LARRY ABBOTT: So I put this slide up so you could walk in and say, I came to this course to learn about high level cognition machines that do amazing things. Surely this guy's not going to talk about a fly. But I am going to talk about a fly.

And I will try in the beginning to explain to you why. And hopefully, by the end, we'll see. I won't declare victory at all, and you can tell me if it applies or whatever. The reason I'm talking about the fly is this quote, really.

Flies are in all sorts of mushroom bodies. I'm going to talk about a part of the fly brain called the mushroom body. And mushroom bodies are in all sorts of insects. And I kind of like this quote. We more like to say the mushroom body is the soul of the fly.

But what I want to point out here is this part of the quote. Flies are not as intelligent as you, but they're intelligence, much of it, comes from this small part of the brain called the mushroom body. If you have free will, they have free will.

But unlike you, we can here point to a part of the brain and say, that's where it is. And I'll try to convince you of that as we go on. And so that's why I'm talking about a fly. You can really say there's this small part of the brain called the mushroom body-- I'll show it to you in a second-- where maybe not uniquely, but certainly is a point at which intelligent behavior arises, in which something like free will, whatever makes different flies do different things on different occasions arises.

Now, not only that, again, compared to what you might know in a mammal, this is a region of the brain where all the cell types are known. There's genetic control over all the cell types. There is a very good optical level anatomy. And very soon, there'll be an EM level anatomy.

So it's a region of the brain that has all of that. And so I thought it would be a good example of where the limits of knowledge of neuroscience are really extended out. You have to put up with it's only a fly. Sure, it's kind of a stupid animal. But I'll show you some behaviors and what

this thing can do, and we'll see how it will go. And maybe just some machine learning.

So these are the people involved. This is a talk in which the fraction of the work that I did is not zero, but it's very, very small. And so much of it is done in collaboration with Richard Axel and members of his lab, of whom all of these-- you can see the names. Ann is a theory student who worked with me.

But an awful lot of the work is done at Janelia in a collaboration with Jerry Rubin's group, and in particular, Yoshi Aso did a huge amount of work here. So I feel just fortunate to be able to kind of correct the commas on the paper. That was my role in this project. So that's the people.

OK, so what is the mushroom body all about? This is a diagram of the olfactory system. I should have said at the beginning, not only is it flies, it's olfaction. Two of the most-- can you pick a more boring sense and a more boring creature? So anyway, let's give it a try.

So flies have receptors along their antenna. The don't have a nose, but that's where the olfactory receptors are. I'll show you in a schematic a little bit later. There are neurons that receive the odors, then send a signal to this structure. This is called the antennal lobe.

And at that point, it gets relayed from these set of neurons to a set of neurons called projection neurons. Those go up here, and they send their signal to the mushroom body. And the mushroom body is this kind of L-shaped thing here that I'll describe in more detail.

And then they also send axons to another region of the brain called the lateral horn. I'll come back to that. I think I'll come back to that. Maybe I should say it now so I don't forget.

So as you'll see, the mushroom body is going to be responsible for learned behaviors, is responsible for learned behaviors in the fly. And the lateral horn is responsible for innate behaviors. You'll see at the end of the talk, actually, evidence of that.

So there's a division which actually occurs in your brain too of the olfactory pathway to an innate pathway and a more flexible learned pathway. So this is a diagram. This is, again, from the Janelia work of real pictures of the stuff. It's overlaid on a fly brain.

You don't see the periphery here, but these are the antennal lobes. So this is this relay station in the fly brain. Here is one of the projection neurons you see here. This is the mushroom body. Again, this L-shaped thing.

And then it goes backwards. So that purple are the cell bodies of the mushroom body. And here you can see it, again, going to the lateral horn. This is the optic part of the fly brain for doing olfaction.

This is obviously a schematic of the stages of olfaction in the fly. These are supposed to be the receptors, so let me start with them. They're called olfactory receptor neurons. There are about 1,000 of them.

They come in around 50 types. And here the types have been drawn in colors. And what a type means is a cell that expresses a single receptor molecule. So it will bind to a set of odors, whatever that particular molecule does.

And so all of these red guys are virtually identical in their responses. All of the green guys are identical, et cetera. And there are 50 types. And I'll show you in a second, they form about a 30-dimensional representation of olfactory space. Not as high as in your nose. Not nearly as high as in a mouse's nose. But that's what you get.

OK, as I mentioned, these project to this structure, which is the antennal lobe in my little diagram. And they have the property that all of the cells of a certain type, in other words expressing a particular receptor, project to the same site, which is called a glomerulus. So you can see all the red guys go to the red one. All the purple guys to the purple one, et cetera.

