Hello everyone, welcome to another episode of Q&A about Future of Science and Technology.
And I see a variety of questions here.
Oh, here's one.
From Crossroads. If we ever find alien life, how much do you think it would actually change science as we know it?
That's an interesting, messy question.
Because the first issue is, what would we mean by alien intelligence, or alien life, for that matter? What,
we have to kind of have an abstract definition of intelligence that goes beyond the form of intelligence that we're familiar with, which is human intelligence. When we think about artificial intelligence, AI,
The things that we're most excited about are the versions of artificial intelligence, the versions of computational kinds of things that are somehow aligned with human intelligence.
in lots of things I've done, just sort of exploring the computational universe of possible computational processes.
We don't think of that really as being something like intelligence. We don't think of it as being… we could think of it as being alien intelligence, but we wouldn't usually call it that, because it doesn't seem like intelligence of the kind that we're familiar with in humans.
So, you know, a thing I often commentated, we look at the physical world, we also see things which one often attributes intelligence to, or at least at a kind of human level. Like, we'd say, the weather has a mind of its own. It seems to be making choices of its own.
That we can't, that we can't predict.
Similarly, in many, kind of, early, kind of, well, in many, kind of, sort of, traditions, spiritual traditions about nature, it's like, you know, the spirit of the wind, or something like this. It is something where that is a way of describing the sort of
complexity of behavior of the wind, it sort of seems like an intelligence. It seems like a thing that's making choices for itself.
where we can't readily predict what it's going to do. We wouldn't say so much it's the spirit of the falling ball or something, because we know that it's… we kind of feel like we can readily predict the ball is going to fall in the way it falls. If, on the other hand, it was the thing that was fluttering around in all sorts of ways, we might more identify that as a thing that was driven by some sort of spirit, some kind of intelligence-like
force.
In my own efforts to understand, kind of, how things work.
I've been big on this idea of the principle of computational equivalence, the idea that when it comes to, sort of, the computational processes that are going on, a vast range of systems are equivalent in what kinds of processes they have going on, which means that fluttering thing that is… has all sorts of complicated dynamics based on its structure and fluid dynamics and so on.
might be computationally just as sophisticated as the brain of some creature that's figuring out whether it should walk left or right, and so on. So it's kind of the idea that there is, in a sense, a lot of
intelligence of that, in that kind, something like intelligence, some abstract generalization of intelligence, insofar as that is like just the process of computation, that's something that's ubiquitous in lots of kinds of systems. Lots of kinds of systems out there in the cosmos, lots of kinds of systems here on Earth, and so on.

Now, the issue is, most of those kinds of systems, we say, well, yes, it's kind of, it might be kind of like a spirit of the whatever, if we were thinking about that in a sort of not very, sort of science-y kind of way, or we might say, well, yes, it's a compu… we can think of it as a computational process, but it's not intelligence of the kind that we're used to with humans.
Again, there's this sort of corner case of AI, where it is something which is fundamentally a computational process, but it has been specifically set up to be aligned with human intelligence.
See, if we looked at a neural net.
of the kind that's used in modern AI, and we just said, let's scramble the weights in the neural net. Let's not have weights that were specifically set up to align with the ways that humans think about things. Let's just have a random neural net. What it does is just some random computation.
That computation is, I think, by any definition, sort of as fundamentally sophisticated as the computation that's done
In the kind of human-aligned neural net that's been trained to work on making images of things that humans are used to, and so on.
So it's… but the raw computation is the same kind of thing. It's just the computation happening in the neural net with random weights is something that is incomprehensible to us humans. We don't resonate with it, it's not reminiscent of the kinds of things that we are used to doing in our own brains, and so on.
So, the question then is, well, what about alien intelligence from the cosmos, so to speak?
Well, the fact is, at the level of computational sophistication, there is alien intelligence
Out there, all over the place.
You know, when we see… I mean, I think one good example of how this sort of sets itself up is a piece of history from about 130 years ago or something now, when,
when radio was new, and people were first starting to kind of listen to natural radio emissions.
So, famously, I think Marconi, had a yacht that he was going back and forth across the Atlantic, on, and in the middle of the Atlantic.
You know, he had a radio mast on his yacht, and he could hear these kind of funny sounds, being picked up by, and, you know, converted to sound from radio waves and so on.
Didn't know what they were. I think Tesla had a theory that those things were communication, radio communications from Martians.
what were they in actuality? They were modes of the ionosphere. They were magnetohydrodynamic phenomena in the Earth's ionosphere that were leading to, it was…
were leading to radio emissions that had sort of sounds that were… kind of seemed like they were kind of eerie, sort of, woo-type… type sounds, that
At the time, it's like, well, what could make those? Oh, it must be, thought Tesla, for example, the Martians, because what else would make those kinds of things? Well, it turned out to be just a, quote, natural process.
Now, of course, if you're really sort of philosophically pure, you would say, well, the Martians, if they're out there, they are also a natural process. They happen to be a natural process that we imagine are creatures that are, you know, red and green, or whatever they are, that are wandering around and seem to be doing sort of human-like things, perhaps, that makes them more the Martians, as opposed to just
some natural process happening on Mars that produces radio emissions.
So, I think it's kind of a subtle thing when you say, is there alien intelligence out there? The alien intelligence… alien processes
That's easy. Alien intelligence, well, it depends what we mean by intelligence.

