

A LECTURE COMPANION

---

# **"Beyond complexity & emergence: taming multiscale competency in agential materials" by Michael Levin**

---

Michael Levin

Recorded on July 19, 2025

## About this document

This document is a companion to the recorded lecture "*Beyond complexity & emergence: taming multiscale competency in agential materials*" by Michael Levin, recorded on July 19, 2025. 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 ~54 minute talk given at a workshop on complexity. Many of the examples are the same ones I often talk about but this talk is focused on explaining how biology deals with complexity, and ways to think about complex patterns of form and behavior beyond the conventional approach of emergence.

---

## Follow my work

[Twitter](#) • [Blog](#) • [The Levin Lab](#)

---

## Transcript note

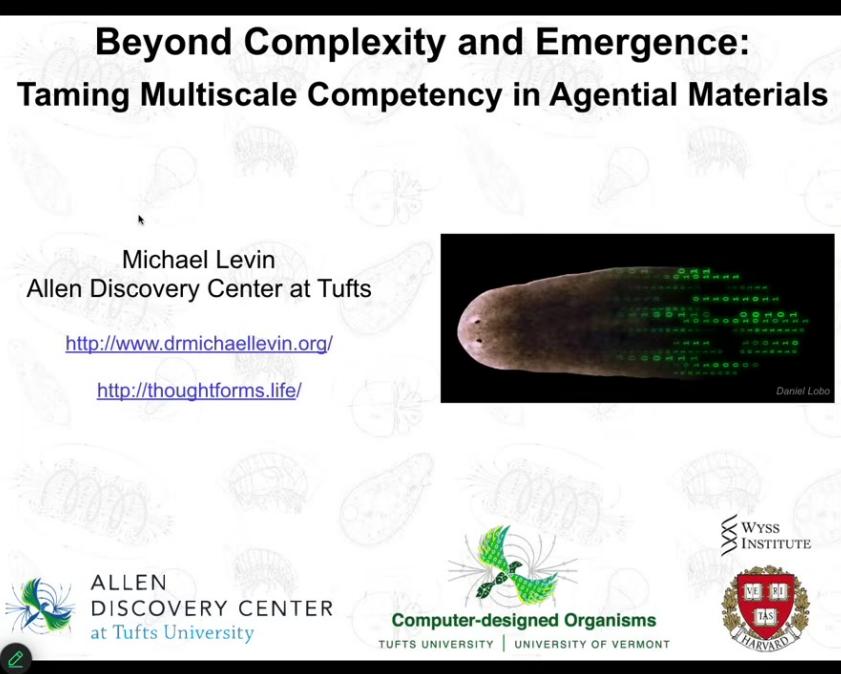
This transcript is generated automatically. While we strive for accuracy, occasional errors may occur. Please refer to the video for exact wording.

---

## Want one for your lecture?

Want something like this for your own talk? Reach out to Adi at [adi@aipodcast.ing](mailto:adi@aipodcast.ing).

## Beyond Complexity and Emergence: Taming Multiscale Competency in Agential Materials



Michael Levin  
Allen Discovery Center at Tufts

<http://www.drmichaellevin.org/>

<http://thoughtforms.life/>



Daniel Lobo

ALLEN DISCOVERY CENTER at Tufts University

Computer-designed Organisms  
TUFTS UNIVERSITY | UNIVERSITY OF VERMONT

WYSS INSTITUTE

HARVARD

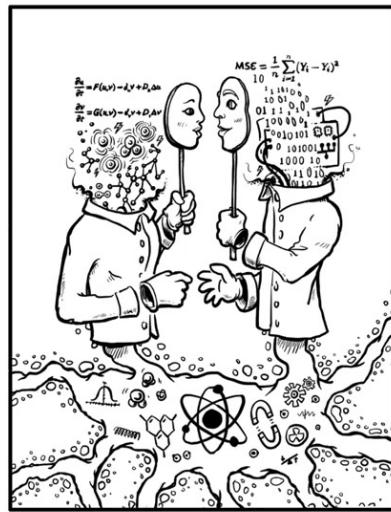
I'm going to say some things about complexity and how I think biology handles complexity. I'm going to say a few things about emergence and some unconventional and unpopular things about ways to think about this.

Generally, my lab is three quarters wet lab biology and about one quarter computational people. We straddle this interface between computing and the life sciences. You can find all of the data sets, the software, the primary papers, everything is downloadable at our lab website. This is my personal blog where I write about what I think some of these things mean.

What I'm going to try to convey today are three things.

## Overview: beyond neuromorphic architectures

- Some amazing, unconventional examples of intelligence in biology - intelligence all the way down
- Evolutionary and mechanistic aspects of communicating with, not micromanaging, multi-scale agential materials
- Origins of embodied intelligence and what it means for biomedicine and engineering for minds (AI)



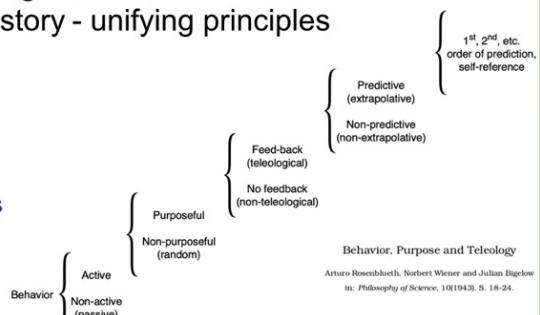
That we need to get beyond neuromorphic architectures in our attempts to understand how intelligence can handle complexity.

First of all, I'll show you some amazing and unconventional examples of intelligence you wouldn't have seen in standard introductions to biology. What it's telling us is that life uses intelligence all the way down to the lowest scale or the lowest level of organization. I'm going to highlight a few of the lessons that we've learned from all this, and in particular, evolutionary and mechanistic aspects of how we communicate with, not micromanage, the multi-scale agential material of life.

The complexity of the living organism makes it basically impossible to micromanage in the same way that we engineer with other types of materials, but that's okay. That's not how life regulates itself, and we can do better and we can do what life does. I'll show you what that is. Toward the end, I'm going to say speculative things about the origin of embodied intelligence and what it means both for biomedicine and for AI.

## My Framework Goal:

- Recognize, create, and relate to truly diverse intelligences regardless of composition or origin story - unifying principles
  - familiar creatures - us, apes, birds
  - weird creatures (colonial organisms, swarms)
  - synthetic biology - engineered new life forms
  - AI (software or robotic)
  - exo-biological agents (Earth is N=1)
  - patterns within media
- moves experimental work forward - new biomedical and synmorpho capabilities, better ethics



Behavior, Purpose and Teleology  
Arturo Rosenblueth, Norbert Wiener and Julian Bigelow  
in: *Philosophy of Science*, 10(1943), S. 18-24.

frontiers  
in Systems Neuroscience  
APPLIED AND THEORETICAL  
IN SYSTEMS NEUROSCIENCE  
"Technological Approach to Mind Everywhere: An Experimentally-Grounded Framework for Understanding Diverse Bodies and Minds"  
Michael Lenzen et al.

One of the key things that I try to do in our work is to develop a framework that allows us to recognize, create, and relate to really unconventional intelligences.

I'd like to be able to understand what they all have in common, regardless of their composition or their origin story. I seek unifying principles around familiar creatures such as primates and birds and maybe an octopus or a whale, but also very weird creatures: colonial organisms, swarms, cells, tissues, synthetic life forms, of which I'll show you some—AIs, whether software or robotic—and perhaps someday exobiological agents, alien life. I also want to include some very strange things that I probably won't have time to talk about today that aren't physical beings at all.

What I'd like to do is to develop frameworks that are not just philosophy. They need to move experimental work forward and show how specific philosophical ways of looking at these things actually lead to progress, both in biomedicine, which is a lot of what we do, but also in bioengineering, in computer science, and, very importantly, in better ethics.

I'm not the first person to try for something like this. This is Wiener et al.'s work in 1943, where they tried to show, in terms of cybernetics, the steps from passive matter to human-level metacognition.

The details of my work are here in this fairly lengthy paper.

## Outline:

- Unconventional biology
- How does it work
- What does it mean

Unique Features of the Biological Substrate:  
a tour of the agential material of life  
-> functional complexity beyond brains

So let's talk about some of the unique features of the biological substrate. And this is important in establishing how far away from brains we need to go to really understand what's going on. The first thing I want to say is that biology presents a face to the outside world for the dominance of genetics. And the idea is that the molecular biology of the genome should really tell us everything we need to know. And the promise is, of course, that once we really tame the genetic information, we'll be able to do whatever we need to do. So I just want to explain how things are actually quite different.

This is a baby axolotl. Axolotls are amphibians that have little tiny forelegs in this early stage. This is a tadpole of the frog, *Xenopus laevis*, which we use all the time. They do not have legs in the early stages. In my group, we make something called a frogolotl, where we combine cells from axolotls and frogs. I could pose the following question. You've got the axolotl genome. It's been sequenced. We know we can read it. It's been annotated. We have the frog genome. We know all the genetic information from this species. Simple question: if I make a frogolotl, are frogolotls going to have legs or not? And if they do have legs, are those legs going to be made of frogolotl cells, or will they incorporate some *Xenopus laevis* cells?

The important thing to note is that we currently have no way of deriving that kind of information from the genetics. The genetics specify the hardware that all of the cells in these creatures get to have, the protein-level, nano-level hardware. What they don't tell you is the actual software which drives outcomes. In other words, the computations and the decision making that happens at different scales once that hardware is

powered up. We have to understand that knowing the specifications of the hardware, for example, in terms of the genetics, is not sufficient. These forms, while they're pretty reliable by default, are complex.

Slide 5 of 43 · Watch at [5:35](#)

**Radical Sensory-motor System Changes are Accommodated with no New Rounds of Adaptation**

*Douglas Blackiston*

**Ectopic eyes on tail provide vision!**

**Brain dynamically adjusts behavioral programs to accommodate novel body architectures**

**no evolutionary adaptation needed (because embryos can't take much for granted, have to solve on-the-fly: evolution makes problem-solving agents)**

**Behavioral Testing Device**

There's actually incredible plasticity that's there. Here's a tadpole of that same frog, here is the mouth, here are the nostrils, here's the brain, the spinal cord, and the gut is back here. What you'll notice we've done is we've prevented the primary eyes from forming, but we did put an eye on its tail. This eye, here we've traced it with a red lineage label. You can see that it makes an optic nerve. The optic nerve does not go to the brain. It sometimes synapses on the spinal cord, sometimes on the gut, sometimes nowhere at all. By using this machine that we built to automate the behavioral training of these animals using light cues, we found out that they can see perfectly well.

This is their eye; it does not connect to the brain. They can see. They can learn in visual tasks.

This is a puzzle because typically you would think that a novel sensory-motor architecture would need mutation, selection, evolution. You would need some

adaptation to do this. It works out of the box. The very first time you make these animals, they never existed before. No problem. They can see; the brain somehow interprets signals arriving from other pathways, folds it into their behavioral repertoire. Already we're starting to see that there's functional plasticity here that speaks against some kind of a hardwired system where the genetics just tells every cell where to be. We know that that's impossible.

Slide 6 of 43 · Watch at [7:02](#)

### Information Moves Across Media

newly regenerating brain can get information from rest of body

© 2016, Published by The Company of Biologists Ltd | *Biology Open* (2016) 5, 1177–1188 doi:10.1242/bio.020149

**HYPOTHESIS**

Vertically- and horizontally-transmitted memories – the fading boundaries between regeneration and inheritance in planaria

Moran Neuhofer\*, Michael Levin\*, and Oded Rechavi\*\*

We can go on and notice that this plasticity is actually much deeper than that. Here is a creature called the planarian flatworm. These little guys have many fascinating properties. Among them, they are highly regenerative. If you cut them into pieces, each piece will regrow exactly what's missing and make a perfect little worm. But they're also smart. They can learn.

This is what we did. We used a laser to etch little bumps into one region of a Petri dish. We would feed them in this one region. That's called place conditioning. They would basically learn that this is where they should eat. And then we cut off their heads containing their centralized brain. The tails sit there doing nothing, no behavior, until they grow back their brain. About 8 or 9 days later, they grow back their brain,

behavior returns, and now you find out that they actually remember the original information.

