Good evening everybody. Hello. First off, some thank yous, the most important one. We, the Santa Fe Institute like to thank the McKinnon Family Foundation who underwrites this entire series that the SFI presents here at the Lensic. It is thanks to them that we get to present these to you, the public, for free. So that is a very great shout out to the McKinnon Foundation Foundation, other crucial support from the Santa Fe Deporter and from the Lensic Performing Arts Center itself. But then also very importantly, very importantly, we would like to thank you, the members of the audience for participating in this series. Again, very, very prosaically, thank you for voting the Santa Fe Institute. This series yet again as the best lecture series in this town of Santa Fe, in the city of Santa Fe. But also more importantly than what you're voting, thank you for your participation. We would like to view these presentations as conversations between us and you. And every conversation involves multiple participants and we would like you for allowing us to engage in these conversations with you on a regular basis. Thank you very much. OK, Now let me spend a little bit of time introducing tonight's speaker, your interlocutor for tonight's conversation. OK, Vijay Balsubramanian. Wow, okay, How do I begin? Let me start talking about his research interests, and I'll be presenting a lot of quotes directly from him to try to get the nuances correct. Even though I'm familiar with a lot of his work, I've collaborate with him, and so on. In Vijay's own words quote, he is interested in how natural systems manipulate and process information. Okay. Well, that level, you know, big deal. Everybody here is interested in how systems process information, right? Well, God is in the details of what one's interests are. And with Vijay, those details are pretty damned impressive. For example, in the way that Vijay construes those particular words, one of the ways that natural systems process information is, well, you know, the fundamental nature of space and time. So, for example, he has done pioneering work in what is known in quantum cosmology as the information paradox. That's the fact that, naively at least, information seems to get lost into a black hole, even though the laws of physics forbid there being any loss of information in the universe. Following this theme, he's investigated, again, quoting from him directly how the familiar smooth structure of space-time you know, that's what we're familiar with, how it can emerge for more complex underlying physical constructs. So, you know, these are the kinds of things that you and I might Mull over while I'm taking a bath or on the way to work this morning. Let's see, I'll have to walk the dog when I get home and buy a cup of milk and, Oh yeah, just what is going on at the plank scale at the event horizon of a spinning black hole that's got a little bit of charge? You're the kind of thing that everybody does on the way to work in the morning. Anyway. No surprise, though, the mind of someone like Vijay cannot be contained in just the fundamental nature of the physical universe. He actually takes a step back from that and asks, well, how is it that we might actually come to our understanding of the physical nature of the universe? He's been interested in how it is that human, that the human enterprise of science comes to try to understand scientific reality, again using his own words. He has also worked on problems in statistical inference and in particular Occam's razor, which is very interesting because all scientific theories involve fitting of models to data, trading it off versus the complexity of that model. Again, you know, just the kind of thing it's maybe this is for the shower rather than the bathroom, and go into work. In short, Vijay is somebody cast in the same mold as the Santa Fe Institute is in general a person of very many interests. Now to wax a little bit more prosaic, though, one difference between the SFI and Vijay was that Vijay was born a little bit earlier than the SFI was. Specifically, he was born in Mumbai, India. Bombay back then moved among many cities in India. His family eventually moved to Jakarta. Then he came to MIT in the US. His degrees at first were physics and computer science. Then he got a PhD at Princeton. After that, among other awards, well, just achievements. He was accepted as a junior fellow with the Harvard Society. Around this time he first formed a connection with the SFI. Since 2000, he has been a professor at University of Pennsylvania. He's a fellow of the American Physical Society has received many awards, over 17,000 citations in the academic literature. So that's a little bit of the more prosaic kinds of, well, he did this and he did this and he did that kind of details. Anyway, now though, the

third one of his major interests, what he'll be talking about tonight, This is one way to phrase it. It's how your brain, the £3. 00 or so of pink mushy Jelly ensconced in your cranium that you know, it kind of wiggles around when you do this too fast. How it is that that the information processing, remember that's his underlying theme here. How it is that the information processing that goes on in that Jelly becomes you, OK. And I have personally been very fortunate to have some interactions. We've had many discussions Vijay and I, on a topic related to that, namely the energetic costs of the information processing in the brain. Anyway, tonight you're really in for a singular treat. And with that, over to you, Vijay.

00:09:31

Well, David, thank you very much. Oh, first of all, can you hear me? Yeah. OK. Thank you very much for that lovely introduction. That was really kind. And it's a delight to be back in Santa Fe. This is one of my favorite cities on earth. I've always really liked coming here. And thank you very much for inviting me back. And thanks so much for coming to my talk today. So great. So let me first of all check that my iPad works. It works great. So the subject of today's talk is. How the brain makes you So we should start by first discussing what I mean by making you so you're an animal like any other animal. So the first thing we got to ask is what do animals do? And you know, here are a number of animals from the OSA Peninsula and Costa Rica. These are photographs taken by my son, who is also one of the animals there. And what do animals do? So the main business of being an animal involves eating, navigating, exploring the world, sometimes resting. Like this one here, manipulating the environment. Look at the nest and this, this web here, interacting with each other, reproducing things of this kind. And that's what animals do almost all of the time. To do so, they have to do certain things, right? They have to be able to sense the world and then they get information from the things around them, the things the other animals, you know, the sky, whatever they need to process this information that's coming to them and remember some parts of it, those parts of it that help them predict what's going to happen in the future. And they after doing this, they need to eventually make decisions, both in the short term. On the long term and be able to plan out sequences of actions for the future. So these are all things that an animal has to do just to be an animal. Now, some animals do additional things. They write poetry. Here's one of my favorite poems by the poet Gerard Manley Hopkins. They paint. This is a painting done by my daughter Aruna. They write music. This is by. Heinrich Injatz. Franz. Juan Bieber. I'm sure I pronounced that incorrectly. He was before the Baroque. And, you know, some people write equations like a guy named Einstein and a guy named Schrodinger wrote this equation. So we theorize about the universe in this way. Not all of this. We being human, seems very impressive to us. And and you know, we're very interested in this. I'm very interested in this. You're all very interested in this. But in fact, the brain devotes very few resources. To such activities, right. It's important to us. But really, mostly that's not what the brain does. Mostly you as an animal, like other animals, mostly you eat, navigate, explore, rest, manipulate environment, interact socially with each other and reproduce. That's what you do. And then there's these additional things that we do on top of it. OK. So I'm going to focus most of the talk today on these sorts of activities that all animals do. Then we'll talk a little bit at the end about these sorts of activities that sort of interest us, particularly as humans. Now, to do all of this, what do we use? The organ. We use the brain. And I'm going to think about the brain today as a sort of computing engine that produces the behavior of animals. So here's the human brain in particular. And what does it do? Well, it does the things that I said animals need to do, which is they it gathers information from the world. It learns from experience, forms memories and makes decisions for future actions. So that's the job of this organ. You know, the liver processes chemicals for you and this organ, and the heart pumps blood. This organ does this stuff with processing information, the kind of computer that you have in your head. Now, the challenge for this computer is that it's a very big, complex and often unpredictable world. And you know, you have very limited resources. It's this thing that you carry around study head. It sloshes, it bumps. You know, stuff like this, That's what you get to use. So myself, as a scientist in this area,

the question that animates me is I'm interested in what are the organizational principles that allow the brain to meet such enormous challenges that animal life faces. So that's what I'm interested in. Now I'm trained as a physicist and a physicist posed to the question about how does this thing, this system work. The instinct of most physicists is that, well you know what I'm going to do is I'm going to 1st workout what? What stuff this thing is made of? What are the parts that make it up? It's kind of an atomic instinct. You know, figure out the parts that make up a system. Then you figure out how all the parts talk to each other, how they interact. And then the idea is if you know that well enough, you can workout what the whole system is doing. That's the kind of method that physicists get trained in. And, you know, it goes back a long way, hundreds and hundreds of years. And that's the tendency. That's the way that's the approach to the scientific question. So what's the brain made of? Well, that brings us to the neuron doctrine. This is Santiago Ramoni Kahal, who lived from 1852 to 1934, and he is widely considered probably the founding figure of neuroscience. What did he do? He was actually. He first wanted to be an artist, by the way, but then he wound up becoming an anatomist. And famously what he did is he founded the neuron doctrine. The neuron doctrine is the idea that the, if you like, the atomic constituents of the brain, the thing that the brain is really made of, the stuff that does the

