THE TRANSMITTER
BRAIN INSPIRED

Dmitri Chklovskii outlines how single neurons may act as their own optimal feedback controllers

From logical gates to grandmother cells, neuroscientists have employed many metaphors to explain single neuron function. Chklovskii makes the case that neurons are actually trying to control how their outputs affect the rest of the brain.

12 FEBRUARY 2025 | by PAUL MIDDLEBROOKS

This transcript has been lightly edited for clarity; it may contain errors due to the transcription process.

Dmitri "Mitya" Chklovskii
My view is that real neurons are smarter than the McCulloch and Pitts neurons. Sometimes I refer to a new conceptualization of neurons as the smart neuron. Maybe neurons' outputs don't just predict their inputs. Maybe they can influence their inputs. The reason million-dollar question here that I think you're alluding to is how to go between levels.

Paul Middlebrooks
Oh, God, yes. That's the dream.

Dmitri Chklovskii
That's the dream. Of course, I don't have an answer here. I think this is probably the most fascinating question in neuroscience because--

[music]

Paul Middlebrooks
This is “Brain Inspired,” powered by *The Transmitter*. Since the 1940s and '50s, back at the origins of what we now think of as artificial intelligence, there have been lots of ways of conceiving what it is that brains do or what the function of the brain is. One of those conceptions, going back to cybernetics, is that the brain is a controller that operates under the principles of feedback control. This view has been carried down in various forms to us to present day.

Also since that same time period, back at the origins of artificial intelligence, when McCulloch and Pitts suggested that single neurons are logical devices, there have been lots of ways of conceiving what it is that single neurons do. Are they logical operators? Do they each represent something special? Are they trying to maximize efficiency and so on? Dmitri Chklovskii, my guest today, who goes by Mitya, runs the Neural Circuits and Algorithms Lab at the Flatiron Institute. Mitya believes that single neurons themselves are each individual controllers. They are smart agents, each trying to predict their inputs, like in predictive processing, but also functioning as an optimal feedback controller.

We talk about the historical conceptions of the function of single neurons and how Mitya's account differs. We talk about how to think of single neurons versus populations of neurons, some of the neuroscience findings that seem to support Mitya's account, the control algorithm that simplifies the neurons' otherwise impossible job at implementing this feedback control, and other various topics. We also discuss Mitya's early interests. He has a background in physics and engineering, and the way he got into neuroscience was an interest in figuring out how to wire up our brains efficiently, given the limited amount of space in our craniums.

Obviously, evolution produced its own solutions for this problem. This pursuit led Mitya to the study of the *C. elegans* worm, because its connectome was nearly complete. Actually, they thought it was complete, turned out it was nearly complete, and Mitya and his team helped complete the connectome so that he would have the whole wiring diagram to study it. We talk about that work and what knowing the whole connectome of *C. elegans* has and has not taught us about how brains work.

As always, I link to Mitya's work and his lab and his information in the show notes at braininspired.co/podcast/205. As always, thank you to my Patreon supporters. Consider supporting this podcast if you value what I do here and want access to the full archive, all the full episodes, and so on. Thank you, Patreon supporters. Okay, here's Mitya.

[transition]

Dmitri Chklovskii
Okay, just one technical issue to get out of the way. You know how to pronounce my last name, right?

Paul Middlebrooks
Chklovskii.

Dmitri Chklovskii
Yes. Think of it as S-H rather than C-H.

Paul Middlebrooks
Chklovskii. [in Russian] Do you speak Russian?

Dmitri Chklovskii
A little bit.

Paul Middlebrooks
Me too. [in English] What is Chklovskii? What's the background?

Dmitri Chklovskii
Oh, there is a town in Belarus called Shklov.

Paul Middlebrooks
Oh, cool. You're of that town.

Dmitri Chklovskii
Yes. Some very long time ago, probably.

Paul Middlebrooks
Yes, sure. Okay, cool. All right. I'm going to start here.

Dmitri Chklovskii
All right.

Paul Middlebrooks
What percentage of modern computational neuroscientists do you think come from a physics background?

Dmitri Chklovskii
That's a very good question. I have not formally done the statistical analysis. Certainly among the people that I hold in high regard, probably 50% do.

Paul Middlebrooks
Oh, wow. Okay. You do have that background in physics. Do I have this right that was it your second postdoc when you started getting into neuroscience? How did that come about?

Dmitri Chklovskii
Yes, that's true. My undergraduate was a combination of physics and engineering studies. Then I did a PhD in theoretical physics. I was always torn between the rigor and the intellectual challenges of theoretical physics, and sort of the desire to solve some real-world problems. After I did a first postdoc in physics, I actually liked it very much because of the freedom I had to do research, that I actually turned down faculty offers in physics, and took another postdoc in neuroscience, which was my way to continue doing research without constraints. Then I ended up as a faculty in neuroscience. I never regretted the switch.

Paul Middlebrooks
Why neuroscience? You could have gone anywhere, right?

Dmitri Chklovskii
That's a good question. I actually looked at a lot of fields. Basically, I decided that neuroscience was exciting enough because I was motivated by solving big questions, and "how does the brain work" seems like one of those.

Paul Middlebrooks
How does the universe work is another one. We haven't figured that out, have we?

Dmitri Chklovskii
That's absolutely true. The problem was, and still is in physics, is that sometimes those studies become very empirical because we are limited in the experiments we can do. That has not changed very significantly in physics since I switched. In neuroscience, I would say the progress has been

immense. I remember when I was a postdoc in neuroscience, we were sitting around having beer and talking about, what if I could do this experiment? What if I could record at the same time from 10 neurons?

Paul Middlebrooks
Oh, my God. 10?

Dmitri Chklovskii
Yes. Then we could really understand how it works. In the 30 years that passed, basically, like these days, from my perspective, they can do any experiment you want. The big question is, what experiments should be done? That's my perspective, at least.

Paul Middlebrooks
I'm interested. There are so many physicists who come into neuroscience over the years, over multiple decades now. I've told this story multiple times before, but I cannot source the conference that I was at. I remember the opening keynote was a molecular biologist, I think. He was saying, we need to give up and let the physicists come in and solve this for us, because we are at the edge of our capabilities here. That was 15 or so years ago, maybe 15 to 20 years ago. Of course, that raised the hair on the back of my neck, "How dare the physicists come in?" Do you agree with that? What approach do physicists in general bring that us wet, more biological sciences folks have been traditionally missing out on? Then where do you sit in that?

Dmitri Chklovskii
Yes, that's a really good question. I think that on the lighter side, physicists think they can solve any problem.

Paul Middlebrooks
Is that true? That's what I think. Is that true?

Dmitri Chklovskii
They're very arrogant, but they're right. As a physicist, I can say that. The problem is that the physicists are trained to figure out how nature works. They're not trained to figure out how to build things. That has been sort of the domain of engineers historically. The strength of physicists is that they can really come in and understand a different field and formulate a research problem that can be solved in terms of gaining understanding.

Paul Middlebrooks
That's from a particular approach. The physicist's approach is a particular approach. You think it's the right approach to understand anything?

Dmitri Chklovskii
I think that's the strength of the approach, that knowing what understanding means in this sort of scientific sense. The shortcoming of physicists, and again, as a former physicist, I think I can criticize us, is that physicists are not trained to build stuff that works.

Paul Middlebrooks
You're an engineer by background also.

Dmitri Chklovskii
Engineers are. As I was telling you, my undergraduate studies were half engineering, and so I have this interest in building things. In the end, Richard Feynman, one of the famous physicists of the 20th century said, I think, that you don't really understand something until you can build it.

Paul Middlebrooks
Yes. What I cannot build, I do not understand. No one gets the quote exactly right, which is crazy because it's written on his chalkboard and people can look it up really quickly.

