NARRATOR: The following content is provided under a Creative Commons license. Your support will help MIT OpenCourseWare continue to offer high quality educational resources for free. To make a donation, or view additional materials from hundreds of MIT courses, visit MIT OpenCourseWare at ocw.mit.edu.

LEYLA ISIK: So I'm going to just go over some very basic neuroscience, mostly terminology, just for people who have very little to no neuroscience background. When you hear the rest of the talks, you would think like, what does it mean that they're talking about spiking activity? Or what is fMRI measuring? So that's like the level at which this is.

So my disclaimers are one, like I said, that it's very basic. And two, that it will be CBMM and vision centric, because the goal is to get you ready for the rest of this course. So please don't think that this is an exhaustive, or what I think is an exhaustive, summary of basic neuroscience. So just to give you a brief outline, first we'll talk about the basics of neurons, and their firing. Basic brain anatomy.

How people measure neural activity in the brain, both invasively and non invasively. And then a brief rundown of the visual system. This is a neuron. And it has dendrites and axons, and the signal is propagated along the axon, and the axon terminates on another cell. And when one neuron terminates on another neuron, they form what's called the synapse. So here are some pictures.

Sorry, it's hard to see on the projector, of neurons synapsing on other neurons. And that is how neurons communicate. They send electrical activity down their axon, and it reaches the next cell. And the synapse is both an electrical and chemical phenomenon. We're not going to get into the details of that, but if you're interested, I encourage you to Wikipedia it.

Neurons have a ion gradient across them. So there is a different concentration of certain types of ions inside and outside of the cell. And there are ion channels along the cell. And these ion channels are voltage gated. So what happens is when these ion channels open, the voltage inside the cell changes, and that eventually leads a neuron to fire. And they fire what is known as an action potential.

So it's possible for neurons' voltage to change a little bit, and that is known as potentiation. So they can get either excitatory and inhibitory potentiation. So that means either higher or lower

activity, as shown here. And then once it reaches a certain threshold, they fire what's known as an action potential. So action potentials are all or none firing, and that's what is referred to as neural firing, or neural spiking. It's this actual spike in the voltage.

That is all you need to know, like when people are talking about neural spiking, they're talking about the actual action potential. But oftentimes, we're not measuring things at the level of single spikes. So I'll get into it in a little bit, about what people are actually measuring, and what they're talking about when they're talking about different recording techniques.

So some basic brain anatomy. This is a slice of the cortex, and just to orient you, I'm going to put these online, just so you know the terminologies. But there are different lobes. The occipital lobe is in the back, that's where early visual cortex is. Temporal lobe, parietal lobe, and frontal lobe. And if people are talking about the inferior part of the brain, they mean the bottom, superior top, et cetera.

And this is a rough layout of where different sensory-- can people see that? Kind of, different sensory and motor cortexes, where they land on the cortex. So Nancy is going to give a really nice introduction to the functional specialization of the brain. This is just some basic anatomical terms to familiarize you all.

Right, so neural recordings. So when we're talking about invasive neural recordings, the first type that we'll talk about is electrophysiology. So single and multi-unit recordings. And what that means is that somebody actually sticks an electrode into the brain of an animal and records their neural activity. So this can either be a single unit recording, which means you are recording from a single neuron. And either by sticking the electrode inside or on top of the neuron, or very close to the neuron.

And that means that you're close enough that you're only picking up the changes in electrical activity from that one neuron. But what's more commonly measured now is multi-unit activity. That means that you stick an electrode in the brain, and it's picking up activity from a bunch of neurons around it. So you can either take that data, and get what's known as the local field potentials. So that is the changes in potential, in general, in that whole group of neurons. And people often analyze that data.

Or, you can do some sort of preprocessing to figure out how many neural spikes you're getting. So that's typically trying to look at the neural firing. So from that activity, you can either

get the spiking pattern, or what people refer to as the local field potential. And then you probably heard, you will hear a lot about ECoG data, from Gabriel and others this time. So this is really exciting.

It's the opportunity to record from inside the human brain. From patients who have pharmacologically intractable epilepsy. So sorry this is kind of gross. But when people are having seizures, if surgeons want to resect that area, they first have to map very carefully where the seizures are coming from, and what else is around there, to make sure that they're helping the patient. So to do that, they place a grid of electrodes on the surface of the subject's cortex.

