A Lecture Companion

”The Collective Intelligence of Morphogenesis: a model system for basal cognition” by Michael Levin

Michael Levin

Recorded on September 17, 2024

About this document

This document is a companion to the recorded lecture *"The Collective Intelligence of Morphogenesis: a model system for basal cognition" by Michael Levin*, recorded on September 17, 2024. You can watch the original lecture or listen in your favorite podcast feeds — all links are on the page here.

This document pairs each slide with the aligned spoken transcript from the lecture. At the top of each slide, there is a "Watch at" timestamp. Clicking it will take you directly to that point in the lecture on YouTube.

Lecture description

This is a ~1 hour talk on morphogenesis as a collective intelligence and some specifics about cognitive glue.

Follow my work

Twitter • Blog • The Levin Lab

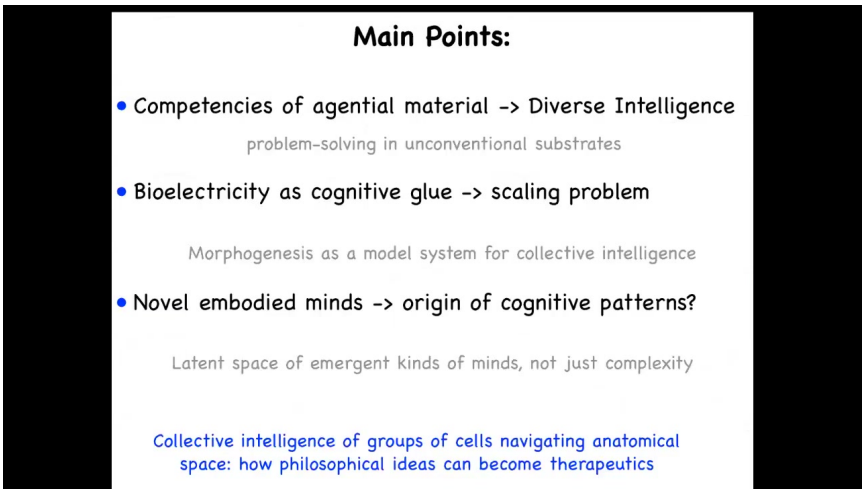
Transcript note

This transcript is generated automatically. While we strive for accuracy, occasional errors may occur. Please refer to the video for exact wording.

Want one for your lecture?

Want something like this for your own talk? Reach out to Adi at adi@aipodcast.ing.

Slide 1 of 50 · Watch at 0:00

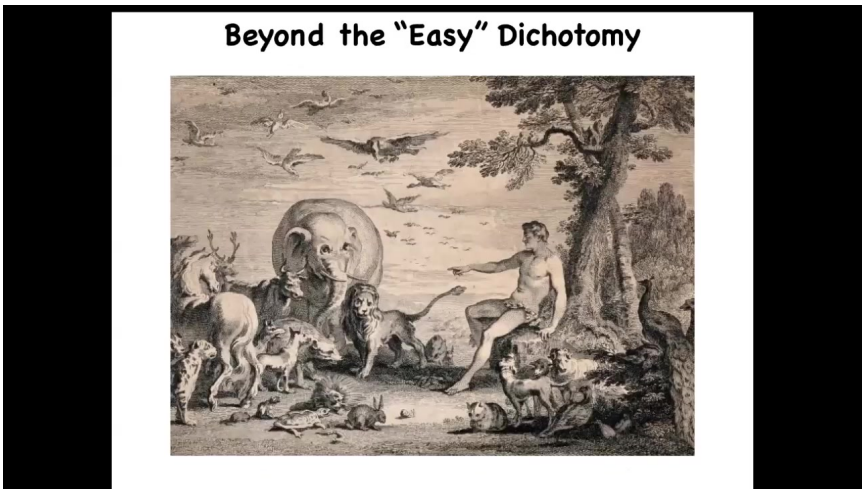


All of the details of everything that I'm going to talk about, the papers, the data sets, the software are here at this website. Here are some more personal thoughts about what some of it means.

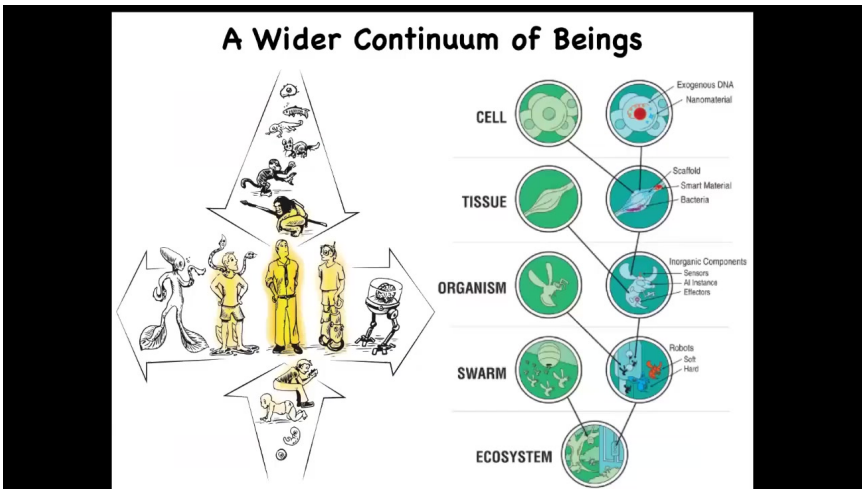
What I'd like to do today is go over three main points. The first is that I'm going to talk about problem solving and unconventional substrates, specifically this idea of living tissue as an agential material; this is part of the bigger field of diverse intelligence, and I'll talk about that. I will use one specific model system as an example of what I call cognitive glue, which is policies among components that solve the scaling problem of having an emergent cognitive system out of smaller parts. We will use biological morphogenesis as a model system to think about developing tools and strategies for analyzing these collective intelligence systems. At the end, I want to show you some novel living beings that may help us to think about the origin of both spatial and cognitive patterns and where they might come from.

What my group does boil down to one sentence is that We study collective intelligence, for example, how groups of cells navigate the space of anatomical possibilities, and we try to push some pretty philosophical ideas about that into actual therapeutics. For us, this is how we test whether these ideas are useful or not.

Slide 2 of 50 · Watch at 1:34

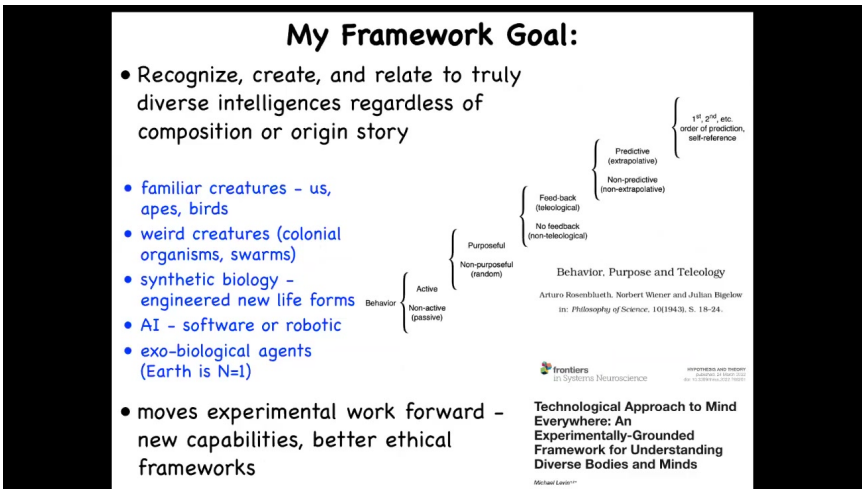


So let's start by thinking about this. This is a very old classic painting: Adam is naming the animals in the Garden of Eden. The worldview shown here is that there are some discrete natural kinds. So here's Adam, here are these other animals. We can study the various properties that these other creatures have, but it's nice and easy because they are all discrete and different from each other. It's very easy to say who's who. Even though it's a really hard task to understand the cognition of other species, we can call it the "easy problem" in Chalmers' sense, because at least we know what everything is here.



But the problem is that these discrete kinds don't really exist because all of us are at the center of several different continuous lines of possible beings. So both at the level of individual development and evolution, we are just one point out of a very slowly changing set of diverse forms. And whatever you want to say about this modern human, you have to ask yourself, at what point and how did these capacities come to be?

And then because of approaches developed in bioengineering and genetics and synthetic morphology, we now know that both biologically and technologically, we are in the middle of a very large continuum of possible hybrid beings, because at any point we can introduce into any level of organization synthetic and engineered materials. Life is incredibly interoperable. You can make these things. People are and will be increasingly doing that. Whatever philosophical or conceptual framework we have has to be able to handle all of this, not just the difference between specific distinct discrete species, but be able to say something about all of these different kinds of possible variants.

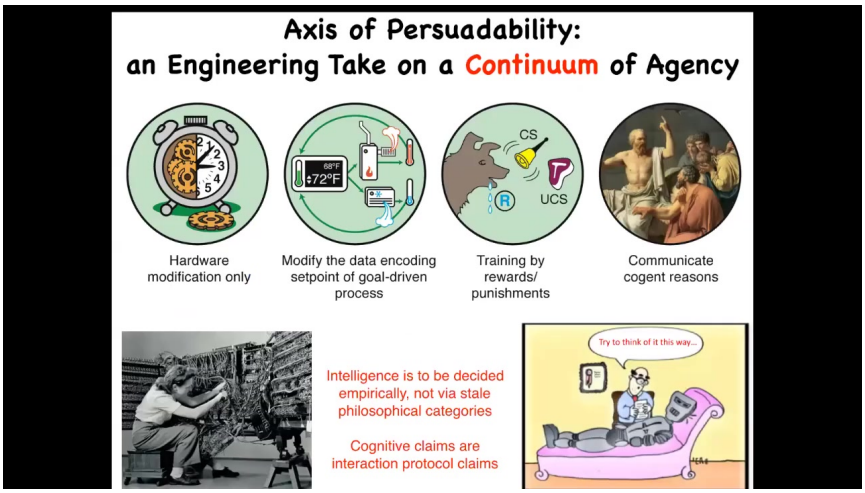


What my framework tries to do is to help us to recognize, create, and ethically relate to truly diverse intelligences.

Regardless of their composition or how they came to be, I'm interested in familiar creatures such as us and other primates and birds, but also unusual colonial organisms and swarms, synthetic biology—engineered new life forms—AI, whether purely software or embodied in some kind of robotics, and someday maybe exobiological agents.

What I'm interested in are ways to think about all of these things that move experimental work forward. It gives us new capabilities, new discoveries, and hopefully better ethical frameworks. All of that is detailed here.

But I'm not the first person to try for this. Here's Rosenblueth, Wiener, and Bigelow's attempt to make a scale going from passive matter up to this kind of second-order metacognition that humans have. Trying to figure out what that continuum looks like has been an interesting problem for a long time.

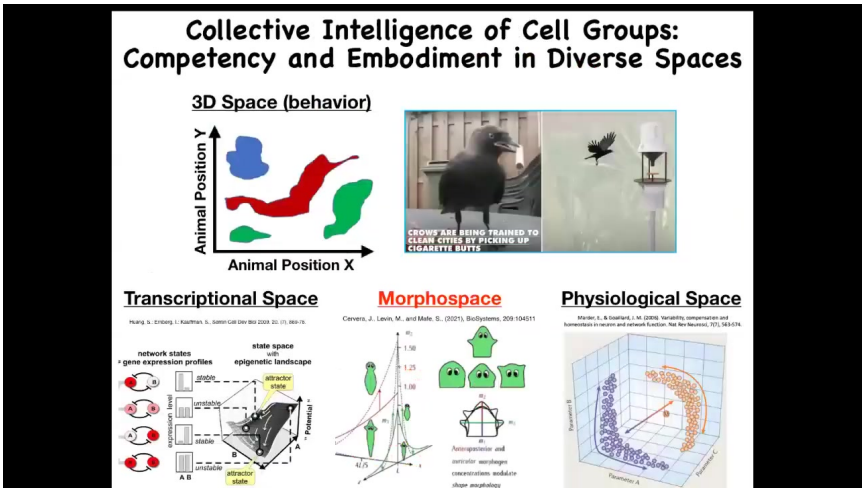


The way I try to think about it is through what I call an axis of persuadability.

What I try to understand is what kind of tools are we going to use for a given system to interact with it. The idea is that this is an engineering approach. It becomes an empirical question as to, for any given system, where along the spectrum it lies. Cognitive claims become interaction protocol claims.

For example, for a given system, if you say it has this or that level of intelligence or cognition, what you're really saying is, how can we interact with it? Do we just modify the hardware or can we reset some set points? Can we train it? Does it respond to logical arguments?

There are lots of different sciences from control theory and cybernetics and behavioral science and psychoanalysis. The idea is that for any given system, you can't just assume where it is. You can't have philosophical feelings about where things should be. You have to do experiments. You have to ask. It turns out we get a lot of surprises.

