EPISODE 71

[INTRODUCTION]

[0:00:10.8] SC: Hello and welcome to another episode of TWiML Talk, the podcast where I interview interesting people doing interesting things in machine learning and artificial intelligence. I'm your host, Sam Charrington.

A big thanks to everyone who participated in last week's TWiML online meetup last week, and to Kevin Tee from SigOpt for presenting. You can find the slides for his presentation in the meetup Slack channel, as well as in this week's show notes. Our final meetup of the year will be held on Wednesday, December 13th. Make sure to bring your thoughts on the top machine learning and AI stories for 2017 for our discussion segment.

For the main presentation, prior TWiML Talk guest, Bruno Goncalvez will be discussing the paper Understanding Deep Learning Requires Rethinking Generalization by Chiyuan Zhang from MIT and Google Brain and others. You can find more details and register at twimlai.com/meetup. If you receive my newsletter, you already know this, but TWiML is growing and we're looking for an energetic and passionate community manager to help expand our programs.

This position can be remote, but if you happen to be in St. Louis, all the better. If you're interested, please reach out to me for additional details. I should mention that if you don't already get my newsletter, you are really missing out and should visit twimlai.com/newsletter to sign up.

Now, the show you're about to hear is part of our Strange Loop 2017 series, brought to you by our friends at Nexosis. Nexosis is a company focused on making machine learning more easily accessible to enterprise developers. Their machine learning API meets developers where they're at, regardless of their mastery of data science, so they can start coding up predictive applications immediately and in their preferred programming language.

It's as simple as loading your data and selecting the type of problem you want to solve. Their automated platform trains and selects the best model fit for your data, and then outputs predictions.

To learn more about Nexosis, be sure to check out the first episode in this series at twimlai.com/talk/69 where I speak with co-founders Ryan Sevey and Jason Montgomery. Be sure to also get your free Nexosis API key and discover how to start leveraging machine learning in your next project at nexosis.com/twiml.

In this episode, I speak with Matthew Taylor, Open Source Manager at Numenta. You might remember hearing a bit about Numenta from an interview I did with Francisco Webber of cortical.io on TWiML Talk number 10, a show which remains the most popular show of the podcast to date. Numenta is basically trying to reverse engineer the neocortex and use what the learn to develop a neocortical theory for biological and machine intelligence that they call Hierarchical Temporal Memory. Matt joined me at the conference to discuss his talk, the biological path towards strong AI.

In our conversation, we discussed the basics of HTM, its biological inspiration and how it differs from traditional neuro-networks, including deep learning. This is a nerd alert show. After you listen, I would also encourage you to check out the conversation with Francisco which we'll link too in the show notes.

Now, on to the show.

[INTERVIEW]

[0:03:45.2] SC: Hey everyone, I am here at the Strange Loop Conference in St. Louis. I am joined by Matt Taylor, who is an Open Source Community Manager at Numenta. I am super excited to have you here with me, Matt. You just delivered a talk here at the conference and I'm looking forward to diving into that. But before we go anywhere else, welcome.

[0:04:06.9] MT: Thank you. It's a pleasure to be here.

[0:04:08.5] SC: Pleasure to have you on the show. Why don't we get started by having you tell us a little bit about your background how you got into machine learning and AI.

[0:04:17.6] MT: Yeah. I don't know how far back to go. But I mean, in computers I got interested in computers when I was enlisted in the air force. I was an intelligence analyst in the air force. Then that turned into a department defense job in the same place, and I was doing a lot of simulation like air defense simulations in like Fortran and shell strips. It was our cake, but it was a very powerful simulation. That's where it got me into programming.

I didn't really think much about artificial intelligence until I read *On Intelligence,* which is a book that our founder Jeff Hawkins wrote, I think in 2005. I was working in the software industry at that point. I got away from the air defense industry and moved here to St. Louis to work in software after I got my software degree.