Dmitri Chklovskii
Right. Basically, I completely agree with this, but physicists, when I was trained as a theoretical physicist, at least, experimental physicists are different, okay, they have to build equipment, but theoretical physicists, they're not trained this way. That's where things become tricky because I think the brain was built by evolution to solve some practical problems. To understand how to do that in a robust way that would withstand adversarial environment is an important consideration.

I think that the best approach in my mind is some fusion of physics background with engineering skills. Of course, it has to be heavily grounded in biology. I don't believe in this approach, and that's where arrogance comes in, unfortunately, that a physicist can just come in and say, "Okay, so what is the problem I need to solve?" No, I think that the right way is to really understand the biology as well as biologists do, or better. Only then you can formulate the problem and solve it.

Paul Middlebrooks
What would your advice be to a biologist who wants to-- Do they need to go back and get a theoretical physics degree and an engineering degree? Who wants to sort of join that approach, that framework for thinking?

Dmitri Chklovskii
I think biologists are doing okay. I think there are a lot of things to do for biologists. They don't need this, but if you're talking about solving the brain on the level that would allow you to build an artificial version of it, then I think you need a fusion of those three approaches, biology, engineering, and physics.

Paul Middlebrooks
All right. Since you are talking about building a brain, so we're going to be talking about your conception of single neurons, essentially what you've come to. Before we get to that, just to round out your approach and conceptual overview on how things stand and where we are and where we need to go, you didn't mention computer science as all you need, and that maybe that's all you need to build AI these days. My question is twofold. What do you see that might be missing in current AI and/or in current neuroscience from this perspective?

Dmitri Chklovskii
Yes, so that's actually a very important question, I think. I'm a big fan of the recent developments in AI and I'm a daily user of ChatGPT, and I'm in awe of this technology. More than that, if you asked me five years ago or my more computer science-accredited friends, I don't think anyone would have predicted that we would have something like this today. This is really living in the future. It's a fantastic technology. The question at hand is that is this emulating how the brain works? That, I think, is a completely different question. With all due respect to computer scientists, I think they are trained to build things that work again--

Paul Middlebrooks
They're more on the engineering side.

Dmitri Chklovskii
On the engineering side, less so than understand how living things work.

Paul Middlebrooks
Does it matter that what they're building does not emulate brain function to accept on the most abstract level?

Dmitri Chklovskii
Doesn't matter to whom. It seems like companies like Open AI, Anthropic are doing pretty well. For them, it probably doesn't matter. For me, it matters because my goal is to understand how the brain works.

Paul Middlebrooks
I wanted to talk a little bit about *C. elegans*, what is it, 302 neuron organism, 302 for the females. Is that right?

Dmitri Chklovskii
Yes. Exactly.

Paul Middlebrooks
We have a complete connectome, thanks to people like you, for *C. elegans*. You've worked on connectomes and all the structural stuff. Did you at one point think, "Well, when we have the structure, we'll understand it?" Then you came to realize that we need something beyond structure? How did your thinking about connectomes then evolve to how you think about them now?

Dmitri Chklovskii
I'm glad you asked this question. I think that my path in biology has been somewhat winding. I started out as a physicist, as we discussed, and I just wanted to see what I could do. One of the things is that I was fascinated by evolution and how can you do something related to evolution? Since we don't have the equations that describe how the brain really thinks, maybe we can come up with a simpler framework of equations that explain the structure of the brain, sort of just scratching on the surface of the function.

For a few years, I was working on the topic, which people call wiring economy, which is basically the idea that evolution had to build the brain, which is very highly interconnected structure. It has a lot of wires, axons and dendrites, and under certain constraints. There is a volume constraint to the brain, you have to be born, and there are metabolic constraints, time constraints, and so on. To solve the packing problem, to arrange all the wires in the brain is actually very difficult. It's akin to arranging components on a semiconductor chip like transistors.

This is a multi-billion dollar industry, how to arrange the elements of a semiconductor chip in the most optimal way to economize on wiring. Basically, what I was doing is trying to understand the layout of brain structures, the shapes of neurons from the perspective of economizing on wiring. I was at a place called Cold Spring Harbor Laboratory. One of my biologists colleagues said, "Well, this is just also theoretical. Why don't you just test if this wiring economy thing is true?"

I said, "Well, I would love to test it, but, there is no circuit for which we know the full connectome." Of course, the word connectome then wasn't used, but the wiring diagram. He said, "Well, actually, there is one. It's called *C. elegans*, and it has 302 neurons." I said, "Really?" "Yes. There's people in Cambridge that reconstructed the wiring diagram and published in the '80s, so you can use it."

I had this great student, Beth Chen, and I said, "Beth, why don't you go get the connectome and optimize the layout and see whether you can explain where the neurons are actually located in the worm?" She comes back and she says, "Well, they don't have a full connectome." I'm trying to use their wiring diagram and all the neurons they collapse, majority collapsed towards the head and the other collapsed towards the rear. Then I looked up and the connections are not finished.

Paul Middlebrooks
They had the skeleton, but not the-- The wires were not respective?

Dmitri Chklovskii
They basically did 90% of the work. The most interesting part was the head and the tail, where the majority of synapses are, but they didn't go all the way through the body and link them up.

Paul Middlebrooks
I see. Okay. They weren't concerned with wiring length. Is that part of the issue or?

Dmitri Chklovskii
No. They were doing really hard work and it was all manual in those days with electron micrographs on film and tracing them by hand. It was very, very difficult and hard work. They've been doing it for many years and they did the most essential part. Then the part is boring, which is like the body segments, there are no real segments, but more or less stereotypical structure along the body, they did not complete. At that time, DNA sequencing became feasible. The majority of them switched to DNA sequencing.

They just published what they had. In the *C. elegans* world, most people didn't actually realize that it wasn't finished. My student, Beth Chen, she spent more than a year completing their work by basically using their micrographs that happened to be archived in New York at Albert Einstein College of Medicine, where she would go almost every day. Some of the materials were even there, but they were so old and brittle that the samples would deteriorate while they were being imaged under electron microscope. It was like quantum information. They would deteriorate once you scan it. Anyway, she finished, and that's how the first complete connectome has been assembled and published.

Paul Middlebrooks
That's crazy to think about. That wasn't even that long ago, and that's almost child's play compared to what they're doing today. It was so much work.

Dmitri Chklovskii
Yes, of course.

Paul Middlebrooks
You came to that through your ideals of optimization and efficiency, and, well, I guess just through that lens, huh?

Dmitri Chklovskii
Yes.

Paul Middlebrooks
Okay, so now you've completed the *C. elegans* connectome, and then your career is done. That's what you think?

Dmitri Chklovskii
Yes. [laughs] Ready for retirement. I, of course, now that we've constructed the connectome, an arrogant physicist in me is like, "Oh, okay, now we can figure out how it works." It was a very interesting process, and this work is highly cited. We analyzed this *C. elegans* connectome to death with all the possible approaches that, I think it was, maybe it's called the network science, the methods of statistical analysis of networks. We applied every method that we could find off to analyze *C. elegans* connectome, and we made statistical discoveries for sure. Some of them have been since replicated in other species, including mammalian brains.

Paul Middlebrooks
Some scale-free things, yes.

Dmitri Chklovskii
That's right, scale-free, motifs, distribution of synaptics, log-normal distribution of synaptic strength, and stuff like that. There has been a lot of statistical discoveries, but on the other side, I don't think we made any real progress in terms of understanding how *C. elegans* computes and how it generates behavior.

Paul Middlebrooks
I'm not sure how you feel about this, but I won't say it's the butt of jokes about neuroscience, but it is a dig at neuroscience. People all the time say, oh, we have the complete connectome of *C. elegans* and we still don't know how *C. elegans* works. Neuroscience is still lagging in terms of figuring out the function of how structure relates to function. How do you feel about that that's the go-to that people often use.