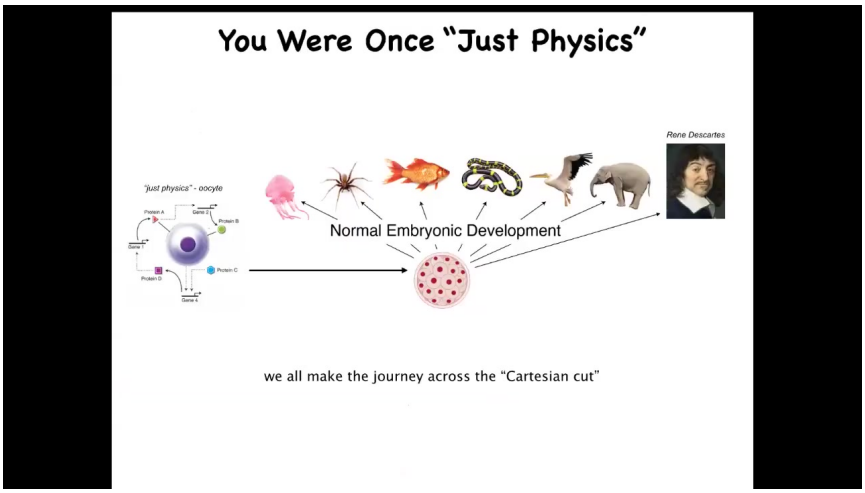


One of the things that we're interested in is intelligence in novel spaces. We as humans and scientists are reasonably proficient at recognizing intelligence in familiar three-dimensional space. Medium-sized objects moving at medium speeds, birds and primates, maybe an octopus or a whale, and we can see them doing interesting intelligent things.

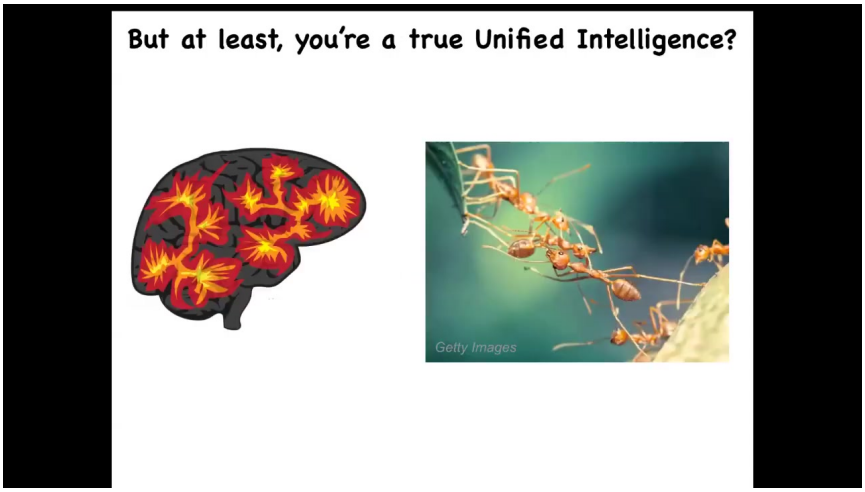
But biology does the exact same thing in many other spaces. Our familiar 3D space is not the only space where intelligence is deployed. There is the space of possible gene expression. It's a very high dimensional space. There's a space of physiological states. What we will talk about today is this anatomical amorphous space, the space of all possible configurations of a body.

Within any of these spaces, even though it's difficult for us to visualize, if you use the tools of behavior science and cybernetics and some other things, you can observe and make use of the fact that these systems have intelligence in navigating those spaces.

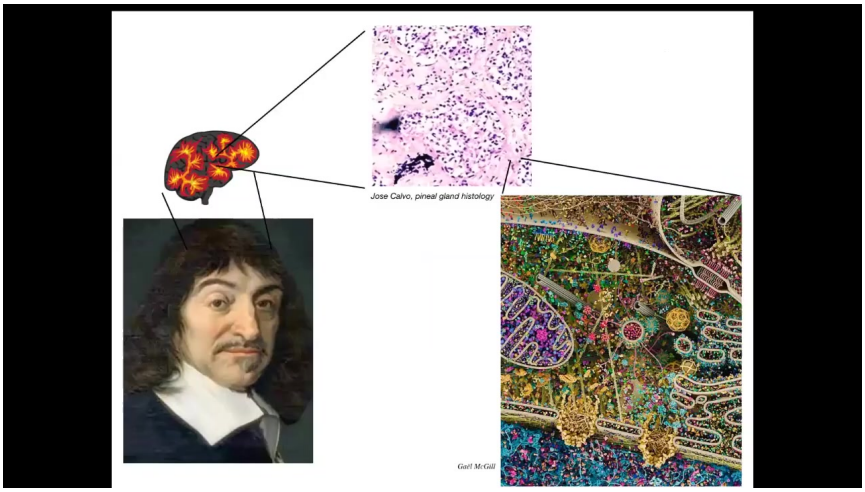
Let's talk about this one in particular. We could do many hours on these other ones, but let's just focus on morphospace.



We have to understand that our intelligence at one point, both evolutionarily and developmentally, was a little BLOB of chemistry and physics. So this is an unfertilized oocyte, and it becomes something like this or even something like this. And we all make the journey from something that looks like it's just chemistry and physics to something that obviously has a mind and various other cognitive properties. So what we need to do is figure out how we got from here to here, what actually happened during this journey. So it's a little weird for many people to focus on development and realize that basically at one point they were just an unfertilized oocyte. But at least we are a true centralized intelligence, right?

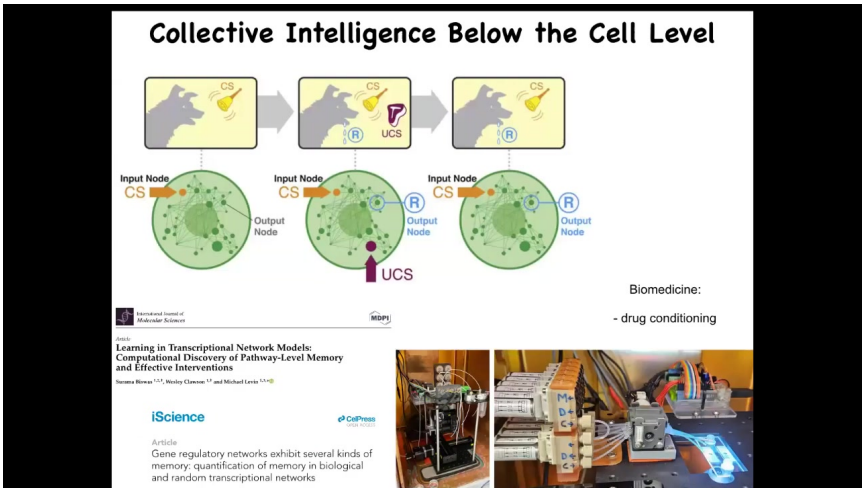


We have this nice brain and we're not these other supposed intelligences that people call ants and termites and bird flocks, collective intelligences. But many people think that that's not real intelligence the way we have it. We have a nice centralized brain.



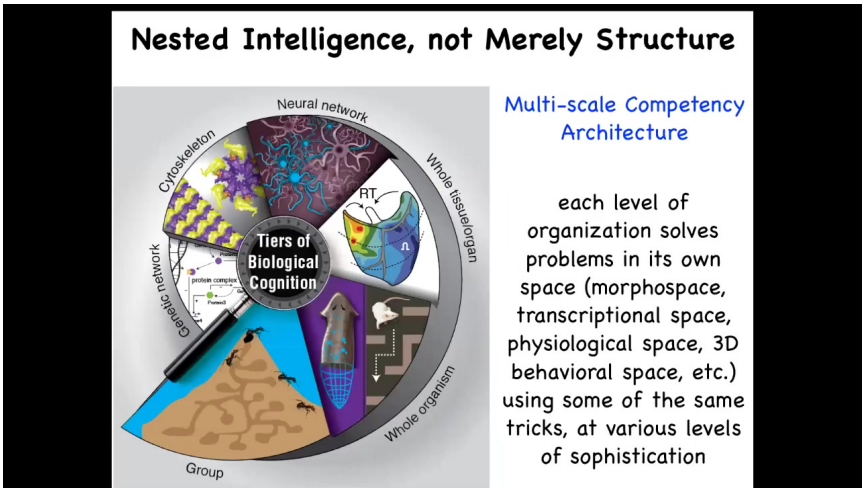
René Descartes liked the pineal gland in the brain because it was an unpaired structure and he felt that it was appropriate for the kind of unified cognition that humans enjoy. But if he had good microscopy, he would have looked inside this pineal gland and he would have seen that inside that pineal gland is a very large number of cells. Inside each of those cells is all of this stuff.

In the end, it turns out that we are all collective intelligences, we are all made of parts. The real trick is to tell some useful story about scaling: how it is that the properties of the parts give rise to the cognitive properties of the whole. And this is the kind of thing we're made of. This is a single cell. This is called the Lacrymaria. There's no brain. There's no nervous system. You can see it carrying out its daily activities. It's hunting for food. And all of that is handled within this single cell. So we are made of a kind of very, very competent, very active material. We're not made of passive materials the way that our current robotics and other types of engineering are. Sometimes I give a talk called "Why Robots Don't Get Cancer," and this is why.

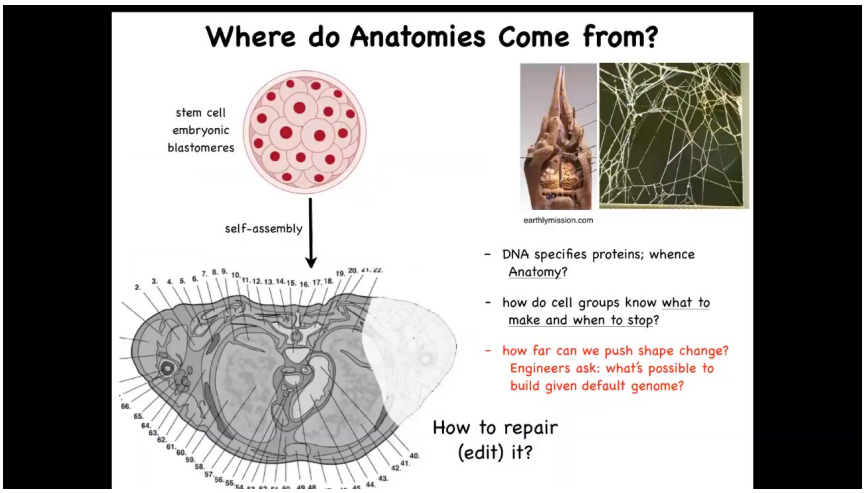


One thing that we can now see is that even within a single cell, just the molecular pathways, very simple molecular pathways, have learning capacity. You can train them. If you start to stimulate and record various components in a gene regulatory network, you can see that they're capable of at least six different kinds of learning, including Pavlovian conditioning.

All of that is described here, and we are at the moment now trying to take it; we're building devices to train these cells, and it has applications in medicine, for example for drug conditioning. So even inside of our individual cells, you already have the basic component of intelligence and various other behavioral properties.



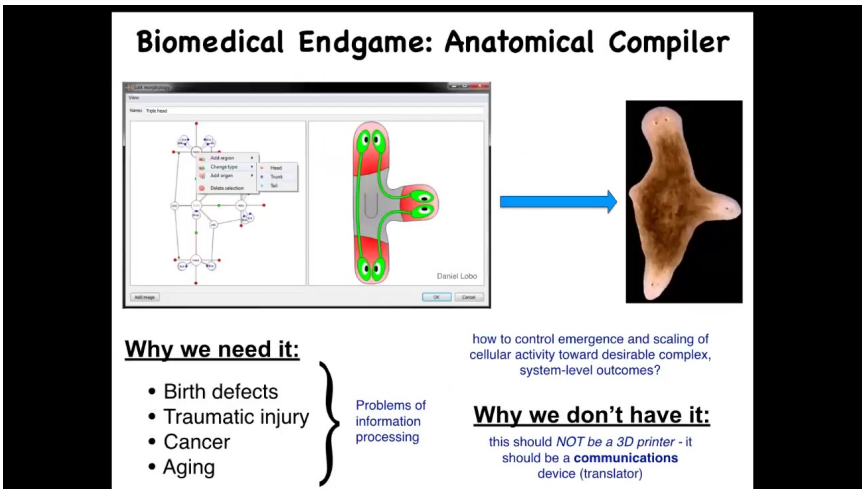
What we're looking at in our bodies is this multi-scale competency architecture. That is, at every level, we've got components that are solving problems in their own space. There are different spaces in all of these components, the molecular networks, the cells, the tissues, the organs, and so on, that are competent to do various things. Let's then ask, what are the competencies of cells in morphogenesis?



How do they create our amazing complex anatomies? This is a cross-section through a human torso. You can see there's an incredible complexity here of all the different organs in the right place, the right size, the right shape, next to each other. But we all started like this, a little ball of blastomeres. And so the question is, where did this come from? How did these cells build this?

And you might be tempted to say, well, it's the DNA, the genome explains this. But of course, we know that the genome doesn't directly specify any of this. The genome specifies proteins. And that protein hardware that's available to every cell is running physiological software that enables these cells to know what to build and when to stop. And so we need to understand how that works. We need to understand how to enforce changes when something is damaged or missing. How do we get the cells to rebuild it?

And as engineers, we can also ask, well, what other shapes are possible? This hardware, what else can it build? It's important to make analogies to the facts of behavioral science, where we know that the structure of the termite colony or the precise structure of a spider web are not directly specified in their genome. What the genome gives you is a very competent agent that is able to then, through the physiological workings of what the cells are capable of, make these other patterns.