So this is an incredibly precise wiring getting these 50 olfactory signals from the 50 types to a point in space, or a region in space. And that's the point at which the next cells-- so this is obviously the input layer. I'm sort of over here, giving you computer language, if you want, for all this. So at this point, you have the projection neurons pick up the signal.

There are about 200 of them, again, in the exact same 50 types, because there are a few projection neurons for each of these different glomeruli. And they send the signal onto the mushroom body, and as I mentioned, the lateral horn, although we won't talk about that whole lot until the end. And this is a one-to-one connection.

So every projection neuron-- let's say there are red type projection neurons that just pick up the red signal, send it onward. There are purple type guys-- they're not really called this-- but they accept these 50 signals, maintain them as separate pathways. So what's this thing doing from a sort of computer science point of view?

Obviously, it's pooling. So these 1,000 cells are pooling their resources into 50 glomeruli. So you're averaging and you're reducing noise. And there's also a normalization process that goes on here. There are lateral connections here that try to even out the responses so that-- let's say at a fixed concentration, one odor that causes a lot of responses in the receptors and another odor that gives much less response kind of get equaled out here, so that the strong odor doesn't overwhelm the weaker odor.

OK, so that's this stage. And I thought I'd show you some of these responses. So here are the 50. This is not all data. This is data plus extrapolation. But the data comes from a beautiful study of Hallem and Carlson.

These would be the 50 ORNs, types, so one of each type. And here are 110 orders that were tested. And the responses in firing rate color kind of look like this. You can see they've been graded here. The responses get stronger as you move from left to right. That's just the way they ordered them.

And you can see they're quite uneven. So here is a kind of weak responding odors and here are much stronger responding odors. So that's what's coming in.

Now, if you look at the PN level-- now, this is not data. This is a model. It's a model really due to Rachel Wilson and members of her lab, but also constructed by Sean Luo and Ann Kennedy in my group. And you can see the argument.

So basically, what's happened is these inputs come in and have gone through a model that reproduces what we think the PNs are doing. PNs have not been tested with this whole panel of odors. But you can see the normalization effect. You notice that the activity is spread much more equally across these odors than these odors. And that's reflected in the fact that if you measure by various ways the dimension of this representation, you get about 30. And here it goes up a little bit to 35 because of this kind of equalization effect, and also some decorrelation effect that goes on.

So there is that. So what I've described here is sort of the front end of this olfactory system. And it is completely stereotyped. It's a precise wiring. I've described it to you. It's the same in every fly.

If you look at two neurons of the same type, they look virtually identical. So this is a hard-wired system. And you would not say there's any free will or intelligence in this system. It's just

getting the signal in. And you'll see a little bit more of that later.

OK, so what about the next level? The next level is the mushroom body. So these yellow things are the mushroom body neurons. They're called Kenyon cells. There are about 2,000 of them. They come in only seven types.

So already, we sense something's happening here. There's something changing about the representation. The representation is getting much higher dimensional. It's something like 1,000 dimensional.

So there's a projection out to a high dimensional representation. And this is where the free will comes in. And in anatomical terms, the reason it does is because-- I'll try to persuade you with the data-- that this acts exactly like a random, high dimensional, hidden layer in a machine learning system.

So this guy is suddenly a new beast. Within one synapse, the system's gone from completely stereotyped to, you know, crazy. Completely random. I would say there's lots of evidence that it's different in every fly, that every one of these neurons is different. And you've completely given up the stereotypy.

So now, how do you get back to sense? Because you've built this beautiful olfactory representation here, and it's as if you've thrown it out. You've just gone crazy. And so now, I put the box around here just to remind us, this is a different beast all of a sudden. And it's a very unusual beast in the fly brain. I'll come back to that.

But now you have output. So these yellow neurons, as you'll see, do not leave the mushroom body. They don't send any signal out. They're completely intrinsic to the mushroom body. But there are neurons called mushroom body output neurons that do send the signal out.

And again, now it's a new ballgame. First of all, look at the numbers. You've gone from 2,000 neurons to 34 neurons of 21 types. You've got about a 20-dimensional representation. There's been a collapse of the representation.

So I would argue you can just see right away from this slide that this is an olfactory representation. This is an olfactory representation cleaned up a bit. This is a crazy, random olfactory representation.

This is not an olfactory representation. The dimension is lower than what you started with, so

there's no way you can represent the full thing. This is already, somehow, making a decision about olfaction. It's well on the way to a behavior.

And again, the great thing about the fly here is that you get there very quickly. If you went to Jim's talk today, I'm sure he talked to you about the long pathway in the visual systems of monkeys, in which these stages take up a good fraction of your brain. These more complicated stages do it.