I just consulted around St. Louis, going to a bunch of different places and did a bunch of different jobs. I read that book. I remember reading it *On Intelligence,* and another book called *The Singularity is Near* by Ray Kurzweil. Reading those two books at the same time really flipped the script for me. It made me start wondering all these big questions, like what is consciousness, what is intelligence, how do we even define these things? Is it really possible that we can build intelligent systems out of non-biological materials?

At the time, I was just working here doing mundane software programming and stuff, but I don't have a math degree, I didn't have any experience in deep learning or artificial neural networks. At that point, deep learning wasn't even a big deal yet.

I just gave it up as a pipe dream. But at some point, I got a job at Yahoo as a frontend engineer, which is odd because I had never done frontend engineer before. But I got a job at Yahoo and moved out to the Bay Area. Worked there for a couple years and out of a blue I got a call from a recruiter for a frontend position at Numenta. I was like, "What? Okay. Sure."

I jumped onboard at that point and started doing web stuff. Eventually moved up to do web services. I was the manager of web services. When my boss Jeff decided he wanted to take all of their algorithms open source, I was like, "I want to help with that." Because I like open source. I've always been an advocate of open source and been a part of different communities. I was like, "Sign me up." So I did.

[0:06:46.1] SC: That's awesome. That's fantastic. Maybe for folks that aren't familiar with Numenta, you can walk us through the company and its position in the machine learning space. Because I think the company has a unique approach to machine learning and folks that have been around with the podcast for a while and listened to Francisco Webber's podcast might recall Numenta and Jeff Hawkins were coming up in that context, because the work that Cortical is doing is related to what Numenta is doing.

[0:07:21.7] MT: Yeah. That podcast, I think was a great primer for us, for Numenta. Of course, they're a partner of ours and Francisco is a brilliant guy. He's knows exactly what he's talking about.

Our mission at Numenta is different, I think with most companies. It's always been this ever since that I've been at the company. It's two things; understand how intelligence works in the neocortex, the second thing is implement those things outside of biological systems, like try and build a – but basically reverse engineer the neocortex is our mission. Hopefully we'll make money off at some point. Honestly, we're really R&D focused right now. Very small company, very focused on the research.

[0:08:05.8] SC: Is it primarily just funded by Jeff?

[0:08:09.6] MT: Yeah. It's privately funded –

[0:08:11.9] SC: Jeff and Donna.

[0:08:12.9] MT: Yeah, a group of contributors that had been long time associates of Jeff and Donna. They built Palm and Handspring, and so there is a crew of board members that I think help with the funding, but I don't know the details of all that.

[0:08:25.5] SC: Is the implication of that mission though that the company is not under your traditional venture commercialization pressures? Is it better to think of Numenta as like an open AI than machine learning company X?

[0:08:40.1] MT: I guess. I've never thought of it in comparison to open AI, but I guess it would be similar in that we're not building products, we're not selling services. What we're doing is we're trying to make discoveries. All of our discoveries are based on neuroscience research. Our research engineers are always reading the most recent neuroscience papers that comes out, they're interacting with different neurosciences in the community, trying to answer questions that are relevant to how we understand intelligence in the cortex., What we do is as we make these discoveries and we test them out, we prototype them in software. It's like, "Oh, this is how it works." It actually does. Our theories seems to work in software the way we thought.

Then we will create patents around those discoveries. Very specific ones about things that we've discovered about how the brain is working and how we've implemented it and currently in software, but it could be implemented in hardware too.

The idea being the monetization strategy is in the value of the IP itself. We don't want to be distracted by consulting, by providing services or by creating applications at this point. We really want to focus on the discovery, on the brain trying to figure out how it works, and we think that good things will come of that.

[0:09:58.9] SC: Okay. For your talk, one of the big things that I think you talked about at least from the perspective – from what I got out of the abstract was you premised it on, "Hey, there's a lot of excitement out there about neural nets and deep learning and things like that, but these are all based on a model of a neuron that is rather outdated." I presume you then walk through some of the new things that we've learned since then. Can you us walk us through your talk and the ideas that you wanted to share with folks?

