it is possible to discover additional crucial properties of the CPU. This includes how transistor states are coordinated through time (that is, circuit dynamics), how the system functions during normal operation where the same exact set of transistor states may seldom if ever be seen twice, and to postulate the essential features of the software. For example, it is possible to understand the fine-timing coordination principles among a set of transistors responsible for adding two numbers, as well as to understand how faulty coordination between transistors (that is, a timing "glitch" caused by a design "bug") leads to arithmetic mistakes, all without needing to have the same two numbers added repeatedly and all control circuitry in precisely the same state, which may be essentially impossible. This is possible by virtue of simultaneous measurements, dimensionality reduction and dynamical systems analysis methods and modeling, and, again, levels of abstraction—which assures that detailed knowledge of atoms, transistor sizes/types, and wiring that are not essential to proceed with analyses well suited for this final level of investigation do not cloud the investigation or answers. Similarly, we anticipate that measuring from hundreds to thousands of neurons simultaneously and analyzing these data with methods capable of revealing fundamental operating principles (for example, dimensionality reduction, dynamical systems, network modeling) should now be possible and insightful. For example, it may now be possible to understand how populations of neurons in the brain make decisions based on a constant, and seldom if ever repeated, flow of sensory and goal information experienced as part of everyday life.

It is important to note that while, for simplicity, the four broad levels are described from "bottom up" and the importance of levels of abstraction is also emphasized in this unidirectional fashion, this is only half the story (see figure 2). In the CPU analogy is it equally important to apply levels of abstraction starting at the fourth level (for example, what general classes of software/algorithms need to be supported) and proceeding toward the first level (for example, what materials are needed to support a certain type of transistor performance). This also completes the design cycle, as well as moves closer to a comprehensive understanding, by relating the software/system requirements all the way to the materials and transistor choices and tradeoffs. One would expect this to also be the case with neural systems. A better understanding of the key neural computational principles should help deepen understanding of anatomical connection patterns, single neuron computation, molecular underpinnings and their various design trade-offs.

Summary

We are currently in the midst of a neurotechnology revolution that is making it possible to measure (and stimulate) thousands and potentially millions of neurons simultaneously. This unprecedented access to neural data is on the one hand extremely exciting and on the other hand profoundly humbling. What will we do with all of these data? How will we make sense out of it all, and how can we even begin to think about what details matter to each level of understanding and question being posed? While it is tempting to carry on with inherently single-neuron-oriented analyses, or to treat this unique neural dataset as just another "big data" dataset and unleash somewhat generic machine learning algorithms on it, both would likely limit the full extent of insights that are believed to be possible. We discussed here just two of the key challenges moving forward, and we offer two possible approaches—dynamical systems analyses and the levels of abstraction philosophy.

Further Reading

Churchland, M. M., B. M. Yu, M. Sahani, and K. V. Shenoy. 2007. "Techniques for Extracting Single-Trial Activity Patterns from Large-Scale Neural Recordings." *Current Opinion in Neurobiology*, special issue on new technologies (17): 609–18.

Diester, I., M. T. Kaufman, M. Mogri, R. Pashaie, W. Goo, O. Yizhar, C. Ramakrishnan, K. Deisseroth, and K. V. Shenoy. 2011. "An Optogenetic Toolbox Designed for Primates." *Nature Neuroscience* (14): 387–97.

Gilja, V., P. Nuyujukian, C. A. Chestek, J. P. Cunningham, B. M. Yu, J. M. Fan, M. M. Churchland, M. T. Kaufman, J. C. Kao, S. I. Ryu, and K. V. Shenoy. 2012. "A High-Performance Neural Prosthesis Enabled by Control Algorithm Design." *Nature Neuroscience* (15): 1752–57.

Shenoy, K. V., M. Sahani, and M. M. Churchland. 2013. "Cortical Control of Arm Movements: A Dynamical Systems Perspective." *Annual Review of Neuroscience* (36): 337–59.

Yu, B. M., J. P. Cunningham, G. Santhanam, S. I. Ryu, K. V. Shenoy, and M. Sahani. 2009. "Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity." *Journal of Neurophysiology* (102): 614–35.