And then it's very difficult to see where this transition is to decisions and things like that. Here, the transition from orderly input representation, sort of retinal-like, to IT-like, if you want, in the visual system occurs in one synapse. And then the return to a decision, a behavior, in another synapse.

It's very quick. And I would think of that in computer science terms as a readout layer. As you'll see, it's actually a layered system, but it's the readout.

OK, and this system here, it goes back to being completely stereotyped. There are very few neurons per type, if you notice. There are almost as many cell types as there are neurons. And they're the same in every animal.

So you've gone from stereotypy at the input stage, a wild and crazy random thing in the middle, and then back to stereotypic to get to the output. Which of course, you have to do, right? Your motor neurons have to go to the right muscles. You can't randomly wire your motor neurons.

And thinking occurs between those. Same thing with your retina. It has to be wired to give you the basic visual signal. But between those two extremes, that's where we do our thinking. And as I say, that you can see here, but it's in this one layer, OK?

All right, and the key is going to be exactly as in a machine learning system. As you'll see, the key to the whole system is the plasticity and modulation that occurs at that set of connections. There's no evidence that these connections, these connections, and these connections are at least very plastic. They may be modulated a little bit, but the business end of this thing, just as in many machine learning networks, is that the readout unit's being adjusted. And I will come back to that.

All right, so here, the mushroom body, it started out as it was in the fly. And as it turns, you'll see why it's called the mushroom body. Yeah, now it looks like a mushroom.

So these are the cell bodies. They receive-- you can't really see very well here, but they receive their input right under the mushroom. And then they send axons down. And these axons form the load.

So this whole thing is made out of Kenyon cells. That's the Kenyon cells all together forming this structure.

How many cells?

A couple of thousand. OK, now here you can see one of the projection neurons. Here's where it gets its input from the antenna lobe, goes up to the mushroom body, goes over to the lateral horn. And here you can see-- it's sort of hard to distinguish that neuropil from the cell bodies here, but here you can see that sort of under this layer of cell bodies, it's making its connections.

And what I want to stress here is this idea that the projection neurons occur very few cells per cell type. Now, these cells types-- I guess I'm going to get ahead of myself a little bit. But through work at Janelia Farm in particular, there have been these intersectional strategies for expressing various markers in these cells. And they've been supremely successful.

So typically, when you get a cell type in this business, it's often two cells, one on each side of the fly. They're perfect mirror images of each other. And they're identical in all flies. So that's what you mean by a cell type.

And in much of the fly, there are very few of them per-- this is per side. There will always be an even number. And you can see, there are 50 types, a couple hundred cells. That's part of the specific wiring.

Now, if you look at the Kenyon cells, so here they are, there are, as I mentioned, about a couple of thousand of them. And there are up to 600 of them per type. It's much more like what we think of as cortex. We don't think of the cortex as having millions and millions of cell types. Maybe thousands, but there are many, many cells per cell type.

And that occurs here. Very small number of cell types relative to the other things. And here's one of them. It's superimposed. So these Kenyon cells, they have their cell body here. They make their connections. So they get the input from the projection neuron, send an axon down, which in some cases splits.

And there are five lobes here. There is an alpha lobe or an alpha prime lobe here, a beta lobe or a beta prime lobe here. And then some of them send a single axon down to a gamma lobe. You will see that a little bit more.

Then that's it. That's how the mushroom body's built. And if you notice, they do not send anything out of the mushroom body. So the first thing I want to ask, then, is what happens at this junction between the orderly world of the fly, characterized by these PNs, and the wild and random world of the fly, characterized by these Kenyon sets? Here's where they meet in this calyx of the mushroom body.

So the experiment that I was involved in the data analysis of came from Richard's lab and was done in the following way. First, a single Kenyon cell-- so here you can see all these cell bodies of Kenyon cells. There are zillions of them up there, thousands of them up there. But one of them has been-- the GFP in one of them has been activated, photoactivated.

So you can see this single Kenyon cell comes down. Here it's making connections to get the olfactory input from the projection neurons. And then the axon's going to go down through the floor into the other parts that I showed you.

So the trick in this thing-- you can't see very well, but I think I maybe made a circle around one. You can't see it very well, but the terminals of this guy, the postsynaptic terminals, are like claws. They're called claws. And they grab hold of one of the terminals of the projection neurons and make a synapse. So that's how they work.

This guy has about seven of these claws, so there are very few connections per Kenyon cell. The trick was for Sophie to inject the die right into the claw here, which is a very tightly sealed little microglomerulus. And that die is taken up by a projection neuron, the one and only one projection neuron that has a terminal there.