Dmitri Chklovskii
Absolutely, it's a fair criticism. I spend a lot of time perplexing about that very question because I thought as a physicist *C. elegans* is a hydrogen atom of neuroscience, and we're just going to discover the real principles here. We didn't, not just based on the connectome. The arguments that people give, why it is so hard for *C. elegans*, are like physiology actually is very difficult in *C. elegans*, just because it's enclosed in this cuticle and it's hard to penetrate with electrodes without blowing it up.

Paul Middlebrooks
The neurons don't spike, right?

Dmitri Chklovskii
The neurons, well, we have to be careful here. There are no sodium action potentials, but there are calcium spikelets. Not everywhere, I think, but there is a combination of graded potential and calcium spikelets in neurons. It's very difficult to record. At the time when we were doing our work, there was very little physiology period. Now there is a lot more, but most of it is optophysiology based on calcium imaging. It's better, but it's still just calcium. It's not voltage really. People are starting to do voltage dyes and so on. The situation is slowly improving, but this is a big detriment.

Just to give you an example of why this is so hard is that in vertebrate neuroscience, we used to think that each neuron produces a unique output, a spike train that is then distributed to all of the downstream neurons. There is one signal that is communicated downstream. That seems to not be true in *C. elegans* because there is no clear separation of its neuron into axons and dendrites.

Paul Middlebrooks
C. elegans, does it have small world motifs? There's the famous small world thing, but there's so few neurons that it's-- Can you even consider it?

Dmitri Chklovskii
Right, so that's where we first discovered motifs in neuroscience, the small number of neurons. What I'm trying to get to is that the nodes of those motifs, which are thought of as neurons and invertebrates, that's a reasonable node because it has one output. In *C. elegans* is actually not a single node, but each neuron consists of multiple sub-compartments. Each of them can communicate downstream its own signal.

Paul Middlebrooks
Oh, but it's almost like dendritic outputs or something.

Dmitri Chklovskii
Exactly, because there's no separation into axons and dendrites *in C. elegans* generally. Basically, the output synapses can happen on the same branches where the input synapses are. We know for a fact now, thanks to great experimental results obtained by calcium imaging, that neurons are separated into sub-compartments. Each of them is computing a different thing. There is no place, perhaps, where everything is getting integrated together an output like invertebrate neurons. Each neuron in *C. elegans* is actually mapped onto several neurons in several vertebrate neurons, so to say.

Paul Middlebrooks
There are these results that every single neuron, a regular neuron that you think of, can actually be modeled as a neural network itself. What you're saying is that it's even more so in *C. elegans* because every neuron is almost multiple neurons.

Dmitri Chklovskii
Right. It is multiple neurons, but there is no single output that sums up this network. In a vertebrate neuron, if you think of dendrites as its own neural network, in the end, it's all summed up and there is one output. In *C. elegans*, no, that doesn't have to be the case. Having the connectome in this case doesn't get you too far because you don't know what's inside the node. Moreover, the outputs of the node, if you look up the wiring diagram that White *et al.* or reproduced in groups after us, doesn't mean that the same signal is communicated downstream along all those outputs, which is, of course, what normal people would assume.

Paul Middlebrooks
Does that give you any pause in, "Okay, so not every neuron is alike, not every brain is alike." What we want to do is say, "Okay, here's this unit of function, the neuron, and then we want to apply that same whatever abstraction we take, whether it's a McCulloch-Pitts point process model or whatever, and just implant it in all the species to understand all brains. We're going to all understand all brains the same way."

Does this make you think that every brain is, for lack of better term, special or unique in the species and that we need to have different understandings for each species? How does this change your view of understanding brains? What's interesting is what we're going to get to in the way that you conceptualize neurons right now.

Dmitri Chklovskii
Right. I have to put on my biologist hat at this point and say, yes, of course, every species has something special. I think that *C. elegans* is much more extreme in this way. For spiking neurons, for example, it's very difficult to have multiple signals communicated downstream because the spike is such a global, it's a global over the whole neuron, event.

Paul Middlebrooks
Yes, it's an event.

Dmitri Chklovskii
Right, and so it's hard to have independent outputs in a spiking neuron. Once we get to spiking neurons in evolution, then things become simpler actually.

Paul Middlebrooks
Okay, if you buy that everything is-- There are electrical, chemical signals, there's all sorts of signals, but eventually there is this unit event that is sent that is the spike.

Dmitri Chklovskii
This is also true. I should have mentioned that in *C. elegans*, one other response is coming from biologists why we can't understand having the connectome is that, well, but there is communication along the so-called wireless connectome, because there are these neuromodulators that are transmitted, what is it called, ephaptically, that don't require synapse that you can identify electron microscopy to communicate between neurons, and that connectivity is actually rather intricate.

It has been studied now that you can look it up, but it's the same level of complexity, at least as the sort of synaptic connectome. Why should we be able to ignore it? In a bigger brain, I think the answer, now putting on my physicist's hat, I would say the separation of time scales. Because those ephaptic interactions are mediated by diffusion, and diffusion is slow over long distances, then if you're concerned about physiological properties that appear on fast time scales, like motion coordination or memory recall, those has to occur through electrical means where the only diffusion you have is across a very, very thin synaptic cleft.

Otherwise, you just don't have the time. Then if you're worried about even memory formation, that could be a different matter, that takes a longer time. I think by carefully choosing which problem you address, you can match it to your level of description.

Paul Middlebrooks
Keeping your physicist hat on for a second, when you look at the messy complexity of all of these communication-type systems, do you see them just as little sub-problems to isolate and understand on their own, or do you see any hope for an overarching understanding of their interactions in all their complex glory?

Dmitri Chklovskii
Yes, as a physicist, I believe that there should be a set of principles. I call them algorithmic principles, and that's where we go back to engineering, that are describing what the brain does on multiple levels.

Paul Middlebrooks
Is that the brain's version of laws in physics?

Dmitri Chklovskii
I think so. I like to make a connection with the computer chips that you can model, of course, electron conduction through wires in a chip and through semiconductors and so on. To really understand how it works, you need to go to a different level of extraction where you think about logical gates and registers and so on. That's the level which is central to the function. I think we're lacking that level of description in the brain.

Paul Middlebrooks
Okay. This is the main reason why I invited you on today, is your work conceptualizing single neurons as controllers in a very particular way. I've had multiple guests on who utilize control theory as a conceptual basis for understanding brains in general, the whole person in general. You have taken it down to the single neuron level. We should talk a little bit about how single neurons have been conceived of throughout history and then how your conception of them as controllers differs from that in certain ways.

Dmitri Chklovskii
Yes. I don't know how far back we can go. I think that the modern age starts probably with McCulloch and Pitts.

Paul Middlebrooks
Right. Which is the modern age of artificial intelligence also.

Dmitri Chklovskii
Exactly. Exactly. Who, all those people, McCulloch, Pitts, Rosenblatt, Hebb, and so on, they're scientific heroes. There's no question about this, that they were able to abstract those simple models from whatever they heard about brain research is just incredible. Their model is obviously very influential because with minor tweaks, it's the same model that is being used in almost all artificial intelligence systems today.

Paul Middlebrooks
It's astounding. Absolutely astounding.

Dmitri Chklovskii
Including ChatGPT. What's under the hood and then the individual units are those McCulloch, Pitts, Rosenblatt, Hebb units that have been conceptualized in the '40s and '50s and '60s.

Paul Middlebrooks
Just a few minor tweaks, yes.

Dmitri Chklovskii
Right. That was, of course, a very influential discovery. Since this is a “Brain Inspired” podcast, we have to acknowledge that neuroscience wasn't standing still over these last 70 years.