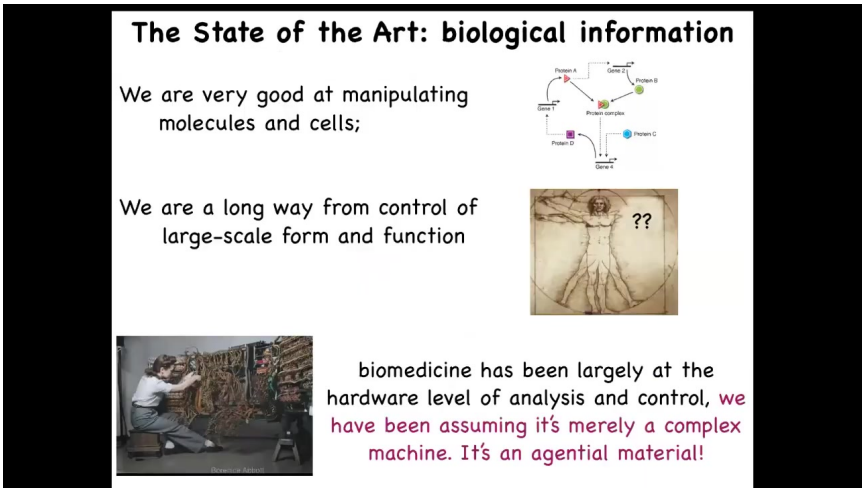
The way that we can formalize this problem definitively is to think about the end game of regenerative medicine. What is it that we want in the end? In the end, we want something like this. I call this an anatomical compiler.

The idea is that someday you will be able to sit in front of a system like this and you will be able to draw the plant, animal, organ, or biobot that you want. In this case, we've drawn a three-headed flatworm. The idea is that if we had this anatomical compiler, what it would do is take this shape and compile it into a set of stimuli that would have to be given to cells to get them to build exactly this.

The idea is if we knew what we were doing, we would have complete control over growth and form. You would know what stimuli should you give to cells to get them to build anything, whatever you wanted.

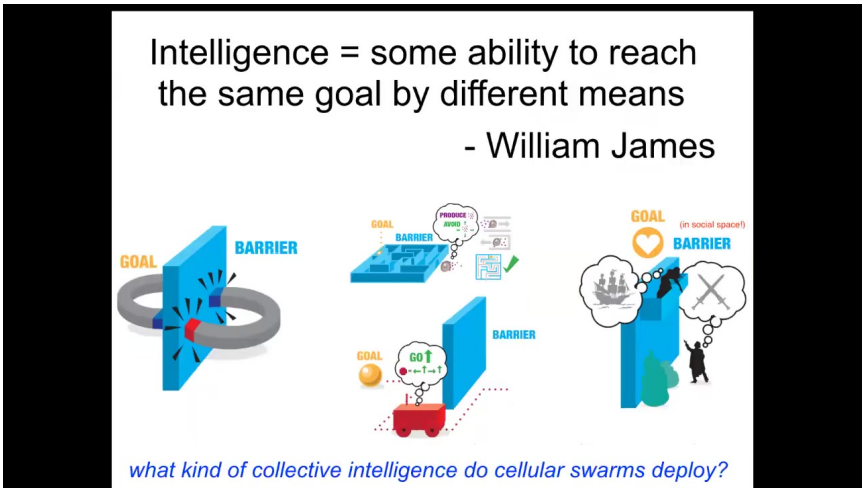
The reason we need this is pretty obvious. If we had something like that, then birth defects, traumatic injury, cancer, aging, degenerative disease, all of this would go away because we could tell the cells what new tissues to build. The reason we don't have anything like this, even remotely, and despite the fact that molecular biology and genetics have been making great progress over many decades, is that we've been thinking exclusively at the level of hardware.

What I think we need to realize is that this system is not something like a 3D printer that micromanages the position and the gene expression of the cells. It's a communications device. What you're trying to do is translate your goals to the goals of the cellular collective. That requires thinking about this in a very different way.



So in modern molecular medicine, we're very good at the hardware. We're very good at figuring out which genes interact with which other genes and proteins and so on. We're actually really far away from large-scale control of form and function. If somebody's missing a limb, has a birth defect, or has a tumor, we're still really very primitive in all of this.

And I think that's because we've been assuming that the biological material is merely complex. And it is complex, but that is only the beginning of the story. It's not just complex, it's actually an agential material that has its own competencies and agendas. And I think regenerative medicine is going to be completely transformed when we adopt some tools from the cognitive and behavioral sciences to deal with this material.



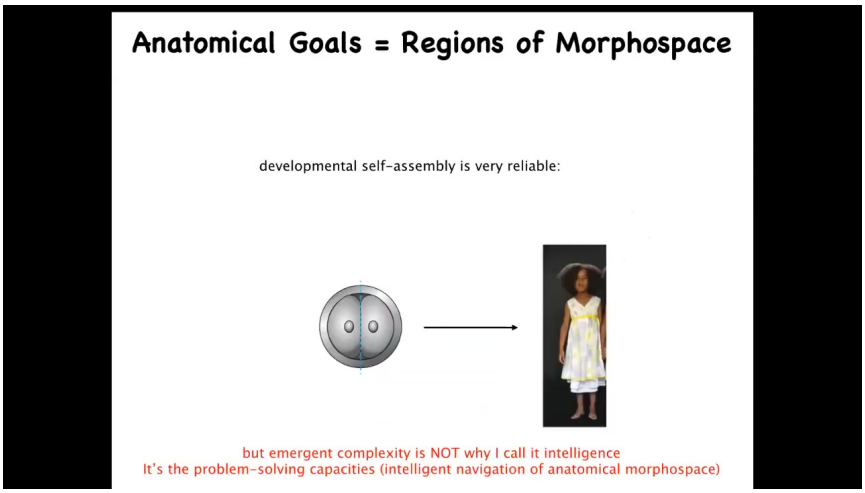
I've used this word now a bunch of times. Let me define it for the purpose of today's talk. I do not claim that this is the one definition that subsumes all of intelligence. This is just one slice of it that I want to talk about today.

I like William James's definition, which is that it's some degree of ability to reach the same goal by different means. The idea is that when you want to understand the intelligence of a system, you have to hypothesize a problem space that it works in. You have to hypothesize what the goals are. You can hypothesize what competencies it might have to reach those goals despite various interventions. And then you do the experiment and you see how well that worked out.

We can test it in very simple systems like two magnets separated by a barrier. They're very stupid. They will never go around the barrier to meet each other because that requires them to temporarily get further from their goal. All they can do is go down a smooth gradient. Here, Romeo and Juliet, in contrast. They have all kinds of skills and planning and memory, and ways to avoid and go around social and physical barriers. In between, you have all sorts of things. You have autonomous vehicles, you have cells and animals that have some degree of navigation ability.

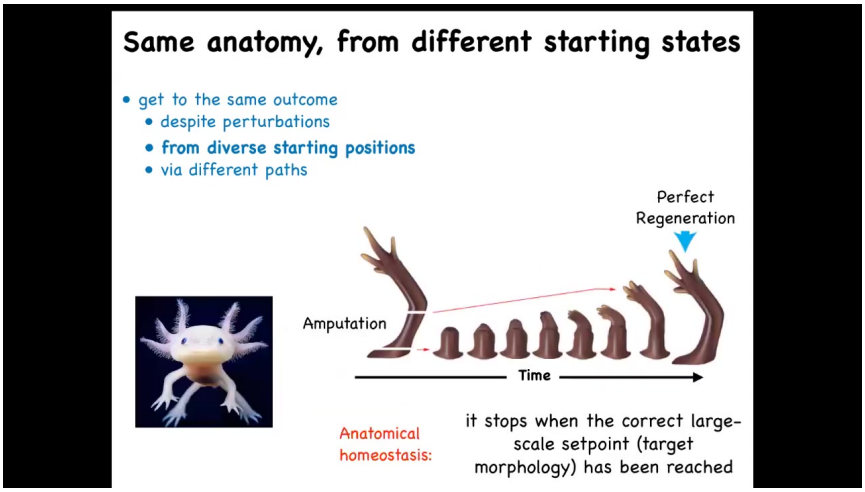
This is what I'm talking about when I say intelligence. I mean specifically navigating a problem space with novel barriers to get your goals met.

Now let's think. What kind of collective intelligence do cellular swarms have?



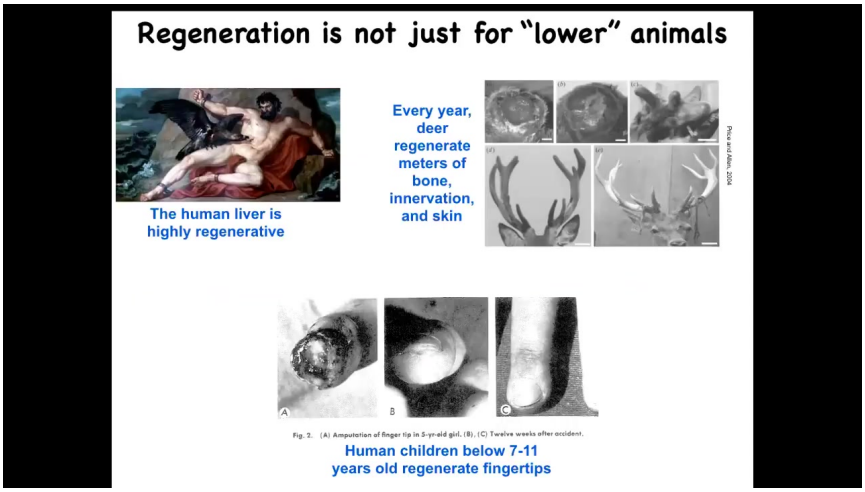
We know that development is incredibly reliable. Almost all of the time, a few embryonic cells will rise to a normal human anatomy. We also know that this process involves a huge increase in complexity, but that is not why I'm calling it intelligence. Intelligence is not simply being reliable or simply generating complexity. Those are very easy to do, and there are lots of very simple systems that do that. This is not about just complexity and reliability. It's really the problem-solving capacity. You start to see that when you perturb the process, not merely observe what it does normally.

For example, what happens if I cut this early embryo in half? Now half of this thing is missing. We know what happens if you cut any of our machines in half. What happens if you cut this? You don't get two half bodies. You get two perfectly normal monozygotic twins because this system can start from many different starting configurations. In fact, you can scramble these cells up to a certain point. You can take cells away. You can add cells. There's all kinds of things you can do. They will navigate this space. They will go around local minima and maxima in that space, and they will reach this ensemble of goal states that represents the normal human target morphology. The system can tell that half of it is missing. It knows what it needs to do to fix it and eventually gets where it's going.

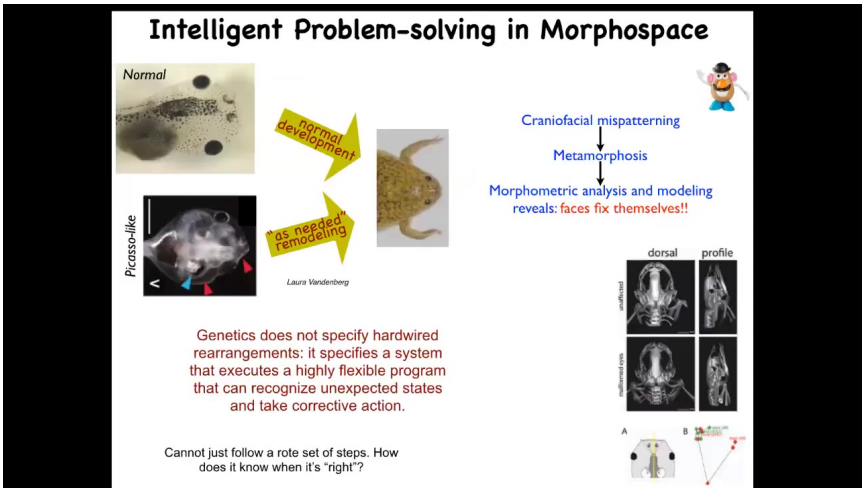


Some animals can do this throughout their lifespan.

This is an axolotl. These animals regenerate their eyes, limbs, jaws, portions of the heart and brain. For example, in this limb, if it's amputated at any particular point, the cells will rapidly grow to regenerate exactly what's needed, and then they stop. The most amazing thing about regeneration is that it knows when to stop. How does it know when to stop? It stops when the correct salamander arm has been completed. This is an example of a kind of anatomical homeostasis. The system is quiescent, and when it's deviated from its correct position in anatomical space, it will work hard to get back, and when it gets back, it stops.

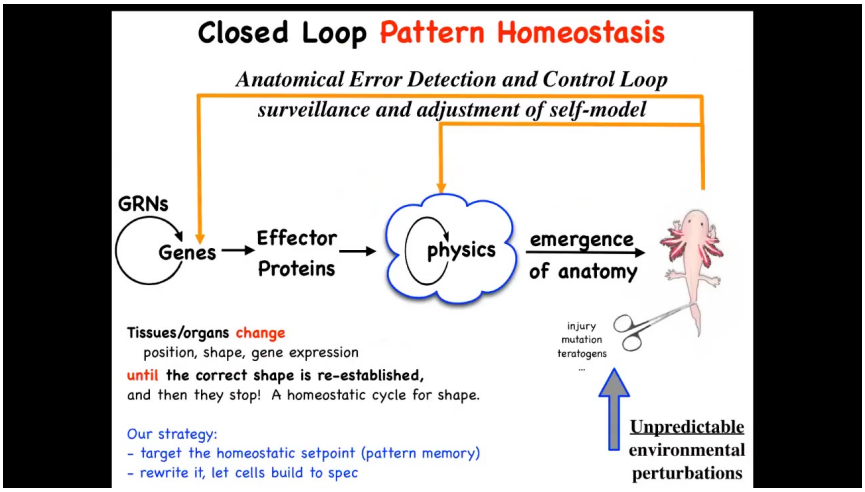


This is not just something for worms and salamanders. Humans have it, and other mammals. We regenerate our livers. Deer regenerate massive amounts of bone and vasculature and innervation. They grow about a centimeter and a half per day when they're regrowing their antlers. Even human children can regenerate their fingertips below a certain age. This is a very generic capacity that's widespread throughout the tree of life.