[0:10:32.9] MT: Sure. I'd like to say for stuff that I don't have anything bad to say about artificial neural networks or deep learning. I think that that's necessary technology that we needed to build. But one of my main points is that – that it's not going to naturally evolve into what people call strong AI. The first thing I say in my talk is weak AI is not intelligent and won't become intelligent. There's not going to be some – this like exponential growth, and suddenly dissentient.

There is some core things about the AI and endpoint neuron models, specifically that don't have the capacity for intelligence as we understand it. Those are –

[0:11:14.0] SC: What are some of those core things? Let's just dive right in.

[0:11:16.8] MT: There is two main things. One is that the neuron needs to have three states, and current neurons have two states, active or not. We add the idea of a predictive state. The neuron goes into a predictive state to indicate that it thinks based upon the context of its input that its going to be active soon.

That prediction is core to everything about our theory. We take that from understanding how the brain works. Your brain is constantly making predictions about what it's going to see next, what it's going to feel next, all the time. You can see that by investigating this depolarized programmable neurons; in the neuroscience they call these cells depolarized, which means they're primed to fire. We're missing that in the AI and neuron model. There's no concept about it. There is that.

The other thing is pyramidal neurons have different integration zones. They don't just have one group of connections to other neurons. .They've got apical dendrites that provide feedback from layers that are either above it or different parts of the cortex. There is distal, a distal zone that's lateral. That's getting connections from – it could be from another layer, it could be from the layer itself, but that provides context.

These both provide context for the proximal input. The proximal input is really the driver input that's typically coming from the direction of the senses. That's like the sensory input that we need to understand, we need to process. The pyramidal neurons do that in the context of these other zones and the context of distal input and apical input. Those are the two things I think we're really missing from that point neuron model.

[0:13:02.4] SC: I get that the neuroscience research has identified these things in human biology, but it's not clear to me how we've demonstrated that those are required for intelligence, or even that those things can't be approximated with artificial neural networks as we currently know them. Like the last thing, the different zones. It made me think of, "Well, we just have different inputs in different ways." Then as far as predictions are concerned, if we're able to predict that a network level, who is to say that the neuron itself has to have that predictive state in order to create intelligence?

[0:13:41.2] MT: Well, it's true that current artificial neural networks and deep learning could potentially put together models that replicate the parts of the things about the neuron that we're saying are required for intelligence. I think it's possible.

[0:13:53.6] SC: I mean, we use them for prediction all the time.

[0:13:56.3] MT: Yeah. But I don't know that that's – It doesn't feel natural to me. Think about this, recently there has been this big discussion in the deep learning community about back propagation, because Geoff Hinton has recently said, "Let's give up on back propagation. Go back to the drawing board and try and figure out what's really going on."

We did that 12 years ago, so we never tried back propagation. We've always tried to do this, but we don't see back propagation happening in the brain. For the longest time, Hinton and Bengio were insisting that back propagation is happening in the brain. We just don't see it.

That's a move in our direction. Even from the deep mind crew they recently had this blog post about how important neuroscience is to contributing to artificial intelligence. It feels to me like the community is starting to move in our direction. Maybe they will be able to hack these properties that we're saying we need in the neuron model into deep learning systems. That could happen, but I don't think that it will happen without them doing something to incorporate those ideas.

[0:15:01.4] SC: Bengio just this week published a paper that talked about – I forget the exact title. Something about consciousness. I don't know if you saw that.

[0:15:09.1] MT: I did not see that notes.

[0:15:11.7] SC: It was controversial might be strong, but it raised a lot of questions, because he proposed that somehow we need to take into account some notion of consciousness in our models, but the paper didn't present any experimental results or whatever. It was just like a prod to the community.

[0:15:29.5] MT: Anything about consciousness is going to be controversial, because what is consciousness, Sam?