And then leave that there often for a week, for several days, while they do different types of mapping in that area. So this provides the opportunity for scientists like Gabriel to then go and test the neural activity in those humans. Which is a very rare opportunity to be able to record invasively from humans. And again, since we're on the surface of the brain, this is not single unit activity. So you get something that is more similar to the LFP type signal.

And then what I and many other people in the center do is also neuroimaging. So this is noninvasive, often in humans. Although people also do it in animals as well. And the main types you'll probably hear about at this course are MEG and EEG, which are very similar, and functional MRI. So when many neurons fire synchronously, so the neurons in your cortex have the nice property that they're all aligned in the same orientation.

So when they fire at the same time, you actually get a weak electrical current. And that electrical current causes a change in both the electric and magnetic fields around it. And EEG and MEG measure the changes in electric e, and magnetic m, fields from those neural firings. But it's usually on the order of like tens of millions of neurons that need to be firing. So we're now at a much larger scale than we were with the invasive recordings.

And because the neurons all have to be firing at the same time, usually they're not all firing an action potential. Because if you remember, it was just this very brief spike. You're just measuring kind of the changes in the potentiation of that whole group of cortical neurons. So this is a very coarse measure, but it's a direct measure of neural firing. So it has very good temporal resolution. So the question was about, I don't know if everyone heard, the temporal scale of MEG. So it's a millisecond temporal resolution. I think you can maybe even get higher.

fMRI, on the other hand, usually has a temporal resolution of seconds, a couple of seconds.

But this spatial resolution of fMRI is on the order of millimeters, whereas it's more like centimeters in MEG. And actually, so the problem in MEG and EEG is you're recording from-- here's a picture of the MEG, scanner subject sits in, and there's this helmet that goes around their head. And that helmet has 306 sensors.

If it was an EEG, they would be wearing a cap. You've probably seen an EEG cap before, and the electrodes would be directly contacting their scalp. So you're measuring activity from 100 to 300 sensors, and often you're trying to estimate the activity in the cortex underneath. And so that is on the order of like 10,000 sources. And so it's a very ill posed problem, meaning that there is not a unique solution to go from sensors to cortex. And so because of that, we don't actually-- that's why they say that the spatial scale is so poor. But actually it's not a well-defined problem.

So it's hard to even know where the activity is originating from. But that's a very active area of research for how you can constrain that problem with anatomy, and other measurements, to get better resolution. But still, I think people typically think of it as being on the order of centimeters. So the other main type of noninvasive neuroimaging we'll talk about is functional MRI. So here's a picture of an FMRI scanner. Subject's laying there, and often if we're doing a visual task, they look at stimuli on a mirror that reflects from a screen where we're presenting the stimuli.

So fMRI measures the changes in blood flow that happen when neurons fire. And so as a result, this is not a direct measure. So this is not a direct measure of the actual neural firing. So it has a longer latency for the blood flow effects to occur. And so that's why it has the temporal scale that's more like a couple of seconds. But it has quite good spatial resolution. There's structural MRI, which if any of you have ever been injured, you may have had an MRI, and that measures the actual-- it doesn't measure the blood flow, it measures the actual structures underneath.

I mean, often people will do an MRI and a functional MRI, and co register the two, so you have a very precise anatomical image that you can then put the brain activity on. OK. So I got into this a bit. So invasive electrophysiology is the highest resolution data, both spatial and temporally, I think, that most scientists collect. But it has some advantages. One, that it's invasive. So it's hard to test questions in humans.

And just more difficult in general. And two, you're limited by brain coverage. So you can only

stick a grid or an electrode in a couple of brain regions at once. So you really can't get information from across the whole brain, at this resolution with the technologies we currently have. fMRI, on the other hand, has broad coverage and good spatial resolution, but lower temporal resolution. And EEG and MEG have high temporal resolution, broad brain coverage, but low spatial information.

All right. So a bit about visual processing in the brain. So this is a diagram. Can you see the colors? OK. A little, sorry. Of roughly what people think of as visual cortex. So the blue in the back is primary visual cortex, or V1, that's the earliest cortical stage that where visual signals originate. And then there is what's known as the ventral stream, which is often called the what pathway, or where people roughly believe object recognition occurs. And the dorsal stream, which is often known as the where pathway, which is thought to be more implicated in spatial information.

However, this is an extreme oversimplification. I think Tommy put up this wiring diagram the other day. This is still a simplification, but a more realistic box diagram of all the different-- each box represents a different visual region. You can see that there's connections between all of them, between the ventral and dorsal stream. And while we roughly think of it as feedforward, which means that the input from, the output from one layer serves as input to the next, often there's feedback connections. Meaning that information can flow between areas.