00:15:08

interesting things in the brain are the neurons of the brain. So here's a picture of a neuron. This is a so-called perkinji cell in your cerebellum. That's in the back over here. This is drawn by Romani Kahal. And you know, he was a trained artist, so he drew really beautiful pictures. So his pictures to this day are wonderful artworks. In their own right now he got the Nobel Prize in 1906, along with another scientist named Camilo Golgi for their work on anatomy. And Golgi did not propound the neuron. Doctor, the reason is they used different experimental methods, different stains to stain cells. And the nature of Golgi's stain was he had the impression that all the neurons were kind of connected together and they formed a web work and there was sort of one thing that made-up the brain. And it was a very careful anatomy done by Ramoni Kahal, which identified that no, no, no. There are many discrete separate objects in the brain called neurons. And the neurons talk to each other in some form or fashion in order to produce all the effects that we call animal behavior. So let's talk about neurons some more. So you'll see that I'm not a trained artist, so you'll see my drawings do not compare to this in either precision or beauty. But let me explain. What is a neuron, right. A neuron is a cell like any other cell in your body. So it has a cell, body. Let me see. Is that big enough? Can can you see that? Can you see the words? OK, great. So, well, see, the nice thing about iPads is you can sort of increase it. See, Is that nice? OK, so if you can't see it, tell me and I'll just increase it. So a neuron is a cell like any other cell in your body, and has a little membrane that contains within it all the molecular machines. That make a cell work. But neurons are particular cells that are intended to help your brain compute, and as such they have very specializations that allow them to do that. So in particular, you know a computing element like a transistor or a diode or a resistor needs to take in inputs from other stuff. So indeed, if you look at the neuron, it has this set of branches coming out from one end, which is called the dendritic Arbor. And on the dendritic Arbor, there are the dendrites, Those are the little fibers at the end. And on the dendrites, there are the famous synapses that most people have heard of. The synapses are little junctions where one neuron makes a connection with another neuron that's a chemical connection. One neuron would dump something called a neurotransmitter, then sort of diffuses over to the other neuron is eaten, taken up by the synapse. And then that leads to communication that we'll talk about later. Now the output. You know, if you're a little computing device, you need to have an output, and the output of the neuron goes through a wire called the Axon, and the Axon then makes an axonal Arbor, which also has some synapses at the tip. That's where the neuron communicates the next neuron whatever message it's going to send. So your brain, the human brain, contains about 100 billion neurons and it contains about 100 to 1000 trillion synapses. So. There's not just one kind of neuron. You know, if you go look inside a computer or silicon computer, there are few kinds of circuit elements, right?

There are transistors, There are diodes, There are capacitors. There are, you know, resistors. There's a repertoire that that engineers use to build up computers in general, right? So in fact, in the brain there are very many types of neurons, and Kahal identified them by how they looked. And for example, here is what's called a cortical pyramidal cell because it looks a little bit like a pyramid. And these pyramidal cells are the workhorses of your brain. All over the cortex you find them layered upon layer of cortical pyramidal cells. This is this cerebellar Perkinji cell that we mentioned. Here's this huge thing. It has a huge dendritic Arbor which it uses like a giant fan to catch inputs. That are passing perpendicular to it. So it does that as part of controlling your body and allowing you to move your body in ways that you want. And then in addition, here's another cell, kind of cell called a cortical stellate cell. So in fact, as we'll see, there's a great diversity of forms of neurons and usually in biology, there's a rule, a working rule. Which says that form follows function. So what's going to happen is that they look different like this versus this. They're going to do different things. They'll have different biochemistry. They'll they'll perform different functions within this computer. So it's like you've got many circuit elements that are stuck together to produce your behavior. Now, the key thing for us is that neurons, literally like the components of a computer, are electrical circuit devices. They're electrical devices. So how did this work? So once again, here's a picture of a neuron. And if you go measure its voltage, literally, I mean its voltage, right, you'll find that its voltage. Inside a brain, the voltage will be something like -70 millivolts. Typical neurons, they're like at rest, they're like at -70 millivolts. Then if you have input coming into it, and the input could be, you know, signals, neurotransmitters, which I mentioned earlier, sent out by the previous neuron. Or if this is a neuron in the eye, it might be light landing on it because there's a little sensor that leads that senses light. Or if it's in the nose, if it's an olfactory receptor neuron, it's a little molecule floating in the air which binds to a receptor and that produces a signal. You know, whichever you do, if a signal comes in, here's what happens.

00:20:56

The voltage of the neuron begins to increase, and when it reaches some critical value, there's a sort of runaway feedback process. It's a little bit like feeding the output of your speaker back into the microphone. And so it goes around and around, so it goes boom, and so it explodes rapidly to some very high potential. And then there's another feedback process and negative feedback process, as they call it, which causes this to dip back down and flip back to rest. OK, Vijay Balsubramanian. Wow, okay, How do I begin? find that it goes like this and then boom, there's a very sharp pulse of electricity. So if you sort of look above this level, it's like this silence and then boom, silence, boom. So it's essentially a digital signal. It's like looking at the digital computer where you have this device going every once in a while sending out ones hold up the spikes or action potentials or zeros where it's silent. So it's very much like a digital computer. That's what you have in your head. So just to drive home that this is absolutely a mechanical, physical process beautifully studied in biophysics. We can even delve in in the exact, you know, the instinct in many of many scientists is you see a phenomenon like this. You say well, well, how does that exactly happen, you know, and you can work that out in complete detail. So for example, if you look at the membrane surrounding the cell, that's the membrane right there. To find embedded in this membrane, various molecular machines such as ion channels which allow specific ions like sodium and potassium mentioned here, or chloride or calcium, you know, they let certain ions pass back, back and forth. And then you have these things called pumps, which have the job of when the ions go back and forth changing in this way, the voltage of the neuron, something has to be done to restore it, right? You send a lot of current back and forth and the voltage changes you got to restore everything. And these pumps basically burn energy. They burn the, you know, the, the gasoline of the cell, which is a molecule called adenozine triphosphate, ATP. And it burns ATP and goes chunk, chunk, chunk, chunk, chunk for a while and then sort of restores the levels of ions inside and outside. So at rest there's a certain proportion of sodium and potassium inside and a certain proportion of sodium and potassium outside. And then you have the kinds of processes physicists talk about stuff. You know,

it's like if you have more salt outside a membrane than inside, you know how the salt comes in, right? In the same way if there's more potassium sodium outside than inside, the sodium wants to come in and the potassium because there's more inside than outside wants to go out. So that these processes like this basically stuff you can understand perfectly well and then these pumps are working against it. And then what happens is that you have external inputs but change how open or closed these channels are to allow the ions to pass before through them. And that's it. It's a completely, I mean just like a mechanical device with stuff flowing back and forth. And you can work this out in detail and show that if you have a system like this, it'll produce these voltage spikes. And then if you make a voltage spike over there, it'll then propagate down the axle and it's literally making you give it the correct inputs. It's going to make a digital pulse and the pulse is going to move down. In fact, this is so well understood that a Nobel Prize went to it. So Hodgkin and Huxley in 1963 worked all of this out, and here's a set of equations describing their model of the cell. You don't have to read the equations. I just want to show you you can write it on a damn page, right? This means you can put it on a computer and you can certainly simulate it. You can totally describe how a single cell describes in complete and utter detail. It's actually a little bit more complicated than this because, you know, for the purposes of this talk, I'm talking about sodium and potassium and things like that. But there's more things, you know, this chloride. And there's calcium. And actually, there's more than one kind of ion channel for potassium. And some details, details, but basically, you know exactly how to do this. So in principle, you could just build yourself a brain. If you knew enough about all the other neurons and all the types of neurons and all these problems, you could just build yourself a brain on the computer. That's the promise of this kind of result that won this Nobel Prize, right. But what's the problem, right? The problem is, oh, wait, actually, before, before I tell you the problem, let me just sort of say a word about this. So this idea that, you know, the brain really just functions in electricity, just an electrical device, a computer of some kind, goes back a long, long way. And the discovery that somehow there's an electrical basis to animal behavior goes back all the way. I don't know why I said this is a 19th century idea. It's even an 18th century idea. It goes back all the way to the 18th century with the discovery by Luigi and Louisa Galvani. We're in Bologna that, you know, you can make frogs, legs for example, twitch for putting electricity into them. So at the time, you know, Luigi got most of the credit because Louisa, being a woman, couldn't be a professor in the university and couldn't also take credit for the discoveries. But well, you know, times have improved a bit. They're perhaps not completely, anyway, So they were the ones who discovered this notion of animal electricity. Then immediately after that, this is the time when Benjamin Franklin, you know, is discovering positive and negative charges flying his sky and all this kind of stuff that's all happening at the same time. And then Volta, also in Italy, invented the battery, which then allowed controlled injection of currents into living tissues. And then you could see how they moved. And people had this idea that somehow electricity itself was the vital spark of life. It's not that it made neurons work because they didn't know about neurons really, right. But they thought it was life itself. So, for example, so they thought, no, maybe we can reanimate the dead. So Aldini reanimated a criminal using who was executed using electricity. And then it was done again in 1818 with another criminal. And in those days, there were the laws that allowed you to experiment with the dead bodies of, you know, of murderers and things like that.