[0:15:34.1] SC: All right. What is intelligence?

[0:15:35.6] MT: Right, exactly. That's where I started my talk off with was asking people in the audience, who believes humans are intelligent? They all raised their hands. Who believes chimpanzees are intelligent? I just go down the evolutionary ladder and see hands going down. By the end of it I'm asking, who thinks plants are intelligent? There's still one or two people that think plants are intelligent. They may be right. We don't know. Paramecium, whatever.

There's a lot of disagreement. The thing is everybody believes humans are intelligent. At least we can start with that. I think by association we can include most primates in that too, because they have the same neocortical structure that we have. We focus that on that neuroscience, on what we think – what we all know is intelligent and that's the neocortex of the mammalian brain.

[0:16:23.4] SC: You started off talking about, like level setting on intelligence and just how open-ended that is, then talked about the evolution of the neuron. How do you get from there to systems?

[0:16:35.3] MT: Okay. You think about the pyramidal neuron as I said as these integration zones. It's hard to visualize without a picture.

[0:16:43.8] SC: Francisco said the same thing. You did a good job.

[0:16:46.4] MT: I know that he is. But my talk will be online at some point, so he could find Matt Taylor's talk in Strange Loop. But I got a bunch of drawings and stuff. If you look at a pyramidal neuron and it's got –

[0:16:56.9] SC: We'll link to it if you shoot us a link.

[0:16:58.6] MT: Okay. It has these integration zones; distal, which is lateral to the side, proximal which comes from below, apical which comes from on top. The cortex has this homogeneous structure. If you took your neocortex and you unwrinkled it and unfolded it and flattened it all out, it's a sheet of cells. It's about the size of a dinner napkin, about the thickness of a dinner napkin.

It's homogeneous throughout. It has the same structure. They're this computational unit in the cortex called the cortical column. This is something that is more recent of a neuroscience discovery. We've known for a 100 years that the cortex had layers. Like there is this distinct little layers in the sheet, and that their structure was different enough that we thought, "Well, they're doing different things. Not exactly sure what they're doing."

Now that we know they're not just layers. There's also columns. We can take that each column and say, "Okay, that each one of these is some individual computational unit. Maybe they can share their computation, or the output of their computations with their neighbors and stuff.

This idea that a column can have layers within it, and every layer is full of these pyramidal neurons. Okay, so imagine a column that's cut up into layers and this is a cylindrical column, cut up in layers. Each one of those layers is full of pyramidal neurons that had these integration zones; apical, up and down, to the north sort of and proximal to the south and distal to the side.

Each layer itself has the same integration zone properties as an individual neuron, because they're all oriented in exactly the same way. You can treat that layer as a computational here. A layer gets proximal input, a bunch of proximal input and all gets piped into its neurons in different ways. From some space that's representing generally some spatial sensory features changing over time, or something like that.

You can think of the layer itself as a computational unit. Depending on where it gets its proximal input, where it gets its distal input and its apical input, it does different things. Also, there is a bunch of different layers in the cortex, somewhere between six and 10, depending on which neuroscience you talk to. But each one of those layers is structured a little bit differently too. There's some minor deviation in the organization or those pyramidal neurons within the layers. It also give them a little bit of different computational aspects.

[0:19:23.7] SC: Organization in what sense?

[0:19:25.5] MT: For example, we have these algorithms that we're saying are happening in these layers. One is called a special pulling algorithm that takes some input and spreads it, normalizes it while retaining the semantics of the input. These create these mini-column structures of neurons, and some layers have this. Typically the distal connections from each one of those neurons as its perceiving proximal input, they start connecting to each other over time.

When you take that distal input to a layer and you say, "Okay, we're not going to get that distal input from somewhere else." We're going to have all of the pyramidal neurons within the layer give each other distal input.