So that's why it's been so challenging to probe with physiology. OK, so like I said, there are many layers and they are thought to be roughly organized hierarchically into the first level primary visual cortex. In that area, you have cells that respond to oriented lines and edges. So a cell will-- I'll show an example of this, but fire for stimuli that it sees, that are in a certain orientation, in a certain place.

And that is known as the cell's receptive field. And so it's often thought of as an edge detector. It's very analogous to a lot of edge detection algorithms in computer vision, for example. But then at what's thought to be the top layer of the ventral stream, inferior temporal cortex, cells fire in response to whole objects. And it's not just a specific orientation that they like. They will see this-- they will fire whether they see this object at different positions, and also have some tolerance to viewpoint and scale as well.

So a lot of what we know about the visual system stem from Hubel and Wiesel's seminal work in the 1960s, looking at cells and cat V1. This is the stimulus that they're showing to the cat.

It's an anesthesized cat, and they're recording. So you'll hear a popping, and those pops are the neural activity that they're recording.

[POPPING NOISES]

So they're recording from a single cell right now. So you see, you can hear anytime they present that light bar, in that specific position, the cell fires. And then as soon as they move it out of the bar, the cells stop firing. So that specific cell really likes this bar in this orientation. And they called this a simple cell. We can fast forward a little bit.

They also show, OK. And then they found that there are these other types of cells.

[CAR HONKING]

Sorry. They showed that if you rotate it, doesn't fire at all. And then they show that there were these other types of cells. This is maybe not the movie we want. There are other cells that fire not only to that specific position, but to slight shifts in that position as well. And so it seems like those cells formed an aggregate over the simple cells, and they called those cells complex cells.

And then people did similar things in mostly macaque IT. And so they found that in contrast to simple lines and edges, cells here fired in response to hands. So this is showing the cells' response here. So this is the number of spikes over time. So it fires a lot to hands. And it fires to that hand. This cell likes that hand, no matter what position you show it in. But it doesn't like these kind of other more simple objects, and this one is not selective for faces.

So in IT, there are cells that are selective for very high, you would think of as high level objects. And they're tolerant to changes in those objects. So people have done many more sophisticated studies. This is an example from Gabriel and Jim DiCarlo, and Chou Hung, where they showed neural decoding. So applying a machine learning algorithm to the output of many cells, that these cells were again very specific for certain objects, but invariant to different transformations.

So in particular here, they showed this monkey face at different sizes. And they showed that the cell fired. There was information present in the population of neurons for this specific monkey face, regardless of what size we showed it at. So these cells are often thought to be-- so it's often thought that as you move along the visual hierarchy, cells become more selective.

So meaning, they like more specific objects. And more invariant, so more tolerant to changes in different transformations.

And so the other thing I wanted to talk about was hierarchical feedforward. So computational models of the visual system, because Tommy mentioned this briefly, and I think it will tie into a lot of the computer vision work you'll hear about. So these are inspired by Hubel and Wiesel's findings in visual cortex. So meaning-- and I'm going to talk both about the HMAX model, which is the model developed by Tommy and others in his lab which is a simpler, more biologically faithful model.

But this sort of architecture is also true of deep learning systems that you heard a lot about recently, and that have had a lot of success in computer vision challenges. So if you have an input image, you can then have a set of simple samples. Again, these are inspired by Hubel and Wiesel's findings, so they are oriented lines and edges. So this cell will fire, if you have an edge that's oriented like this, at that part of the image.

And so again, it's just a basic edge detector. And so these perform template matching between their template, which is in this case an oriented bar, and the input image to build up selectivity. And then there are complex cells. And these complex cells pool, or take a local aggregate measure, to build up invariance. And so what that means is if you have, say this red cell here, this complex cell would look at these four simple cells.

So you are now selective to that oriented line, not just at this position, but at all of these positions. And that gives you some tolerance to changes in position. So you'd be able to recognize the same object, whether it had this feature. Whether it was presented at this corner or in a local area. And so the way you do that is you take a max over the response of all those input cells.

And then you can repeat this for many layers and, it's essentially the same thing as a multilayer convolutional neural network. And at the end, in this HMAX model, you take a global max over all scales and positions. So, in theory, you have all these more complex features that you can now respond to, regardless of where in the image and how large they're presented.