00:27:22
Anyhow, this led, for example, Mary Shelley to write Frankenstein because it led to the idea that life and living things and animals could be just built, that they're they're sort of machines that could be built. And that's what's partly animating us. Still, this notion of artificial life and artificial intelligence goes back, excuse me, goes back to these discoveries, you know, hundreds of years ago. So now, of course, we know that electricity, shocking a dead body, doesn't make it alive again. What it really does is following what I told you a moment ago. It causes the nerve cells to fire, and when they fire, they cause muscles to move. And then you know the body will move in the way you

know. It looks as though you've reanimated the dead, but you haven't, really. Anyway, so suppose we know how a single neuron works in all of its detail. Then what's the problem with understanding the rest of the animal behavior by just putting enough neurons together? Well, the problem is, well, you know, single neuron can't produce a poetry, a single neuron doesn't feel love. You know, things like this, right? Really, neurons produce the brain through the circuits that they work in. It's really the interaction, the collective of all the neurons that makes that makes the brain. So what do these circuits look like? So perhaps the best studied circuit in the entire brain of any animal is the retina. So this this is a drawing done again by the Moni Kahal back in 1917. And what it's drawing is a picture of the retina like this over here. These are the famous photoreceptors that many of you will have heard of. Photoreceptors are neurons which take in light and convert the light into an electrical signal that comes out the other end. So that's the first layer of the retina. Now, the second layer of the retina consists of cells called bipolar cells. These are like analog computing devices. Namely, they're not digital. They have voltage levels that go up and down, and they're all continuous. It is ones and zeros, if you like, but they do all kinds of calculations and computations which are studied. Then these things form. The second layer feed into a third layer and the third layer are cells called ganglion cells. And these ganglion cells are the output cells of the retina, right. And these output cells, you know, the axons come out and the bundle of axons is a thing you call your optic nerve, right? So very often people have a tendency to think about the retina like you know, like a, like a, like a camera. It just takes the pixels and writes down the amount of light in each location. But nothing could be further from the truth. This is your primordial example of what would be called a three layer neural network. It's a piece of your central brain that got put on a stock and then sent out to the front of your head during the during development, during embryonic development, so that you could see in the direction in which you're going. It's really a piece of your central brain. And indeed in keeping with that, you know central brain is a very complicated place. This picture drawn by Kahal is actually a cartoon. We now know that more than sixty kinds of cells, different types of cells, just so different computational elements in the retina alone. So for example, if these are the photoreceptors and then these are the bipolar cells describe the second layer and then these are the ganglion cells described in the third layer. In fact in between there is an entire zoo of other so-called interneurons that sort of communicate laterally and what they do is the following. What they do is, you know, most of the light coming to your eye is not actually useful to your behavior. There's only certain things that are useful. And indeed the ganglion cells, what they do is they report the things that are useful. Like there are ganglion cells called on cells that respond and tell you when there's a bright spot in the world at some location. There are off cells that tell you about dark spots. There are so-called local edge detectors that detect local edges that are direction selective cells that will respond when a little BLOB moves left to right or right to left or up or to down. You know, there are all of these cells, these different kinds. So how do you extract the so-called visual features from the scene? Well, this huge diverse body of neurons is charged with removing from the visual input everything that's irrelevant and leaving what is relevant So famously, for example, you know, in analogy, someone once asked Michelangelo, how did you sculpt the sculpture of David and he said, well, I well actually David was already in the marble. I just took out everything else, right. And so that's what these things are doing. They're taking out everything else and leaving the important stuff to go through to the brain. So this is how it's organized. So you can see it's a large collective effort of many, many types of neurons arranged in a very specific circuit. For example, you will always find that the slow bipolar cell here connects to the local Ed cell and the fast bipolar cell here connects to the brisk transient cell. So it's like a circuit that Intel would design, you know, the engineers designed something. This isn't designed, it's selected out the evolution. But anyway, there you go, it's like a circuit that you'll find in every vertebrate eye. Now even that is not actually a sufficient picture of the degree of cooperativity in the brain that produces all of your behavior. So if you pop out from the scale of neurons and instead look at the whole brain, here is your brain, you know, looking this away, and in the back of your head, there are circuits that are associated with vision. Over here, there's a strip going down the side of your head that helps you control your body.

Over here, there's an area called Broca's area, which is associated with the production of speech. And you know many other areas of this kind, right? The hearing, etc. And in the frontier, there's regions of your brain that are associated with planning and personality and things of this kind, right? So This is why, by the way, you know, decades and decades ago there was a treatment for schizophrenia which involved basically sticking an icepick up your nose and scrambling this area to because there were there's some issue with it. And the result is, you know, the person won't be schizophrenic, but they won't be themselves anymore because you scramble the circuits that make you really should be very careful with any kind of treatment that involves sort of messing with bits of the brain. OK. So anyway, so there are all of these things, and we

00:34:04

know that function is localized in this way, that this is called localization of function for many reasons. So one reason is, for example, if you bash the back of your head, you fall backward. You know how in the cartoons you see stars, right? You know Bugs Bunny, You see stars when they bash the back of when he bash the back of his head. It's because Bugs Bunny has a bruise on this part of the brain, so the circuits aren't working quite right, right? That's what's going on South. When or when someone has a stroke, Sometimes you'll see one side of the face sag, but everything else is completely fine. That's because there is an issue. There's a problem, let's say, in this region of the brain. This is the motor cortex. And within the motor cortex there are circuits here that control the knee and circuits here that control the face. So if you have a stroke there, this thing sags, but the rest is fine. But it's all collective, right? There's a collective circuit, which of course has little areas that concentrate in different things, but there's a collective effect. And this is not just on the surface of the brain, by the way. So deep in the middle of the brain, there's an area there called the hippocampal formation, which is associated with, for example, memory, right. So if you have damage there, you will not be able to form long term memories, right? And it's also associated with spatial cognition, with spatial, you know, animals need to map space and be able to make plans and navigate. And so all of those circuits are hidden here. OK, so actually human beings have known about this kind of thing that there's function, localized different parts of the brain for a very long time. This is a picture of what's now called the Smith Papyrus, after Smith, who bought the papyrus in late nineteen 1800s. This was a papyrus written in 3000 BC. And the author of this papyrus describes various kinds of traumatic brain injuries and then explains that you know, something will stop working, but everything else works. So this author knew that function was localized in various spaces of the brain. The modern version of this idea is due to Paul Broca, a French Doctor Who basically discovered Broca's area because he had patients who had impairments in their speech. And he discovered they had lesions in this area of his brain, of their of their brains. And there's another area called Wernicke's area which is associated but later discovered having to do with speech understanding. So if you go to Paris, you can go to the 13th and you know, admire Rue Broca, the name for him, for all of his great discoveries. There you go, there's the picture of the street sign of Rue Broca. All you need to do to get a street sign is discover localization of function in your brain. OK, very good. So now that's also not enough in describing how complicated this processing machine is. You know, you all know if you ever looked at a wiring diagram of something like a chip that powers your computer, you know, it's got lots and lots of pieces with intricate circuits, and all those pieces connect to every other piece. Otherwise, of course, they couldn't produce their collective effects. And indeed, all of these brain areas connect. So today in the 21st century, there are all of these extraordinarily beautiful techniques. Now for tracing out the wires. It used to be the people at Cajal. You know, they have to take the brain of an animal, slowly slice it up very, very carefully, trace everything, draw it out by hand. You know, can you, Can you imagine the level of dedication and sheer labor it it took? Now we have all these amazing techniques. You can do things like you can make neurons grow, glow in different colors and get the color to propagate along the Axon. You know the output wire so you can see where it's pointing to. You can inject viruses into bits of brain. The virus will go backward along where you know the the the the the the neuron connects