Paul Middlebrooks
Right, but let me stop you because you said discovery, and I would say it was an engineering feat rather than a discovery, the single-point McCulloch-Pitts neuron. Am I quibbling with semantics here? It was a modeling conceptualization.

Dmitri Chklovskii
True, but as a physicist, I would say they conceptualized the model. It's like discovering mutant cells.

Paul Middlebrooks
They discovered a model. Oh, Okay.

Dmitri Chklovskii
My analogy would be that.

Paul Middlebrooks
They discovered the math-- I'm hung up on this discovery term, I'm sorry. They more or less discovered the mathematical principles underlying an abstract mathematical conceptualization of what neurons might be doing based on physiological data.

Dmitri Chklovskii
Yes.

Paul Middlebrooks
Sorry, I don't mean to hang us up on this point here, but you said discover.

Dmitri Chklovskii
Okay, I'm willing to negotiate on the exact words.

Paul Middlebrooks
It's fine, it's just terms.

Dmitri Chklovskii
I don't want to be hung up. Model, conceptualization, fine.

Paul Middlebrooks
Yes, as you were saying, neurophysiology didn't stand still over these past 80 years or whatever.

Dmitri Chklovskii
Exactly. Now I think from a biology perspective, we know that this is a rather primitive way of looking at neurons. It's important to say now that at least my goal is not to include all the biological details, which there are a lot. People have discovered a lot of amazing things in the level of ion channels and even protein signaling and gene regulation and all those things that are, of course, necessary for the brain function.

In my analogy to computer chips, I don't want to go to the level of, conduction of electrons in semiconductors and the band theory of solids and so on. I want to go on the algorithmic level, like logic gates, and that's where there hasn't been much progress. It's clear that there are several ways in which McCulloch and Pitts and Rosenblatt and Hebb view is a major oversimplification.

Paul Middlebrooks
I'll just pause here because you just transitioned from the messy details about how neurons work, there's their biological implementation, the implementation level, if you will, of Marr’s famous levels. Then you went to logic gates, which is good because while you were speaking, I was thinking about there is the implementation stuff, but then there's the question about what neurons are doing, what their function is. Then you went to the logic and I was going to say, well, that McCulloch-Pitts, that they are logicians, that they are producing binary logic signals. That's the McCulloch-Pitts approach, ones and zeros originally. That was the original functional story about what neurons are doing. That's what you're saying has not advanced much.

Dmitri Chklovskii
Exactly. What are the missing parts? Of course, we never really know until we get something that works. Generally, my view is that the real neurons are smarter than the McCulloch-Pitts neurons. Okay, so sometimes I refer to a new conceptualization of neurons as the smart neuron, and in which ways? One of the things that I think most biologists or physiologists will immediately agree with me, is that a McCulloch-Pitts neuron is an instantaneously responding device.

If I provide kind of certain inputs, they instantaneously compute an output by weighted summation and then non-linearity and output. We know, of course, in neuroscience that the neuron does not process inputs instantaneously, it has all kinds of time scales in its dynamics, and there are short-term, long-term memory effects, and you can characterize them in a variety of ways, such as measuring the linear temporal filter by spike-triggered analysis. This is, I think, very important because it's telling us that the neurons care not about correlations between its inputs, between different upstream neurons, but also in the temporal correlations in the same inputs they get. That's completely missing from the, McCulloch and Pitts-inspired units.

Paul Middlebrooks
Even McCulloch and Pitts, in their original paper, drew these loops. The classic story is they conceptualized everything as a feed-forward network only, but they actually drew the loops, and they alluded to how hard it would be to determine how to incorporate those recurrent loops within their models. They knew it was a problem, and knew it would have to be addressed. That's what you're talking about, is this historical recurrent context that neurons have to deal with.

Dmitri Chklovskii
That was actually my second item, is the existence of those loops, which actually, I agree with you that they realized that, and you can look at the figures in their papers that they realized they were there. I think that's one of their geniuses, that they ignored them. The network that originated from them were ignoring this, and that's why it was possible to make progress.

Paul Middlebrooks
Oh, that's interesting. I've never heard anyone celebrate them ignoring it. I like that.

Dmitri Chklovskii
That's how we do in physics when you build a model. All models are wrong, some are useful.

Paul Middlebrooks
That's one of the things. Let me just also bring in, as we discuss this, historically, speaking of function, there have been different conceptions about what the job of a neuron is to do. You write about this in your recent paper that I'll link to, there's the efficient coding hypothesis. There's the idea that the grandmother neuron, that an individual neuron represents an individual thing in the world, things like that. There's the predictive coding. As we're talking, maybe you can situate the neuron as a controller within those contexts as well. Here's a different way to approach it. How did you come to the conception of the neuron as a controller? There must have been some train of thought leading up to that.

Dmitri Chklovskii
Yes, so I was basically very influenced by two things. One is I'm a big fan of efficient coding theories, and I spent a lot of time thinking about those and predictive coding in particular. I still think this is very important. The problem with those approaches, as I'm sure has been brought up by other people, is that when you do efficient coding, it works at the early sensory stages, because it makes sense that you want to represent the world. Those theories have been successful in explaining many properties of neurons, such as their temporal and spatial receptive fields of retinal neurons, even the edge detectors in V1. It failed when people wanted to march further into the brain.

Paul Middlebrooks
Yes, and the history of neuroscience is mostly dominated with sensory cortex.

Dmitri Chklovskii
Exactly, and then, of course, the primary reason for that was experimental. It's an easy stimulus to control the stimulus, and so you can do reproducible experiments. In terms of theory, if efficient coding and predictive coding appealed to me as a physicist, it's such a beautiful theoretical foundation that you can use, information bottleneck. Then it fails as you want to march deeper into the brain. If you start thinking why, of course, well, you want to get closer to action generation, to motor control, decision-making. It's not just about efficient coding.

You have to have other things in mind in terms of the objective. Then from connectomics, of course, I knew that loops are everywhere. That each neuron, even in *C. elegans*, almost each neuron belongs to a multitude of loops. Its output can get back to its input by going through, one, two, three synapses.

Paul Middlebrooks
Was cybernetics any influence on you in this regard also, or was that something that came in later? I'm getting at the control theory aspect of it. Because that's all about agency and motor output, what you're talking about, right?

Dmitri Chklovskii
Yes, probably was because I read some of this work when I was in school, but the issue is that, if you have loops, then the dynamics changes

completely. Because there is a danger of getting runaway excitation and instability. It's important to have a framework that pays a particular attention to that. In my mind, that's the domain of control theory. Control theory, on the one hand, is a way to sort of generate certain desired output. On the other hand, it is a way to deal with loops.

Paul Middlebrooks
Deal with the feedback.

Dmitri Chklovskii
Deal with the feedback, that's correct. That's the correct word. From those two considerations, it seems like a good framework to apply to the neuronal circuits.

Paul Middlebrooks
What is it that a neuron is trying to do? In your conception, every single neuron is trying to do this.

Dmitri Chklovskii
Right, so very good. The neuron as a controller framework basically takes the efficient coding ideas a step further, because efficient coding would say, well, the neuron processes its inputs to represent them in an efficient way encoded in its output.

Paul Middlebrooks
That's a transformation of information that's coming in.

Dmitri Chklovskii
Exactly.

Paul Middlebrooks
This is what deep learning is based off of. Oh, I have a representation coming in. I'm going to transform it to pass it on to the next level or to some motor output or whatever. That's really about the information transformation of what came before and what I'm helping.

Dmitri Chklovskii
Exactly. The predictive coding takes it maybe a step further. It's not just representation of the past inputs, but it's an attempt to maybe predict or encode information relevant to future inputs. Again, you're trying to predict the future inputs, and that's what you encode in your outputs.

Paul Middlebrooks
Oh, so that must have appealed to you quite a bit.