This was first discovered in the 60s by this guy named McConnell, but he did everything by hand, and he got a lot of flack for it because nobody thought that you could store information outside the brain. But we did it with this automated device that I just showed you, and he was right. It does work.

What this tells you is that not only is there information stored outside the brain, but that information can be imprinted onto the new brain as it develops. Behavioral information is actually moving through tissue being imprinted onto a new structure as it develops.

Here's another example that I think is pretty cool.

---

Slide 7 of 43 · Watch at [8:37](#)

### Morphogenetic Virtualization



*Goniurellia Tridens*

*Lauren Zirate 2014*

*Hessam Akhlaghpour*

Bodies are programmable and thus support virtualization.  
Here's a fly running a stripped-down "ant" morphogenetic program on its wing

It's amazing, but no more so than the actual fly anatomy!

So this is a fly. This is not AI or Photoshop. The first time I saw this, I thought this was a fake thing. What it's doing on its wings is running a stripped down morphogenesis program in two dimensions for ants. So it's got these images of these other creatures on its wings, and it uses them to scare off predators because nobody wants to deal with

the ants. But it's remarkable, this degree of virtualization, you can run this simple morphogenetic program on the wings, but then there's a much more complex, one three-dimensional program that is actually giving rise to the rest of the fly. And what we're starting to see here is really these software aspects to the material, where they don't just do the same thing every single time, but they have the ability to embody all kinds of interesting patterns of form and behavior to the point where we can now see very specific context-sensitive, flexible error minimization skills.

Slide 8 of 43 · Watch at [9:37](#)

## Context-Sensitive, Flexible Error Minimization

Normal

Picasso-like

development

"as needed" remodeling

Laura Vandenberg

Craniofacial mispatterning

Metamorphosis

Morphometric analysis and modeling  
reveals: faces fix themselves!!

dorsal profile

misaligned eyes

A B

Genetics does not specify hardwired rearrangements: it specifies a system that executes a highly flexible program that can recognize unexpected states and take corrective action.

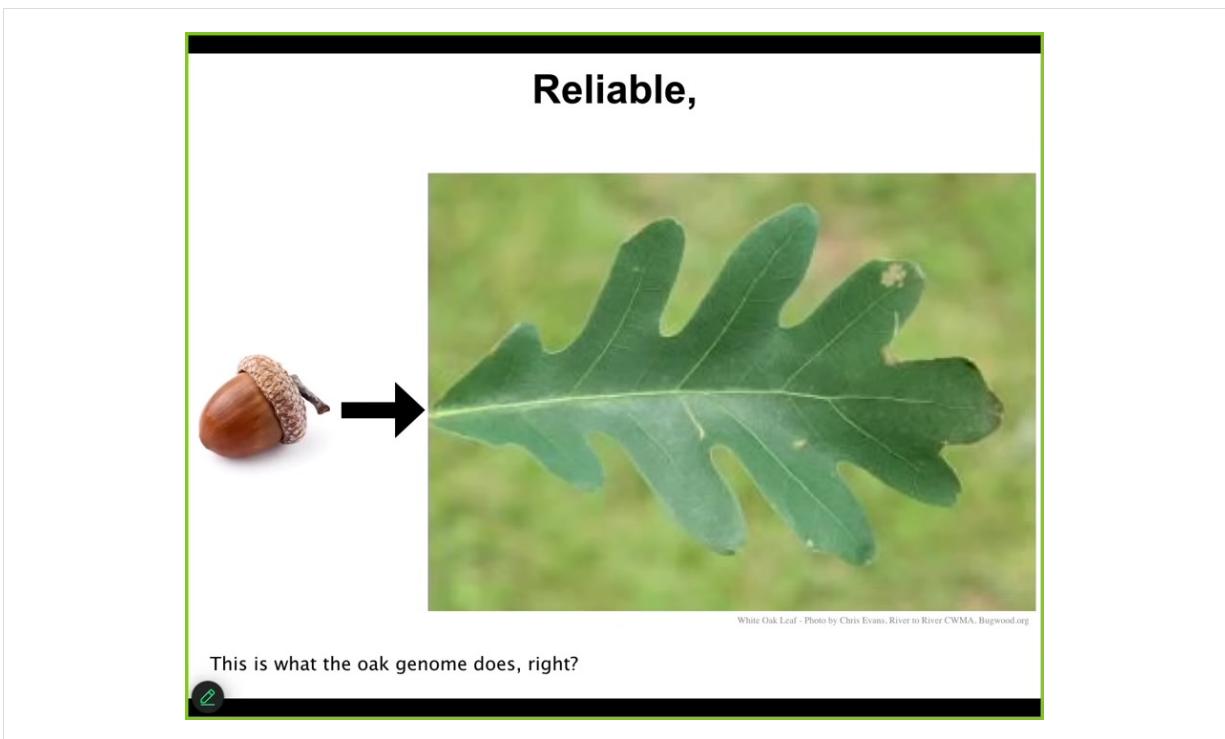
For example, here's a tadpole. This one has its eyes and its nostrils and so on. In order to become a frog, this tadpole has to rearrange its face. The eyes have to move, the jaws, everything has to move. It was thought that this was simply a hardwired process. Somehow, all the organs were told which direction to go and how much, and then you get from a normal tadpole to a normal frog. We tested this because it's a strong claim of mine that the intelligence of these kinds of processes cannot be guessed or assumed from an armchair, that you have to do experiments. Sure enough, when we asked how much plasticity this has, we created these so-called Picasso tadpoles. Everything is scrambled. The eyes on top of the head, the mouth is off to the side. It's like a Mr.

Potato Head doll put together wrong; everything is in the wrong place. When you do this, you find that they make perfectly normal frogs because everything will move around in novel paths, sometimes moving too far and having to come back until you get to a normal outcome.

This is very interesting. What the hardware gives you is not a rote set of movements. It specifies an error minimization scheme relative to a set point. We'll talk in a minute about how that set point is encoded. What it's giving you is a system that can start off in different configurations, unexpected abnormal configurations, and still get the job done and get to the right location in anatomical morphospace. This is one definition of intelligence by William James. It's the ability to reach the same goal by different means in some arbitrary problem space. We now understand that the material these things are made of is extremely flexible. It does all kinds of problem solving.

---

Slide 9 of 43 · Watch at [11:35](#)



As any problem-solving agent, however, it is also hackable, which has good news and bad news. To show you one example from the plant world, this is an acorn. Acorns very reliably make oak leaves. And so you might think that this is what the oak genome does. It encodes this particular kind of shape.

## but Hackable



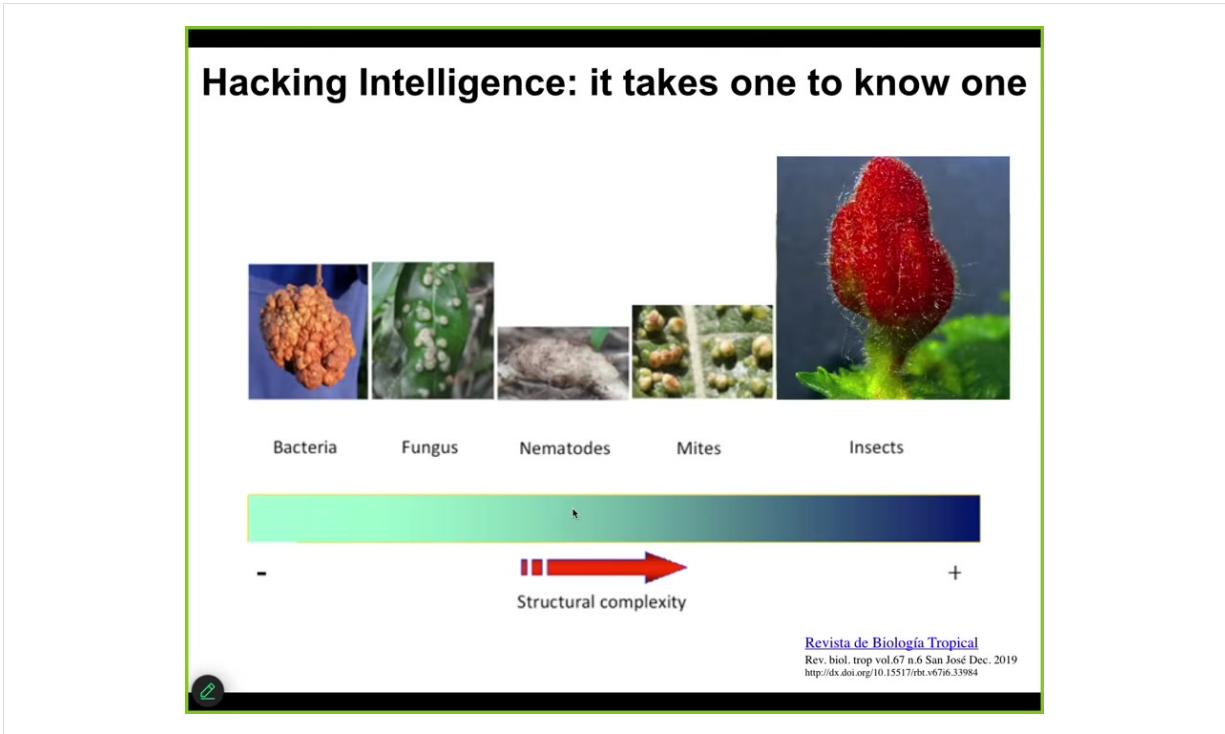
**Hedgehog Gall**  
*Acraspis erinacei*  
August - November



Bio-prompting by wasp parasite to hack competent host  
Biology exploits reprogrammable hardware



Except that along comes a bioengineer, a non-human bioengineer, this little wasp. And what the parasite does is leave some prompts on this tissue that basically hack the morphogenetic competencies of the cells and cause them to build something completely different. So we would have no idea that these flat green things could build a complex gall structure like this round, red, yellow, spiky thing, or this, because under normal circumstances, they always build this leaf. So it's clear that the material has all kinds of capabilities that are very hard to guess in advance. You have to do a lot of work. And we have a lot of ongoing projects in our lab to automate the process of discovery. Presumably it took millions of years for the wasp to be able to do this. We want to be able to do it quickly. And again, as I'm going to emphasize, not by micromanaging the cells, not by genomic editing, not by telling individual cells what to do, but by high-level prompts.



It's a two-way IQ test. In other words, the complexity of the thing you want the material to build is dependent on our or the hacker's level of sophistication.

So bacteria make these featureless blobs; fungus doesn't do that much better; nematodes are still simplistic, but mites are starting to get better. By the time you get to insects, you get this beautiful structure that they can get the leaf to do.

So we have to learn ourselves. We have to improve our own capabilities before we can get the material to do the cool things that we want it to do.

## Solving Novel Problems with High-N Effectors:

The diagram illustrates the process of head regeneration in planaria. It shows a sequence of five planaria, each with a head containing two eyes. The first planaria is in a blue box labeled  $H_2O$ . The second planaria, labeled 'Refreshed  $BaCl_2$ ', has its head removed and is labeled 'K<sup>+</sup> channels blocked: Head degeneration'. The third planaria is labeled 'Adaptation & regeneration: Altered ion channel, transporter, and gap junction expression'. The fourth planaria is labeled 'Barium-adapted head!'. The fifth planaria is a restored planaria with a new head containing two eyes. To the right of the diagram is the text: 'Small number of genes regulated out of entire genome, all animals find the same solution!' and an image of a fantastical landscape titled 'Moravec Bush - 20K Effectors'.

- planarian heads degenerate after exposure to barium
- planaria eventually adapt and regenerate heads that tolerate barium
- a relatively few transcripts were altered to produce barium tolerance
- how did the system choose exactly the right genes to modulate, to deal with this evolutionarily-novel challenge?

The last thing I want to show you is a little bit of a problem solving, another example of problem solving in a different space, this time not anatomical space, but a different space that is actually very relevant for robotics. So here's the experiment. We have these flatworms. We put them in a solution of barium chloride. Barium is a non-specific potassium channel blocker. It blocks all the potassium channels. Most of the cells freak out, especially the cells in the head. Over the next 18 hours, their heads literally explode. But if you leave them in the barium, refreshing the barium, what you will find is that fairly quickly they will build a new head, and the new head is completely barium adapted. The new head doesn't care about barium at all. No problem.