to. And then you can find out which neuron you one neuron down, you can tag it in that way and then it'll stop working. The virus stops working at the point you can engineer all these things. So in this way we now have far more detailed maps of the brain. Here's a picture from a paper by these authors. It was a review paper showing some of the nerve tracts, as they're called, the wires that pass between and now. Because of that, we just have a much more refined understanding of how collectively all these bits of brain produce functions that we care about. South, for example, following what I said a moment ago in Broca's era or shortly after thereafter, we might have said, well, you know, here's brokers area which you know somehow is responsible for articulation, production of speech. Here's Wernicke's area that is somehow associated with comprehension of speech. Let me explain the difference between those two If you damage Wernicke's area, you'll produce speech. It just won't make any sense. And if you damage brokers area, you can understand speech perfectly well. People can talk to you but you won't be able to make it. You can't control, you know your muscles to make the speech, you know stuff like this. There's also another area, Gershwin's area for concepts, you damage that your problems with concept anyway. So there are these areas that people identified. So a more refined model you might derive by looking at the way the nerve tracts, you know, go is that there's clearly a pathway from here to here, so from the comprehension area to the production area. And there's a secondary path. You know, there are two pathways like that. But now we have even more fancy techniques. You can put a person in the fMRI machine, right, functional magnetic resonance imaging machine, and have them do stuff. And while they're doing stuff you can record, you can find a way of measuring which areas of the brain are active. So which areas of the brain are active together when you cell, that's the membrane right there. To find embedded in this membrane, various molecular machines such as But now you get a much more refined picture that all of these areas are there. It's true that they're associated with articulation, concepts, and comprehension, but depending upon exactly what you're trying to do, they connect in different ways. So there's a sort of flexible computational engine wherein wherein you

00:40:18
know the different bits of it work together for one task or not, depending upon what's necessary to do the task. Is this amazingly flexible engine inside your head for producing all of animal behavior. So that's the task, to work out how all these parts connect and how they reconnect and reorganize themselves when they need to do things. OK, so myself, my interest is to understand how all of that happens. I think of the brain as a computing device. I want to know how this computer, this computer works. So there are many ways of approaching this, but one way of doing that is, well, so the way I do that is I take a view of the circuit repertoire of the architecture of this computer as a kind of memory of the computations that have predictive value for your behavior. Because, you know, animal life is about predicting what's going to happen and taking action appropriately. And that these computations have been learned over evolutionary time and then encoded in your genome and in the developmental program, right. So that's the way I tend to think of it. You've learned these are computations that are useful to you and then what you use something and then you use them. So the question that animates my work is what are the organizational principles or laws, if you like, that control the collective computational computation and information processing of the brain? So, OK And you know, if you're interested, there's a popular book by Peter Sterling and Simon Loughlin. This book here, Principles of Rural Design, that you'll probably enjoy, which lays out many possible principles that may be operative in the organization of the of the circuits in the brain Now. So what principles might be relevant? What kinds of laws or principles? So there are several that you could try to name. For today's purposes, I'm going to name two, and then I'm going to illustrate them right in in in the operation of some of these circuits. So, for example, one principle involves the costs of computation. So this is your brain, and this is your laptop. And you know your brain is only 2% of your body weight, but it's actually 20% of your metabolic load. What that means is it's 10 times more expensive than muscle. This is a seriously expensive thing to own,

right? And you know, it's a it's also packed solid, right? Every millimeter cubed contains 4 kilometers of wire. This is like, this really dense thing, right? So on the one hand, this sounds like, wow, you know, that's a really expensive thing. On the other hand, you know your, your, your, your brain consumes something like 12 to 20 watts of power. Your laptop consumes 80 watts of power, right? And it sure can multiply fast, but you know, it can't give this talk. Well, at least not yet so. But even things like, by the way, you know, the AI seems to have the promise of, you know, getting making machines. That'll do me. But actually they have to train their machines on so much data, you know, planetary A planet's worth of data. They use a city's worth of power to train these things, right? Whereas I can learn stuff with one example, I don't need the city's worth of, you know, the planet's worth of data. So there's a real difference in the way the brain works. It's just much more efficient, right, in the way it operates than your typical silicon machine. That's also one reason why computer scientists are interested in this kind of thing, because they'd like to make more efficient machines. So you could ask yourself, how does the brain achieve this sort of efficiency relative to the engineered systems that we've been able to produce so far, Right. I've described this baroque architecture with lots and lots of pieces that interact in some way. So one idea that's very powerful and has been used powerfully in neuroscience is that the brain achieves its efficiency by adapting its circuits to the structure of the world and then using learning and self organization to further adapt the changes. Right. You first adapt to the world and then if the world is not quite what you expected, you change the circuit so that it's well adapted to the world. So now why would this make the system more efficient? Well, the idea is roughly like this. So when if you talk to, you know, sociologists or something, they'll tell you that, you know, back in the day we were all, you know, hunter gatherers, we all dug the ground, we hunted the animals, we built the house, you know, whatever we did, everything and society got along. And then as societies progress, the civilizations develop, etc. You know, there's a segregation of function. What happens is that the Shoemaker makes the shoes, the Baker breaks the bread, you know, the Mason builds the houses, and then each unit in the society develops efficiencies. They do their job really well because they're well trained, they're adapted to it, etc, etc, etc. And the whole system, by communicating between all the individuals, works better, works smoother, it's more productive, etc. And it's more efficient. So you could try to apply a similar idea to the organization of these circuits in our head, which have evolved over time. So that's a very important idea in neuroscience. So that's principle 1, the efficient use of resources. So the idea of this principle is that brains exploit the average structure of the world to efficiently allocate the limited resources to the tasks that they have to do to maximize gain for the Organism. Now if I were to draw a cartoon of something like that, it might be something like this. Remember how in the retina we said, well, you know, there's all the light coming in, but at the back end of the eye you have a certain number of different neurons, right? Certain there are twenty types of ganglion cells, we call them, but you know, they're each charged with doing a thing, right? One thing we'll look for bright spots. That's the on cell, the one thing. For dark spots, it's the off cell, you know, and so on and so forth. There's a certain repertoire of things that are apparently needed for vision. So you get those, you make those things, and then each of those things is one of these blobs. And together they cover the space of those aspects of the visual world that you need for your behavior. And they have to allocate those resources in an in an effective way. And you could keep doing this with the sense of smell, with your sense of place, you know, the whole brain. And you could try thinking in this way. So this is a very important principle that's used often by neuroscientists to try to understand the architecture of circuits in the brain. Another important principle is this one, it's learning in self organization, these individuals who might know well did a lot of that.