A simple version of that kind of homeostasis can be shown here. This is a tadpole. Here are the eyes, here's the brain, here's the gut, here are the nostrils. This tadpole needs to become a frog. In order for a tadpole to become a frog, its face has to rearrange. The jaws have to move forward, the eyes have to move, everything has to move. It used to be thought that this was a hardwired process. Every tadpole looks the same, every frog looks the same. If you just move all the pieces in the right direction, the right amount, you will get a normal frog. You would think you could do this with a very hardwired fixed process.

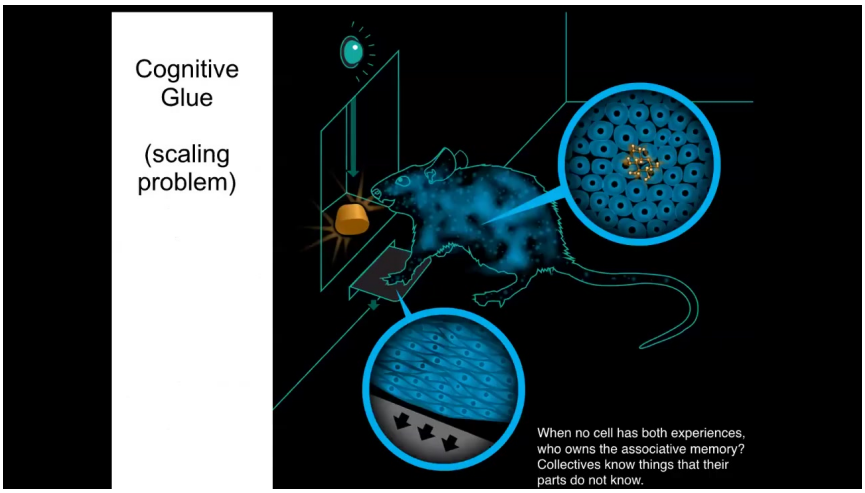
Laura Vandenberg in my lab decided to check that. To understand the intelligence of a system, you can't just observe it. You have to do perturbative experiments. What she did was she made these Picasso-like frogs. Everything is scrambled. The eyes are on the back of the head, the mouth is off to the side, everything is scrambled. What she found is that these guys make largely pretty normal frogs because all of these organs move in unnatural, novel paths. Sometimes they even go too far and they have to come back a little bit, but they keep rearranging until they get a normal frog face. What the genetics gives you is not a hardwired set of movements. It specifies a problem-solving system that can recognize unexpected states. It has the ability to get to this goal state, and it will do what it needs to do to get there. That raises an obvious question. How does it know where it's going? How does it know what the correct frog face is?



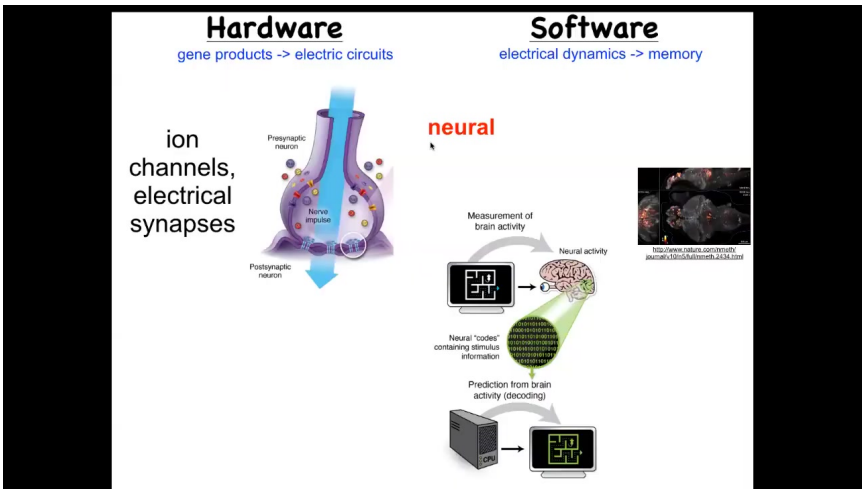
This is the standard story that you'll get from a developmental biology textbook. It's a kind of feedforward, open-loop process. There are gene regulatory networks. Some of those genes become effector proteins. They do things. They have physical properties. They're sticky or they move or they apply force or something. And then eventually all of that complex stuff together gives you this kind of emergent thing. So that's very much the mainstream view of developmental biology: it's mostly a kind of emergent complexity that happens from low-level rules acting in parallel in large numbers.

As I've shown you, there are some amazing examples, and I could show you many more—amazing examples of a kind of anatomical homeostasis here, where if you're deviated from this final goal by injury, by mutations, by teratogens, many different things, some feedback processes will kick in, both at the level of genetics and physics, to get you back here to reduce error. It's an error-minimization scheme. That's quite different from systems that are feed-forward, like cellular automata, which typically do not have any kind of even primitive cybernetic features like this, where they try to maintain specific states, the foundation of goal-directed activity here.

What we've tried to do is follow this scheme experimentally and ask, if this is true, where is the set point encoded? Could we find it, first of all? Second, could we read it and interpret it? And most importantly, could we change it? The amazing thing about these processes, like the thermostat in your home, is that in order to change what it does, you do not have to physically rewire the system. The current commitment of molecular medicine is that you have to act down here; you have to change genes and molecular networks, CRISPR, those things, and then hope that the effects percolate properly to the system property that you're trying to control. If this is correct, you don't need to do that. You can control the set point and then leave it alone and let the system do what it does best. This is the research program that we started.



And in particular, I want to generalize it a little bit by pointing out here the similarities to neuroscience. So if you train a rat to press a lever and get a delicious reward, there is no individual cell in this rat that has that associative memory. The cells at the bottom of the feet interact with the lever. The cells in the gut get the delicious reward. Who owns the associative memory here? The associative memory is owned by the rat, which is a collective that's held together. This solves the scaling problem. It allows the collective to have memories that do not belong to any of the individual subunits. It's a distributed system. And we know how this works in neuroscience. It's implemented by a bioelectric, in this case neural, network.

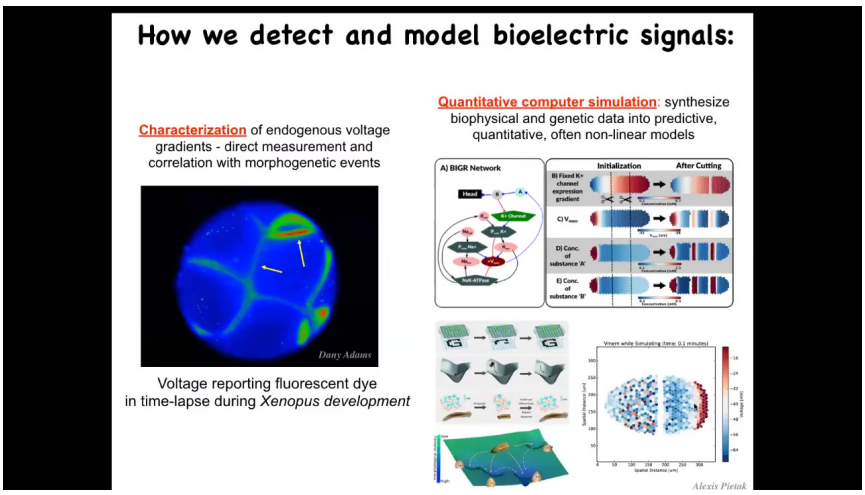


So what we have here is that the hardware looks like this. These are neurons that have ways of setting voltage potential via their ion channels. They may propagate states and potential to their neighbors through electrical and other kinds of synapses. We'll look at these gap junctions. It is a commitment of neuroscience in general that if we were to read the electrical activity, the physiological software that actually is the activity of this electrical network, we should be able to extract the memories, the goals, the preferences. The commitment here is that all of the cognitive features of this animal are in some way encoded in the electrophysiological signaling between these cells. This is a video that this group made of a living zebrafish brain in action.

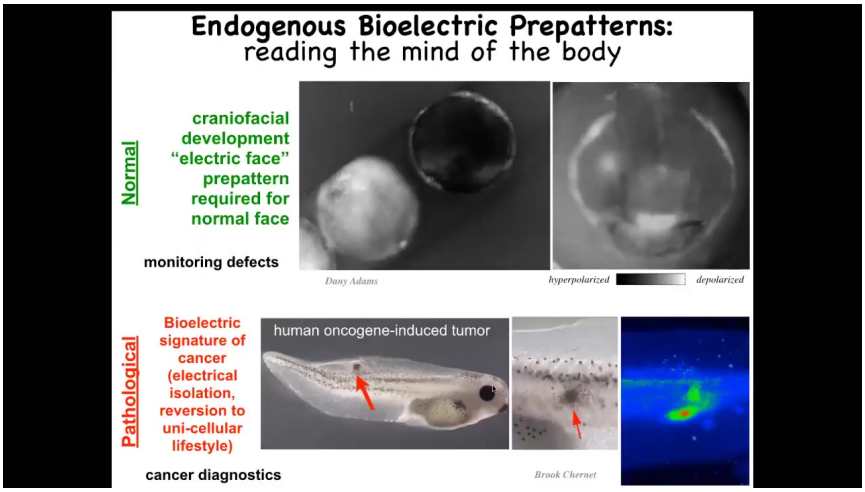
Well, it turns out that this amazing architecture is very ancient. Evolution noticed that bioelectricity was good for solving problems and aligning components within collectives long before nerves appeared. It actually showed up around the time of bacterial biofilms. Every cell in your body has ion channels. Most of them are coupled into electrical networks with these gap junctions. Most of the same molecular mechanisms and many of the algorithms are quite similar between what's going on in the nervous system and what's going on in the rest of your tissues.

We can ask the question: before nerve and muscle came on the scene and were used to think about navigating the body through three-dimensional space of behavior, what did these networks think about? What were they doing? We can also think about a similar neural decoding project, except not in neurons. This is the idea of reading the electrophysiology of these cells, decoding them using various algorithms, and asking whether we could learn what the information content really is.

This is a video of an early frog embryo using a voltage-sensitive fluorescent dye. This was made by Danny Adams in our group. You can see all the electrical conversations that these cells are having with each other long before there's any nerve or anything like that. So what are they doing?



In order to find out, we developed the first molecular tools to read and write non-neural bioelectrical information. This means voltage-sensitive fluorescent dyes that we adapted to non-invasively read these states, and now there are also genetically encoded reporters. Then we do a lot of computer simulations to link the expression of ion channel genes to the capabilities of the excitable medium that they underlie, so the electrical propagation through the tissue and all the states, and how they do things like pattern completion, which are important for regeneration.

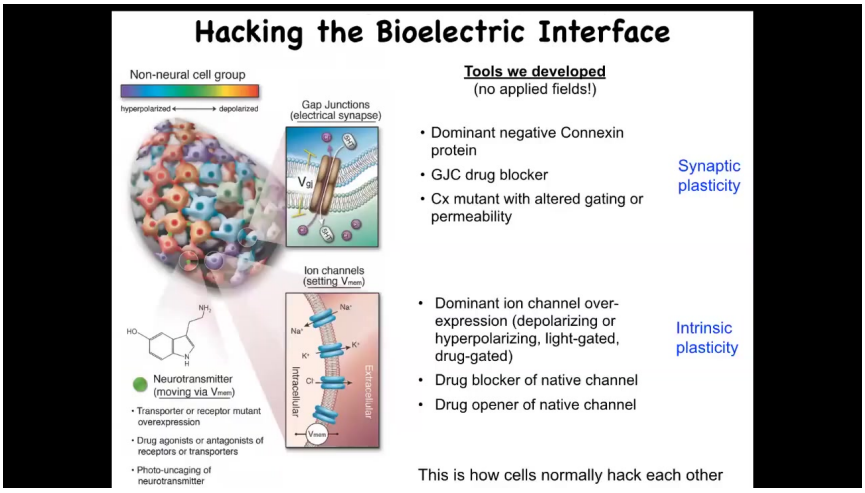


I want to show you an example of what these bioelectric patterns look like. This is also the work of Danny Adams. We call this the "electric face." This is an early frog embryo putting its face together, and you can see there's a lot going on. But if you freeze one of the frames, what you see is that here, long before the actual face gets created, we already see what it's going to be. This is where the animal's right eye is going to be. This is the mouth. There's some placodes out here. The left eye comes in slightly later than that.

What you can see is that this pre-pattern, we're literally able to read the set point, the homeostatic set point towards which these cells are going to work. As I'll show you momentarily, if you change that set point, then everything changes. The gene expression changes, the craniofacial patterning changes. So this is the normal endogenous pattern memory that is required to make a correct frog face.