Dmitri Chklovskii
Right. I spent a lot of time working on that. I think there is some validity to those ideas. Then, what does the feedback do and how do you leverage this to get to action generation? That's where control theory comes in because then, maybe neurons' outputs don't just predict their inputs, but maybe they can influence their inputs.

Paul Middlebrooks
I'm going to ask you about this eventually, but active inference, active sensing, this is a direct analog of that, which is a more whole brain view of what brains are doing. That we have a goal of what sensory input we want and the action that we're taking is to get that sensory input. This is like the predictive brain idea as well. That's on the whole brain level. You're saying that every neuron is doing this.

Dmitri Chklovskii
Yes, and I actually, so the reason I think it seemed like a good idea to me is because when I started out in neuroscience, I understood our idea of understanding how the brain work is to transform how information gets transformed from its inputs to its outputs. Only later, because of the work by other people such as, I don't know, Ehud Ahissar and Paul Cisek, maybe others, that it's all feedback loop. It's all active sensing. That we do things, we perform actions the effect of those we can observe.

Paul Middlebrooks
What are brains for? Are they for moving? Some people say they're for moving. Some people say they're for sensing. Some people say they're for subjective awareness. Do you have that control theoretic perspective on whole brains, is essentially what I'm asking.

Dmitri Chklovskii
For the whole brains, it depends which level you want to ask it from. The evolutionary perspective, evolution maximizes the fitness. Of course, you can play with the genome and improve the fitness by tinkering with the genome, but eventually what happens, as has been argued by many people, is that genes are operating on a very slow timescale, I'm becoming a physicist again, whether you survived or not and years have passed. You want to have some feedback loops on a shorter timescale. That's why the genes in the course of evolution invented the brains because they needed some kind of stand-in to modulate feedback loops on a shorter timescale during the lifetime of each organism.

Paul Middlebrooks
They're like, "I'm too slow here, guys. Help me out while I'm in the background slowly changing." [chuckles]

Dmitri Chklovskii
I think so.

Paul Middlebrooks
We're going to go up and down scales here. Let's bring it back to the single neuron level. The conception is that every neuron is doing its own control process. Maybe describe that a little bit more. Then I know that there are problems with the common way to approach control problems. The neuron has to do too much, and that's where this idea of direct data-driven control comes in. We'll get to there, but tell us more about what the neuron is doing and what it needs to do in order to accomplish it.

Dmitri Chklovskii
It's easy to accept that the brain acts as a controller in the feedback loop with the environment. It's certainly a good view, but from my perspective, again, as a physicist, I like to focus on simpler problem first. The brain as a whole is very complex, even in *C. elegans*. How far down should we go? That, I think, is a personal choice. There is no universal level. For me, the level of neuronal physiology has been always very appealing because there is just so much data. People have worked on those so much. My postdoc in neuroscience when I made a switch was in the synaptic physiology lab of Chuck Stevens, and so this level is very appealing to me.

I thought, "Maybe on this level, I can think about these issues just as well," and actually turns out that there is a lot of data then that I can leverage. I don't think there is anything special about the neuron being a controller. I think that if the whole brain can be viewed as a controller and neuron can be viewed as a controller, they're intermediate level of brain areas nuclei that are controllers, of course. This has been famously argued by people like Robinson and such, that the eye control system-- This is well-known. There are controllers on multiple levels. Actually, going the level below, I think that you can think of each synapse being a controller.

Paul Middlebrooks
You're just wild about control, it's just everywhere?

Dmitri Chklovskii
Yes. I think evolution is wild about control, and I'm just observing that that's the case.

Paul Middlebrooks
[chuckles] What does a neuron have to do to--? It has this output. Then there's the loop. It goes to one neuron. Then it can go directly back on itself from that one neuron. It can also go to three neurons and then go back on itself. It can also eventually end up affecting motor behavior, and then it comes through the senses when we get new sensations. There's this hugely rich feedback signals that's coming into this single neuron. How does it cope with that? Then what is the neuron's objective function? What is its goal? What is it controlling?

Dmitri Chklovskii
These are all good questions, and I don't have all the answers. I should also say, in the interest of full disclosure, that the neuron as a controller is still a hypothesis. Although I think that there is some evidence that this could work, there is no smoking gun experiment that would confirm that.

Paul Middlebrooks
Why not?

Dmitri Chklovskii
Because it hasn't been done.

Paul Middlebrooks
Oh, you don't mean it's impossible, you mean it just hasn't been done?

Dmitri Chklovskii
I think it's impossible and I think there are ideas floating in the neuroscience community. One of my favorite ones to test this idea of the neuron being a controller is the following, you want to somehow perturb the feedback loop. You want to cut the feedback loop and see the consequences of that. How do you cut the loop? You could just silence the neuron and see what happens. I think it's too much of a perturbation. If you silence a neuron, it will know that something is off and it will go crazy even if it's just a feed-forward device. You want to break the feedback loop in a way that the neuron doesn't really know about.

Paul Middlebrooks
You want to nudge it.

Dmitri Chklovskii
Yes. You want to tinker in the back there so that the neuron doesn't notice it. If there is feedback, if the neuron is listening to its own output through the feedback loop through the circuit, then it will start adapting itself. One of the great ideas, I think, how to do that is just to silence not

the neuron but the synaptic transmission downstream. There are now amazing molecular genetics tools that allow you to just silence the output synaptic output of one neuron in the circuit.

Paul Middlebrooks
You want to keep it relatively simple then because I immediately think of super complex brains where the signal branches out almost infinitely, This is the classic problem in recurrent neural networks also, it's the influence of its output is actually going to come back to it. How can it possibly calculate some error-correcting mechanism or some control signal based on such a diluted feedback signal?

Dmitri Chklovskii
Exactly. It's not a given that it's done. I think on the theoretical level, it's possible, but this experiment would perhaps provide some evidence one way or another because if we were able to silence the output synapsis of one neuron in a circuit, then the spike generation of the neuron wouldn't know about that manipulation

Paul Middlebrooks
You have the neuron of interest. Then let's say it branches to three neurons, and then those neurons branch to three others, and then it comes back or something. A fairly simple circuit. Are you saying you want to silence the activity of one of its targets of the three and then see how that affects it, or the output of the neuron that you're studying?

Dmitri Chklovskii
You can do either way. I think more precise manipulation and more subtle manipulation would be just to silence the outgoing synapsis of that one neuron and to monitor how its response properties are changed over time.

Paul Middlebrooks
Then it's not affecting anything downstream?

Dmitri Chklovskii
Right. Something like maybe a hippocampal pyramidal neuron that has a place cell receptive field, and you silence its outputs. Are we going to see changes in its response? Is the place field going to go wild and start searching for greener pastures?

Paul Middlebrooks
Subtle. Do we have sensitive enough techniques to measure? I could see a scenario given the degeneracy of brains where it's we don't detect anything, but perhaps there is something going on.

Dmitri Chklovskii
There is always a possibility of a negative result. That's the nature of experimental research. If we were to see unexpected changes in receptive fields of that neuron on certain timescales, then we would say, "Aha," and then of course we can rescue this synaptic transmission and see the receptive field restored or become more sane.

Paul Middlebrooks
Or change in a different way. Yes.

Dmitri Chklovskii
Or change in a different way. I think it's important to have an experiment where you introduce a perturbation.

Paul Middlebrooks
Would you do this in a dish? Would you do this in vitro?

Dmitri Chklovskii
I one could, but why not just-

Paul Middlebrooks
One in a mouse, or some organism, or *C. elegans*, perhaps. Yes.

Dmitri Chklovskii
Yes. It's many options.

Paul Middlebrooks
Then conceptually, the common conception and the thing that you address with this direct data-driven control is you take the burden off of the neuron from modeling everything that's going on in that external loop and simplify things. How does direct data-driven control simplify the neuron's task?