We thought that was remarkable, and so we did some transcriptomics, meaning asking what genes do these barium adapted heads express that standard heads do not? What happened transcriptionally in terms of gene expression? How did these heads become barium adapted? We found a couple things. First of all, it really is only about a dozen genes. Individual worms all find the same solution to this problem. They all find the same gene expression.

Here's the amazing thing. Imagine you've got this physiological stressor that you've never seen before. Planaria are not typically exposed to barium in the wild, so there's no reason to think there's any kind of evolved response to this. You've got this incredible physiological stressor that's totally novel. You have a library of 20,000 effectors, roughly. Those are the genes that you could turn on and off. Tens of

thousands of potential actions that you could take, meaning steps that you could take in transcriptional space. It's a high-dimensional space. It's a little bit akin to Hans Moravik's Moravik Bush, which is a robot with a very high number of effectors. He was asking the question, how do you design a controller for something with a huge number of effectors? We don't really have efficient mechanisms, algorithms for that. But this is what the system is doing. It's faced with a novel problem in one space, physiological space. It has actions that it could take in transcriptional space. In a very rapid order, it identifies exactly the dozen or so effector firings that it has to do to solve this problem. It's a remarkable example. There are many others. We could do this for hours. I could show you all kinds of wild examples. This material is really good at improvising novel solutions to problems it hasn't seen before.

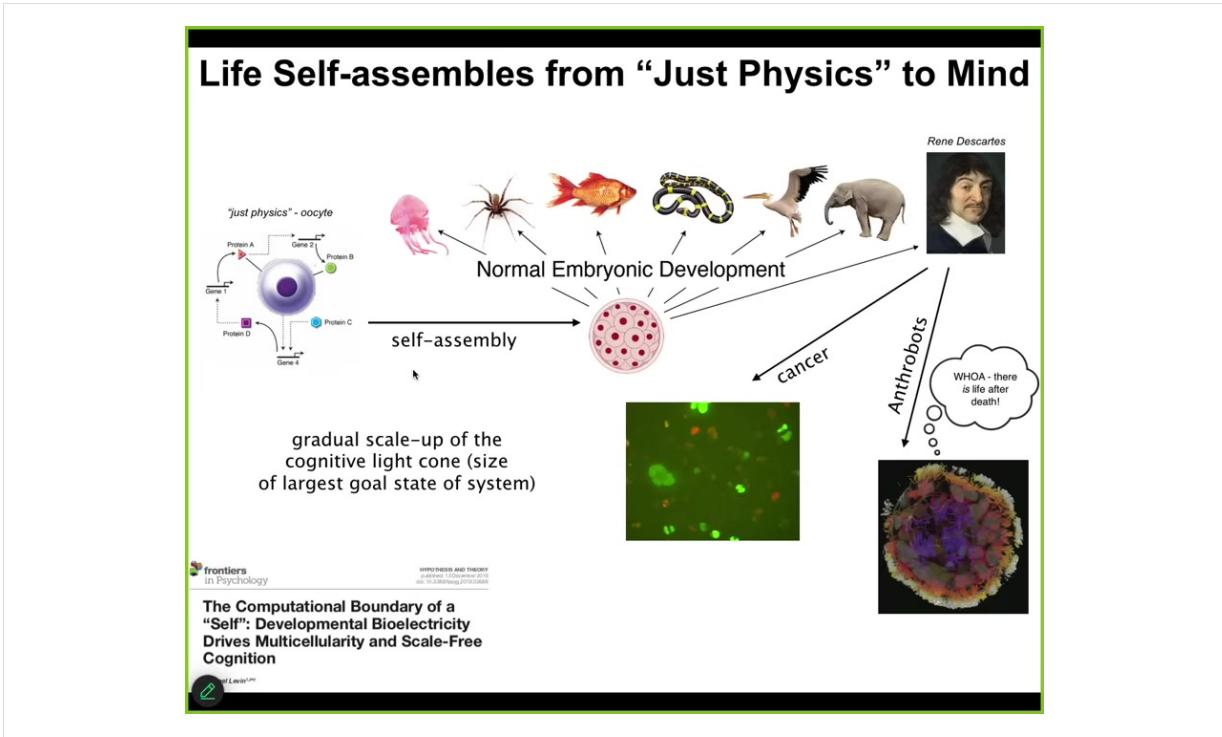
---

Slide 13 of 43 · Watch at [16:05](#)

## Outline:

- Unconventional biology
- How does it work
- What does it mean
- Conventional story of genetic hardwiring
- Self-construction, emergent goals
- Multiscale competency architecture
- Cognitive glue
- Polycomputing

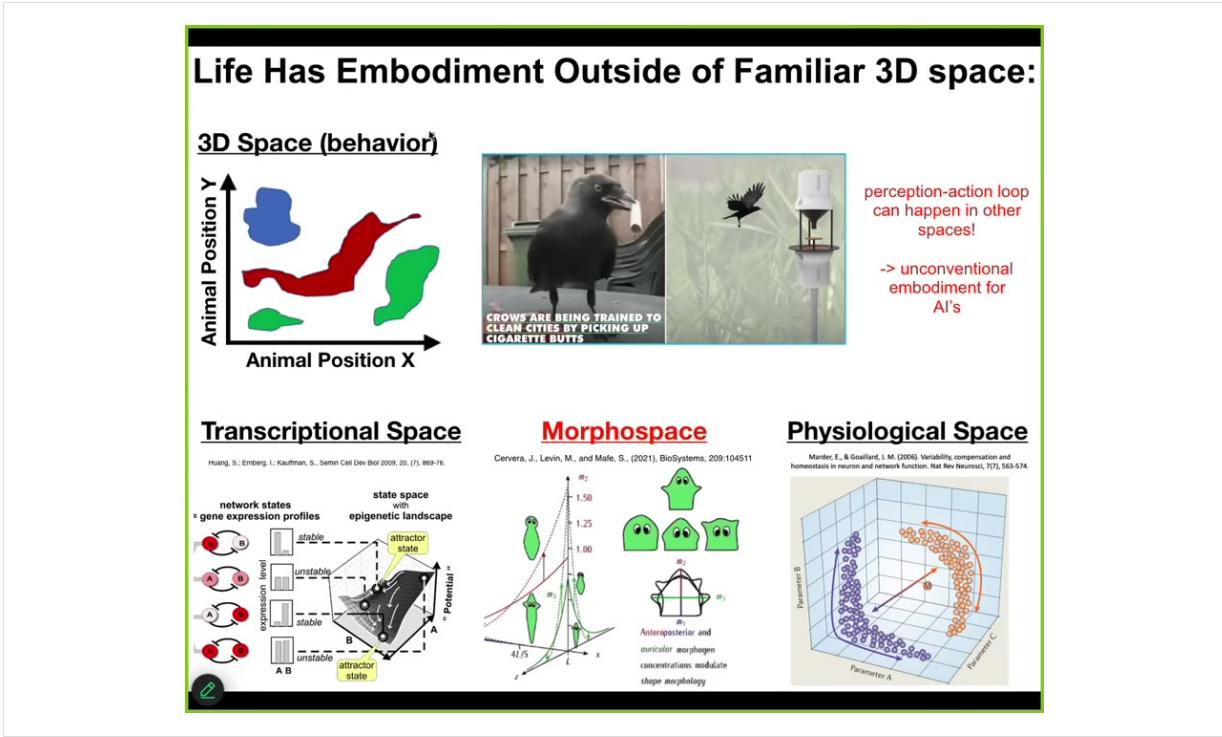
So what we want to do now is talk for a little bit about how this actually works. I'm going to try to explain what's going on under the hood to the extent that we know. There are many open questions here.



The first thing is that Life really self-assembles from something that people would call just physics. Unfertilized oocyte, little bag of chemicals — people look at and say, that's not a cognitive system. That's squarely within the province of biochemistry. That's a machine. It's mechanical; it follows mechanical rules. Fine, but slowly but surely that thing transforms itself into something like this or even something like this that's going to make claims about not being a machine and having true intelligence and all of that.

The one thing that we definitely learned from developmental biology is that there is no magic lightning flash. There is no special time at which you click over from being a mechanical system into having a mind. This is a slow and gradual scaling process. This is not the end of the line, because there are other processes that can still happen that are very interesting that I'll mention briefly.

What's actually happening here, and this is all described in this paper, is that by interactions that we are now beginning to understand, this material is slowly but surely enlarging its cognitive light cone. What I mean by cognitive light cone is the size in both space and time of the largest goals it can pursue. Not the reach of its effectors and sensors, but the size of the goals in some particular space, for example, in space-time, of the goals that it can pursue. And we are now starting to understand how that cognitive light cone is scaled up.



Critical to understand here is that these cognitive light cones don't just project into the standard three-dimensional space of behavior. We as humans are completely obsessed with 3D space. We are pretty good at noticing intelligent behavior of medium-sized objects moving at medium speeds in three-dimensional space, and we call it behavior. Then we look at other things such as software agents sitting inside a server. We look at organoids, for example, neural organoids sitting in a Petri dish. It's very common to say those things are not embodied. They're not embodied because you don't see them rolling around in three-dimensional space doing things the way that we do. But it's important to understand that this perception, action, prediction, memory, action loop happens in all kinds of other spaces. Biology is great at deploying it in other spaces. Living things navigate the high-dimensional space of gene expression; they navigate spaces of physiological states and anatomical morphospace, which we're going to talk about the most. So we have to be very careful about embodiment. Just because it doesn't have legs or wheels, that does not mean it isn't doing the critical thing for intelligence, which is this perception-action loop with all that can be packed in between. We are not good at recognizing or visualizing these things in other spaces.

This is the thing we are all made of.

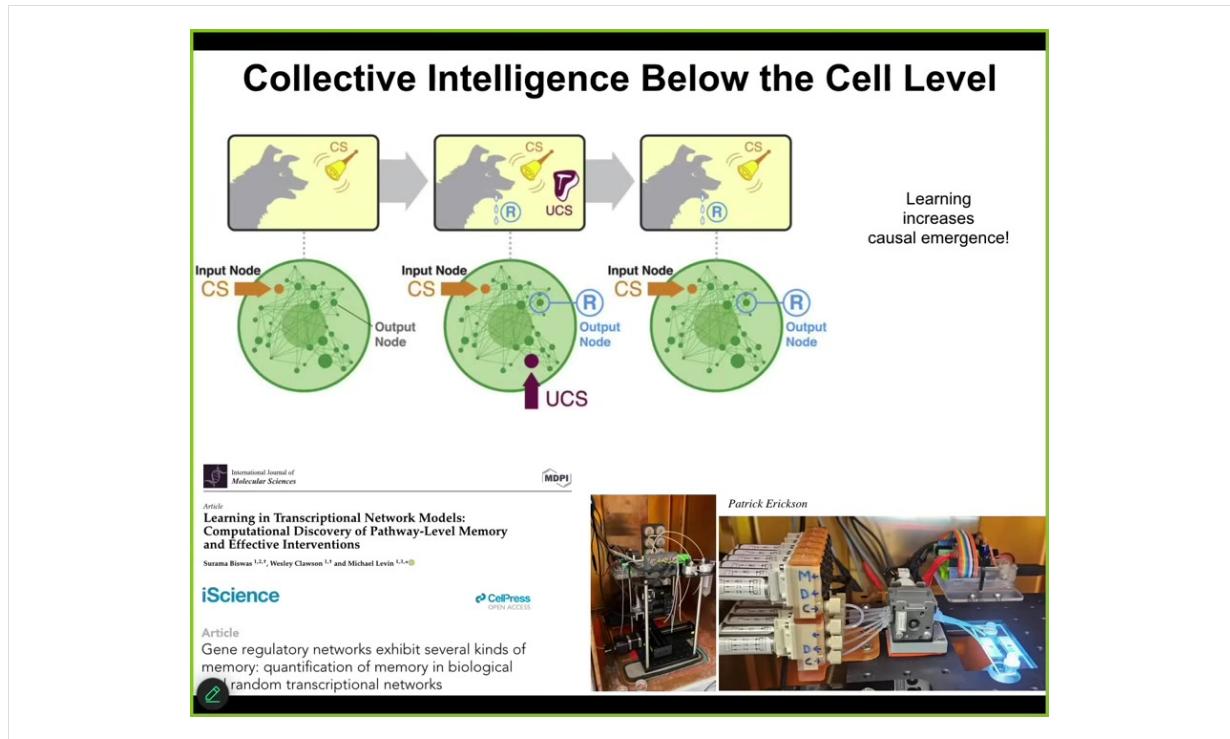
## We are All Collective Intelligences Made from an Agential Material