What you're doing is just naturally creating a temporal context, because your only context to some input is what state you've been in the past, then that's the temporal context. If you're getting that input from somewhere else, who knows? That context could mean any number of things. But if you're just giving yourself context, that's – you're looking at your own paths. That layer has context of its own history when you leave them back to itself.

That's one of the core things that we discovered. We call this the temporal memory algorithm and it relies on these little mini-column structures that takes that input, the bits of input that are coming in from some sensory organ or perhaps from another part of the cortex, normalizes it into these column activations and then activates cells within each column based upon the distal context that it's going.

What you get is it's starting to tie sequences together. When you see a pattern repeating over and over, over and over, you get these distal connections that are being reinforced, because they see the pattern and the distal connection will create that connection to the active cells that it just saw that represented the previous spatial input. Then we get another input and there may be a prediction. I saw that last time, I'm going to be next. It makes a prediction. If it's right, and the next input activates the column that that cell is in, then it becomes active. It was a correct prediction.

[0:21:31.4] SC: The context you're creating for me is how I felt when Francisco was explaining some of the stuff for me. I was like, "Wow."

[0:21:31.4] MT: It's a lot easier with visuals. Hence, that's why I created these bunch of videos on our YouTube channel to try and explain it all visually.

[0:21:47.4] SC: You explain the micro structure, then the macro structure. Then what's next? Strange Loop is a developer conference. How do you get from there to, okay how do I build something?

[0:21:59.4] MT: Well, there's two questions there, I guess. Strange Loop is a developer conference. However, it's also like a weird conference.

[0:22:06.1] SC: Self-granting. Granted it is. Eclectic.

[0:22:10.4] MT: Yes, it's very eclectic. You can get in, if you have something that's like on the fringe but very interesting, you can get in and talk there. I think that's why they kept this talk.

[0:22:21.1] SC: I mean, in addition to developing IP and all of that, as I understand it, Numenta is a company offers tools that allow people to actually use these stuff. Is that correct, or no?

[0:22:34.2] MT: Open source. All of our code is open source and anybody can try and use it if they want to. I've created a lot of tutorials and code samples, and I try and make it as approachable as possible for our community. We've got a very active forum with lots of discussions about the theory and about the code and that stuff.

[0:22:53.6] SC: As a user of these open source tools and things like – do I need to think about columns and dendrites and all of that stuff, or am I thinking about other representations?

[0:23:06.2] MT: It could go either way. It depends on what you're trying to do. We have a pretty diverse and eclectic community. They're interested in this. Typically people who are really interested in how the brain works, or – yeah, I could say they can be a little off. But I mean, they're always very smart and inquisitive and curious.

It amazes me, the types of things that people try and do with our stuff, and I always encourage it. I'm always like, "Yeah, try it. Give it a try. Who knows? We don't know what's going to happen." Our software that we open source is called NuPIC, the Numenta Platform for Intelligent Computing.

We just released 1.0 of that a few months ago. That includes up to what I just talked about, the temporal memory part of it. A few years back after we went through this research cycle and made the temporal memory discovery. That was a big discovery for us to see how sequences were memorized in the brain, in the cortex.

We just dumped it all open source and we started building these potential sample – We just brainstormed by what could we make with this that people might want to use, and we made all these sample applications. There is one that was like rogue human behavior detection, which is something you can install on a computer and it monitors the different metrics that are coming out of the computer over days and weeks and can give an indication about a user's behavior.

Are they behaving oddly or differently based on the time of day and the thing that they're doing and the metrics that are coming out of here. That's a sort of thing you can do.

We also had a IT analytics program that hooked up to AWS. We actually licensed that to another company called Grok. They are actively selling that to IT companies that have a bunch of service on Amazon. It will automatically through CloudWatch connect to all the different metrics coming out of your servers and it will create models for all of them and they'll just streaming the data into them, and you don't really have to do anything. They're all preconfigured.