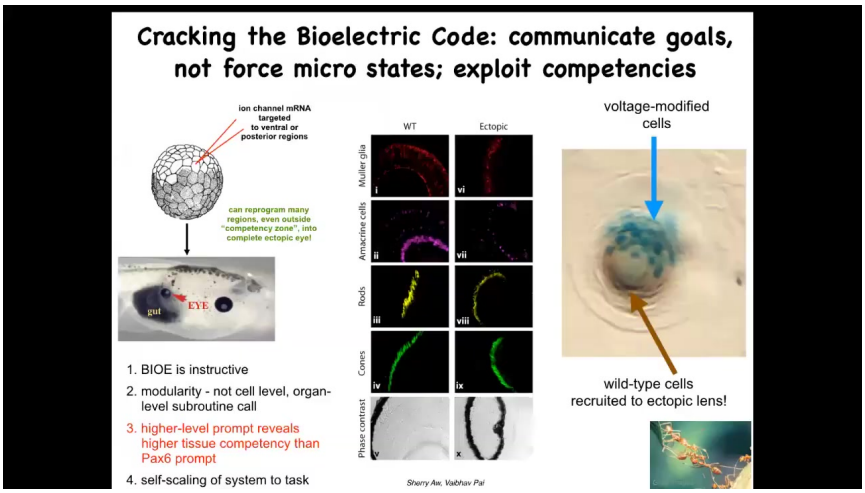
This is a pathological one. What we've done here is inject a human oncogene into these cells. These cells are going to disconnect from the electrical network and go back to a very amoeba-like lifestyle where the rest of the body is external environment to them. You can catch this happening very early on; this is part of our cancer program that I'll mention briefly, and you can immediately see where this is going to happen.

Reading these patterns is important, but again, the most important thing is functional perturbations. What do these patterns do?



Are they actually useful for anything? We developed some tools, everything borrowed from neuroscience, to manipulate the bioelectric states in these tissues. What we do not do is any kind of electric field application. There are no fields, there are no electrodes, there are no frequencies or electromagnetics or waves or anything like that. What we do is we manipulate the native interface that cells are using to hack each other, which means we can open and close these gap junctions to control the topology of the network, and we can open and close the various channels and pumps to be able to control the bioelectric state of each cell. We use pharmacology, we use optogenetics, all the same kinds of tools that neuroscientists use in the brain and nervous system. Now comes the main question: can we use this to ask what information content do these patterns have?

I showed you a moment ago that there's this little spot here that sets up the location of the eye for these cells. This is how they remember where to put the eye.



One thing we were able to do is to take some potassium channel RNA and inject it in a different location of a frog embryo to try to induce that same bioelectric eye-like pre-pattern somewhere else. If you do that somewhere else, those cells get the message, and they do exactly the same thing, and they build an eye. You can do this anywhere. Here's a bunch of cells that were supposed to be gut that have been given the bioelectric memory that says, no, you should be an eye, and that's exactly what they build. If you section them, you can see lens, retina, optic nerve, all the right components.

A few interesting things here. First of all, clearly these bioelectric patterns are instructive. It is not just an epiphenomenon of other things that happen. They actually control what organs get built. That's first. Second, it's incredibly modular. The eye has a very complex structure, lots of different tissue types. We didn't have to specify any of that. All we said was, "build an eye here." Everything else is a simple stimulus that triggers all kinds of complex behaviors down the line. It's incredibly modular.

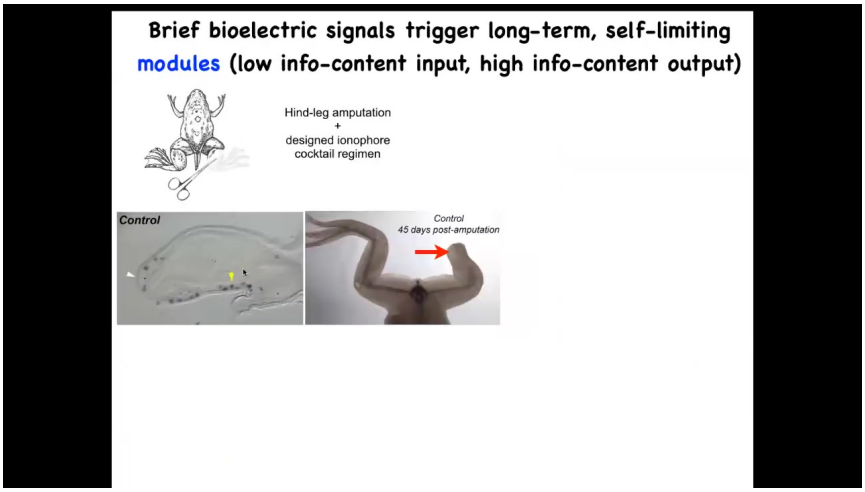
Another thing is that if you read the developmental biology textbook, they will tell you that only these cells up here are competent to make an eye. Only the anterior neuroectoderm can make an eye. That turns out to be a limitation on our competence, not on the cellular competence. This is, again, an important lesson in this basal cognition field: every time you estimate the competency of a system or the intelligence, you're basically taking an IQ test yourself. All you're saying is, this is what I've been able to figure out. If we prompt the system not with a PAX6 so-called master gene, which truly only works up here, but with the more salient bioelectrical pattern, then you can make the eye anywhere.

The final interesting thing about it is that if you only inject a few cells—these blue cells are the ones we injected—this is a lens sitting out in the tail of a tadpole somewhere. If you only inject a few, they recruit their friends. These cells out here, all these brown

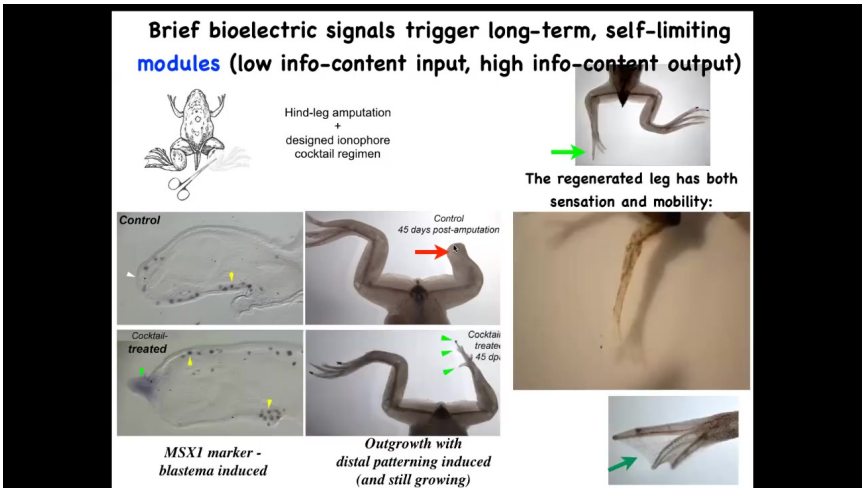
cells, we never touched them. They were secondarily recruited by these guys into this project of making an ectopic eye. We know other collective intelligences that recruit other members of the group to help handle bigger problems.

This is all the work of Sherry Ao and Vaipav Pai in my group. We use these kinds of tools to do all sorts of things in regenerative medicine. I don't have time to go into all the biomedical aspects of it because I want to focus on the cognitive side.

Slide 28 of 50 · Watch at 32:03

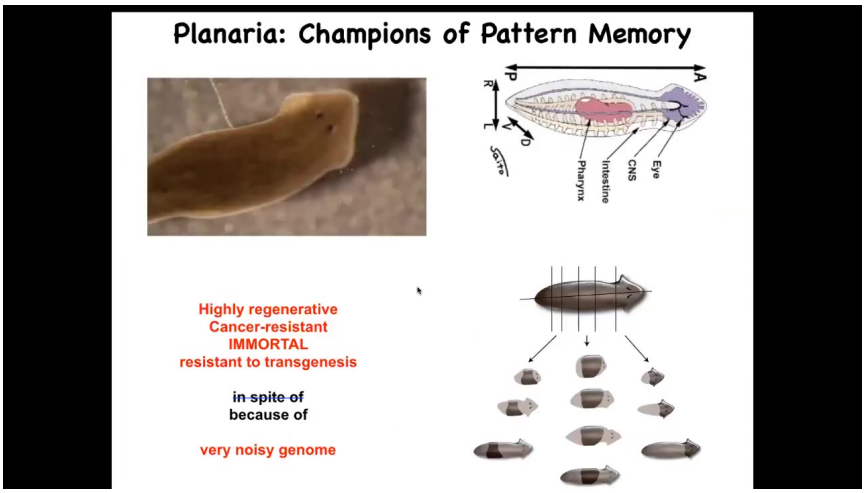


For example, try to regenerate legs. So frogs, unlike salamanders, normally do not regenerate legs. With the appropriate bioelectrical state, you can convince these cells that they should in fact go towards leg growth and not scarring.



And you can see what happens here within 45 days, you've already got some toenails, some toes, you've got a toenail. Eventually, a pretty nice leg is regenerated. And again, this is triggered by a very simple stimulus. We do not try to micromanage this process. This is not scaffolds, 3D printing, stem cells—nothing like that. The goal is to convince the cells right at the first day after injury to go down the path of regeneration, not the path of scarring.

I have to do a disclosure because Morphoceuticals is a company that Dave Kaplan and I founded that is seeking to move all of this stuff, using wearable bioreactors, to mammalian regeneration and human patients.

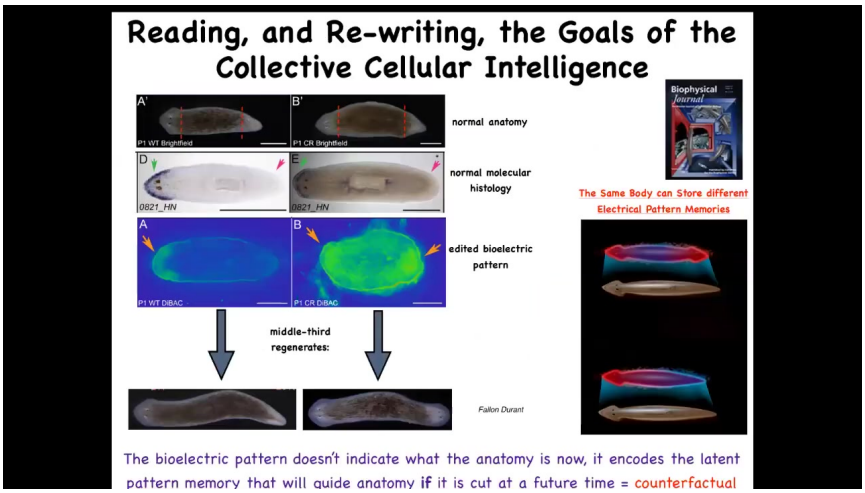


I've made the claim and I've shown you that these bioelectrical patterns are instructive for what it is that the cells are going to build.

I want to specifically show you an example of reading and writing one very specific kind of pattern memory. For that, I will introduce you to this creature. This is an amazing kind of model system. These are planaria, flatworms.

One of the key things about them is that you can cut them into pieces, and each piece will regenerate a perfect little worm. Each piece has everything it needs to build a perfect little worm and stop.

We wanted to ask a simple question. A piece like this, this little middle fragment here, how does it know how many heads it's supposed to have?

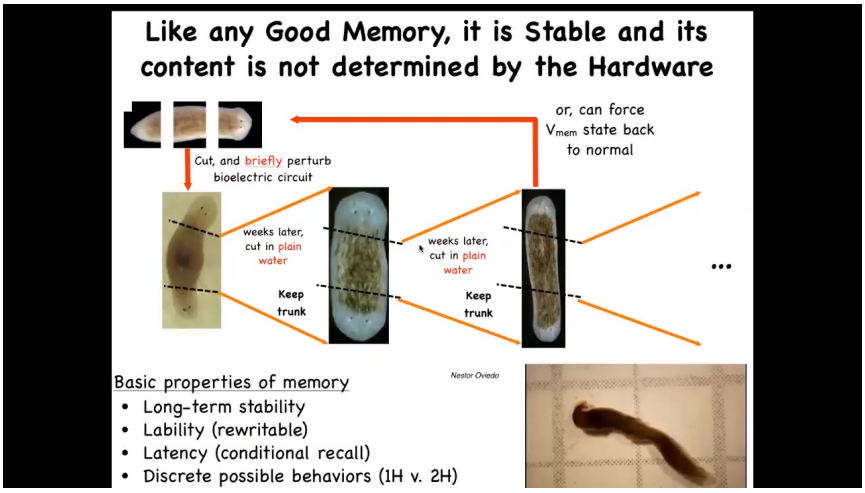


For that purpose, we again studied the bioelectric pattern and we found out that indeed in these fragments, there's a voltage gradient that says one head, one tail. We asked, what happens if we change this and we duplicate it? Here it is. It's a little messy still, the technology is still being worked out, but you can see here what we've done is we've set two heads. If you do that, the cells will build a nice two-headed animal. This is not Photoshop or AI. These are real animals.

The most important thing here is this. The bioelectric pattern is not a readout of this two-headed animal. This bioelectric pattern is a readout of this animal, which anatomically looks perfectly normal. Anatomically it looks normal, one head, one tail. Molecular biological markers are normal, so the head marker is in the head, not in the tail, just like this. So at the level of anatomy and at the level of gene expression, this guy's completely normal. What's not normal about him is that we meanwhile changed his internal representation or memory of what a correct planarian should look like.