00:47:02
And So what do they do? So as far as they can tell, their little heads self organize a lot of the architecture within it through dynamics and learning. And what do they do? Well, this learning and organization, learning and self organization adapts the brain to ongoing variations in the world. You

know their world is different from the world I was growing up in. It allows them to learn new tasks and environments. At this age they definitely didn't know how to ride bikes and you can correct inaccuracies in pre-existing dynamics. Anybody's interact with small children knows perfectly well that you know when the baby first tries to reach for stuff, it doesn't quite you know doesn't get there right, you can't quite do it. And then and then they learn how to control their muscles to get there to reach it. Or you watch a kid learning a musical instrument. Well, you know, the first you know, 500 times doesn't sound so good. And then eventually it does because there's something that happens where you learn how to operate the bits of you to do the right thing. So that's that's the learning that has to happen in a variable world. OK. So what I would like to do in the remaining, what, 1520 minutes, 20 minutes or so is that I would like to give you examples of how in the field people use these kinds of ideas to understand the organization of neural circuits. So far. I gave you a description of the architecture right, of these circuits from the neuron level, you know, the atomic constituent of the brain all the way to the whole brain, right? But in science, we want more than that. We want an understanding of why they're organized the way they are, how they produce the dynamics and effects that they do. So there is an enormous amount of interesting work that has happened since the era of Kahal in these directions. So I want to give you a little flavor of some of it by giving you a few examples from my work. I'm going to pick simple examples because you know, there's 20 minutes left, so we should go. So it has to be simple and precise enough. So I'm going to give you examples in four domains, right? One domain is seeing vision, the second domain is smelling olfaction. So these are both sensory domains. Then I want to talk about learning to sing. This is in songbirds. So that's what the professionals would call motor control because you need to control all the muscles in the, you know, the songbirds, you know, apparatus in order to produce the sounds. And then I'm going to talk very, very briefly about navigation, which is actually the subject that I spend probably most of my time on thinking about right now. So I'll just have to say a little bit about that and then we'll end. So I'm going to give you 4 examples of the sky. So let's start with one. So first I'm going to talk about vision and and my goal. By the way, what's my goal with these examples? Remember I talked about these two principles, adaptation to the world and learning as two mechanisms that are used to, well, to organize neural circuits. You can understand why they're organized the way they are, but thinking about adaptation to the world and about learning to reorganize. So I want to give you 2 examples of adaptation. Those will be the sensory examples and an example of learning. OK, so first let's talk about vision. So why do we, why do I choose vision? Well, we're very visual animals and it's a sense we use most of so that naturally, you know, neuroscientists gravitate often to thinking about the visual world. So I want to ask a very basic question. Here's the question. So we drew this. Ramani Kahal drew these pictures with all of these different neurons and then I showed you another picture, more modern one of all the cell types. And there are like sixty types of cells, God help us in this structure, but what do they do? Let me just look at the output layer for now. These were the so-called retinal ganglion cells. And I told you that these cells detect features, visual features, the world, bright spots, dark spots, color, local motion, etc. And then they put the signal, these digital signals that I talked about go down the optic nerve. And that little bundle of wires, that's like an Ethernet cable worth of data. And that data goes to your brain and you do all the stuff that you do with vision. So following a kind of question that immediately comes to mind is you have a certain repertoire of these cells. Well, humans, we have about a million of them. If you're a Guinea pig, you have about 100,000 of them. So yeah, it's not an infinite resource. So if you spend all of your money on detecting bright spots, there's nothing left. If you spent all of your cells on detecting bright spots very well, you will have nothing left for detecting motion, for example, right? So there has to be something interesting in the way this sensory device, right? This is like you have a robot and has this little eye sitting in front looking at stuff, right? So you've got to decide how you can decide the design. That eye, it's only yay big. It's got only so many units in it. How are you going to allocate the different pieces of the circuit? So here we should there should be some. There should have been some choice done by evolution for the allocation of these features. So for this talk, let's take the specific example. It's because it's easy

to imagine of bright and dark spot detectors. So bright spot means you look for a little bright spot relative to the near surrounding region, a dark. So that's called an on cell. And an off cell does the reverse. It looks for a little dark spot relative to the near surrounding region. So we're going to ask the question, well, suppose you have a certain number of spot detectors, how many on and off cells should you have? How many bright and dark spot detectors should you have? And we're going to try to use this to understand the architecture of the retina, just incidentally. It's also important therapeutically, right. So if somebody's, you know when the day comes when we can replace the retina of some, if somebody's retina degenerates and we want to replace it by silicon device, it

00:53:10

would be better if we could understand how many of these and how many of these detectors we should put in into that silicon device, right? So if the idea, if this kind of organization involves adaptation to the world around us, a question we have to ask is what is the world around us? So for each sense or for each mode of interaction with the world adapted to the world, then one question you have to ask is what does the world look like? So in this case you can go collect a large database of images and study its properties. How is light organized in these images? And one thing you find is that in any color channel this is supposed to be red, blue and green. You'll find that the distribution of the intensity of pixels, you know if you look at the brightness of different pixels is very very skewed. There are lots of pixels which are pretty dim, but then there's a very long tail with some pixels that are enormously bright. So it doesn't matter where you go. You do this inside a city, you do this at the North Pole, You do this in a jungle and do it wherever you want. You're going to find that this is the case, right? It's a very universal feature of the visual world, and it has an important consequence. The average intensity of a pixel is lower then the median. Median means the median intensity is the intensity at which half the pixels are dimmer and half are brighter. So the median exceeds the mean in the world. There's another fact which is that there are long range correlations. What do I mean by that? This is some technical figure showing that. But what we really mean is, you know, if it's kind of red here, it's going to be kind of red here, It's going to be kind of red here. And then after a little while, it's improbable that the color is going to persist forever. But there's a very particular structure of those correlations which you can measure. It's a regularity of the natural world that you can measure. So you would think that, well, if I'm going to detect brightness, dark spots, maybe that's related to these statistics. And actually, without even doing anything, we can immediately argue. And let me argue to you now that these two facts that light is correlated over distances and that you have the median light intensity exceeds the average. That immediately means that there are more dark spots in the world. Here's how to see that, right. So suppose what you do is you compare the light in a small region here to a small to a large region there, right? So what's going to happen is that if you measure a A, the if you compare the light in those two regions, because of these two statistical facts, it's going to turn out that the spot in the middle is generally going to be darker than the surrounding region. That's just because the average, if you see the average here, the average is less than this. So that turns out to be the case. So, for example, you can take a bunch of images, and you can put down here a thing that detects a bright spot, a thing that detects a dark spot. And you can just count how many bright and dark spots do you detect in every image. And you'll find that no matter what angular size you look at, there are more dark spots than there are bright spots. OK, so that's what this figure is supposed to be showing. So we're going to ask a question based on this. So let's agree, but because of these statistical facts, there are more dark spots than bright spots in the world. So then we can ask, suppose I give you an allocation that you can buy or you can put in the eye a certain number north of bright and dark spot detectors, namely on and off sells. So you get to divide them. You get to decide how many dark spot detectors and how many bright spot detectors you have. And my challenge to you is workout what would be best given the structure of the world. So it's immediately clear that if you have only money to buy one cell, it's better to buy a dark spot detector. Why? Because there are more dark spots, so it's more likely to respond, so you're more likely to get useful

things out of it, right? If I give you 2 steps, so then you buy some dark spot detectors and you can cover the eye with them, but then they're going to start overlapping. So they're going to start telling you the same thing because they're overlapping. They're telling you redundant information. So then you should buy some on cells, bright spot detectors, and you can continue with this game and workout the mathematics of it about how you should tile the retina with bright and dark spot detectors in order to get the most information about the brightness of spots in the world. So that's a mathematical problem that you can write down and solve. And it turns out that the answer is about you should have 1 1/2 to about two times as many dark spot detectors in your system if you want to say as much as you can, given your budget of the number of cells about brightened dark spots. So we can ask, right, How does that compare with the actual eye? So over the last two decades there have been a very large number of experiments on this. So it used to be in the 1990s, for example, if you ask people, they would talk about, you know, yin and Yang, bright and dark, equal and opposite. But around the year 2000, a series exactly as you predict if the circuits of the eye were adapted to the statistical structure of the world so exactly in that way. So the collective organization that suggests so as I was saying according to the principle I mentioned, that the collective structure of the retina, the object which is charged with processing light and turning it into an electrical signal to the center of the brain, that that collective structure is partly controlled by adaptation of the circuitry to the statistical structure of the visual world. OK, so that's an example of how neuroscientists try to understand how these circuits work. So let me give you a second example. Similarly, in the sensory world, let's talk about smelling, and we care about smelling, about olfaction, because it's in some ways the most primordial sense, right? All living things, sense molecules, right? So all living things smell in that sense that every animal, most animals, you smell much more importantly than vision to do their daily jobs. So smell is actually a sense that's very different from vision. Why is that? Because there's an intimidatingly diverse space of smells. So what is smell? You know, when you smell a thing, what you're actually sensing is volatile molecules, right, Molecules that float around in the air. And there are many, many of these. You know, there are catalogs of molecules that float in the air.