Then it will give you anomaly indications over time. After it's seeing that server data for a while, it gets an idea of what's normal and what's not normal, then it notifies that something is wrong with the server. It doesn't know what's wrong with the server, but it can tell you that something abnormal is happening. Even with this server and this server, and the combination of those. Anywhere that there is streaming analytics that you need anomaly detection, I think that there is a potential application for what we have right now with NuPIC 1.0.

There is also this really interesting thing that I think is still a big opportunity for people who want to try and build something novel with this. We figured out a way to encode geospatial location into a format – When Francisco and you talked, you talked a lot about STRs, about sparse texture representations of corticals, it's all about. They come in semantic fingerprints.

We found a way to encode location information, like latitude, longitude, altitude into an STR, so we can take something that moves through time and space and give the algorithms, the intelligence algorithms a way to understand the patterns in the movement of that object.

For an example that I always do, is I go walk my dogs on the same dog walking route every day. If I take a tracker with me and then I go put all my points back through the algorithm, the first time it sees the walk, it's like all anonymous. It doesn't think – none of it is familiar, because it's brand new. The second time I do it, it's a little bit less familiar. The third time I do it, it's like no big deal. This is normal.

As soon as I deviate from the path that I've taken and even if I just go walk on the other side of the street, or if my dog has decided they don't want to stop at that tree, they want to stop at some other tree, I get anomaly indications coming from my path. I think this has big applications in fields like logistics, air traffic control, human tracking, pet tracking, stuff like that where you've got normal routes of things that normally happen.

You don't necessarily to the tee went to and say, "Oh, if they deviate right now, or if they're not at this point at this time, there's something wrong." You just want to get an idea of their general movement, whether it's strange or not, or whether it has been seen or not, then it can do that sort of thing. That's really interesting.

[0:27:19.8] SC: Certainly for the network and server anomaly detection and the example you gave before that, they're things that you could do with a variety of different techniques. Are there things that you found that either the approach you take, because of the approach you take, just best in class, or like it's – if you need to do X, Y, Z, this is the best way to do it? Either from a complexity of creating the solution or computational costs, or some other metric, like –

[0:28:00.1] MT: Well, we wanted the same thing. But the problem we had and this was several years ago, is that there are no standard benchmarks for streaming temporal anomaly detection, or just from temporal anomaly detection. Most of the benchmark is around spatial data. Most of the machine learning techniques work on spatial data.

We didn't find anything that we could compare what we did with what – like LSTM for example has some abilities to do temporal analysis on things. It would be in sort of in batches that move along. We created a benchmark we call the Numenta Anomaly Benchmark. We've setup ours as one of them in the running and we set up analysis TNY and we set up – there is one from Twitter, there is one from Etsy, that does streaming anomaly detection, like they've got open source projects that do that sort of thing.

We've created these input data sets, things like how many calls were there in New York City over an entire period of time or something. You can look at that data and you can say, "Something weird happened there for sure." You go look it up and they're like, "Oh, there was a big game in town, or stuff like that." You can find that data.

We'd find data sets like that that had a good amount of data and had obvious anomalies that were labeled and marked, and we'd run all of these algorithms against them and score them based on how well they detected the anomaly. Waiting it, I think we waited it pretty heavily on not providing false positives, I think. I can't remember exactly, but it's open source. It's on Github. It's at numenta/nab for Numenta Anomaly Benchmarks.

We have at least. Of course we're the winner. We always win. We had this contest, we're like, "If anybody beat us at this," and somebody came and beat us at it and we're like, "Okay, we're going to fix it." We fixed it and we're like, we're beaten again. There's always some tweaking that you can do to try and get that last few percent.

[0:29:54.9] SC: It leaves me with an impression that this is a tool that we've – or a set of tools. You have a strong feeling that closely models the inner workings of the brains as we understand it, and that over time that will lead to – I'm assuming you're banking on order of magnitude capabilities over current approach. Like the things you could do using Numenta, and the things you can do using other things will diverge over time.