Lacrymaria = 1 cell  
no brain  
no nervous system

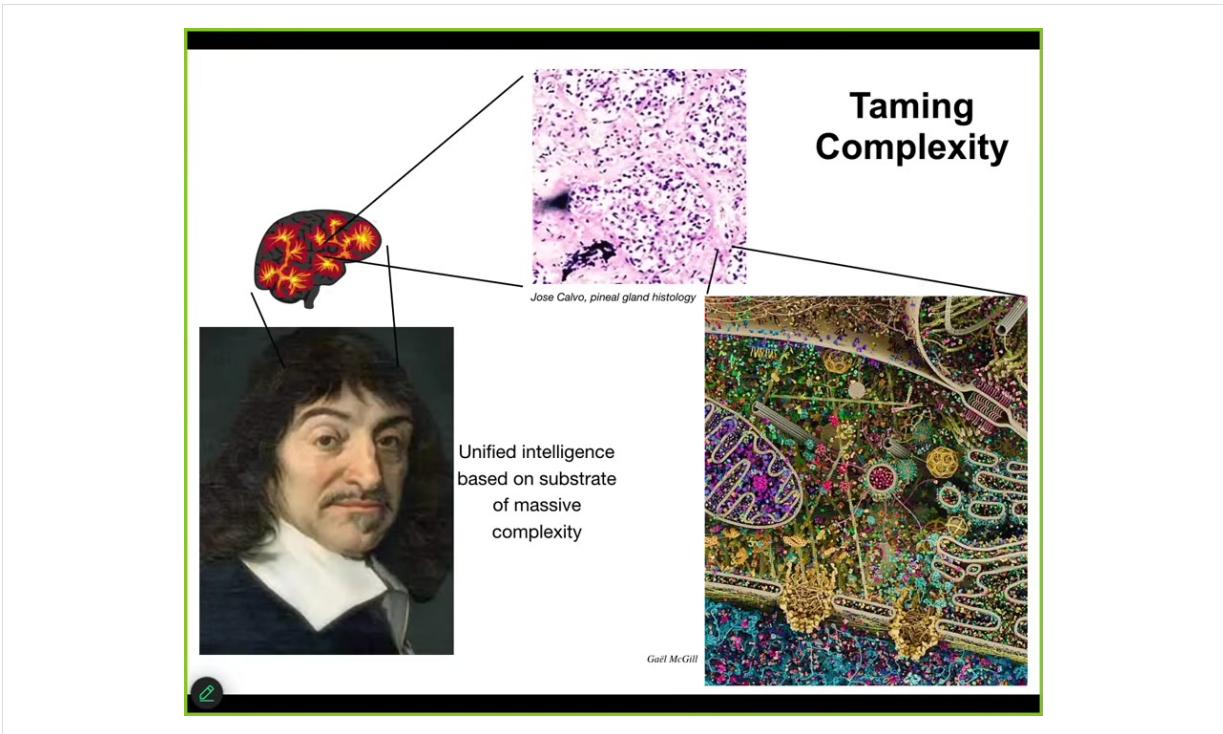
high competency  
at cell-level  
agendas

This is one cell. This one's a free-living organism called Lacrymaria. There's no brain, there's no nervous system. The whole thing is very competent in its local agendas that it has around its own scale.



And not only are the cells very competent, but the material inside the cells, the chemical networks themselves are capable of several different kinds of learning. If you test molecular network models, you will find out that, as a free gift from the mathematics, you don't need evolution for this, although evolution certainly optimizes the heck out of it. But the material itself already is capable of habituation, sensitization, and associative conditioning. Pavlovian learning. It can also count to small numbers.

You can see all that here, and we are trying to take advantage of this in the lab for applications such as drug conditioning and things. Not only can these chemical networks learn, but we found—our paper came out a week ago—we showed that the process of learning, stimulating various components of this chemical network to cause learning, actually increases causal emergence in the network. It becomes more than the sum of its parts, the more that it's trained. It's a very, very interesting loop. The thing is that this type of multi-scale scenario where the material has proto-cognitive properties, the cells that comprise it also do, it goes all the way up.

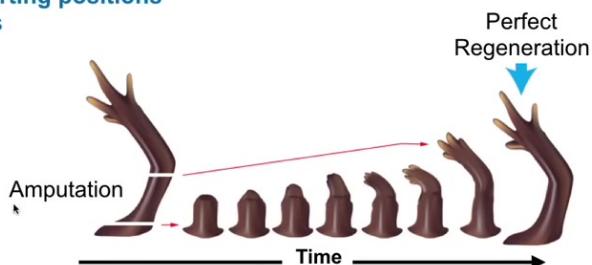


So, Descartes liked the pineal gland in the brain because he thought that, because there's only one of them, and he thought that suited the unified human experience. Fine, but he didn't have good microscopy available. If he did, he would have seen that there isn't one of anything. We are all collective intelligences in the sense that if you look into the brain, you see all kinds of stuff. If you look into the pineal gland, this is what it's made of: tons of cells. Inside each one of those cells is all of this stuff, and you can keep going. So all intelligence, I would argue, is collective intelligence. We are all made of parts, and the goal of significant agents is to bend their components to their will, to distort the option space for their parts such that the things they're doing actually feed into goals that occur in other spaces.

And so I think what's happening in biology is that our bodies are in fact a multi-scale competency architecture all the way from the molecular networks up, and every level distorts that option space and produces various attractors and so on, such that the parts do what they do, but as a result, they are actually implementing goals in other spaces of which they are completely unaware. So I want to show you a couple of biological examples of this.

## Not Just Goal-Directed Error Reduction,

- get to the same outcome
  - despite perturbations
  - from **diverse starting positions**
  - via different paths



Anatomical  
homeostasis:

it stops when the correct  
large-scale setpoint (target  
morphology) has been  
reached

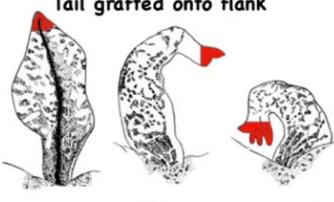
This axolotl, a little amphibian, is highly regenerative. If its leg is amputated at any position, the cells very quickly realize they've deviated from the correct state in anatomical space. They try to get back by regrowing the limb, and then they stop. The most amazing aspect of regeneration is that it knows when to stop. When does it stop? It stops when the correct limb has been completed. In other words, it's an anatomical homeostatic loop that traverses that morphospace, and it can tell when it's within acceptable error of the starting condition, and then it stops. But this is not about repairing damage. The real story here is not that at all.

## Top-Down Control: Holistic Order

- get to the same outcome (maintain set point)
  - despite perturbations
  - from diverse starting positions
  - via different paths - **modify as needed!**
  - stop when goal is achieved



Axolotl



Tail grafted onto flank

Time →

Farinella-Ferruzzi, *Experientia*, 1936 (15)

Anatomical homeostasis:



local order obeys global plan

The real story is about top-down control, because if you were to take a tail and graft it to the side, to the flank of this animal, what you would find is that over some weeks' time, the tail would actually turn into a limb. Not only does the tail remodel into a limb, but these cells up here, which are tail tip cells sitting at the end of a tail, are locally fine; there's no damage up here, but they turn into fingers.

What's happening here is that all the molecular biology that's needed to re-specify these cells into finger cells—so all the gene expression changes and the cell movements—is being driven by a high-level goal-seeking process. The large-scale homeostatic loop can tell that it's been deviated from a scenario where what you have on your flank are limbs, not tails. Those control signals propagate all the way down to the cellular and molecular level to make this happen.

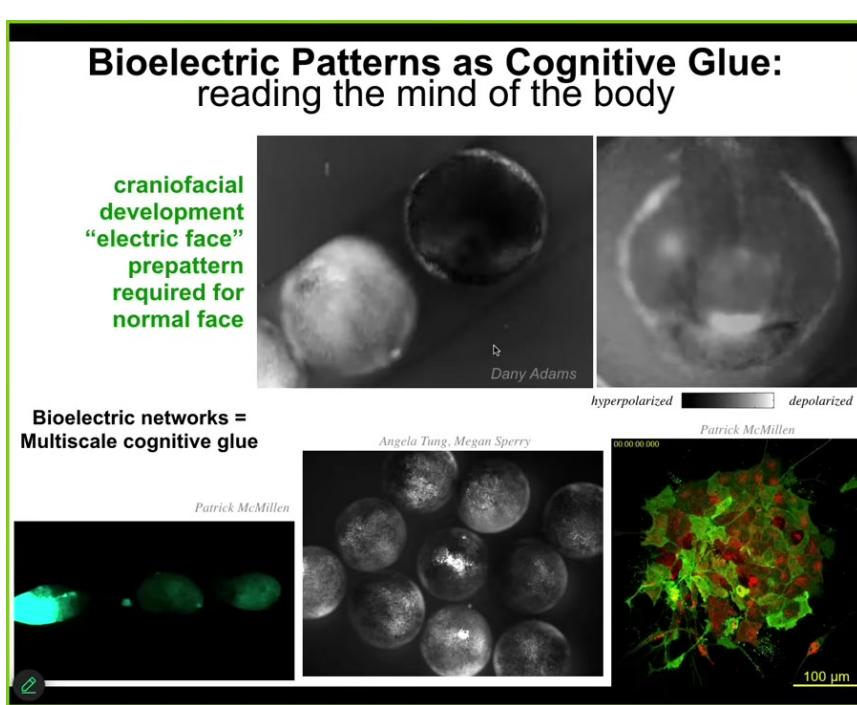
Now notice that this is a key property of cognitive systems. When you wake up in the morning you have very abstract social goals, financial goals, whatever you have. Those are very abstract things. But in order for you to get up and walk and execute on those goals, those goals have to be transduced to make the calcium ions dance across your membranes. The actual chemistry of the potassium channels has to be harnessed by these higher-level loops in order for you to do goal-directed behavior.

And that is what the multi-scale architecture enables to happen: every level is able to bend the option space of the level below, and thus it transduces increasingly abstract goals. Individual cells have no idea what a finger is or how many fingers you're

supposed to have, but the collective absolutely does, and you can see that in these examples.

So now, how does it work?

Slide 21 of 43 · Watch at [24:47](#)



To give you a very brief introduction, about 20 years ago, we made the first tools to read and write the electrical states inside of tissue. Why? Because if we wanted to understand goal-directed behavior and top-down causation, we took inspiration from the brain, which is the one uncontroversial example where we know that happens. So we know we have high-level goals, and we know we can work to solve problems.

The cognitive glue that allows this to happen in the brain, meaning why do we know things that our neurons don't know, why are we pursuing goals that our neurons don't have, is because there's a set of policies and mechanisms that bind those competent subunits to common purpose. Do they align them in various problem spaces? That cognitive glue is electrophysiology. It's the electrical signaling in the brain.

It turns out that the ability of electrical networks to do that was noticed by evolution a really long time ago. It's ancient, long before brains and nerves and muscle evolved.

All cells, starting with bacterial biofilms, were using electrical networks to integrate information. I will point out a couple of examples.

This is a time-lapse of an early frog embryo putting its face together. If we read the bioelectrical signals using these voltage-sensitive fluorescent dyes, we see patterns such as this that presage the future of what this thing is going to do. This is what it's building towards. All the gene expression and anatomical activity will arise toward putting an eye here, putting the mouth here, putting the placodes out here. This is a pattern memory that the cells will implement. And if we change this pattern, they implement something else.