01:00:12

There are at least 10,000 different molecules to make matters worse. You can go consult with the people who do fragrances and flavors. You know, people who make artificial flavors, and they'll tell you that if you look at natural odors of different kinds of foods you know from other animals, etc, typically an odor will contain 50 different kinds of molecular species. So you can estimate from that that in terms of possible natural smell that it's like 10,000 to the 50. This is such a vast number, it's even hard to conceive right. That's the number of different things, like older objects, if you like that you might have the sense in the world. To make matters worse, the older environment changes, reflecting seasons, circumstances and new opportunities, right. So what is, what do you need to do? Well, to detect a molecule, actually, you need to feel its shape, right? The identity of a molecule is basically defined by how all of its atoms are arranged in space. So, so oldfaction, the sense of smell, is a method used by the brain to sense the shapes of little things floating in the air. So what do you need to do this? Well, so it turns out that odors are sensed when molecules bind to receptors in the nose. So there's basically another molecule in a neuron in the nose so that it's an olfactory sensory neuron. The molecule will be sitting there, and then in will come some other molecule, an odorant as it's called, floating in the air. And if this molecule fits into the binding pocket of your sensor and sticks there, doesn't fall off, then the neuron will respond. So this binding pocket, your olfactory sensory neuron, is literally feeling the shape of the molecule, whether or not it fits. Now, as it turns out, every receptor, olfactory receptor that allows you your sense of smell is encoded by a separate gene. That's because you know, to make your receptors have to have different shapes because they're trying to feel out different shapes. So to make a molecule, a protein that has a different shape, you need a different gene to encode it. So the genetic level, you know, they're all different genes. In fact, it's the biggest gene family in most animals. So flies of about a hundred of these

humans will have about 300 and even the elephant only has about 2000. So this leads us to feel, looking at this, that somehow, you know, apparently animals can sense all the molecules in the world and all the combination of the molecules in the world with just a few 100 receptors. And the reason this is very, very strange is you could phrase the difficulty of the problem in the following way. You could say, suppose an odor is defined somehow by the concentrations of all the molecules in it. Then it's possible that there are up to 10,000 different molecules. So you put the first concentration this psi on this axis, you put the 2nd concentration this axis, the third concentration of this axis, etc. There are 10,000 different axes on which you can draw all these concentrations, and somehow you have to represent all of this information in the responses of justice. 500 neurons, let's say, or 300 neurons. And so it's the response of neuron one and the response of neuron two. Response of neuron 3. You just still have. It would feel like there's no way to represent so much information about all the molecular concentrations with so few neural responses. You could, just to make the problem more easy to visualize, consider the following thing. Suppose I told you that you have positions in three-dimensional space, right. XYZ like that there are positions in three-dimensional space and I insist to you that you tell me where you are in 3 dimensional space. But telling me just you know coordinate, if you like a location in one dimension, I mean that's just really not doesn't seem like it's very sensible. It's be very hard to do to do that right? Because you've got, you know, three directions. You could go up, you could go left, you could go right. And somehow encoding that in just One Direction seems very hard. OK. And what's more, you'd like to do that in a way that somehow preserves similarities and differences. You need to be able to tell that this odor is the same as this odor or similar to, and these two odors are very different. So somehow the brain has to solve the problem of representing all of this information with very few neurons, right? So very high dimensional space of odors, as they say, in a low dimensional space of responses. That's like going from three dimensions to 1 dimension, and needs to do it in a way that preserves distances. So an insight that's happened in neuroscience in the recent past is that actually there's a way in which you can do this by adapting to structure in the space of orders. The structure is the following. We said already that if I look at, let's say, this odor, it'll contain only a few molecules, let's say 50, and all the other things are absent. This odor contains this molecule and this molecule and all the other ones are absent. Each odor in the natural world contains about 50 odorant molecules, right? Fifty types of odorant molecules. So it turns out that there's a theorem in mathematics that you could take signals like this if they came to you and store them with many fewer responses than the number of things that are coming in. If every sensor bound to all the molecules that are coming in kind of randomly, it's rather than being structured like, you know, in vision, you have the cell that pulls out a bright spot and you have a cell that pulls out the dark spot or left to right motion and things like this. So instead of doing that, if you just said here's an olfactory receptor, I kind of wanted to bind kind of randomly to everything. And Justice give me a sort of mishmash of numbers. There's a theorem in math that says if you do that, you could store all the information that's there in the older world with a few 100 receptors. So there's a theorem like that. So the question is, does the olfactory system use randomness in that way? That's a very different kind of organization, mind you, than the visual system. That's why I'm giving this example, right? So the visual system, everything is highly structured, right? Because the visual world is highly structured. And here the idea would be the math would tell you that it should be kind of random what it does, and indeed so Here's a picture of different receptors in the fruit fly binding to many different odorant molecules, and the darkness of the color tells you the strength of the response.

01:07:09
Now, this doesn't look entirely random, and it isn't entirely random, but there's a sense, a mathematical sense, in which this is sufficiently random. It's pretty close. Random, so much so that the represent. So this is what so this is. So. In other words, basically what happens is that every olfactory receptor here is binding to many, many odorant molecules, that every odored molecule is

binding to many, many receptors. So this is highly multiplexed thing. It's not like 1 receptor tells you this molecule is there or this receptor tells you this molecule is there. It's like every receptor binds to everything at some level and every molecule binds to everything. So this is highly multiplexed kind of description of the odor world that appears in the nose. So this is what's called a kind of combinatorial code in the, in the, in the, in the jargon of the field. And this continues. So this is a drawing, a diagram of your early olfactory system. So you have all the molecules in the world. This is a vector. This is a a figure showing concentrations of which molecules are present. They bind to different receptors, You know, they're all colored differently. And then what happens is that the signals from the receptors are collected at a second stage. That's what's drawn here in the picture by Camila Golgi. So the schematic that I've drawn here is actually this thing that I've, you know, simplified. And at this stage it's all cleaned up. And then it turns out that the signals from here project to the your central brain, to the cortex also in a random way. So once again, you get this sort of randomness and the information is kind of spread out all over the place and you can and you can show. And so that's what I was trying to argue to you is that. So you might be surprised if you, if you study vision, you might be surprised by this, right? Why is vision so structurally organized with all these little feature detectors pulling out different things in the world and sending the brain and so on? And why this olfaction looks so random? So what I tried to argue to you is that the collective organization here in the olfactory system, the system of smell, is via randomness as opposed to structure, because that's what you need mathematically to be able to process signals of this kind. So in both cases, in both vision and olfaction in one case, the structure in the other case, the sort of randomness in the circuit are associated with adapting the circuit to process efficiently the information in the world. So that that that's what I'm trying to convey here. So and then by the way, I'm trying to give you some of the technical details here because often I find that in these kinds of lecture, when I listen to public lectures, you know, there's lots of pretty words, but I don't have a sense that something can actually be done right. There's something concrete I'd like you to go away with the sense this is completely concrete. You know, you can do these calculations, you can do these measurements. So this is a completely concrete thing and and that's in the details of of this kind of analysis. That's why I'm giving you some of these details. You have a sense that of how this works. So I'm going to give you one more example and then stop. So the example I'm going to talk about is now learning, right? So I'm going to talk about, So we've so far talked about adaptation and evolutionary selection of circuits to be adapted to the world. And now I want to talk about something that isn't evolutionary selection for adaptation of the world but rather involves learning. So there are songbirds. Here are two kinds of songbirds, zebra finches and the brown headed cowbird. And famously, songbirds learn to sing by imitation, practice and innovation. By the way, the reason why we should care is humans do the same thing. Here's a human learning by practice, right? So that is a very important part of animal activity, to learn by practice and to get good at something. So how do songbirds learn? Well, it turns out the juvenile male bird babbles, and this goes on like that. And then over weeks it learns to sing clearly, partly through imitation, because its song will partly imitate the songs it hears often of its parent, with some innovation. And then the adult male bird sings to female birds, often seeking a mate. And then the female birds seem to evaluate these songs and then decide which one or which ones accept. And this is all really important, right? It matters because, for example, in social species like this cowbird, the males and females work out a kind of social hierarchy based on their singing to each other, and then they produce pair bonds. And the success of those pair bonds is absolutely critical for later success. In egg laying, you don't make the pair bonds, the colony doesn't produce very many eggs. So it's actually really important for the survival of the species that this is all well organized and that the males learn to sing. So how do they do this? So now let me tell you about how all of these circuits have talked, have talked about reorganize themselves. They reorganize themselves when neurons which are connected to each other autonomously rewire their circuits, right. They just just there's a rule, they'll use the rule to reorganize. There are various sorts of rules and one of the rules, one of these mechanisms is called spike timing dependent plasticity. So let me give you an example.