But today, it doesn't sound like there is a bang on the table like, if you need to do X,Y, Z, these tools will get you there a 100% faster and a 100% times cheaper, or even 10. It's an interesting approach and something that's worthwhile for people to learn and take a look at and to understand the thinking around.

[0:30:47.0] SC: What kind of a killer act. There's no cure.

[0:30:50.0] MT: I guess that's what I'm getting at. Yeah. There's no killer act. But we're patient too. There is a lot of things about the brain that we don't – we still don't understand. What we have currently in Numenta 1.0 is just temporal memory stuff. All of our other work that we're doing is in research repositories that attach on top of that.

We're taking those core algorithms, which aren't going to change and we're building new and different things with them, because the core algorithms in your brain don't change. But we discover that it can do lots of different things with those core algorithms. We're building structures now, because we think we understand how sensory motor integration happens with sensory input and movement. But it is the integration of two layers in one of those columns. Remember I told you about the layers having these integration zones?

We could have one layer that is running the same temporal memory algorithm that I described earlier with the many columns and everything, but we don't send it its own distal input. We don't give it a temporal context. We can pipe in the context, the distal connection comes from somewhere else in the brain, comes from a different layer down.

If we assume that that layer, or the output of that layer is providing us with location information associated with a sensory input, that's proximal coming up to the layer from the bottom. That's the driver signal is the sensory input. The distal signal is going to represent the object of being touched and what location on the object that sensory feature was sensed.

Then we can have a layer that can represent every object we've ever touched and what sensory input we've felt we're on it. That layer now provides that information to another layer, which we call an output layer. This output layer has a little bit of a different structure, because it doesn't have the mini columns like the one underneath it, but it represents over time a library of every object we've ever learned.

So we can train this thing and say, "Okay, this is a coffee cup. Touch it all over the place. Okay, here is a banana. Touch it all over the place." We can build a library of objects that that top layer represents. The bottom layer is basically just going to represent all the sensory input you felt on every location, on every object that you've touched.

[0:33:08.3] SC: This is the temporal memory concept?

[0:33:08.7] MT: It's the temporal memory concept and it's not doing temporal memory anymore. It's doing sensory feature and location association. Just because we've changed the distal input, so it's no longer giving itself distal input, it's getting it from somewhere else and it does something entirely different.

[0:33:24.0] SC: It sounds like the idea there is if you think about using deep learning, object recognition, our best guess at the way the different layers work now is you've got layers and figure out edges and layers that figure out colors. When that inputs the banana, we'll get the curvy layer firing and the yellow lane firing, that kind of thing.

[0:33:46.1] MT: Yeah. There's nothing that deep learning couldn't do.

[0:33:50.8] SC: What you're describing sounds like maybe in the internals is capturing a richer representation of these various things.

[0:33:58.0] MT: At the time the big difference is our model incorporates movement. That's the big difference. Can you name anything that is intelligent that cannot move?

[0:34:09.0] SC: Nothing comes to mind.

[0:34:10.8] MT: Nobody ever does. Because there's nothing intelligent that can't move. We believe that's a core feature of intelligence, the ability to interact with your environment has to be baked in to the architecture of the intelligent system. It's not something that you can just add. You can't just add behavior to a system that you're building.

That's to be baked in to the flow of information. Like I said, when you move your finger to touch an object, you know where your finger is going to move, because you just commanded it to move there. That information is available to your brain. That loop has to be baked in, so that every time you touch something you know where it's going and you know what you expect to feel. If you don't feel that, something is wrong.

[0:34:50.8] SC: Matt is demonstrating all this with a glass of water, and we've experimenting with a video camera set up here. We may be able to show the visual aids with the motion.

[0:35:02.4] MT: It helps with the visual aids. Like I said, if you want visuals go to numenta.org. I got lots of stuff.

[0:35:09.3] SC: Nice. Nice. Awesome. Anything else that you covered in your talk, or last – final thoughts that you want to leave us with?