And here you can see the axon of that projection neuron as it makes terminals in other parts of this calyx and makes connections with other Kenyon cells. So there it is. So that's not the important part. The important part is you can trace back this projection neuron to the antenna lobe and see where it got its input.

And now, because the antenna lobe is a totally stereotyped, structured thing, you can now read out. If you know the antennal lobe, you will know that this input is of a certain type. It's

from a certain set of receptors. So you know right away that this guy is getting input from receptor number three, or whoever sends projections to that thing.

Furthermore, you can repeat this with other terminals of that cell, get a whole lot of projections, and find, essentially, all of the inputs-- sometimes not all, but most of the inputs-- that go to this Kenyon cell and figure out what they are. So in other words, the result of this, without doing EM or all that, is a connectome. It's the connection matrix between the glomeruli, or if you want, these olfactory channels.

And there are 50 up around here, plus some hot and cold and some other stuff. But basically, the 50 glomeruli are at the top. And 200 Kenyon cells that were measured going down the side. Not all 2,000 Kenyon cells were measured. These are not measured from the same animal. But you basically get this connectivity matrix.

A red little square here means that this connection was found for this Kenyon cell. And a yellow one means a double connection. There were actually two connections between that Kenyon cell and that glomerulus. So there's the matrix.

So then my job at this point was say, well, what's the structure of this matrix? And that's a trickier problem. I mean, you look at it by eye, you say, well, it looks random. It just looks like a bunch of dots.

But what you have to remember, and I think this is a really important thing to remember in connectomes, is connectomes don't come labeled, all right? So this matrix is arranged in the following way. This is alphabetical, which probably is not of fundamental neuroscience significance. And this is the order in which the cells were measured, which is also probably not of neuroscience significance.

So the question is, is there any way to permute the rows and columns of this matrix to get a structure? That's the question you have to answer here. And just let me show you an example of that.

So here's a matrix that I've shrunk the size a bit, but it's exactly the same kind of matrix. In fact, it probably looks to you pretty much like the data. It doesn't have the colors, but other than that.

So here's a data matrix. But this one I made up. And it turns out, of course, I knew the trick that if you re-sort, if you permute the rows and columns, it looks like this. So just because that

looks random does not at all mean there's no structure there.

So you have to do a lot of analysis to convince yourself that there's no structure. So one of the first things you could do-- random doesn't mean uniform. So one thing you can do is just sum down the columns here and ask, how many connections does each of the glomeruli make? And it's not uniform. It's quite uneven. Here's the histogram.

But really, the question we ask is, is there something more to it? For example, if a Kenyon cell gets one of these inputs, is it more likely to also get one of those inputs? Are there any correlations here? And that's really the question.

And we did a whole lot of analysis. And the answer's no. I'm not going to take you through it. That all the tests we could possibly do are completely consistent with just randomly selecting from this probability distribution without independent ID, or whatever it's called.

OK, so there are other papers, an earlier paper and a later paper, that essentially come to the same conclusion. What's interesting about the Murthy, Fiete, and Laurent paper is they actually provide some evidence that, in fact, it's different in different animals. This doesn't prove that because this is already taken from different animals. I'm not going to present that evidence. But there is evidence that this is different in different animals. So this looks like a random structure.

Now, it's interesting, you guys, why seven connections? Seven seems awfully small to us cortico-centric people. And so why seven? Well, you can do a following little exercise.

You can say, suppose that the Kenyon cells only had one connection. Then how many duplicate Kenyon cells would there be? Well, there are only 50 possible types of input, right? There are 50 types of Kenyon grand cell. So if you only have one connection and you're making 2,000 cells, you're going to get tons of repeats, hundreds of thousands of pairs that are identical. So that you would not spread out.

Now, you can do this calculation for two connections, three connections, four connections. And it goes down. And if you look at the line where you'd only expect one pair to be the same, the mushroom body-- in fact, if you average, it's between six and seven. It's right in there. The mushroom body is right at the point where you convince yourself that most of the time every cell will be different. And then why go any further?

Some of you may know something about the cerebellum. These are like granule cells of the cerebellum. Granule cells in the cerebellum typically have four or five inputs [INAUDIBLE]. They're small cells with claws with very few inputs. And their axons form parallel fibers and then can get connected by Purkinje cells.

This system is the same, if you notice. The parallel fibers are forming the trunk and that L-shaped region in the mushroom body. So these are like granule cells.

OK, so where are we? So we've got to get a signal out of this thing or it's completely useless, right? So we've got this random signal into this beast. So what do the output neurons look like?