Suppose you have two neurons, one and two, and they're connected, and here's the synapse between them. Suppose for some reason they both fire, right? They both produce these voltage spikes that I talked about at the beginning of the talk. Suppose neuron 1 fires before neuron 2. You could have a rule that says the synapse strengthens, or you could have a rule that the 1st and you also the second part of the rule says that neuron 1 fires after the second one, the synapse weakens. So this is just a local rule, right? I mean, it's just a dynamical rule. You've got two dumb objects, neuron one and neuron 2, talking to each other, and they have a rule, right, that they strengthen the synapse or weaken the synapse depending on who goes first. You could modify this a little bit.

01:13:29

There are things called neuromodulators, which are sort of chemicals that the brain uses, and these can be used as global knobs to affect how this happens. So for example, there's a thing called dopamine, which if dopamine is released here, it can reinforce changes that have happened that happen to have led offline to some sort of reward. So there's a reward, You know, they'll do something to help cement some change that's happened. The synapse. There's other things that can norepinephrine that also, you know, allow emotion to control how the synapses changes and everything. And you can put all of these rules in computers and see what they do and so on. It just seems really implausible at first sight that something so stupid, if you like, can actually lead to things like actual learning of things you might care about. In the case of the birds, for example, allowing them to learn songs. But let's see, so how does song learning actually work in the actual brains of birds? So here's the bird brain. And what's written down here is all the different major brain areas. Remember we talked about brain areas. These are the brain areas that sort of cooperate collectively to help the bird to learn songs. Three important ones are HVC, R&L, Mann acronyms with some historic provenance. What's very interesting about the system is that this region RA has neurons in it, which are the things that control the muscles. So these are the things that actually produce the song. So sequence of firing of the neurons in the region RA will make a whole bunch of muscles go back and layer are cells called ganglion cells. And these ganglion cells are the output cells of the retina, right. And these output cells, you know, the axons come out and the bundle of axons is a thing you call your optic nerve, right? So very often people have a tendency to think about the retina like you know, like a, like a, like a camera. It just takes the pixels and writes down the amount of So that's the thing that makes song. So to make a song, of course you need to control those muscles in a particular order. Control these muscles first, then these muscles and these muscles, these muscles, and that makes the song. So that's what you've got to learn. So this area of so RA also gets input from another area called HVC and HVC acts as a conductor. So the neurons in this area produce patterns like this and this and this and this and this if you like. But it's actually patterns of neuron firing and they provide a time base. Every pattern of neuro of firing of neurons in this area, HVC marks out a different time in this song. It's literally conducting the song. So when HVC produces this pattern, certain neurons here in the area RA should fire, thereby pushing certain muscles. When neurons in HVC fire in some other pattern, then some other neurons in RA should fire, producing another push, a different kind of push in the muscles. So that's what should be happening. So that's what it's got to learn. As the conductor marks at time, different muscles get pushed. So how does it learn this? Well, it turns out there's a third area called L Man, which is a tutor area. And this tutor area does two things. First, it drives exploration. You see, part of the problem is that the bird doesn't know which neurons control which muscles. You don't know that, just like Baby doesn't know how to move its neurons in its head in order to control their muscles. It's something you learn. So here L Man does two things. One is it drives exploration by injecting a little bit of randomness in the firing of RA that allows the neurons there to sort of just do slightly different things every time to explore what can be done. The other thing it does is provides guidance. So this area gets back a message that says was the how did the output compare to a remembered song? And based upon those differences between what you wanted and what you got, it's able to provide some guidance about which synapses to change. Or well, that was a good thing

you did. That was a bad thing you did so able to provide guidance of this kind. So that's at the level of these brain areas collectively talking to each other. What about at the level of the individual neurons and the synapses? Because somehow here at the end, every neuron is a dumb little thing, and all it knows is a dumb little rule. But how to change? So it turns out that right here there are synapses between the conductor neurons and the student neurons which have the following property. So if the conductor sends a message to the student before the tutor, the synapse strengthens. If the conductor sends a message to the student at the same time as the tutor, then then the synapse weakens. But that's a measurement in these in this system and otherwise basically nothing happens. And this is for this particular synapse and this particular bird. This is the zebra Finch. Other birds and other synapses can have different rules, so it would be a legitimate question to ask whether can such a stupid rule actually allow a bird to learn a song. So you can check that. And the way neuroscientists, or at least theoretical neuroscientists will do this is they'll build a computer model of the system. So you make a model of the conductor area as a bunch of neurons. You allow the conductor neurons to send signals to all the student neurons. Then you say there's a tutor and you get the tutor to send messages to the student neurons too. And then you say, I want the student to learn a particular pattern of firing and you press go. And you say, can't such a stupid rule, this spike timing dependent plasticity rule, where depending upon who fires first, you strengthen A weakened synapse. Can that work? So amazingly, it learns. This is capable of learning the correct sequential outputs, right? You just tell it that I want this sequence of firing in this area. It can learn. It can learn that output. But there's an interesting finding. You can find out this way, right? And that is, the learning works better if the teaching style of the tutor matches the learning style of the student. The kind of thing you hear in kindergarten, right? So basically you know it could be that the neuron has this spike timing dependent plasticity rule. You know if A goes before B then you strengthen, if

01:19:51

B goes before A you weaken. Or it could have this spike timing dependent plasticity rule you strengthen so long as both of these far close to each other. Right now there are many different rules different neurons implement. So it turns out that for some of these rules of learning that an individual synapse could implement. It turns out the teacher the tutor, should teach the student, give it signals so that it learns early segments of the song before late segments of the song. That's the way it's going to work best, given that rule of how the learning at the synapse works. But for other students, by which I mean other kinds of synaptic plasticity or learning rules at the synapse, you'll find that the teacher should just teach the whole song all the time. Just keep giving corrections everywhere and it'll work better. It'll just work faster. That's a very interesting finding. And this is intended to illustrate 2 things. ask is what is the world around us? So for each sense or for each mode of interaction with the world adapted to the world, then one question you have to ask is what does the world look like? So in this case you can go collect a large database of images and study its properties. It's a collective effect. The whole brain cooperates and the synapses have to do things, so it's a collective of everything that's doing this effect, The first statement. The other statement is this kind of theoretical computer modeling. That's now possible given, you know, I told you earlier you could build a model of a neuron, and why not put it on a computer and Justice, see what it does? You can now do that kind of thing and learn a great deal. In fact, this is a prediction for things that ought to be measurable in different kinds of birds and different kinds of synapse. So we're now reached the stage where you can have a sort of the style of of science that was in physics, which is that you can have theory, you can have experiment the theories, make predictions, you can check experiments, experiments refine the theories, etc. That is a mode of doing science that's been basically absent for much of neuroscience and much of biology. And that is now becoming possible because we have so many tools and so much knowledge of all of this. I'm basically out of time. So I'm not going to talk about the circuits that underlie a sense of space on this. You ask me during Question Time, which can tell us more, tell you more. And instead, I'm going to conclude. Right. So