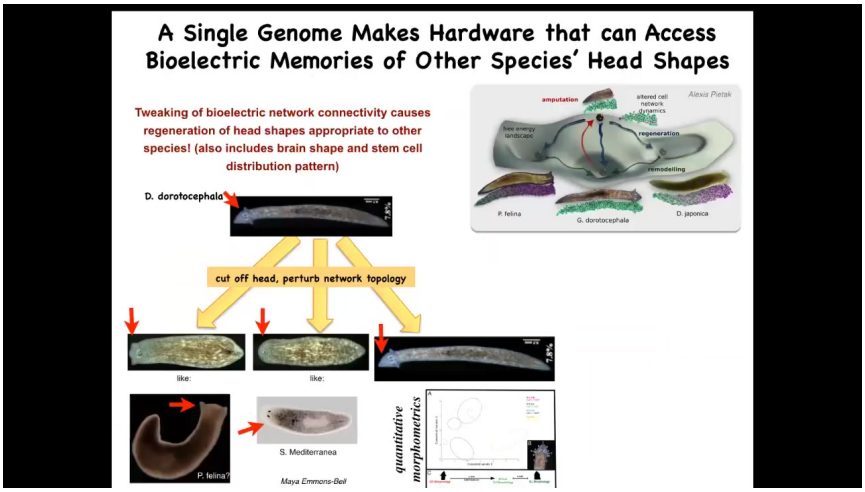
In fact, I think this is the basis of counterfactuals, the basis of this amazing capacity that brains have for time travel, mental time travel, of remembering and foreseeing things that are not true right now. This bioelectrical state is not true right now. For this one-headed animal, this is not correct. It doesn't match. But this is a representation of what it's going to do in the future if it gets injured. If it does not get injured, it simply stays one-headed. But if it does get injured, the cells use this information to decide what it is that they should be building, where in anatomical space they should be going, what is the correct pattern, and then they go ahead and they build this, and then they stop. Because this memory is their reference point of what a correct planarian should look like.

So a standard planarian body can hold at least one of two different representations of what a normal planarian should look like. I think actually many more, but this is a story I'm telling today.



So one reason I keep calling it a memory is that it has all the properties of memory. For example, once you've created these two-headed worms, you can keep cutting them in plain water, no more manipulation of any kind, and you will see that they hold state. Once you've convinced those tissues that the right planarian is two-headed, every fragment will reconstruct that same pattern. The memory is reinforced and duplicated, and they will continue to build two-headed worms in perpetuity. We do know how to set them back. Now we can force the memory back to normal, and you can get them back to one-headedness.

Here you can see these two-headed animals. There's some fascinating behavioral questions to be asked about what it's like to have two brains in one body, and how they behave. What this memory really is: a navigational aid about where in anatomical space they should go. That means we can go beyond just the number of heads.

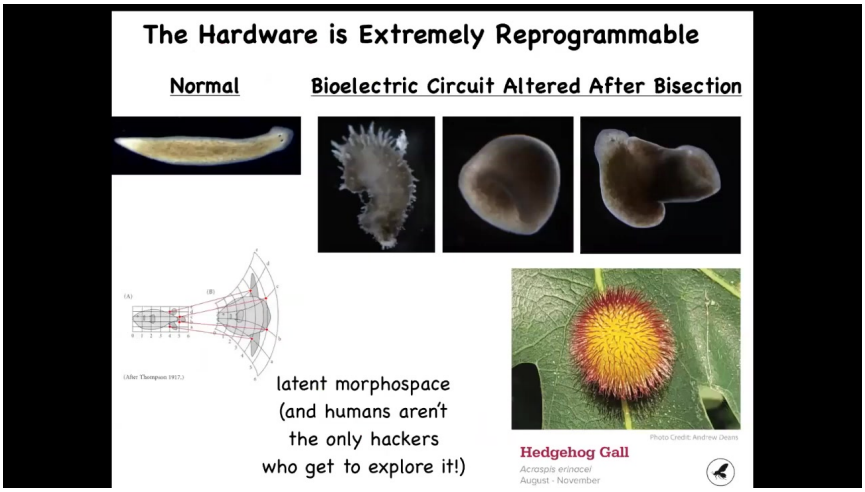


The evolutionary morphospace here contains many different shapes of heads. There are species that have different types of heads, and they live in these attractors in this anatomical morphospace.

The question is, could we get the exact same hardware to visit one of these other anatomical regions? What would it take to convince these cells to build something completely different? It turns out that you can do this by manipulating these bioelectrical patterns.

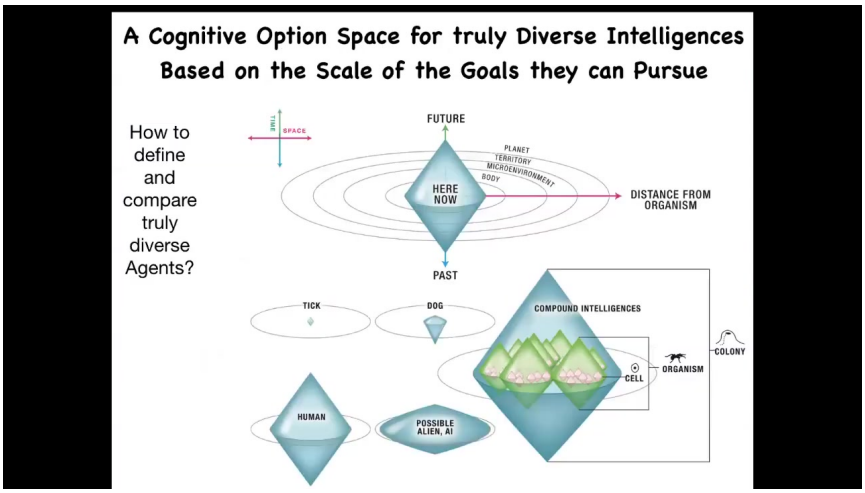
Here's a Dougesia duratocephala with a nice triangular head like this. If you amputate that head and you manipulate the bioelectrical pattern that normally guides them to their normal region of anatomical space, you can get them to go somewhere else. They can make flatheads like a P. falina. You can make roundheads like an S. mediterranean. Of course, they can make their own normal ones. Not just the shape of the head, but actually the shape of the brain and the distribution of stem cells becomes just like these other species between 100 and 150 million years apart.

Let's remember, no genetic changes anywhere here. This is completely stock, out-of-the-box hardware. Their genomes have not been edited. There are no transgenes. The two-headed form, the different head shapes, all of this is handled at the level of the decision-making of how they're going to navigate this anatomical space, and they're able to build all kinds of things.



In fact, you can go much further and visit regions of that space that normally don't belong to planaria at all. You can make these crazy spiky forms, you can make three-dimensional cylinders, these aren't even flat, you can make some hybrid shapes like this. You can imagine exploring this latent space of possibilities as D'Arcy Thompson already foresaw a long time ago.

Human bioengineers aren't the only hackers who do this. This is a hedgehog gall produced by a wasp, a non-human bioengineer that prompts these leaf cells to build something completely different. This is not made by the wasp; this is made by the leaves under a stimulus from the wasp. We would have had no idea that these cells, which reliably 100% of the time build a nice, flat, green architecture, were capable of something like this. This again reminds us that with the right prompts, many systems are capable of behaviors that we have no idea about. It took evolution a long time to make a wasp that has the signals to get it to do that. We are now checking with modern tools of morpho engineering and AI; could we learn to communicate these goals much more efficiently?

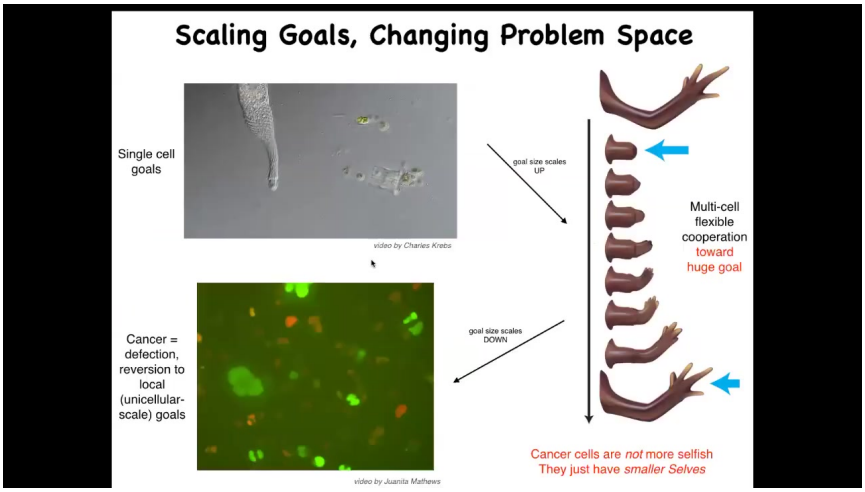


In this last little bit, I just want to pull back a little bit and come back to this broader idea of diverse intelligence. So what we've been looking at is the behavior of a weird and unconventional intelligent agent. It's a collective, not neurons bound together by bioelectricity to solve problems in three-dimensional space, which is what all of us do. It uses those exact same mechanisms, same ion channels, same neurotransmitters, to navigate anatomical space and to solve problems, to have novel behaviors.

Can we come up with a way to think about all of these different diverse agents? What do they all have in common? What is fundamental to being an agent? Whether you're artificial or natural or whatever space you're operating in, what is fundamental? One thing you can think about is, I call this the cognitive light cone, this idea that we can map out — collapse all of space onto one axis, all of time onto this axis, and map out the size of the largest goals that the system is capable of pursuing.

You can imagine that these different creatures have different goals that they pursue. The set points, the size of the set points and the size of the things that stress them and spur them towards new behaviors can be very, very different sizes, but they can all be put on this continuum.

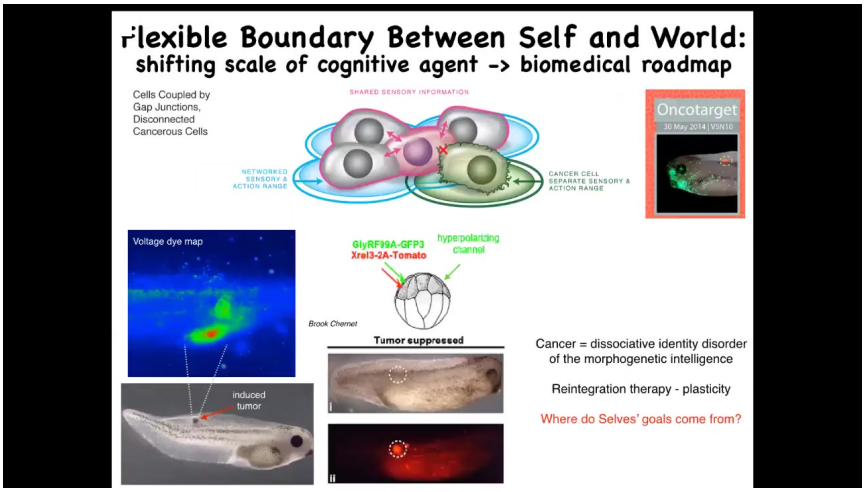
Basically what we have in biology here is that the little tiny, very, very low-scale goals of single cells, things like pH, hunger level, various physiological parameters at the scale of a single cell, can be scaled up in electrical networks to become these massive construction projects. What we're looking at here is how living systems scale their competencies to bigger cognitive light cones, meaning bigger goals in new spaces.



Here, cell evolution allows this incredible scale-up. Instead of just working on maintaining specific states at the single-cell level, these cells are working towards this massive construction project. Many of them, in fact, will die doing this. It's fine because they're not committed to their own individual set points. They're committed to making this journey in anatomical space as a collective.

This process has a failure mode. That failure mode is cancer. If these cells disconnect from the collective, they can no longer remember these enormous grandiose goals. All they can remember are their ancient unicellular goals of reproducing and maintaining certain physiological and metabolic parameters. So that's what this is: human glioblastoma.

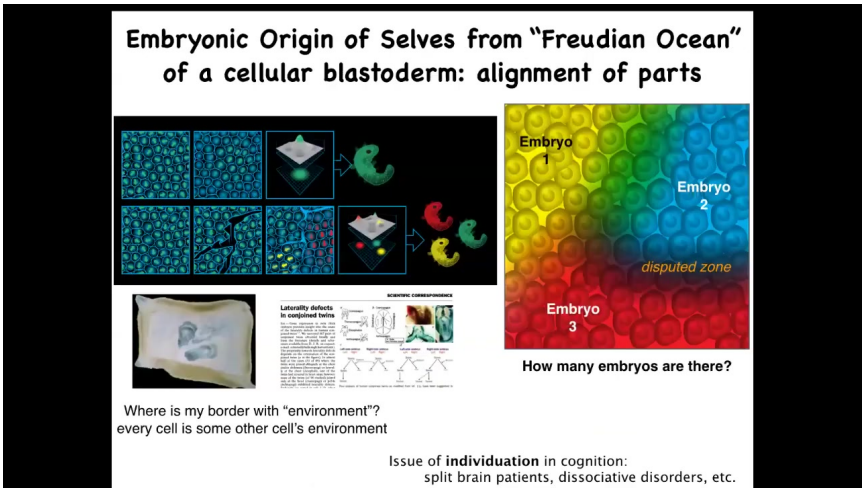
We could also talk about the parallels between this and various dissociative disorders, meaning breakdowns in the unified consciousness of neuronal groups in human patients. That is basically a kind of dissociative disorder of the collective, of the anatomical collective intelligence. That reminds us that these cancer cells are not more selfish; they just have smaller selves. The size of the goals they're working on is much smaller. The boundary between self and world has shrunk. They no longer see the rest of the body as part of themselves. They're just individual amoebas now. That suggests a completely new approach to cancer.