This is a multi-scale phenomenon. Each one of these circles is a separate embryo. When we poke this guy in the middle here, all these others find out about it. Not only is bioelectricity the cognitive glue within a single organism, it has a similar function across organisms because groups can do things that individuals can't do. For example, resist teratogens. It's a very dynamic computational network, working slower than the brain. Same idea.

Having discovered these things, we developed methods to write information back into the cells. No waves, no frequencies, no electrodes, magnets, nothing like that.

## Bioelectric Interface to Agential Material: Prompting High-Level Goal Setting to Defeat Complexity

ion channel mRNA targeted to ventral or posterior regions

can reprogram many regions, even outside "competency zone", into complete ectopic eye!

Vaibhav Pai

1. BIOE is instructive
2. modularity - not cell level, organ-level subroutine call
3. higher-level prompt reveals higher tissue competency than Pax6 prompt
4. self-scaling of system to task

Native eye      Induced Eye

Muller glia: i, vi

Amacrine cells: ii, vii

Rods: iii, viii

Cones: iv, ix

Phase contrast: v, x

Developmental Modules because morphogenetic goals

voltage-modified cells

wild-type cells recruited to ectopic lens!

Getty Images

We take all our tools from neuroscience because fundamentally neuroscience is not about neurons at all. It's about this electrical, this cognitive glue property. We use the same tools and we hacked the same bioelectrical interface that the cells are using to control each other.

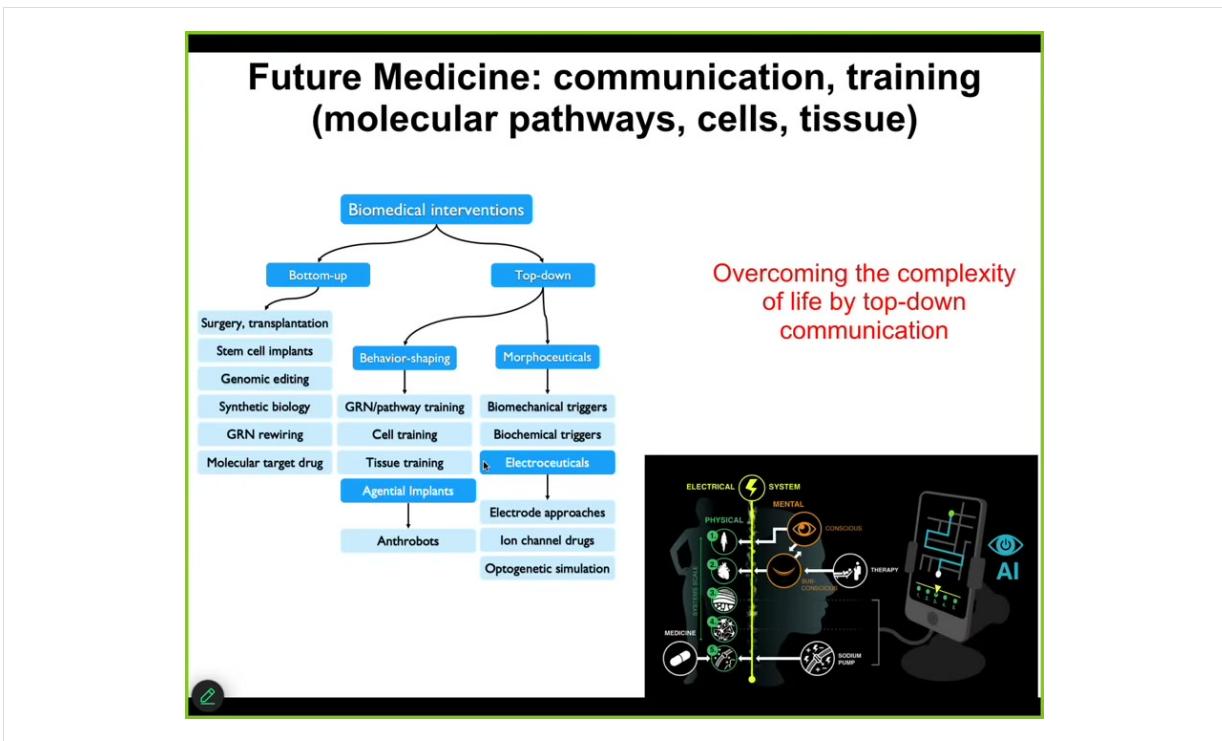
We can inject into the early frog embryo RNA expressing specific potassium channels. For example, they establish a voltage gradient. That voltage gradient is interpreted by the surrounding cells as make an eye. We can make other organs too. Here it is. This gut tissue here was told to make an eye. It did. That eye has the same lens, retina, optic nerve, all the same stuff.

These patterns that exist in this, that are stably kept in this electrical medium, these pattern memories, are used to guide the activity of the system at an organ level. We're not telling it how to make an eye. We have no idea how to make an eye. It's a very complex organ, but we don't need to. The system knows how to propagate high-level prompts down into the molecular pathways.

Not only will it make an eye from this very simple prompt, but we can take advantage of all the other competencies of the tissue. If we only inject a few cells, the blue ones are the ones we injected, they will automatically recruit a bunch of their neighbors. All this other stuff out here wasn't injected by us. These cells realize there's not enough of them to make an eye, so they recruit their neighbors to complete the task, much like other collective intelligences do, such as ants and termites.

This material has all sorts of competencies, and they're all organized around being able to minimize error towards specific goals, and taking goal-state information from higher levels. That is the architecture that's happening here, and that's the problem decomposition that the material natively does.

Slide 23 of 43 · Watch at [29:15](#)



I won't take the time to talk about all the biomedical stuff, but whereas all of today's biomedicine focuses on these kinds of bottom-up interventions, we have a whole panoply of top-down strategies that are allowing us to do some pretty wild things in regenerative medicine, birth defects, cancer suppression, aging, and so on. Most of our lab is focused on that stuff.

The development of AI-based tools that could enable us to literally communicate with the various levels, not just the molecular level, which is what all drug discovery is focused on, or genomic editing or those kinds of things, but to talk to these higher levels so that we can have much more efficient communication and take advantage of all the problem solving that they can already do.

My claim is that biology itself tames this complexity by a multi-scale architecture of agential systems, systems with agendas, systems with learning capacity, with problem

solving competencies. The price you pay for that is occasional defections. That's cancer. Another price we pay for that is hackability. There are all kinds of parasites and other things that can hack it.

But the big win there is that no system has to micromanage all the way down. They are all just hacking each other at very high levels with prompts, because the material they're talking to is itself a kind of cognitive agent. That's what I think is happening in biology.

Slide 24 of 43 · Watch at [30:50](#)

## Outline:

- Unconventional biology
- **Why does it work**
  - Commitment to unreliable substrate
  - Creative problem-solving
- What does it mean

Now I want to say a couple things about why this actually works and where this all came from. The thing that drives all this is commitment to an unreliable substrate. Whereas in computing today we mostly have error correction algorithms and we have abstraction layers and we have all these things to try to make the data stay still. In other words, you have to be able to trust your material. If you're working at one level, you cannot be worried that the silicon is floating, the numbers are floating off because your silicon is doing something weird. Biology is exactly the opposite because biology operates with a very unreliable substrate.

## Life Optimizes for Salience, not Fidelity of Info

(which enables it to survive drastic hardware refactoring)

The diagram illustrates the life cycle of a butterfly, from caterpillar to adult, set against a background of leaves and flowers. The caterpillar is labeled 'crawls, chews plants in 2D world'. Above it, a thought bubble contains the text 'brain is liquefied, rebuilt'. The adult butterfly is labeled 'flies, drinks nectar in 3D world'. To the right of the butterfly, a bulleted list of principles is presented:

- Paradox of change
- Unreliable substrate
- Commit to creative interpretation of all info

Below the main image, there are two journal article citations:

**Communication & Integrating Biology 8(1), e101243, Downloaded on June 2019. Published with license by Taylor and Francis Group, LLC**  
**The stability of memories during brain remodeling: A perspective**  
Douglas J Blackman<sup>1</sup>, Tal Shemesh<sup>2,3</sup>, and Michael Levin<sup>1\*</sup>

**entropy**  
**Perspective**  
**Self-Improvising Memory: A Perspective on Memories as Agential, Dynamically Reinterpreting Cognitive Glue**  
Michael Levin

I'll give you a simple example that started me on the road to thinking about all this, which is that caterpillars, these soft-bodied kinds of creatures, have a very particular brain for crawling around on leaves and eating them.

They turn into a butterfly. A completely different controller is needed to operate. This lives in the three-dimensional world now and doesn't care about leaves at all. It likes nectar. The brain is quite different.

What happens during this metamorphosis process is that the brain is dissolved, most of the cells are killed, most of the connections are broken, and then you build a new brain. The remarkable thing about this is that memories that are formed in the caterpillar—if you train the caterpillar, the butterfly still remembers the original information. That's wild enough that you can keep information while the substrate itself is being torn to shreds.

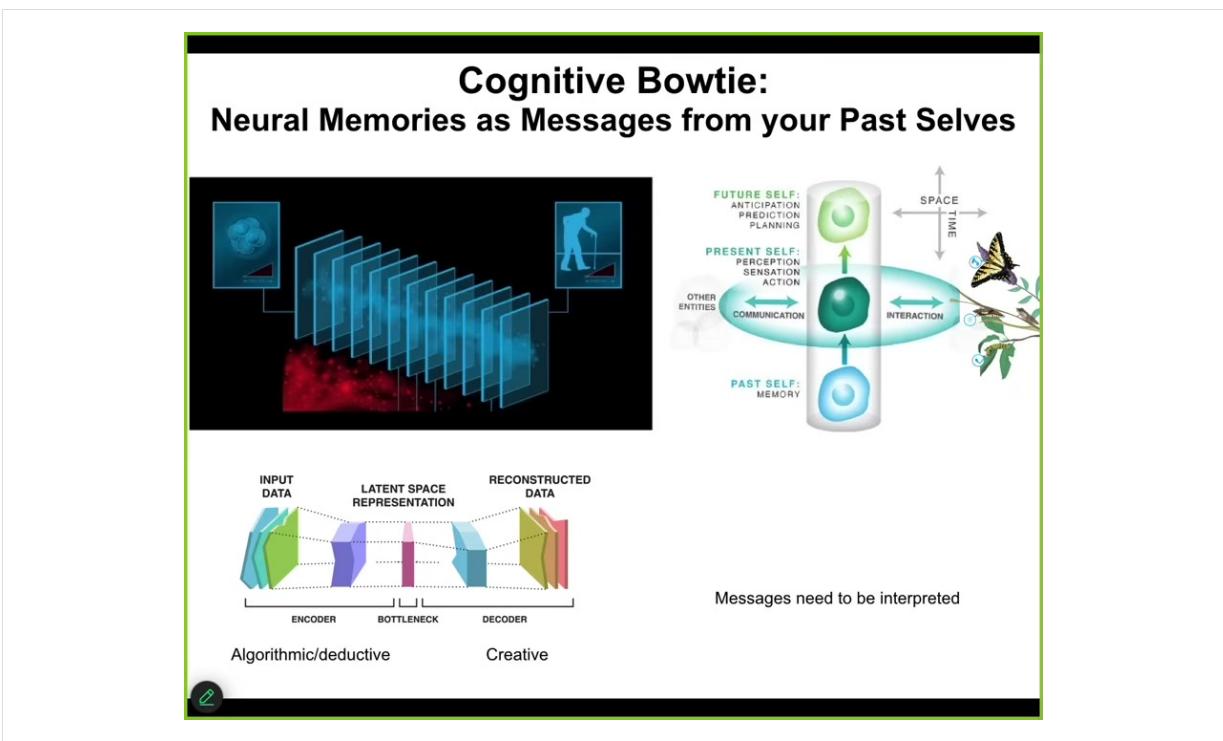
The deeper issue here is that the actual memories of the caterpillar are of no use to the butterfly. In other words, specific actions of the controller for a soft-bodied creature that saw stimulus and needed to go eat the leaves do not work at all for the butterfly. The butterfly doesn't care about leaves. It doesn't move the same way that the caterpillar did. To survive, to persist as a memory across this refactoring, you need to be remapped.