the point of my talk today was to try to suggest to you say something about how our brains make us. And the kind of point I want to try to make to you is that inside our head, we're all collectives. It's not like there's a thing which is you. There's just lots of neurons. Lots of neurons are connected, lots of circuits. The circuits are connected into brain areas. And collectively, the emergent effect of all of this is you. I tried to illustrate that by looking at various examples, you know, in special cases of different kinds, in particular, for example, this learning. Another point that I'd like to emphasize is that essentials, all animals are the same. Like this turtle and my daughter at an earlier age are basically in all essentials. They're the same. All of the stuff we like about, you know, the the fancy thinking we do is a veneer on top of that. Nevertheless, there's a lot more to discuss. So, for example, we're very interested as humans in decision making, social behavior, things like curiosity and creativity, language, right? Which humans have a particular faculty for language, abstract thought, which matters to us. And we'd like to understand the origins of all of these behaviors in neural circuits. And actually, there's progress on this just in the last 20 years or the last 10 years. There are more and more tools for studying all of these things. And I anticipate that the next decade we'll see lots of progress in understanding the neural circuit origin of these sorts of behaviors. Then usually most people are interested in the even broader question, right? One of the things you prize about ourselves is sentience and consciousness. So we'd like to know what are these states? They're clearly states of brain. You know, we don't all say that rocks are sentient or conscious, the states of brain. But we don't quite know how to define even what sentience and consciousness even mean, sort of in general, let alone bunch of neurons. You allow the conductor neurons to send signals to all the student neurons. Right. And we'd like to know from this, how does mind emerge from brain? But I actually think that we will make progress in these questions in the next 100 years, right? That might seem like a long time, but it isn't really. You know, people live to 100 and and if you think about it, we have come a heck of a long way since Kahal. Kahal had pictures of lots of different neurons and suggested that neurons are central to how the brain works. So there are these different things. They're like the atomic constituents of brains. And think about all the things I've been able to say today. And I even only just scratched the surface. So 100 years from now, I think there's every chance that we'd be able to give some sort of definition of what we want, we mean by something that's conscience or sentient and understand how those states arise from neural circuits. So I'm going to stop there. So we'll go ahead and open it up for Q&A. I think we have room for about two questions. That's OK. So in your opinion, what would be the best way to learn a new language? You're asking the tutors should start at the beginning of the sentence only versus that. That's really not clear. I mean, there are different clearly learning styles that people have, you know, So some people, my wife for example, is very good at languages and always talks about how when she learned Arabic, the grammar was what really attracted her. There's this very regular structure to that language and learning that grammar helped her to understand the language and everything. My dad is very good at languages. He speaks 7 languages and from him I get a impression of a kind of on the one hand, he just listens to it and reads it, and then somehow on the way on the side studies grammar.

01:26:14

It's a very different mode. It's also the case that the young and the old learn language in very different ways. So you know, children just listen to it, right? And they just pick the whole thing up somehow, holistically. And there's a certain stage, there's a critical period where you can do that. And after that critical period, the mode in which you learn language is known to be different, right? So it's much more useful when you're older to have structures like grammar and things like that and consciously be aware of them and to learn them. So there used to be this idea. Well, there is this idea due to Noam Chomsky of something called a universal grammar that's sort of embedded in these kinds of circuits that I've talked about that it's just built in, that all humans have the faculty for this and that. In early childhood, when you learn languages, what you're doing is pruning away bits of that that you don't need. It's a little bit like all of those neurons, those interneurons, the retina that

sort of take away all the stuff you don't need, right? So the idea is that you sort of prune those away and you're left with the structure pertinent to the language that you're going to speak in your local area. Now, you know, with the advent of things like ChatGPT and this sort of and their remarkable linguistic abilities, there's some discussion. I mean, there's debates about whether this is still whether people agree or disagree still with this proposition. So I think that's a little bit up in the air because of this particular, if you think about ChatGPT as a development in computational linguistics that's raised some questions about this notion of universal grammar, whether it's really there or not. I mean personally I think they're going to find that and I think the reason why things like ChatGPT work is that there is a underlying regular structure in all human language and this is an engine that's built in such a way that it's able to extract that structure in the language. And so I think people are going to want to understand how to peace support these artificial engines and workout what they're actually doing inside their in arts. I think we will find that there is such a structure so but so you know so it's different strategies for young adults also clearly different for different people. I'm not sure that's a sufficient answer to what your question, but there you go, I'm coming, I'm coming, I'm coming. Oh, thank you. Yes, You touched on something which was related to my question, which is the effect of age on, you know this, This is such a vast number, it's even hard to conceive right. That's the number of different things, like older objects, if you I'm a senior and I think there are a lot of seniors in this room. So can you, can you speak to that? Are we? Yeah, yeah, yeah. So you know, age has produces sort of many effects that on the structure of the brain, you know, in part because adults basically don't get new neurons. There are two regions of the brain where you do so in the olfactory system, so part in this structure drawn by Golgi, this one over here, you actually get new neurons as you age, which is very interesting because nobody understands why, right. The other area where you get new neurons even as even even when you're older is the hippocampus, which I mentioned here but didn't really talk about, which is the area associated with the learning, memory, navigation and things like that. So in these two areas you do get new neurons and their appearance and their placement within the circuit and where they synapse is known to be associated with the production in the hippocampus anyway, of new memories. So one of the things that happens in people who have memory loss is that system isn't working quite as well. You're not getting enough neurons not snapping properly and things of that. So that is so. But we know the locus of that maybe there is a day will come when you know, you can encourage the production of new neurons. So yeah, you know, you could make the argument that part of the reason we developed these kind of age-related problems is maybe humans weren't original, didn't originally live that long, right. Most people died young. They were eaten, They died of disease. You know, the died of war young. So at an earlier time in human history, we wouldn't have needed well in in the general populace, populace to kind of maintain all of these structures to that one. So maybe that's been changing and we're unfortunately stuck with an earlier evolutionary program. That's what's going on. So it's possible that those things can can can help. And also you know in general aging produces effects on all kind on all your cells. So there's a, there's things called telomeres that are at the end of and the chromosomes. And they as you make copies of cells, the telomere shortened. And if they get shortened so much so they're not really delimiting the ends of the chromosomes, then you start running to problems because the, you know, the program for transcribing genes to make proteins and things won't work as well. So that's another issue. But not all creatures have that. So there are creatures that are essentially immortal, that don't have this telomere shortening. They basically die when they're eaten or they don't get food. There are there's a jellyfish that gets younger. At some point it seems to get younger and then age again sort of oscillates in its effective age. Very strange, very interesting. We'd all love that. So, and these are all subjects of, of the study of aging. There's actually even weirder stuff still, right? Just to reveal the sort of set of possibilities that's that can happen. So there's a creature called Hydra. You can chop it into two. It's got a nerve net. It doesn't have a brain, it has a nerve net. So there are nerves all over it. It's kind of hooked up in a kind of neural network that operates hydra, right? I'm mostly talking about creatures with centralized brains, but some creatures actually have a nerve net that's kind of spread about. You can

chop hydra in two, it'll make 2 hydras, right? You can take hydra and put it in the blender and kind of separate its cells, all separate. You put it back together, it makes a hydra. All the cells find their location. I'd love to be like hydra when I grow up, right?

01:33:08
Right. OK, Now let me spend a little bit of time introducing tonight's speaker, your interlocutor for tonight's conversation. tricks that we're only just beginning to Plumb the depths off. And so I think in time we will learn to do these tricks. Heck, you know, you lose a finger in a, you know, in a circular saw. There's no reason why you can't grow it back, right? I mean, the program is there, right? So I'm not talking about neurons now, right? I'm talking about the rest of the body. But all of these things are in principle possible. We just need to know how to unlock within the, in that case, the circuits. I mean, you can think about the organization of your body in terms of the circuits of what cells communicate with what, what turns, what on, et cetera. It should be possible. We just haven't figured out how. And likewise with the brain and with the problems of aging and everything, you know, we once we understand how these circuits work, there's no reason why we can't get them to do what we want. These are machines. We're machines and if you can repair a car, you can repair the head. Thank you so much Vijay, and be sure to join us for our next set of lectures. We have our ulams coming up in September, September 19th and 20th. Have a good night.