There's an output neuron, one of them. And what you might notice is it's going to a very compact region right at the head of this alpha-- I don't know if this is an alpha or an alpha prime. But it's going to one or the other of those lobes. So it's very restricted in its dendrites. And then off it goes carrying the signal wherever it's going.

So in fact, I tried to argue earlier that the output neurons in a mushroom body have gone back to this other mode. Very few cells per type. Practically as many types as output cells. And a very small number of cells. And if you took a picture of this cell in another animal, it would look exactly the same.

All right, so it was known before this Janelia work, that if you take the mushroom body lobe-- so this is this L-shaped structure at the bottom, or it's really at the front of the mushroom body-- and you peel off the gamma lobe, the gamma would sit there but it would kind of block your view. So it's been peeled off here. So you have this alpha beta lobe. That's one set of axons that have bifurcated. And they come in sort of here and then bifurcate.

You have the alpha prime beta lobe bifurcating. And then you have this third gamma lobe. That divides up each of these into five sections. They're numbered like this, but there are five of them, OK? Alpha 1, 2, 3 and beta 1, 2.

So each of these guys gets divided into five compartments. And then there's an extra compartment right here called the peduncle, where-- here's the mushroom head. Here's the stock. And then you get this. And right at the base of the stock, there's another one.

And so what the Janelia collaboration figured out by genetically targeting these cells very precisely. It is summarized by this picture. So this shows different types of these output

neurons in different colors.

And what you can see is that they are respecting the compartments. That you have basically one type of output neuron going to each compartment without overlap. And there really is-- there's now EM level data, and they really don't overlap at all.

Here's kind of what it looks like in an anatomical diagram. Here are the Kenyon cells. Here's the calyx where they get their input. Here's this L-shaped structure. And these output neurons respect each other's territory.

Here is the 16 compartments where they do. And then they're very well organized in another way. You notice these colors here. These colors refer to the transmitter of the output neuron. So all the glutamate guys are over here. All the GABA guys are down here. The cholinergic guys are over here. Now, again, you get this extreme order returning to the system.

Here's a theorist's version of this. Here are the compartments, the 16 compartments, 5 per lobe, plus the peduncle, which kind of belongs to the alpha beta lobe. Here they are, the different compartments. And then here are the output cells assigned to them.

They're not necessarily one cell per blob here. Sometimes there are a few cells. But basically, those are the cell types. And as I mentioned, they respect-- they only go to one compartment each.

And then those are the transmitters, which in this diagram, they don't cluster nicely. But in the other diagram they do. Now, you can ask, why bother to do this? Because there are axons going down. The parallel fibers that the Kenyon cells make, they go that way.

So all of these guys have access to exactly the same input. So what would it matter if this guy decided to send a branch over and pick up the axon over there instead of over there? It would make no difference at all. So at this point, you would sort of wonder, why are they respecting these compartments so faithfully?

And that's answered in this slide. So these are the output neurons, as you can see, kind of tiling the thing in these compartments. And this is a set of dopamine neurons, which were also genetically isolated in this way and labeled, that target these compartments.

And you notice the perfect alignment. So the reason these guys are compartmentalized is so they can be individually modulated by dopamine. And you can see that here. So the dopamine

neurons come, again, in slightly more numbers of types. But they align and exactly innervate these compartments without overlap.

So the reason the beta 2 guy's in here is so it can be innervated by these particular dopamine neurons. The dopamine neurons are divided into two classes. And again, if we go to the anatomical-- oh, I should mention. If you notice, there were some missing compartments there, but some of the dopamine neurons go to 2, so everybody gets covered.

If you go back to this anatomical diagram, what you see is everybody over here gets modulated by these, what are called, PAM dopamine neurons. And they're associated with reward. So when good stuff happens, you hammer this part of the mushroom body. When bad stuff happens, you hammer this part of the mushroom body with a different set of what are called PPL1 dopamine neurons. So again, this beautiful structure.

All right, so let me finish elaborating this for you. This is the basic structure. Again, I didn't put it on at first, but some of these guys actually conduct two compartments. So it's not quite true what I said. But basically, that's the output stream from the mushroom body.

And then there is a layered system put on. These are the connections, but I kind of depicted it down here more schematically. What you have in this output system is a one layer system down here, a two layer system, a three layer system, and a four layer system.

So the output is actually a four layer network, feedforward network. There's no recurrence up to this point. And all of the action occurs on the alpha beta lobe. The alpha beta lobe is responsible for long-term memories. You could think of this as the most sophisticated lobe.

Gamma lobe is more for short-term memories, has a simpler readout. Alpha prime beta prime lobe is, to me, kind of God knows what. But probably somebody knows. Anyway, but it's, again, a simpler output system.