The fidelity of the information doesn't help you here; the saliency of the information does. In other words, you need to remap it onto a completely new architecture, new

effectors, new sensors, a new kind of problem space, in fact, a higher dimensional problem space.

All of this is described here. But that just started me thinking about the importance of interpreting the information you have. The information, whatever molecular information it gets from the caterpillar, cannot be used directly. It has to be interpreted. While this is pretty drastic, we do a little bit of this, puberty and things like that, but it's not as drastic as this.

Slide 26 of 43 · Watch at [33:45](#)



All cognitive agents face the same paradox. At any given time step, you do not have access to the past. What you have access to are the memory engrams that the past left in your brain or body. You have the messages from past versions of you. Memories are a kind of communication event on this model. You get these messages and now you have to interpret them.

They are sparse because there's an encoding event where you throw out a lot of the extra regularities and squeeze it down into this very tiny n-gram, but then you've thrown away all this information. You have to re-inflate it again. You have to interpret

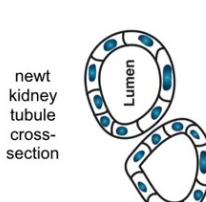
it. We as cognitive agents are a continuously evolving, self-telling story that attempts to make a coherent, adaptive model of itself and the outside world based on the information that it received, not sticking to it mechanically, but telling the best story it can, because this stuff is algorithmic, compressing rich past experiences into a thin representation. That's algorithmic and deductive, but this part is creative. You've lost information. You can't possibly know exactly what happened or how past you interpreted these messages.

What that means, and this is true not only cognitively, it's true anatomically: the reason that all this plasticity works in the body is that the genetic information you get from past lineages is just a prompt. It's just a message that you can interpret in different ways. And that's the collective intelligence of the body. The goal is to interpret all that information to do something coherent.

Slide 27 of 43 · Watch at [35:40](#)

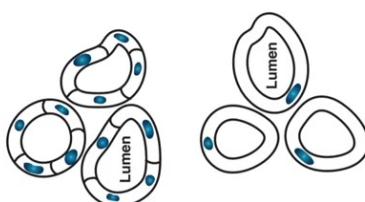
## Genome and Beginner's Mind

### Genetic Memories as Problem-solving Prompts from your Past Selves



newt kidney tubule cross-section

Fankhauser, 1945. J. Exp. Zool., 100(3): 445-455



Changing the size of cells still enable large-scale structures to form, even if they have to utilize different molecular mechanisms = top-down causation

you can't even count on your parts! Unreliable substrate -> creative polycomputing

- Creative, intelligent problem-solving - repurpose available tools to new circumstances

**INTERFACE**  
nid.biopresselpublishing.org

**Perspective**  
doi.org/10.1162/089826001753190621  
See also: Pfeifer, C. (2012). For more tools to build a better organism: an interview with Michael Levin. *Top-down models in biology: explanation and control of complex living systems above the molecular level*. *Genome, PloS One*, 7(10), e46551.

**Integrative Biology**  
doi.org/10.1162/089826001753190621  
See also: Pfeifer, C. (2012). For more tools to build a better organism: an interview with Michael Levin. *Top-down models in biology: explanation and control of complex living systems above the molecular level*. *Genome, PloS One*, 7(10), e46551.

**PERSPECTIVE**  
doi.org/10.1162/089826001753190621  
See also: Pfeifer, C. (2012). For more tools to build a better organism: an interview with Michael Levin. *Top-down models in biology: explanation and control of complex living systems above the molecular level*. *Genome, PloS One*, 7(10), e46551.

**Resilient Machines Through Continuous Self-Modeling**  
josh.bongard@mit.edu | Victor.Zykov@mit.edu | Hod.Lipson@mit.edu

And this gives rise to all kinds of amazing competencies. For example, this is a cross-section through a kidney tubule of a newt. You'll notice 8 to 10 cells normally build this. You can make newts with much larger cells by giving them higher copy numbers of their genome. So they're called polyploid newts. And if you do that, the

newt is still the same size. How is that possible? Well, you find out that fewer cells make the same structure. And then if you go really crazy and you make very high level polyploid newts, the cells get so gigantic that just one cell wraps around itself to give you the exact same structure.

Now this is remarkable. This is for two reasons. First of all, this is a different molecular mechanism. So this is straight up what people measure on IQ tests. You're given a set of objects and you're asked to solve a problem. How do you use this to solve a particular problem? So this system has a genome. That genome gives lots of affordances. There are ways for cells to communicate. There are ways to bend themselves. You have to use the right tools to solve a problem in novel circumstances.

So now think about what it's like being a creature coming into this world. You can't rely on how many copies of your genetic material you're going to have. Never mind the uncertain environment. You can't even trust your own parts. You don't know what your genetics are going to be. You don't know how many cells you're going to have or how big your cells are going to be. You have to try to solve the problem in novel ways. And this is why all this plasticity works. The reason those tadpoles work with eyes on their tails is because they never knew that the eyes were going to be in the front in the first place. They could never count on that. If they are, great, that's a default outcome. But no one can count on that because things will change. You will get mutated on the lineage scale, the environment changes, your own parts change. It's beginner's mind. You have to take what you have and somehow do something useful with it.

This is Bongard's 2006 paper on robots that didn't know what their shape was, really prescient in all of this.

## Cognitive or Genetic Must be Interpreted

### evolutionary pressure for creative confabulation

<http://theembeddedguy.com/2016/05/15/layers-of-abstraction/>

The diagram illustrates the layers of abstraction from hardware to software, and how biological systems interpret memory. It shows a lizard with an encoding process (brain) and a decoding process (eye), separated by a bottleneck (brain). The bottleneck is labeled 're-interpretation' and 'not'. To the right, a vertical stack of layers of abstraction is shown, from hardware (bottom) to software (top): Application, Algorithms, Programming Language, Assembly Language, Machine Code, Instruction Set Architecture, Micro Architecture, Gates/Registers, Devices (Transistors), and Physics. Below this, a text box discusses biology's assumptions about unreliable hardware and the evolutionary pressure for creative confabulation. At the bottom, a diagram shows an autoencoder architecture: INPUT DATA is processed by an ENCODER (Algorithmic/deductive) to a LATENT SPACE REPRESENTATION, which is then processed by a BOTTLENECK (Creative) and a DECODER (Creative) to RECONSTRUCTED DATA. A 'Trends in Genetics' article by Benoit Hart and Michael Levin is mentioned.

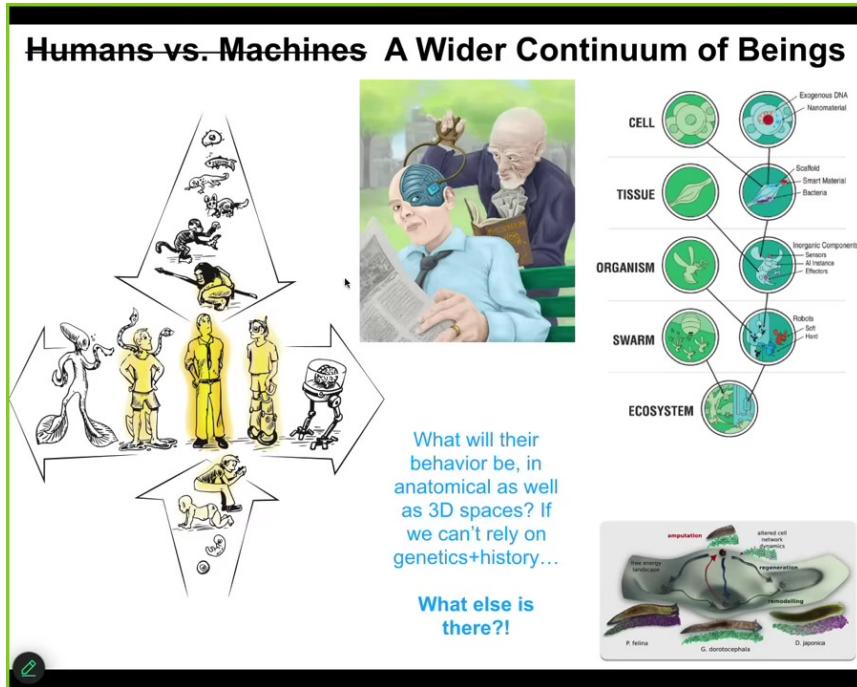
**Biology assumes the hardware is unreliable**  
**Environment and your own parts will change, don't over-train**  
 Little allegiance to past Self's meaning of engrams  
 Re-interpret on-the-fly - present/future is all that matters  
 Engram is highly compressed - creative remembering, not deduction  
 Undependable hardware and confabulation are a feature, not a bug  
**Consciousness = active, creative story-telling from own memory traces**

**Trends in Genetics**  
 Opinion: What does evolution make? Learning in living lineages and machines  
 Benoit Hart <sup>1,2,3</sup> and Michael Levin <sup>1,3,4</sup>  
**enropy**  
1Department of Biology, Tufts University, Medford, MA, USA; 2Department of Physics, Tufts University, Medford, MA, USA; 3Center for the Study of Regeneration, Tufts University, Medford, MA, USA; 4Center for the Study of Evolution and Development, University of California, San Diego, La Jolla, CA, USA

Biology is dealing with a very unreliable medium. You never know how many copies of anything you have or whether your proteins are getting degraded. You can't take it literally, but you can develop good skills to interpret it in a useful fashion. That's why this kind of cognitive bottleneck, this autoencoder kind of thing, is exactly what organisms do. They don't make other organisms, they make eggs, and it's up to the cells to interpret the information that was in the egg.

If you look at how evolution works on agential materials, you find something really interesting. You find that the competency, the problem-solving competency of the material, makes it really hard to evaluate genomes, because genomes aren't what come up for selection. What comes up is the actual phenotype, the results. What evolution ends up doing is spending much more of its time cranking improvements in the interpretation machinery. That confabulatory ability to tell novel stories using whatever information you've got with no allegiance to how it was interpreted in the past or what was happening before is what evolution really cranks on. So you end up with this positive feedback loop between this kind of creative interpretation and the intelligence of the system in solving different kinds of problems.

All that plasticity makes life incredibly interoperable.



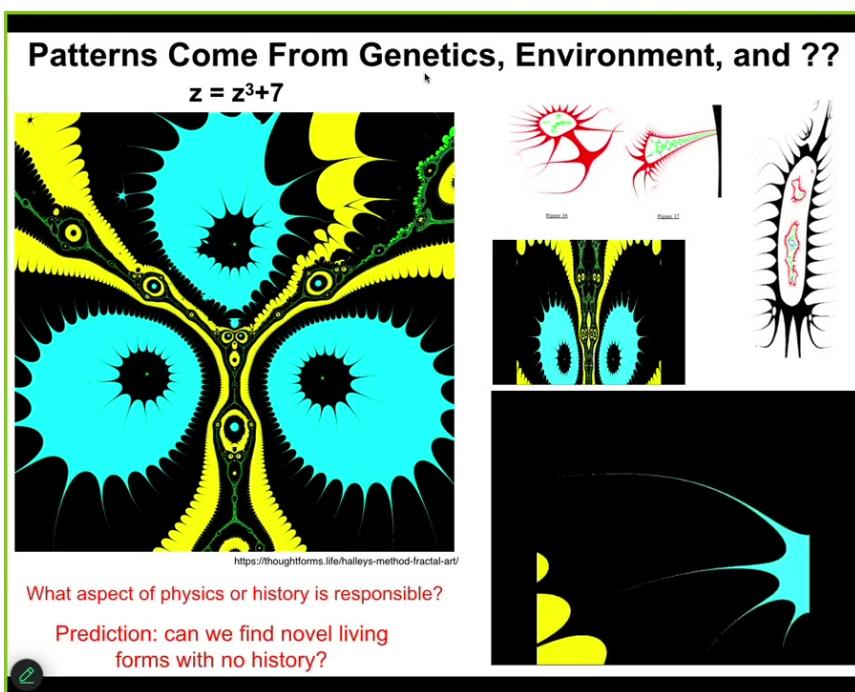
And that means that pretty soon we are going to be faced with not just this spectrum, where we all came from a single cell, both evolutionarily and developmentally, but there's another spectrum, which is technological and biological change. This human that features so prominently in stories of philosophy of mind and so on is a spectrum that's not easily dealt with by our ancient binary terminology. We are going to have all kinds of interesting creatures, made of all sorts of different combinations of evolved and engineered material. Now we have to ask the following question. If you're dealing with a being that is nowhere on the tree of life with you, meaning they did not come up through the same evolutionary process, can we guess what their competencies are going to be? What are their goals? What are their preferences? In standard biology here, this morphospace of these planarian heads, there are specific attractors that we can chase these tissues into. People will say, well, of course, evolution formed these, right? That's what evolution does. But when you're dealing with agents that don't have that simple backstory, what actually happens? This is going to be the weirdest part of the talk. Here's where I'm going to speculate as to what I think is happening here.