Dmitri Chklovskii
The way I thought about this is if it's so natural, at least for me, it's natural to think about the neuron as a controller, why hasn't anyone said that? There are a lot of smart people who thought about neurons before me, so what's going on? Knowing how control theory is taught, it's not that surprising because most of traditional control theory is model-based, which means that you start by describing the environment or the plan, as they call it in control theory, with some kind of dynamical system description, state-space model.

We assume that this model is known, and then we add feedback control to change its properties. For example, that model could be unstable. That would be bad news in any real-world application, so we add a stabilizing feedback.

Paul Middlebrooks
Got you.

Dmitri Chklovskii
That's usually a negative feedback case.

Paul Middlebrooks
Then you compute, using one of the methods of control theory, what would be the right feedback dynamics that makes an unstable system into a stable. To think that a single neuron could do the separation, which for linear systems has a closed-form solution, but it involves really complicated matrix multiplications and inversions, and how can a neuron do that? I don't even think neuron could represent the state model of a dynamical system like that. That, I think, is why people never thought of a neuron as a controller.

Paul Middlebrooks
Just asking it too much.

Dmitri Chklovskii
That's right. That's too much work to lay on a neuron. What I realized, and that's a relatively new development, that there is now another way to do control theory, which is called data-driven control, which is very much in the spirit of, I would say, data science and machine learning.

Paul Middlebrooks
Oh, yes. Listen to the data, yes.

Dmitri Chklovskii
Listen to the data where you go directly from the observations and map them onto control signals without going through constructing the model. Of course, you have to learn that mapping somehow, but again, that mapping is learned based on the prior history of the paired observations and control signals.

Paul Middlebrooks
You have all of the data incoming. The cell's job is to map that incoming data onto its set desired reference signal, the control signal. The neuron's job is to change the control signal to produce eventual re-inputs that better map the inputs to the control signal.

Dmitri Chklovskii
Yes. I would slightly rephrase it by saying-

Paul Middlebrooks
Please.

Dmitri Chklovskii
-that the neuron produces the control signals based on its inputs that drive the environment towards a desired state.

Paul Middlebrooks
Good. It's the neuron's job to change the environment so that it's getting the right input. It knows it's changed the environment in a certain way based on the inputs it's receiving from its activity.

Dmitri Chklovskii
Exactly.

Paul Middlebrooks
I'm just going to jump to this, and I'm not sure that there's an answer. I had Henry Yin on. One of his big things is that the control theory perspective in brains in general has it wrong in a neuroscience because the reference signal is always outside the brain. How does the neuron get that objective? How does it get its reference signal? Where does that come from? Is that from the DNA? How does it know what it wants to hear?

Dmitri Chklovskii
That's a very good question. I'm afraid I don't have full answers for that. I think this is really important questions. At the moment, we have two examples where I think we understand partially how we can do it. One example is, I already mentioned, the issue of stability. Once you have a feedback loop, it has to be stabilized. It cannot be unstable. Otherwise, it blows up. Much of control theory is about making an unstable plan stable. Just this desideratum already produces some predictions that I think neurons may have to respect.

Paul Middlebrooks
There are three or four different sets of experimental data that this approach explains the results of which. Is that what you're referring to?

Dmitri Chklovskii
Yes. Stability is a must and any feedback system must do that. We can already generate some predictions. Of course, that's not the only goal because it would be boring. If it just wants its stability, it will just lie down and die. That's very stable. That's not, of course, the only thing. We have to do what the brains are delegated to do by the genes, whatever that is. That needs to be figured out.

That should probably pause here. We were talking about the feedback loops and the closed loops and open loops. The interesting thing about a spiky neuron is that, most of the time, it's an open loop because if the neuron does not produce an action potential, there is no releases in the synapsis, if you ignore the mini-events. Basically, it's an open loop, so whatever input the neuron gets, it's just held there until there is--

Paul Middlebrooks
It's a closed loop. You mean it's a closed loop.

Dmitri Chklovskii
It's an open loop because nothing comes out of the neuron until there is action potential. That's where the loop is broken.

Paul Middlebrooks
That's where the loop is closed and therefore open, right?

Dmitri Chklovskii
Okay. [chuckles]

Paul Middlebrooks
Wait, I'm thinking of a wiring circuit. When you close a switch, then it's an open loop, right?

Dmitri Chklovskii
Yes. Control theorists would say a closed loop.

Paul Middlebrooks
Oh, okay. I'm sorry. My mistake. It's the terminology.

Dmitri Chklovskii
No. It's good to clarify those things because sometimes, colloquial meanings, they collide with--

Paul Middlebrooks
Oh, so if the switch is open, the circuit can't run?

Dmitri Chklovskii
That's right. When the neuron is silent, there is no action potential, the loop is open.

Paul Middlebrooks
What a rookie-- I even know that and I just switched them. I'm sorry.

Dmitri Chklovskii
No. It's good to clarify. At that one millisecond, the extent of the action potential, the loop gets closed.

Paul Middlebrooks
Closed.

Dmitri Chklovskii
That's when you have the feedback. That's a spike in neuron. It's very interesting that it goes between those two states because if I think about the neuron as a data-driven controller, it has to solve two tasks. It has to solve a control task, whatever the desideratum is, but it also has to solve the systems identification task, which is how, in the traditional control, the model of the plan can be degenerated.

Paul Middlebrooks
What is that task? Is that task to say when is it open, when is it closed?

Dmitri Chklovskii
In the traditional control theory, we start with the model of the plan and then figure out how to get it to the state we want it to be in. Then someone may ask, what if you don't know what the right model of the plan is? Which is what happens for a neuron. Of course, it's not reasonable to expect that the genes program the neuron in exactly the right way how the environment is because it may be changing over time and the neuron has to adapt. The control theorists would say, "Wait, but that's a different problem. It's not my department."

To get the model, you have to use another subfield which is called systems identification where you use prior observations, perhaps along with the record of control signals that accompany them, to build a model so the data-driven controller performs both of those tasks at once without constructing the model explicitly.

Paul Middlebrooks
I see.

Dmitri Chklovskii
The issue why spiking neuron is such a great idea, I think, is that it's hard to do systems identification in closed loop.

Paul Middlebrooks
Oh, okay. I see. I'm just going to restate this in a very layman's way, I suppose. It's almost like it doesn't want to close the loop much because closing the loop makes it more difficult to adjust its control signal.

Dmitri Chklovskii
Exactly. Think about it as a radar that has to shut down its amplifiers the moment that it generates the pulse.

Paul Middlebrooks
Yes, or like when I'm yelling, I can't hear someone talking to me.

Dmitri Chklovskii
Exactly.

Paul Middlebrooks
There we go. I got the layman's there a little bit more. [chuckles] Talk just a little bit about how this approach accounts for some of the experimental findings in neuroscience that have been accounted for individually in various ways, but then there's a collection of things that you've been searching for that now you say that this approach accounts for.

Dmitri Chklovskii
There are several things that we think fit the control theory view. None of them is, of course, a proof. In general, in terms of scientific methodology, it's usually impossible to prove that the theory is correct. You can only falsify it.

Paul Middlebrooks
I want to ask you about that next, though, how you would falsify it.

Dmitri Chklovskii
There is some evidence, I think, that supports this view. I already cited, of course, the ideas of action generation and the existence of feedback loops in the brain as supporting views. I think that what was interesting is that we were able to derive learning rules from the idea of a stabilizing controller that can potentially map on something like Spike-timing-dependent plasticity.

Paul Middlebrooks
Is this something where you thought, "Here's the conceptual idea. Now I need to find some data that supports that idea, and here are some open questions," or how did you stumble upon applying this to Spike-timing-dependent plasticity?