So it's just a beautiful system. In part, I'm just telling you about it because it's beautiful. OK, so that's the thing. And then these outputs go to various regions. If you don't know the fly brain, you don't care.

But what's interesting is now the loop closes. So the regions that receive output from the mushroom body also provide input to the dopamine neurons. So when the mushroom body acts, the dopamine neurons know about it. And when the dopamine neurons react, the

mushroom body knows about it. So you have this closed system which finally loops together.

And the dopamine system is a reporter of behavior. So it tells the mushroom body what the fly's doing. Or also internal state, how the fly's feeling. I'll show you that in a second. And then these are going to be, obviously, some sort of learned or modulating responses, modulated by this system.

You remember that these cannot have any intrinsic meaning, because they've gone through a random stage here. They cannot be assigned meaning without some sort of learning or instruction. So these are learned outputs. And so that's the system.

OK, so what does this system do? One of the nice things that's happened in parallel with this anatomical advance that I've been describing is a behavioral advance of what's the mushroom for. I'll start with the classic picture. Mushroom body has been studied for a long, long time. That quote was from 1850.

And it's mostly been studied as a classical conditioning system, memory system. You train a fly to be afraid of an odor or to be attracted to an odor through a classical conditioning experiment. And here's a nice, recent version of that that's quite instructive.

So in this experiment, what you do is you put one odor in the end of a chamber, a very small chamber that holds a fly. One odor comes in one end. One odor comes in the other end. You pump it out in the middle.

And then you track the fly. Flies pace back and forth. And so the fly paces back and forth. But frequently, if it doesn't like odor B, it might come to this central region, say, oh, that's odor B, turn around and go back. And so what you do is count electronically how many times the fly crosses these boundaries.

And you can get a measure of its preference for being in the A end or the B end. And this is experiments done in Gero Miesenboeck's laboratory. Now, what you can do then is-- in the first set of experiments that I'll show you, they just look at the innate preference of the fly for an odor, without any training. That's due to the lateral horn, as you'll see.

But then you can associate one of the odors, for example, with an electric shock. And presumably, the fly is going to then associate that odor with danger and avoid it. So here's the data. No, first I guess I built a little model.

So it's been long suspected how this could work. This is quite easy. You have-- there are the Kenyon cells. Here's a mushroom body, output neuron. Here's a dopamine neuron.

So an odor comes along-- that's the conditioned stimulus-- activates some Kenyon cells. Then the unconditioned stimulus comes along, the shock. That activates the dopamine neuron. And where you have activity plus dopamine, for example, you strengthen the synapses.

There's evidence that it actually might work by weakening the synapses, but for this diagram, I strengthen the synapses, OK? Then, later on, when the odor comes along, it activates the same set. Now you have these strengthened synapses. You activate the mushroom body output neuron and you send an alarm signal.

So that's just classical conditioning with this system. And here are the data showing it works. So first of all, this is the innate preference. What's interesting-- the reason I included this later experiment is because they looked at the innate preference as well as the learned preference.

So this is just showing you that this is the distance between the PN activity for these odors. So this is a measure of the discriminability. And they sort of argue that these odors which have a zero preference maybe can't be distinguished by the fly. You don't know that, but any rate, zero means they're equally likely to go to both ends.

So these odors, they don't care. But when the odors are quite different, they can have a fairly strong preference for one odor over the other. Now you train, and suddenly you have a strong preference or a strong avoidance, a preference for one over the one that was associated with shock.

And now what they did was genetically-- I mentioned that we now have genetic access to all these cells. One of the things you can do is block synaptic transmission from all the Kenyon cells. So you just wipe out the output of the mushroom body. That's done by raising the temperature of these flies.

And suddenly, they go right back to their innate preferences as if they'd never learned something. But they still can sense the odor. They still have their innate preference, almost identical to what it was before. But they've lost.

Now, if you cool down these flies, they'll pop back up to there.

OK, that's classical conditioning. Oh, I know what I was going to mention here. Not in these

experiments, but in other experiments, you can replace the electric shock by an activation of the dopamine neuron. So you can show that these avoidance type dopamine neurons really do convey the avoidance message, because you can train them to avoid odor B when all they got was an activation, let's say an optogenetic activation of a dopamine neuron. That's been done tons now.

OK, so here's another example. As I said, I think the classic literature on the fly is that. It's classical conditioning studied in zillions of ways, looking at the molecular basis, et cetera, et cetera. But here's some more newer results.

Here's one from Daisuke Hattori in Richard's lab. It involves this alpha prime 3 lobe, just to show you what it is. And it has the following features. So it's a little hard to see, maybe, but this is a pulse that shows that the odor, which is MCH here, the odor has been introduced. And here is the response of this alpha prime 3 output neuron.