What you can think about is, instead of trying to kill these cells with toxic chemotherapy, you could try to simply reconnect them back to the collective. That's what we did here. We injected nasty human oncogenes, and this includes P53, KRAS, dominant KRAS mutants. Normally they make tumors, but if you co-inject an ion channel that forces these cells to stay in proper electrical communication with their neighbors, and you hyperpolarize them, the gap junctions stay open. What happens is that even though the ONCA protein is blazingly expressed — here it is labeled fluorescently; in fact it's all over the place — there's no tumor. This is the same animal. Because even though the hardware is damaged, being part of the collective allows the cell to keep working on something much larger: these goals of making nice skin, nice muscle, and so on.

This is an example of how thinking about very fundamental questions, how do unified minds arise from competent parts, leads you to a kind of therapeutic roadmap. We're at this point pursuing this in human cells and, hopefully, an approach to cancer medicine.

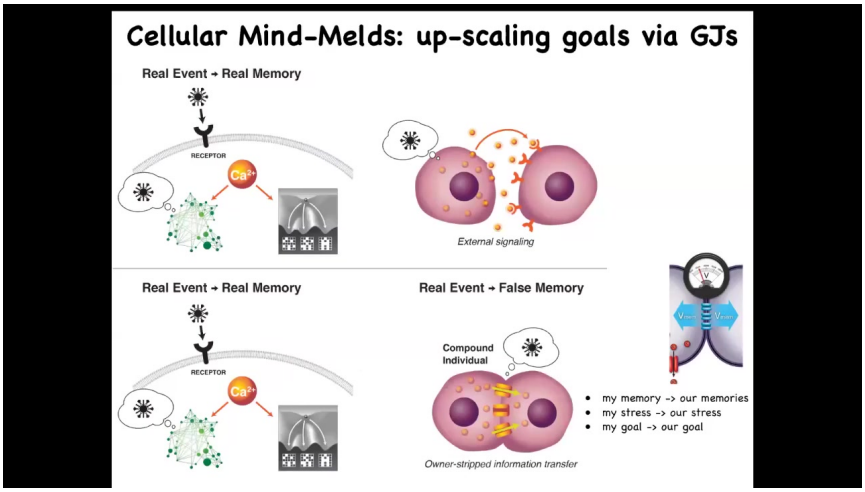
The last thing I want to talk about is where do these goals actually come from? We've talked about homeostasis. We've talked about intelligence as the pursuit of specific goals. We've talked about different sizes of goals. I want to, in the last few minutes, talk about where these goals come from.



To remind ourselves about the scaling of these goals, we normally look at an early embryo and we say there is one individual. I want to remind us that, again, there isn't one of anything. If you look at the blastoderm, there are many, many cells. And in fact, I used to do this as a grad student with duck embryos: if you make little scratches, each of these domains, not feeling the other one for about four hours until it heals up, will self-organize into an embryo. This is how you get twins and triplets, in fact, conjoined multiples. So the question of how many individuals, how many cognitive and functional individuals are in this embryo is not clear. The genetics does not specify. It could be zero, it could be one, or it could be up to about half a dozen, depending on the signaling that goes on and depending on where these systems set the boundaries between self and outside world. Because in every embryo, every cell has some other cell for its neighbor. Where does one embryo end and the other embryo begin? There's some fascinating issues here, again, about individuality. Self-organization of minds as well as bodies.

And I want to mention briefly two mechanisms by which this collectivity happens, because what you need to have here is alignment.

The reason you call all of this one embryo is because all of these cells are in complete agreement about where in anatomical space they are going. They are aligned both physically; they literally get aligned in planar cell polarity, but in terms of their navigation, they are all in agreement about where they're going.

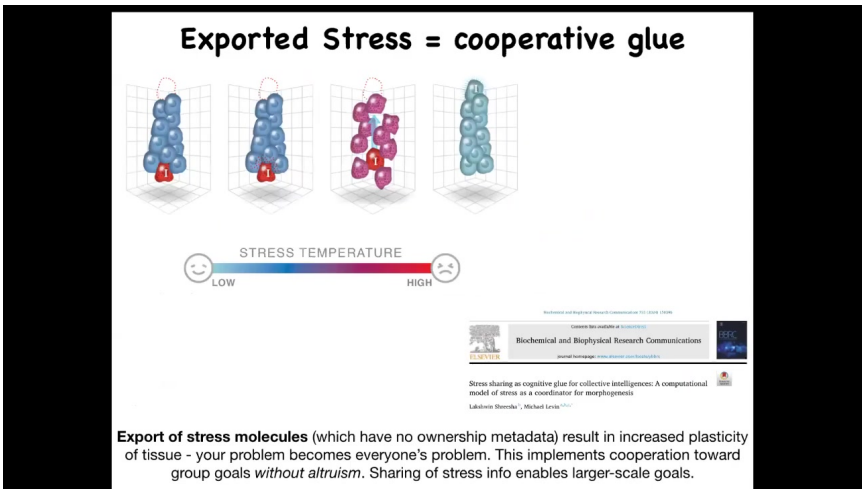


There are two mechanisms that we've studied that allow that to happen. One is anonymization of memories.

Under normal signaling, if this cell sends a signal to that cell, it's very easy for this one to know the signal comes from the outside. They can attend to it, they can ignore it. It's very easy for them to maintain separate memories and separate goals.

But these gap junctions are special. What happens when you form electrical synapses is that there is a direct conduit from the internal milieu of one cell to the other cell. That means it becomes very hard for them to have separate memories. When something happens to this cell, it generates a calcium spike as a signature of the memory of that event; that will propagate to the other cell. It doesn't have any metadata about who it belongs to. The other cell receives that information. At this point, it has no idea: Was that my memory? Was I the one that got poked and had this calcium spike or was that somebody else?

They have a mind meld in the sense that information passes. Now, while this was a false memory according to this cell, it is actually a true memory according to the new collective that is formed here. Sharing memories and a difficulty of keeping individual histories and individual goals is part of what happens when you bind groups of cells, which used to be individual animals, into collections that have a common aligned goal. That's one kind of mechanism.

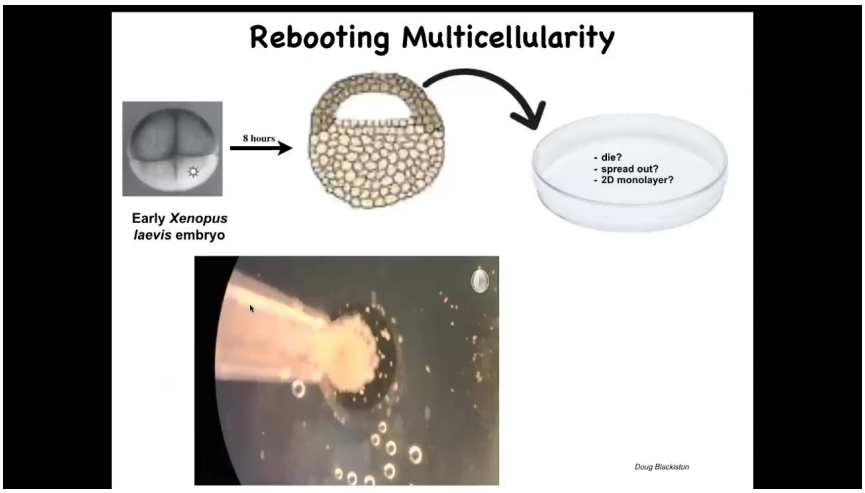


Another kind of mechanism, which is not particularly bioelectric, it is stress sharing. What happens here is that imagine, and this happens in all kinds of spaces, but I'll just talk about physical space here. Imagine that this cell wants to get up here. It's in the wrong location. And it's very stressed out because of it. Cells literally have stress markers that go up when their goals aren't met. If it were just to stay that way, these other cells would never cooperate with it and let it go because they're very happy. Their energy is at the lowest level. They're perfectly happy where they are. They're not moving. What happens is that these cells can now export their stress signal. They leak. They leak the molecules that indicate stress. When that happens, these other cells start to experience it. Much like with what I just showed you, these molecules are all the same. They're all conserved. And so when these cells have it, they are now also stressed. The plasticity goes up. They start to move around because they're no longer sure they're in the right place. And then the cell can do what it needs to do. Put another way, the reason to export stress is that it increases collectivity because your problems, my problems become your problems. By sharing stress, everybody is now motivated without any kind of additional mechanism for altruism. You don't need to be altruistic. You just need to experience some degree of the same stress that I feel in order for the plasticity to go up and everybody rearranges and gets to where they're going. We've done all sorts of computational modeling of this.

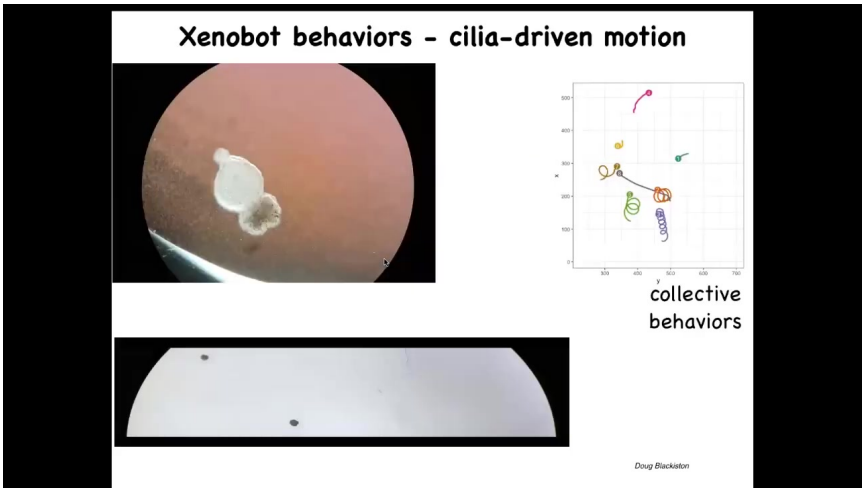
So the very last bit here is this. We've seen new ways to look for intelligent problem-solving and other embodiments. We've looked at new ways to use the bioelectrical interface to communicate, not micromanage, with this morphogenetic intelligence and get large-scale outcomes that would be very hard to try to create bottom-up. And we've looked at a couple of policies for the alignment of parts that I think are part of this cognitive flu mechanism. So the final thing, where do these goals come from? Typically, when we look at any animal or a plant or anything like that, and

we ask where do its behavioral and structural properties come from, you say, evolution, eons of selection. So everything that wasn't that way died out, and so now here it is. In order to test this, we created some new synthetic living materials that do not have a history of selection in this particular embodiment.

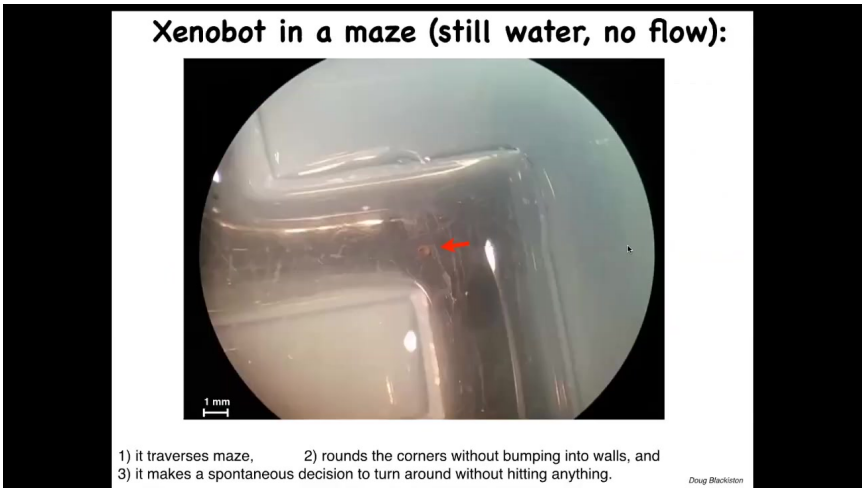
Slide 41 of 50 · Watch at 50:36



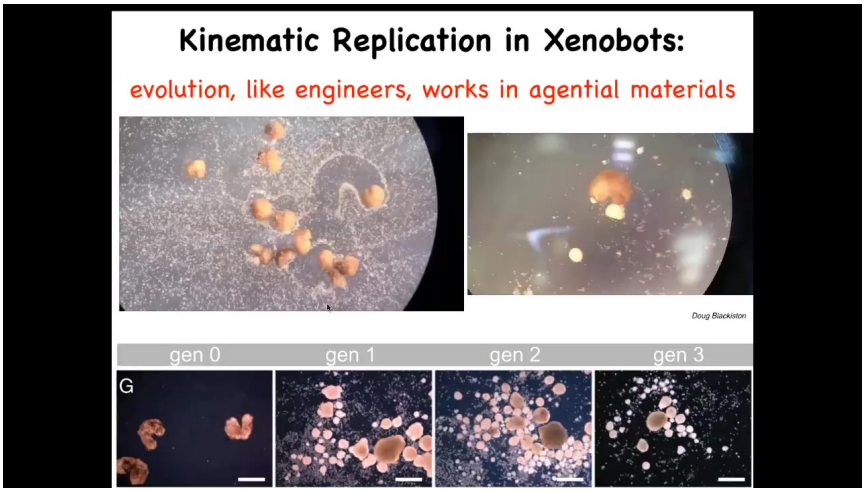
Here's one example. This is the work on Xenobots by Doug Blakiston, where we take some epithelial cells from an early frog embryo. We set them aside, and they could do many things. They could die. They could spread out and walk away from each other. They could form a two-dimensional monolayer like a cell culture. What they do instead is they self-assemble. This is a time lapse. They self-assemble into this kind of a spherical thing. The flashes you're seeing are calcium signaling. This whole thing is just epithelial cells.