[0:35:16.0] MT: I want to emphasize that we have a really nice community. I'm the community manager, so of course I'm going to say that. But honestly, there's some really bright people that have even shown up just in the past year that are doing some really interesting things with HTM. All of our papers are open access, so all the stirring – everything that we theorize about, we write papers about and we put it out there and we do it with code.

We're like, "Here is a paper, here is a simulation, here is the code, you can run it yourself if you want to try and run it yourself." If you don't believe us, you can try it yourself. There is lots of people on our community that have decided they're going to write their own HTM system in their own favorite language, with their own environment.

There's a lot of people doing new and interesting things, creating their own visualizations. Last one was his thesis from this guy in Turkey. He did this amazing sensory motor simulation in a 3-D game environment where he's a got a player trying to find a point and he wrote his whole thesis on it. It's brilliant.

He used our theory, and then attached some stuff on top, like he theorized further and he was like, "Oh, what about got this and this and this?" Trying to create a more complete idea of the brain. Not just the cortex, because we're really just working on cortex right now and he's trying to incorporate some other things like real behaviors or real drivers of what is the motivation for that agent that is running the intelligence.

We're not quite there, but we're focusing our research right now on location. Like that location signal I'm telling you about. We've got a really good idea of how that location signal is generated. It's super interesting. The way that your brain graphs location of things is amazing. I don't have knowledge to explain it, but it's about grid cells, location cells and place cells and stuff like that if anybody wants to go research that. There's some really interesting neuroscience papers coming out about grid cells.

I'll give you a little example. If you put a mouse in a box and it ran around the box and you're monitoring its neurons, you'll see as it runs around the box and you trace where it goes, certain neurons will fire when it's in certain places. Those fire – you've got to identify those cells that are [inaudible 0:37:27.0].

[0:37:27.3] SC: We can whenever and then place X, Y, that specific neuron is going to fire?

[0:37:31.6] MT: Yes.

[0:37:32.7] SC: Wow.

[0:37:32.9] MT: If you look at it, it forms this hexagonal grit. There is this hexagonal pattern of neurons that are firing as you move through space representing where you're at in the space that you're occupying. We think that that interplay of neurons and that idea of neurons representing locations in space plays out at a bigger level to even represent objects in space too.

You have an allocentric representation of any object that you can imagine. Allocentric meaning not related to where you are, not egocentric, but just it's like imagine a cup. That's an object that you have. If you use its center of gravity for its center, you could define it entirely based upon all the sensory input that you've ever received about those objects that you felt or seeing or whatever.

We think that that has something to do with grid cells, how those objects are stored like the plate – how in 3-D space they're defined is linked to the sensory input that we receive about them and what cells are firing in space as we're imagining where we're touching on the object.

[0:38:42.5] SC: Wow. Super, super interesting stuff. I will definitely make a note for folks to listen to the conversation with Francisco a couple of times before this one. Or maybe this one should be the prerequisite for that one, I don't know.

[0:38:56.9] MT: Hopefully it's standalone. Hopefully it's standalone.

[0:39:00.8] SC: Awesome. Well, thanks so much Matt.

[0:39:01.6] MT: You're welcome. I appreciate the opportunity.

[0:39:03.4] SC: Absolutely.

[END OF INTERVIEW]

[0:39:08.4] SC: All right everyone, that's our show for today. Thanks so much for listening and for your continued feedback and support. For more information on Matt or any of the topics covered in this episode, head on over to twimlai.com/talk/71. To follow along with our Strange Loop 2017 series, visit twimlai.com/stloop.

Of course, you can send along your feedback or question via Twitter to @twimlai, or @samcharrington, or leave a comment right on the show notes page.

Thanks once again to Nexosis for their sponsorship of the show. Check out twimlai.com/talk/69 to hear my interview with the company founders and visit nexosis.com/twiml for more information and to try their API for free.

Thanks again for listening and catch you next time.

[END]