So there you're seeing a response. And if you look across time, that response fades away. It even starts to reverse maybe. So there's an adaptation of this response. You say, big deal.

But this adaptation is definitely occurring at the output of the mushroom body. The Kenyon cells are not adapting. It's due to the dopamine, because if you block the dopamine you don't get it. So this is dopamine specific adaptation.

But what's more interesting about it is shown here, that if you take an odor response to MCH, adapt it away, but then present a new odor, benzaldehyde, now you get a response again. Then you can adapt the way the response of the benzaldehyde and introduce a third order, you get a thing. So this is an odor-dependent adaptation, which really suggests that the dopamine is specifically weakening the synapses that are active at the time of the dopamine response.

So it's like the classical conditioned, but there's no conditioning here. Furthermore, when you adapt one odor and then another, the first order remains adapted. So I sort of see this system-- I've always had trouble with the classical condition experiments, imagining where would a fly get into a situation where it smells an odor and gets a shock, or something like that? But this, you can immediately see, would be very useful.

You could adapt to an environment that has a whole set of odors. And then if you come back to that environment and there's a new odor present, you'll immediately know it, because this

neuron is going to respond. Whereas if you come back to the identical environment, or without new odors, this won't respond. So this is a neuron to identify unexpected olfactory features of an environment. Or it's one thing it could do.

OK, I think I just repeated that because I wanted to say that. Here's another example. This comes from Raphael Cohn and Vanessa Ruta's lab. And this is really the effect of internal state. So I argued for you that these dopamine neurons were reflecting the internal state of the animal. And they have a very beautiful experiment on that.

So these are the gamma. I guess I didn't say before, but these are the gamma 2 through gamma 5 compartments that we're going to talk about. What they did was to image the dopamine neurons in those compartments, gamma 2, 3, 4, and 5, and observed that when the fly is-- the fly is in an uncomfortable position, to put it mildly here. It's glued to something, I don't know what. And there's a hole in its head. Other than that, everything's fine.

So it's an unhappy fly. You might want to speculate that this is an unhappy fly. And what they observed is while the fly is flailing about and unhappily expressing its unhappiness, these gamma 2 and gamma 3 compartments have dopamine input. And the gamma 4 and gamma 5 don't.

But they also observed that every once in a while, the fly just chills out, hangs there, like, oh Christ. And when that happens, it reverses the pattern. Now these are not dopamine activated and these ones are dopamine modulated. Although this is not unequivocal happiness. This is mixed.

But then they started manipulating. So here, you take one of these unhappy flies, you give it some sugar, it becomes a happy fly, right? Remember, the red over here, this is a happy fly. This is a sad fly, because they shocked it.

So it's clear that this thing is really reading a-- they'll be a little fanciful-- but happy fly, sad fly. You could take a look at these compartments and say, that is one unhappy fly, or one happy fly. Furthermore, now, so that means the internal state is being represented here. But in addition, it has an effect.

So this is an experiment where they imaged the output term, the dendrites of the output neuron. So they're looking at transmission from the Kenyon cell to the output neuron. They present an odor. And they activate the dopamine neuron themselves.

So when they don't activate the dopamine neuron, this is a measure of synaptic transmission. The odor response here is weak. When they do, it gets much stronger. So now what you have is something-- I mean, I think if I saw this in cortex, I'd go wild-- is a gating effect. You have internal state affecting the thing of this, and it determines where the output goes.

So for example, if you're-- I can never remember which one's happy. This is happy, right? So if you're happy, then odors go to one thing, which might say, approach that order. You know, be sort of a little more easy going. If you're unhappy, then an odor response gets relayed out this pathway. And it might tell you to be afraid of all odors, or something like that. Be very cautious.

So you start to see in this system the routing of sensory information by internal states. That, to me, is a very exciting thing to see. OK, here's another one-- internal state affects memory.

This is from Tanimoto, another collaboration with the Janelia lab. So you can do the same kind of experiment I showed you before with shock, only do it with sweet. So in this case, it's a T maze. A fly comes in this way. And you associate, let's say, odor A with a sweet reward.

Then the fly is going to come in here and most of the time go this way. Flies are never 100% performers, but they'll tend to go to odor A because they associate that with sweet, provided they're hungry. I'll come back to the hungry part. So you take a hungry fly, it goes this way.

Now, here's what they did that's very clever. It involves these PAM neurons, PAM dopamine neurons. So as you all know, you get the hit of sugar in your mouth right away. And then you get nutritional value, or you get fat or whatever later. And the decoupling of those causes us a lot of problems.