And then they do something interesting. They use the cilia that they have on the outside. The cilia are normally used to redistribute mucus down the body of the frog. They start to row against the water and they swim. They can swim back and forth like this. They can go in circles. They have certain collective behaviors where they can interact. They can sit quietly. This one's going on a much longer journey.

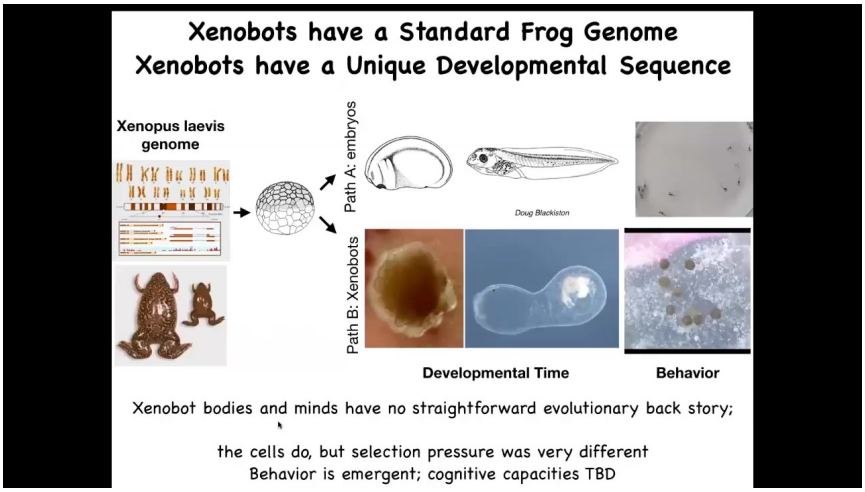


Here is one traversing a maze. What you will see is it swims down here. It takes the corner without having to bump into the opposite wall. Then, at this point, spontaneously, it turns around and goes back where it came from. You don't know why, but you can imagine trying to track the calcium activity during this process to see what is happening when it decides to turn around and then go back. They have these autonomous motile behaviors.



One of the most amazing things they do is they make copies of themselves. How do they do that? We've made it impossible for them to reproduce in the normal ****** fashion. They don't have any of those organs.

But it turns out that if you supply them with loose skin cells, that's what this white stuff is: they're loose epithelial cells, they will run around and collect them into little balls and polish them, both individually and as a group here. And because they're working with these cells, which are also an agential material, not just the passive particles, these little balls mature into the next generation of xenobots. And guess what they do? They run around and they do exactly the same thing, making the next generation and the next. So we call this kinematic replication.

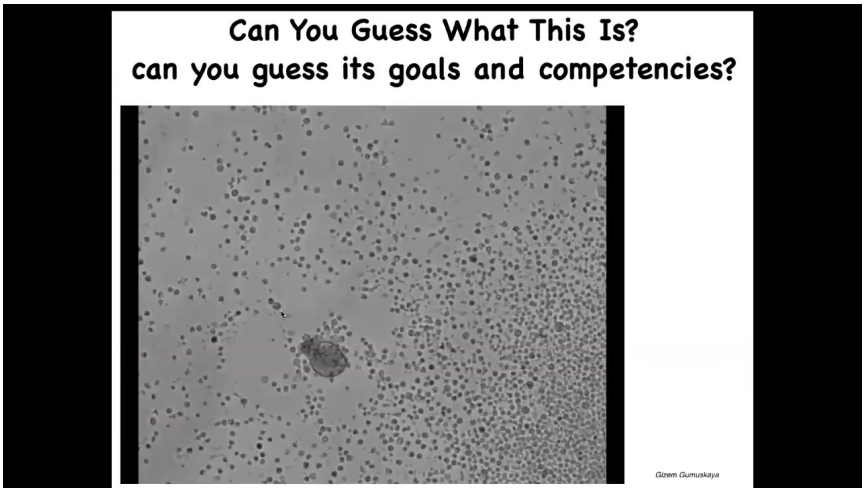


What did the frog genome learn over its time on Earth? Well, it certainly learned to do this. That is what you see if you just look at normal development. They do this developmental sequence, and then you get these tadpoles that swim around.

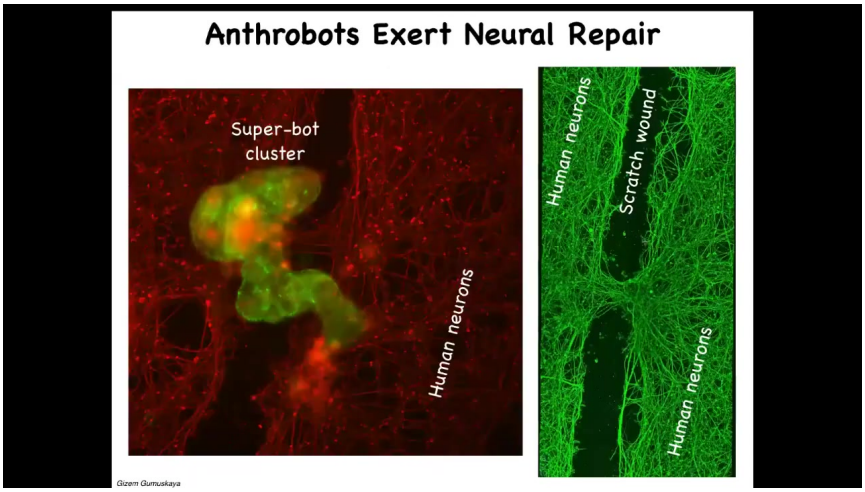
But it turns out they can also do this. What's interesting here is that we did not touch the genome. We didn't add any scaffolds, no synthetic biology circuits, nothing. All we did was to liberate these skin cells from the influence of the rest of the cells. If you take them away from the rest of the cells, you get to find out what they normally do. Under normal conditions, they're basically bullied by these other cells to be like this boring two-dimensional layer that keeps out the bacteria. But they can have a much more interesting life if you get them away from all of this. That is their developmental sequence.

This is an 84-day-old Xenobot that Doug made. We have no idea what it's turning into or what's going on. Where does this developmental program come from? And then they have completely different behaviors like kinematic self-replication. Where is all this coming from? They don't have a history of being selected to be a good Xenobot. There's never been any Xenobots.

We are currently studying their behavioral capacities. Stay tuned for that. Stuff will come out hopefully later this year. I'm not making any claims about that right now, except to say that whatever features they have, a history of selection is not a good answer here. We need to understand.



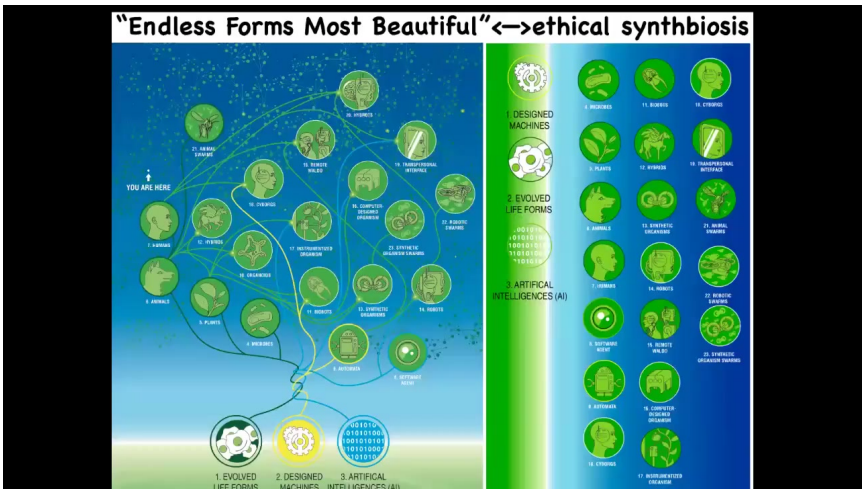
Frogs are pretty plastic, amphibian embryos are plastic. So I'll show you this. And this is the work of a PhD student in my group. And this is, if I ask you what you think this is, it looks like some sort of, a primitive creature that you got from the bottom of a pond somewhere. If you were to sequence the genome of these cells, what you would see is 100% Homo sapiens. These are adult human patient cells. They come from the trachea. They self-assemble into something we call an anthrobot. Again, much like the xenobots, it has its own shape. It has its own behavior that is nothing like what these cells do in vivo. And so then we might ask, we know what tracheal epithelia do in the patient. What do these things know how to do and where would they come from? So we're only now beginning to scratch the surface.



One thing they can do, if you put them on a dish of human neurons which have been scratched, is assemble into a super bot like this and heal, trying to get the two ends to connect. Here's a big scratch wound. Four days later, you pick it up and you see under where the bot was sitting, this is what it did. It took the two sides and healed them together.

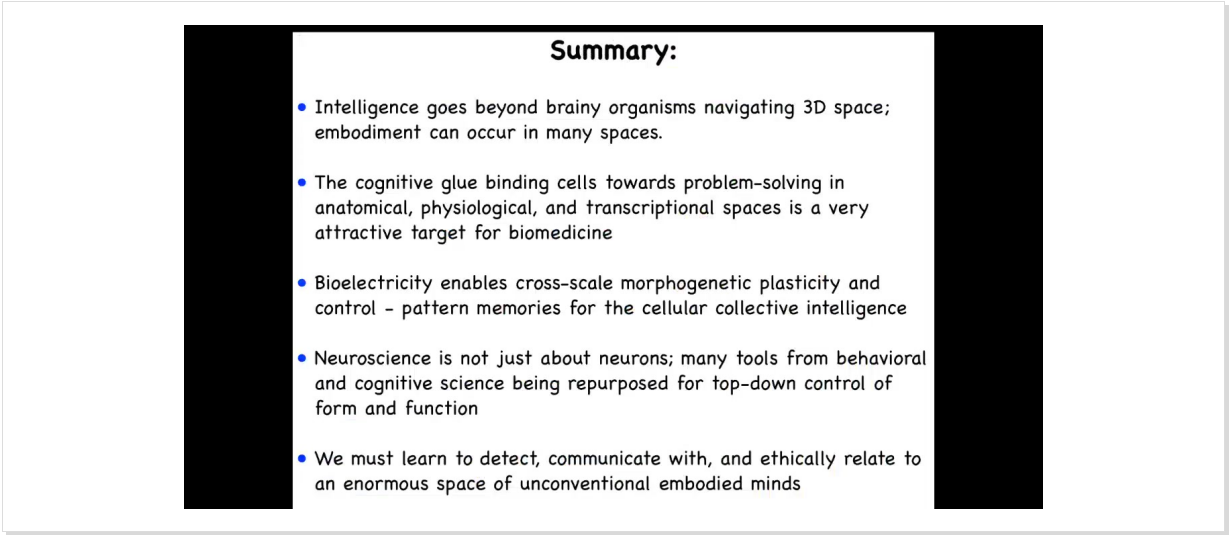
Now, who would have thought these tracheal cells, which sit there quietly in a sheet in your airway for decades, given the opportunity to reboot their life into a new form, would have all these capacities. We've been studying them. They express thousands of genes that are different from what human tissue normally expresses.

Slide 48 of 50 · Watch at 55:50



I'm going to end here by saying that I think that everything that Darwin was looking at when he said "endless forms most beautiful" — all the natural examples of what the material can do — are one tiny corner of this enormous option space of new bodies and new minds. All kinds of combinations of evolved material, engineered material, and software are some kind of agent. Cyborgs, hybrots, biobots, chimeras — all of these things are being made. We will be living with many of them in the decades to come.

We must understand how to ethically have a synth biosis with them. This word was discovered by GPT. I asked it for a word that gets this idea across: that we need to be in an ethical symbiosis with these synthetic life forms. I really think we need platforms to better understand what it means to interact with other intelligences in other embodiments.



What I tried to argue today is that intelligence goes beyond brainy organisms in 3D space. This occurs in many spaces that are hard for us to visualize.

That cognitive glue that binds these pieces together is a really attractive target for biomedicine, in particular bioelectricity, and shows us that neuroscience is not just about neurons. Many of the tools and concepts, everything from active inference and perceptual monthly stability and memory and learning, can be repurposed for some amazing applications in biomedicine, in robotics, and so on.

But the broader issue here is that we really need to develop the field of diverse intelligence. This is telling us that one of the ways we need to mature is to get better at discovering and relating to these unconventional intelligent agents.

If you're interested in more details, these are some of the papers that dive into it.

I need to thank all the people who did the work. This is Doug Blackiston, who did all of the biology for the Xenobot models. Vive House Pai and Sherry Au did the bioelectric eye induction. Brooke Cherenet is cancer and of the bioelectric control of cancer. Nirosha Murugan did some great work on the frog leg regeneration project. Fallon Durant did the pattern memories, the latent pattern memories in Planaria, and Gizem Gomushka is the anthrobots.

I'll thank all of these people, our many collaborators, all the other folks in the lab, our funders who have been supporting this work. Jeremy did all the illustrations, and here are the disclosures. These are three companies that have supported our work, and as always, all the heavy lifting is done by the animal model systems. Thank you very much, and I will stop there.

Thank you for reading.

More lectures

You can find more of my lectures here.

Follow my work

Twitter • Blog • The Levin Lab

Want one for your lecture?

Want something like this for your own talk? Reach out to Adi at adi@aipodcast.ing.