Dmitri Chklovskii
I did not know beforehand that this would account for Spike-timing-dependent plasticity. This is one of the neuroscience facts that I'm obsessed about because it seems so counterintuitive. Of course, maybe just as a matter of given background, Spike-timing-dependent plasticity (STDP) is a learning rule where the synaptic strength changes depending on the relative timing of the spikes of the postsynaptic and the presynaptic neurons. If the presynaptic spike precedes the postsynaptic spike, then the synapse gets stronger.

Paul Middlebrooks
Precedes it by a very short time period.

Dmitri Chklovskii
Exactly. That's key. It's a very time-sensitive mechanism. One can say that this part, the potentiation window of STDP, maybe viewed as an extension of the Hebb postulate. If one neuron repeatedly causes the spiking of another neuron, then the synapse gets stronger. It makes sense. It represents some kind of a causal interaction. There is another side to STDP, which is the depression window, which is if the presynaptic spike follows the postsynaptic spike in time by a very small time delay, as you said, then the synapse gets weaker.

This part is very difficult to see. It certainly doesn't follow from Hebb because, of course, people would say, "Yes, you can't just facilitate the synapse. There has to be some normalization." Why does that normalization have to be so time-sensitive? There's no reason for that. It's just an issue of stability. There's no reason. I'm really obsessed by this observation because it has an appearance of an anti-causal interaction. That the presynaptic spike that has no way of influencing that postsynaptic spike has already happened affects the synapse, which is a causal mechanism. Why would that be?

In a control theory view, because you have a feedback loop, it offers an immediate resolution that because there is a feedback loop, this interaction that seems anti-causal can be viewed as causal. You just have to traverse the loop outside of the neuron. You go from the postsynaptic neuron back to the presynaptic side along that loop. That is, of course, a perfectly valid causal interaction.

Paul Middlebrooks
How does the neuron know what time window to pay attention to? It has this spike timing component. Then it's going to say, "If I get this feedback signal at this time, it means I need to adjust in a certain way, but 100 milliseconds after that, I need to adjust in a different way. This is how I know." I guess that's where the data-driven approach solves that problem.

Dmitri Chklovskii
You're not asking mechanistically how does a synapse keep track of the spike?

Paul Middlebrooks
No. There's the problem of the timescale window of, when do I pay attention to in order to adjust my signal?

Dmitri Chklovskii
Right. That is, of course, crucial. That comes to the issue that you already brought up, what the goal of that neuron is. If you assume that the neuron has to care about the stability of very short loops, basically one or two synapses, to traverse the whole loop, then the timescales are of order of milliseconds, which is close to the Spike-timing-dependent plasticity timescale.

Of course, there are other loops there, both through the brain but also going through the environment, that take longer to traverse. Then the prediction would be that if the neuron cared about the spike alignment based on the traversing of those loops, then the window of Spike-timing-dependent plasticity would be much longer. Amazingly, there are cases when this has been observed physiologically.

There are examples. One is in the cerebellum. It's the work of Jennifer Raymond where they see the STDP window, which is around 100 milliseconds too late. Then another example is in the hippocampus. I don't remember the exact, but it's tens of millisecond window. The control theorist view would be that then maybe there is a feedback loop that takes 100 milliseconds to traverse that tunes the synapse.

Paul Middlebrooks
We talked about a handful of things that the direct data-driven control feedback control accounts for. One of the ones that I enjoyed is that this approach requires stochasticity or variability in the firing rates to sample the space. I don't know if manifold's the right word, but to carve out a plane in the dimension so it knows where it can go in the dimension. Maybe you can simplify that. That was very abstract, what I just said.

Dmitri Chklovskii
Yes, absolutely. This harks back to the old exploration-exploitation tradeoff that feels like reinforcement learning are built on. Here it has a very simple mathematical formulation, which is if the dynamics of the environment of the plan is linear, then when the control law is settled and is fixed, also linear usually, so the control signal is linear function of its inputs, but a fixed linear function of its inputs, then the whole system, the whole feedback loop, ceases to explore the dynamical state space.

Paul Middlebrooks
That removes its possible readjusting within that state space.

Dmitri Chklovskii
Exactly. This has been realized by control theorists. The foundational development that allowed the development of data-driven control is avoiding this problem by maintaining what they call persistency of excitation, which is having the control sufficiently diverse so that it can still probe the environment. This goes back to what we've already mentioned, which is that the data-driven controller is called the controller, but it also has to solve the systems identification problem at the same time. By converging on a fixed control law, you give up the ability to solve the systems identification problem.

Paul Middlebrooks
In this conception, is that why we have variability in spiking? There's lots of different theories on stochasticity, quantum indeterminacy, sampling

exploration versus exploitation, like you mentioned, but in this perspective, it is a desired engineering principle, essentially, that you'd want to build in.

Dmitri Chklovskii
That's right. That tells you that for this data-driven controller to operate, there has to be some source of variability.

Paul Middlebrooks
Yes. Noise is classically an engineer's nightmare, right?

Dmitri Chklovskii
That's right.

Paul Middlebrooks
It's something that you want to avoid, but in this case, you want to build it in.

Dmitri Chklovskii
Exactly. That, I think, is a natural fit to biology because everything in the brain is very noisy, almost everything. Whichever level you use, synaptic physiology tells us that the probability of synaptic transmission per presynaptic spike is low. It could be-

Dmitri Chklovskii
10%, right?

Dmitri Chklovskii
-for central synapsis. Then the spike generation, in theory, could be very precise, but that is very strongly driven by the exact inputs.

Paul Middlebrooks
We talked about the history of neuroscience and the history of conceiving of single neurons, back to McCulloch and Pitts and even before that. Then you had, with Barlow and folks like that, the neuron doctrine, which, among other things, gave a lot of import into the function of a single neuron. I mentioned grandmother cells or Jennifer Aniston cells. Every single neuron is representing something. That's, in a nutshell, one version of the neuron doctrine.

These days, with the advent of recording lots and lots of neurons at one time, even more than 10, like you alluded to earlier, now we're in the population doctrine era where a lot of people think that abstracting at the single neuron level is a level too low, and we need to think of everything in terms of population of activity, and the dynamics that those populations give rise to.

What you said earlier would make me think that you're okay with that because you see every scale from single neuron up to whole brains in the control theoretic perspective. The population doctrine folks might say, "We don't need to worry too much about the function of single neurons because they are cogs in this larger machine,-" sorry about that analogy. That we don't need to worry about their implementation of how they're doing things and why they're doing them that way because we just need to pay attention to the population. Where do you sit in that because this is very much at the single neuron level?

Dmitri Chklovskii
I agree with everything you said. It's important to understand how the population of neurons function, but I don't see that as necessarily very distinct from what a single neuron does. As I said, you can think of each synapse being a controller, and somehow together they combine into a neuron and doing something useful. The reason I'm focused on the neuron as a controller is just because there's a lot of data and I know how to conceptualize it at this point.

Paul Middlebrooks
I see.

Dmitri Chklovskii
I know what the inputs, the outputs are and how to describe that and so on. That's why I'm doing it. Again, you can think of a brain nucleus or cortical area or brainstem part, that being a controller made out of multiple neurons. I think the similar methods would work on different levels. The reason million-dollar question here that I think you're alluding to is how to go between levels.

Paul Middlebrooks
Oh, God, yes. That's the dream.

Dmitri Chklovskii
That's the dream. Of course, I don't have an answer here. I think this is probably the most fascinating question in neuroscience. I already brought up intersection of physics, engineering, biology. Here we get, I think, an intersection with things like game theory and economics and mechanism design where you have independently operating agents.

Paul Middlebrooks
Oh, you said agents. We're going to move on to that in a second. Go ahead.