So you see that here in this system very, very well, because they take a sugar that's sweet tasting to the fly but can't be digested by the fly, that provides no nutritional value at all. And when they do that-- so there's no nutrition in this sugar-- what they get is a short-term attraction to that odor, enough to buy the next Coke, sort of. It's conveyed through octopamine, so the sweetness activates octopamine.

Octopamine activates a certain set of these PAM neurons. That makes changes in the transmission in various compartments. This is not isolated to a unique compartment. Then the next time that odor comes on, it goes out here and it gives you attraction. But it's a short-term attraction, lasts a few hours.

Now, if you make the sugar nutritious-- and they do this in a clever way by using another sugar that has no taste to the fly but that the fly can digest, so now this nutritious-- then it activates a fructose receptor, blah, blah, blah, activates a different set of dopamine neurons. That potentiates or depresses-- we don't actually know-- but it changes synapses in the alpha 1 lobe. And now you get a long-term memory.

So flies are smarter than people in a way. They'll only do this long-term memory if they then also sense a nutritional benefit to whatever they're eating. OK, so very elegant thing.

And then I think this is my final example-- I'll start winding up-- from Scott Waddell's lab, that another feature of this sweet thing is if the fly's not hungry, not surprisingly, it doesn't care anymore. So it's associated odor A with sweet, but now it's well-fed, so who cares. And that's a real effect. So a fed fly will not express this odor preference. But it still has the memory, because if you then starve it, now it will go to odor A.

OK, So what Scott Waddell and his group realized was that this was activated through this dopamine neuron. Now, this was done before all this circuitry was derived. But they figured it out that it's this dopamine neuron, because they could activate this dopamine neuron and simulate the fed state, so the fly would ignore the odor. Or they could silence this dopamine neuron and then they would simulate the hungry state and the fly would be attracted to the odor.

But now, you notice this circuitry, this is a GABAergic neuron that inhibits the alpha beta lobe output. So this is a perfect pathway by which you could turn off the learned response in this during the fed state. And then you inactivate this pathway, and now you turn it on. So again, an internal state gating a memory. But it's a case-- we don't really know that everything I'm saying is true. One should never assume that.

But we now have this pathway. People-- we, I say, but people, we can block this pathway. There's enough known about the circuitry to really work out that what I said is true, that you can start to get at these things. So one-- yeah, I got a minute to do CO2 avoidance.

CO2 avoidance is a really cool one. So CO2 is innately repulsive to a fly. It doesn't like CO2. And the reason that is is in a group of flies that are stressed, they release a lot of CO2. So a fly will sense CO2, know there's trouble in the area, and will avoid it.

So there's a natural avoidance through the innate pathway to CO2. Now, that's kind of a fatal

flaw in the design of the fly, because flies eat rotting fruit that releases tons of CO2. So you don't want to avoid your food source.

So what happens-- it's not completely understood. But somehow, the innate system trains this beta 2 pathway to have, in addition to the innate pathway, a learned pathway for CO2 avoidance. And in the hungry state, the fly channels its CO2-- it still has CO2 avoidance. It channels it through this pathway.

Then if, at the same time, there are fruit odors or fruit tastes, it can modulate this pathway, shut it down, and turn the CO2 avoidance into a CO2 attraction. So again, you start to see the neural substrates of these really quite complex behaviors sitting right before you in this structure.

All right, I'll end here with sort of the lesson for the machine learners. So from a machine learning perspective, this is a simple system. It's not very deep. It's a little bit deep, but not very deep. It contains a random hidden representation.

That's not really anything radical. It contains a set of output neurons. It's actually a layered output. Again, nothing very radical. In neuroscience term, it's kind of interesting that it goes from these highly stereotyped to random to highly stereotyped.

But really, the lesson here is this is a mediocre machine learning architecture. Not very many units and all that. Where does this thing make up for it? It makes up for it in a stupendous, complicated modulation and plasticity beyond any machine learner's dreams.

We don't know about this. I tried to give you hints of the different things it can do. Dopamine can gate. It can induce short-term learning. It can induce long-term learning. It can induce gating of gating, gating of learning.

That's what has to be worked out in this system. But there's going to be a really beautiful effect of dopamine acting in many ways on many time scales. And to me, in this system, that's where evolution has put its money, right there. Not in building 20 layers here or something like that. Not in worrying about a whole lot of back prop.

This is random. It doesn't appear to be back propped. But in putting huge resources into a rich set of modulatory and plastic processes at these output synapses. And I think in the years to come, they will be worked out. And maybe they'll have implications for machine learning once we know what they are.