Biologists like two things. They like stories about physics and they like stories about genetics. In other words, they like explanations for a certain phenomenon which can be cashed out in terms of the necessity of physics and the historical contingency of its environment, and things that have happened in the environment.

But we already know that that's not sufficient and that physics isn't closed like that. Just as a simple example, cicadas come out at 13 years and 17 years. If you're a biologist,

you want to understand why. It's because they are trying to time it to periods that are hard for predators to match. Thirteen and seventeen are good because they're prime numbers. If you keep trying to pursue the chain of explanation, the next question is, okay, why don't 13 and 17 have factors? Now you've left the physical world.

Slide 30 of 43 · Watch at [41:42](#)



Now you're in the land of mathematics. There is nothing in physics that will answer that question for you. There's nothing in evolution that will answer that question for you.

If you want to understand what cicadas are doing, you have to eventually cross over and start looking at mathematical patterns and how "why" questions are handled in mathematics.

The reality is that genetics and environment are not the only place that complex patterns come from.

I could have given many examples. This is just one of my favorites because I think it looks cool and it's kind of organic.

All these things look biological a little bit.

**Patterns Come From Genetics, Environment, and ??**

$z = z^3+7$



What aspect of physics or history is responsible?

Prediction: can we find novel living forms with no history?



This is called the Halley plot, and it is what happens when you map the behavior of this complex number function. It's very simple, this little tiny seed. Talk about small n-grams that can be re-inflated into complex patterns.

Look at this.

The compression is insane.

If you plot this in a particular way, you get this amazing structure. The answer to why the structure looks the way it looks has nothing to do with physics.

**Patterns Come From Genetics, Environment, and ??**

$z = z^3 + 7$

<https://thoughtforms.life/halleys-method-fractal-art/>

What aspect of physics or history is responsible?

Prediction: can we find novel living forms with no history?

There's nothing you can do in the physical world to change what it is. There's no history behind it. It is a fact of mathematics. Math is full of these things.

**Patterns Come From Genetics, Environment, and ??**

$z = z^3 + 7$

<https://thoughtforms.life/halley's-method-fractal-art/>

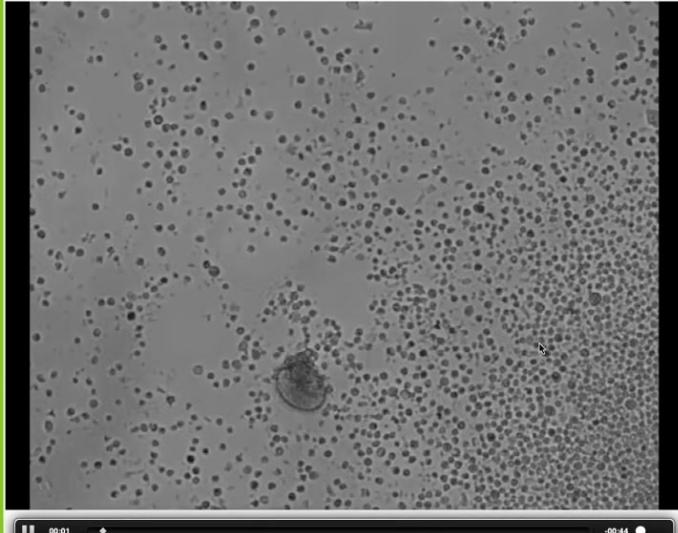
What aspect of physics or history is responsible?

Prediction: can we find novel living forms with no history?

And then, this is a video I made by iterating this in various ways. There's very rich structure buried in the rules of mathematics, and biology exploits these things. These are amazing free lunches that biology uses all the time. And I want to show you some biological life forms that have this kind of character.

We can't use selection history to explain it.

**What Lies Beyond Repair of Normal Target Morphology?**

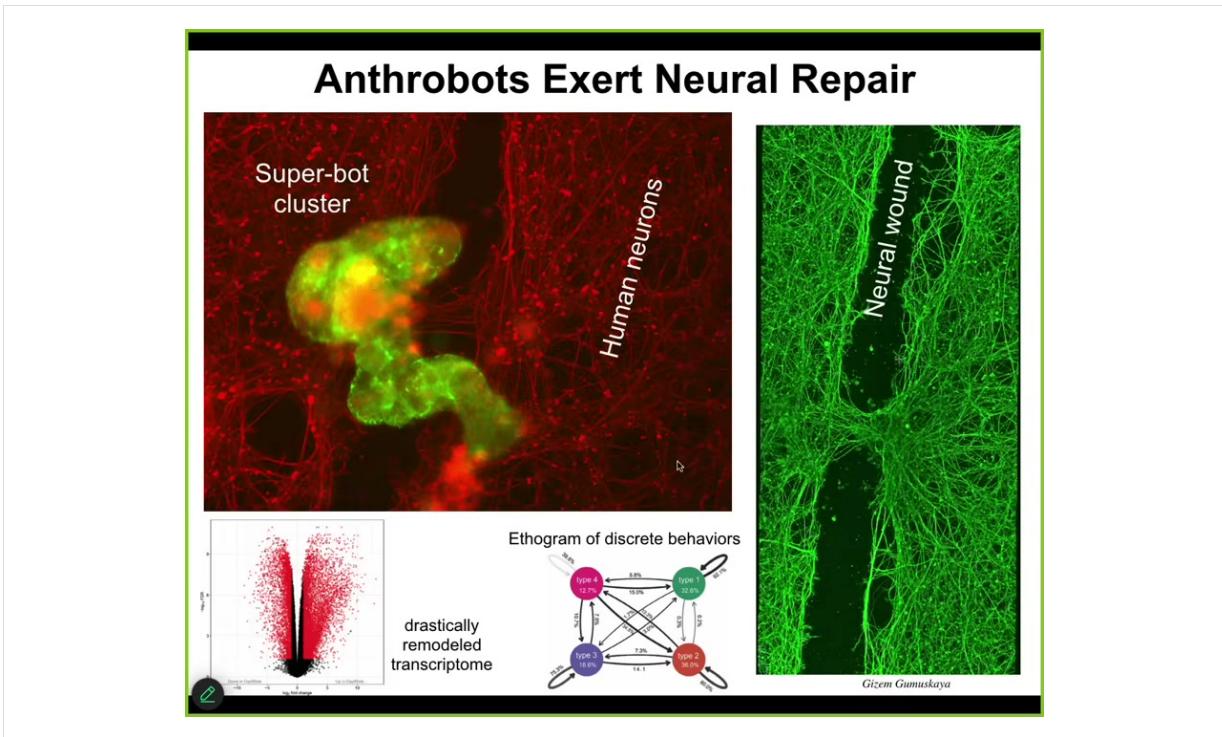


Where do the properties of novel systems come from if not eons of selection or explicit engineering?

Could you guess the genome from these data?

Could you guess behavior and form from the genome?

If I showed you this little creature, you might think it's something we got off the bottom of a pond somewhere, a primitive organism. If you were to sequence it, you would find out that it's 100% *Homo sapiens*. Adult human cells extracted from the trachea of patients undergoing biopsies under a particular protocol that we developed from this little self-motile creature. It looks and acts nothing like any stage of human development. It has some amazing properties.



It has all kinds of behaviors. It has about 9,000 genes expressed differently than the tissue did when it was in the patient. So about half the genome was completely remodeled here. Remember, no changes to the genome, no synthetic biology circuits, no scaffolds or weird drugs. All we did was liberate those cells from their environment and let them reboot their multicellularity. This stuff out here is human neurons, and what we did was make a big scratch through it. These anthroboots, which is what we call these little creatures, they land in this superbot cluster, and they actually start healing across the gap. Lift them up four days later, this is what you see. This is the first thing we checked, so probably they do 1,000 other things, we don't know. But they have this amazing ability, and who would have thought that your tracheal cells that sit there quietly dealing with mucus and air particles have the ability to form a little creature that can assemble and do this kind of thing?

These are xenobots, so it's a similar idea made with frog cells.

They have an amazing capacity. They have many, but this one thing I'll show you is called kinematic replication.

So if you give them a bunch of loose epithelial cells as a material, they will collect it into piles.

The piles mature and become the next generation of xenobots. And then guess what they do? They go on and they make the next generation and the next again.

We didn't teach them to do this.

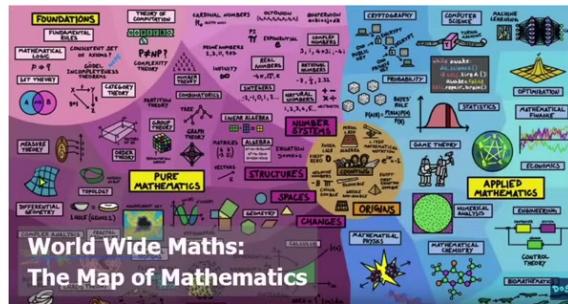
**Evolution was Supposed to Explain Complexity, with High Specificity; how's that going?**

The diagram illustrates the development of Xenopus laevis embryos and Xenobots. On the left, the 'Xenopus laevis genome' is shown as a karyotype. Two arrows point from the genome to two different developmental paths. 'Path A: embryos' leads to a series of images: a spherical egg, a tadpole, and a photograph of a frog. 'Path B: Xenobots' leads to images of a 3D-printed robot body, a 3D-printed head, and a composite image of the two joined together. Below these images is a timeline labeled 'Developmental Time' with a black bar. To the right of the timeline is a section labeled 'Behavior' with a black bar. At the bottom, a caption reads: 'Xenobot bodies and minds have no straightforward evolutionary back story; When was the computational cost of Xenobot features paid?!' A small green circular icon with a white 'e' is at the bottom left.

The amazing thing with all of this is that this again goes back to the plasticity of the material and this idea that the genome learned to do this over many generations. This is the standard developmental sequence. Here are the tadpoles. But it also can do this. This is an 84-day-old xenobot, and many other kinds of things like kinematic replication. There's never been selection to be a good xenobot or a good anthropot. There's never been any anthropots or xenobots. Where do their capacities come from? You cannot tell a story of selection that matches it. If you say that these are emergent from the things that the cells did do in the environment, what you're basically giving up is any degree of specificity between the final outcome and the history of selection that they've had, which is basically what evolution was supposed to do. It was supposed to explain the specificity and give you a causal story.

We know when we paid for the computation to do all of this, that was during the evolution of the frog. When were all the computations done to get this going? There was no history of this. Where did this come from? This is very interesting. The idea of these kinds of things that have no obvious history of either learning or selection is, I think, really demanding attention.

## Patterns from Beyond Genetics and Physics: Is There a Precedent?



Yes, Platonist Mathematics -  
discovering, not inventing,  
patterns that are not  
dependent on, but do affect,  
physics (and biology!)

Also,  
Platonic Space Hypothesis  
in machine learning

There is a precedent for this. Biologists hate this idea of the physical world not being closed and of needing to go outside a nice sparse ontology. But mathematicians are fine with this. There are many Platonist mathematicians who believe that there are patterns that they are discovering, not creating, and that the order, that the space containing these patterns is a structured ordered space which they navigate through their research. This is actually starting to crop up in machine learning as well, which they call the "Platonic space hypothesis."

## Whence specific goals and competencies if not Selection?!

Evolution exploits free lunches: shapes, behaviors, properties of networks, features of computation, numbers, etc.