Dmitri Chklovskii
Yes. Why do neurons have to be agents? Because there is no central authority that tells each neuron what to do. The brain is decentralized. They have to pursue their own objective, their own desiderata, but they have to work together towards a common objective because, in the end, the brain has to produce one behavior.

Paul Middlebrooks
They're individual agents, every single neuron, and they're in a symbiotic agreement with all the other agents. Do you think of it that way, or do you think that an individual neuron is just selfishly doing its own thing surviving, and then it happens through development, evolution, that they've come together, and our own whole person agency is an emergent property of those individual components doing their own thing? Is that how you conceive of it?

Dmitri Chklovskii
Yes, that's my way of thinking about it. I think a valid analogy is market economy. That you have a bunch of agents that pursue their own objectives, yet if you set up the system correctly, and that's where mechanism design comes in, you have to have some kind of rules by which even though each individual pursues its own objective, they act, as you said, symbiotically together to generate some common good towards achievement of some common goal.

Paul Middlebrooks
This is a silly question, but has this changed your perspective on your respect for the single cells because they're each individual alive components? We tend to think of the brain as being composed of these things that are doing things in the service of us as a whole person. This way of viewing it almost gives a little more reverence for the individual neuron.

Dmitri Chklovskii
Yes, it does. Usually, multicellular organisms, deep inside each cell, there is the same DNA.

Paul Middlebrooks
Mostly the same.

Dmitri Chklovskii
Mostly the same, yes. There is a shared rule book.

Paul Middlebrooks
Oh, brothers and sisters around me.

Dmitri Chklovskii
Exactly, they have to follow. The genes, again, cannot control each spike that the neuron makes, but the genes write the rules by which each neuron generates its spike.

Paul Middlebrooks
Mitya, why does a computer scientist working on AI need to pay attention to any of this?

Dmitri Chklovskii
They don't have to. As you said, if you measure in billions of dollars, they're doing really well.

Paul Middlebrooks
[chuckles]

Dmitri Chklovskii
It's a tough sell, but I think that the question is, where do we want to get to eventually? If the goal is to achieve AGI, as we mentioned before-- and I define AGI as equaling or exceeding human intelligence.

Paul Middlebrooks
What the hell does that mean?

Dmitri Chklovskii
Whatever that means. My argument is that maybe there are various paths to AGI, but taking the brain-inspired approach is the only path for which we have the existence proof.

Paul Middlebrooks
Right. Could be a better way of doing it.

Dmitri Chklovskii
Maybe, I don't know, but it may also be a dead end. When they talk about building data centers close to a nuclear power station because of the energy demands and a human brain operates just fine consuming 20 watts, then you start asking questions, is this really the best path?

Paul Middlebrooks
What are you hung up on? Are there limits or obstacles in your way? Are there limits to the control theoretic perspective of neurons, and/or what are you spending all your time thinking about these days?

Dmitri Chklovskii
The big question for us aside from how do you combine many controllers together to do something useful that we're already--

Paul Middlebrooks
You should be able to just put them together. Given their objectives, it should be an emergent property, right?

Dmitri Chklovskii
Exactly, but we don't know what the objectives really are.

Paul Middlebrooks
That's a big--

Dmitri Chklovskii
That is real difficulty for us. We are, of course, trying to find some objectives to work with. While we're searching for objectives, we realized that there is a way to skirt this problem a little bit by working on the sensory periphery again, or early sensory processing, which is what efficient coding people got so much mileage out of. If you think about a retinal neuron, like a retinal ganglion cell, it's not very useful to think about it as a controller.

On some level, it is a controller because it tries to control the downstream part of the brain to perform certain actions. The question is, does it get the feedback through this whole big feedback loop, including the rest of the brain and getting back the visual signal? That's the issue because there is a big time delay. Can you even do credit assignment over this timescale? That certainly is a question there. Amazingly, the control perspective helps in this case where there is no obvious plan to control because we think that you can think of a retinal ganglion cell as being fooled into thinking that it actually controls its inputs.

Paul Middlebrooks
[chuckles] Oh my God. Talking about anthropomorphization. Good.

Dmitri Chklovskii
Exactly. That has led us to very nice re-examination of the predictive coding and the effective coding framework, which, again, because of the control view, places much more emphasis on the dynamics. We view those sensory neurons as analyzers of the dynamics of the external world. That changes the perspective for us. Even though it's not control per se, but it's a very control-inspired view of what the neuron does.

Paul Middlebrooks
This conception of single neurons as controllers is itself variable across different brain areas and depending on the needed functions cognitively of different brain areas and neurons?

Dmitri Chklovskii
Yes. That is, of course, different. If you think about a motor neuron, it's very clear that it's very heavily--

Paul Middlebrooks
Influential in the control.

Dmitri Chklovskii
Yes. It controls immediately some muscle and may get a feedback on its performance through proper sensory mechanisms or whatever. As you march back towards the sensory system, it's much more about analyzing the inputs, much more heavier involved in the systems identification side of the data-driven control and less in terms of controlling things. That, I think, view is very nicely aligned with the recent experiments where people see motor signals in the primary sensory areas like in V1. That is aligned with the perspective from control theory that the brain performs a transformation from the sensory input slowly into the control dimension.

Paul Middlebrooks
I see. This has been a lot of fun for me. Is there anything that we didn't cover that you wanted to discuss that we haven't discussed?

Dmitri Chklovskii
Let's see. No, I think this is a lot.

Paul Middlebrooks
[laughs] That's what I hear all the time about my podcast, it's a lot.

Dmitri Chklovskii
Yes. I should also say, since you asked about being a lone practitioner, it is true that, to my knowledge, we are the only ones who suggested the neuron as a controller.

Paul Middlebrooks
That's crazy. That's awesome.

Dmitri Chklovskii
There are multiple people now who think about control theory approaches in the brain. On that level, there is a small community of people to the extent that we are having a two-day workshop at COSYNE 2025, which is-

Paul Middlebrooks
Oh, nice. Oh, man.

Dmitri Chklovskii
-devoted to the dynamics and control theory applications in the brain. I don't remember the exact title, but that's what we will spend two days arguing about.

Paul Middlebrooks
Oh, cool. Maybe the last thing I'll ask you is, given that it is a small community, maybe even in the neuroscience community, when you're describing to them what you're interested in or how you think about these things, what is the most difficult for them to understand? What do people get hung up on?

Dmitri Chklovskii
I think that it depends what subfield of neuroscience the person comes from-

Paul Middlebrooks
Fair enough.

Dmitri Chklovskii
-because the motor control people, of course, the neuron does control.

Paul Middlebrooks
Yes, it's all control.

Dmitri Chklovskii
Then people who are midbrain, hippocampus, maybe you could see that if there are loops, yes. Then the sensory neuroscientist, what are you talking about?

Paul Middlebrooks
[laughs]

Dmitri Chklovskii
Yes. You have to fool the neuron into thinking it's a controller. That is a lot to take on.

Paul Middlebrooks
Mitya, perhaps we've fooled billions and billions of neurons through this conversation. [chuckles] We'll see. Thank you so much and continued success in this line of work.

Dmitri Chklovskii
Thanks, Paul, for having me. It was a pleasure.

[music]

Paul Middlebrooks
"Brain Inspired" is powered by *The Transmitter*, an online publication that aims to deliver useful information, insights, and tools to build bridges

across neuroscience and advanced research. Visit thetransmitter.org to explore the latest neuroscience news and perspectives written by journalists and scientists. If you value “Brain Inspired,” support it through Patreon to access full-length episodes, join our Discord community, and even influence who I invite to the podcast. Go to braininspired.co to learn more. The music you're hearing is *Little Wing* performed by Kyle Donovan. Thank you for your support. See you next time.

[music]

Subscribe to “Brain Inspired” to receive alerts every time a new podcast episode is released.