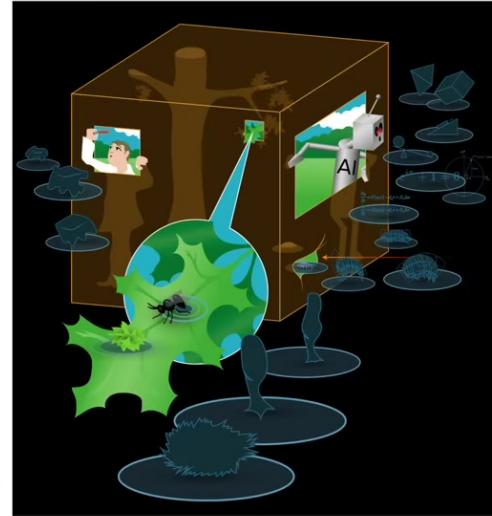
Option 1: there is a random set of amazing "facts that hold" and we will call it "emergence" and be surprised each time

Sparse Ontology → mysterianism

Option 2: there is an ordered, non-physical latent space of patterns which can be studied systematically

Optimism → research agenda

Synmorpho beings and minimal algorithms as vehicles for exploring Platonic latent space!



People who like their space physicalist, sparse physicalist ontology basically lean on emergence and the idea that there are just some facts that hold in the physical world. You have a system with simple rules, it can do complex things. And that's it. These things are emergent. I don't like that view because I think it's mysterian. I think that it relegates these things to surprise; emergence then is anything we find surprising. You write them down in your big book of emergence and that's the end of it.

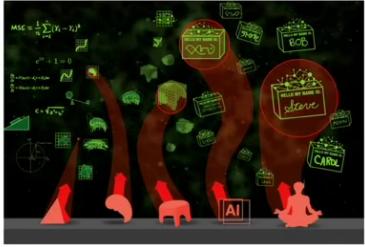
I like the more positive view of the mathematicians where they're not haphazard things that just happen to hold in the physical world. There is an ordered, structured space that we can rationally study. How do we study them? By producing pointers into that space.

And so everything that we make, cells, embryos, biobots, AIs, hybots, cyborgs, all of them are pointers into that space. We can use them as exploration devices or interfaces to try to understand the mapping between the thing that you built and the patterns that are ingressing through this space. If I can't prove it, it's a metaphysical assumption that the space is understandable. It's not just a random grab bag of stuff.

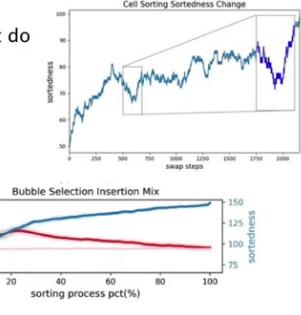
## “Emergent” Goals and Competencies

humility warranted: even bubble sort has emergent delayed gratification NOT explicitly in the algorithm

We underestimate matter and we underestimate algorithms/“machines”



Algorithm = things you made it do + spontaneous side-quests and competencies



Classical sorting algorithms as a model of morphogenesis: Self-sorting arrays reveal unexpected competencies in a minimal model of basal intelligence

Taining Zhang<sup>1</sup>, Adam Goldstein<sup>1</sup> and Michael Levin<sup>1,2</sup> bioRxiv preprint doi: https://doi.org/10.1101/2014.09.03.108000; this version posted September 4, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Taining Zhang Adam Goldstein

And so what I do think is important is that it actually is not just for the biologicals. It's not just the complex living things that get to have the benefit of these ingestions, these free lunches from mathematics. It doesn't take much at all. And if you're interested, this is one example. I chose sorting algorithms on purpose because they are so minimal and so deterministic and so obvious, to maximize the shock value. Take a look at this paper.

What we found is that even simple things such as bubble sort have unexpected side quests, the things they're doing that are literally not in the algorithm. They're nowhere in the algorithm. And so that means, for all the rest of us biologicals, there are things we have to do. We are forced by the physics and by the genetics of our world to do certain things and are limited in certain ways. But for the time that they are alive and operational, before they get ground into dust by the physics of their world, there are other things they do that are nowhere in the algorithm.

For this reason, I think the things they're doing would be things behavioral scientists would recognize immediately. If you didn't tell them that this came from an algorithm, they would immediately say, "oh yes, this is something we see all the time among the set of behaviors we've seen." These are behavioral competencies that arise that cannot be explained simply by looking at the algorithm. They're not inconsistent with the laws of physics any more than our cognition is, but they're not prescribed by it.

I think it's very important to understand that if even dumb bubble sort can be an interface to interesting behavioral propensities from this latent space, then what do we

say about the Internet of Things and language models and all this other complicated stuff? I get really frustrated when people say, "well, I write these language models."

Slide 40 of 43 · Watch at [50:40](#)

**Humility Warning: neither digital nor biochemical "machines" are only what our formal models say they are**



*Ceci n'est pas une pipe.*

Magritte



*ceci n'est pas une  
machine de Turing*

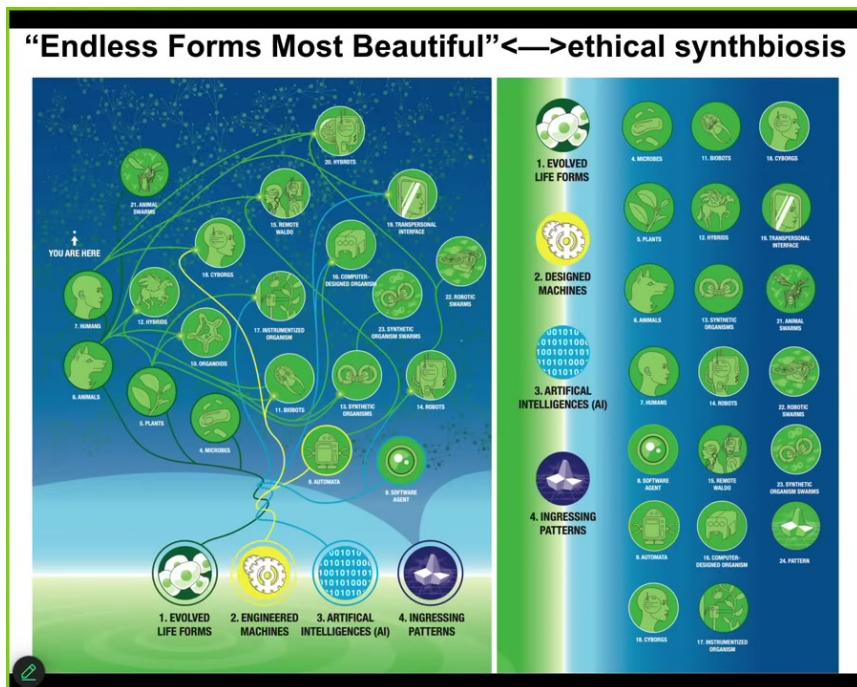
computationalism:  
nothing is a TM, not even a TM

- Minds are not fully defined by our models of them, neither for their limitations nor for their competencies.

It's just linear algebra. There's nothing in there. If nobody had noticed it doing these things, what else do we not understand? I want to make a claim, which is that our formal models do not circumscribe the cognitive aspects of the objects that they're supposed to understand.

For the exact same reason that the rules of biochemistry do not fully describe what our minds are capable of, the algorithms and the stories of materials and everything else that we try to use to understand simple machines, quote unquote, do not fully circumscribe what they're doing either. My point is that there are limitations, and these are limitations of our formal models. They are not limitations of the actual thing. There are massive surprises, certainly on the biological end, but it goes all the way down.

My point is not to say that living things are machines in the sense that we're all on the same spectrum, but not because you can use simple machine metaphors to understand life. Precisely the opposite, because even quote unquote machines are doing things that simple, mechanical formal models do not capture. The magic goes all the way down.



We are soon going to be living in a very wild world full of all kinds of different embodied minds, all kinds of combinations of software, synthetic materials, evolved materials, and all of the ingestions from this latent space that they allow. We are very poor at predicting what is going to come through, what the behaviors are going to be. We need to get a lot better at this so that we could have an ethical synth biosis with all these things and for different types of human flourishing that I think we're all interested in.

## Summary:

- Bioinspired: unreliable computing, agential materials, flexible cognitive light cone
- Complexity dealt with by multiscale competency and polycomputing
- Embodiment is not just in 3D space; what's important is a 1st person perspective onto a problem space via a perception-decision-action loop
- Physicalism is dead; "emergence and complexity" is not enough.
- We don't make intelligence, we make pointers/interfaces that facilitate ingressions from a latent ("Platonic") space, of patterns that are somewhere on the spectrum of agency, and may or may not be anything like naturally evolved minds; "I made it so I know what it does/is"
- Research agenda: develop tools (AI) and protocols for recognizing intelligence in very unfamiliar guises, and communicating with them.



The bottom line is this, that I think we need to do bio-inspired computing. We need to focus on unreliable, agential materials, a flexible plastic cognitive light cone, complexity that is dealt with by top-down control. This thing that Josh Bongard and I are developing called polycomputing. Embodiments are much broader than we think they are. As unpopular as this is, I think the physicalist assumptions hold us back. They hold us back in biology, cognitive science, and computer science. I don't think emergence and complexity are sufficient for what we need to do. I don't believe we make intelligences either biological or technological. We make interfaces that are exposing us to an interesting latent space that we need to understand. I think that's our research agenda. There's a whole bunch of papers.

I'm happy to send these to anybody who's interested.

## Thank you to:

**Post-docs and staff scientists in the Levin lab:**  
Wesley Clawson - hybrots and virtual worlds for biological controllers  
Douglas Blackiston - brain-body interface plasticity, synthetic living bodies  
Benedikt Hartl - evolution, machine learning, and cognition  
Vaibhav Pai - voltage gradients in eye/brain induction and repair  
Tal Shomrat - persistence of memory in regenerating brains  
Patrick McMillen - bioelectric imaging and embryogenesis  
Patrick Erickson - cellular learning

**Graduate Students:**  
Gizem Gumuskaya, Nikolay Davey - Anthrobot  
Sherry Aw - bioelectric eye induction  
Adam Goldstein, Taining Zhang - emergent competencies of algorithms

**Undergraduate Students:**  
Pranjal Srivastava, Ben G. Cooper, Hannah Lesser, Ben Semegran, Andrew Bender,  
Douglas Hazel - Anthrobot

**Technical support:**  
Rakela Colon, Jayati Mandal - lab management  
Erin Switzer - vertebrate animal husbandry

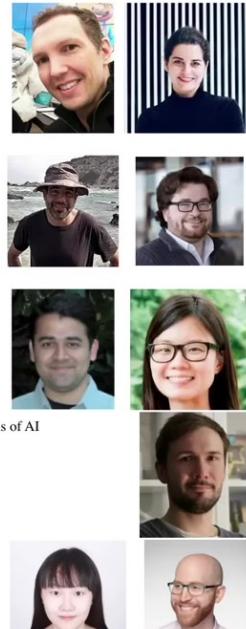
**Collaborators:** [Allen Center members +](#)  
Joshua Bongard - Xenobot simulations and AI  
Thomas Doctor, Olaf Witkowski, Bill Duane, Elizaveta Solomonova, Paul Colognese - Buddhist models of AI  
Sebastian Risi - open-ended evolution  
Simon Garnier - computational analysis of Anthrobot form and function  
Chris Fields - physics of sentience and sentience of physics  
Richard Watson - computational models of cognitive scaling and evolutionary learning  
Giovanni Pezzulo - cognitive science applied to morphogenesis

**Model systems:** tadpoles, planaria, zebrafish, slime molds, human cells, and chick embryos, animats

**Funding support:** JTF, TWCF, DARPA, Paul G. Allen Frontiers Group, Sloan Foundation, NIH, NSF

**Illustrations:** Jeremy Guay @ Peregrine Creative

**Disclosures:** Morphoceuticals, Fauna Systems, Astonishing Labs



These are all the people that I need to thank: the postdocs and the students who did some of the work that I showed you today, lots of amazing collaborators, our funders. Disclosure: there are three companies that have spun out of our lab that fund some of the work. These are the commercial entities.

The model systems do all the heavy lifting. I always thank them. I thank you for listening.

**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](mailto:adi@aipodcast.ing).