Now, if you say, so, to the question about whether… what we would learn… so, again, this… this… the sort of… the thing that one
obviously wonders is, of all the computational processes that are out there, are there ones that are sort of very closely aligned with what we humans do.
sufficiently closely that we would say, oh, we met an alien species type thing, rather than that's just some weird natural process that's nothing like us, that is producing radio emissions, and so on. And I think this question of, well, are there…
What's the chance that there are, sort of,
Computational processes that are like us, that we identify as being very like us out there.
That's an interesting question. I think it's… it's this sort of question in… framed in terms of the kinds of things I think about, of, if we think about physical space, we can ask, you know, how dense are the planets on which we can have certain chemical processes going on?
If we think about what I call ruleial space, the space of possible, kind of, rule systems that we could, for example, attribute to the behavior of the universe, what is the density of, kind of, what's…
what's the density of kind of mind-like things in ruleial space? I kind of think of different minds as corresponding to different points in ruleial space, and the question is, how sort of far apart are minds in ruleial space? If the minds are very far apart in ruleial space, it takes a huge amount of effort to translate the kind of computations going on in one mind
to the computations going on in another mind. It's very similar to when we think about physical space. If we say, how easy is it to get from here to the next galaxy? Well, it takes a lot of effort to get from here to the next galaxy. There's a lot of, in that case, it might be energy expended and so on. There's a lot of underlying computational processes that have to go on
To get from here to the next-door galaxy.
So, that's a sort of… that's what is involved in going across physical space. Across Ruleial space, it's sort of directly computational effort. It's, can you take this kind of computational device and make a compiler, effectively, that converts it from this tower, interpreter that goes from this kind of computational device to another? What… how many computational steps are involved in doing that?
By the way, in our models of physics, there's a pretty direct correspondence between these different stories, because it turns out that energy you can identify as kind of the density of activity in this computational network that represents the structure of space. So it really is the same kind of thing of how much
computational effort is going on to, for example, send a spacecraft from here to there, we interpret that as physical energy in that particular case, or how much computational effort is necessary to translate from this way of thinking about things to that way about thinking about things. That might actually, if we implement it on a computer with current computers that aren't reversible and so on, it might also take energy
But the main point is that, sort of, at the level of counting computations, we can do it in both of those cases.
So then the question is, well, what's the chance that in ruleial space, there are critters out there close to us in rural space? We kind of know what the chance is that there are critters out there close to us in physical space. We know, you know, the next-door solar system is 4 light years away, so that gives us some sense of how spaced out things are in physical space. We don't really know how spaced out things are
ruleial space. My guess is that they are much more spaced out

in some sense of conversion than they are in physical space. In other words, the chance that there are minds out there that are sort of identifiably aligned and similar to ours is even much lower than the chance that there are things that are physically nearby in physical space.
So, in other words, the chances are that the things that are out there in the span of our universe, in the size of our universe, the chance that there are kind of minds aligned enough with ours that we would say, oh yes, that's, you know, a science fiction-style alien intelligence that we can kind of recognize and interact with in that way, I think that chance is rather low.
I think the universe just isn't big enough that we're likely to have that much alignment in rural space between the entities that are out there.
Might not be true, but that's… that's what I suspect. And I suspect that a lot of things where you'd say, well, you know, we've identified this… this surprising phenomenon out in interstellar space.
We've seen a pulsar. People imagine when pulsars were first observed in the first week that pulsars were seen, where there were these periodic radio signals
coming out every few seconds or whatever, the first thought was that might be a beacon from some extraterrestrial intelligence. Of course, it very quickly became clear that, well, actually, you can think of it as just a rotating neutron star that's just a piece of physics. But, you know, making that distinction between the rotating neutron star with its complicated magnetosphere and so on, and the thing, and thinking of that and saying, well, that's just a physical
process, as opposed to, oh, there were these brain-like things that figured out that they should make radio pulses that work this way. How do you make that distinction? Because the brain-like thing ultimately also just has a bunch of physical processes of electrochemistry, of neurons, and so on, in our particular case, inside it.
So, again, it's this question of can we identify enough alignment with us and the ways we do things to be able to kind of say, oh, that's an extraterrestrial piece of intelligence or life or whatever, as opposed to just that's an extraterrestrial piece of physics.
Now, what makes this even more subtle
I mean, I realize this is sort of a rather philosophical answer to this question, but I think one has to get the philosophy straight before one can actually make a serious answer to these kinds of questions. What makes this more subtle
Is that science, its mission, in a sense, is to make this bridge between our thinking
and what's actually going on in the natural world. Science is about creating kind of a narrative about what happens in the natural world that is a narrative that kind of works, that we can operate with in our minds. So, in a sense, to say how aligned is this thing out there with
Our minds is, in part, how far can our science reach?
Is it the case that we can sort of make this bridge from what's happening in that thing out there to what we think about in our minds, or not?
And I think that's… that's… so that, you know, if we were to observe, for example, observe the pulsar magnetosphere, and ask the question, is that an example of alien life?
Well, maybe there's an interpretation of that that we can give scientifically, where it will look an awful lot like alien life.
Or maybe not.
To give an example, in life on Earth.
We know it's evolved in all sorts of complicated ways, you know, 10 to the 40th organisms have lived, they've… they've… we've got certain kinds of structure that emerges from all of that

computational adaptation and so on that's gone on in the course of life on Earth. Let's compare that with geology.
A certain amount of
similar kinds of processes have happened. You know, certain rocks survive, other ones don't. One kind of rock turns into another kind of rock, and so on. How do we compare, kind of, the computational processes, the… is it… is geology lifelike or not? How… how similar to, you know, this kind of rock begats that kind of rock? After all.
If you start with a crystal, the crystal will try and get… as the crystal grows, it'll move atoms to be sort of like the crystal as it already was. That's a bit similar to what happens in self-reproduction in biology, where the creature takes in things from its environment, it takes in molecules from its environment, and it turns them into molecules that are arranged
In the way that they are in that particular organism, and so on, to make a copy, a rough copy of itself.
So… In any case, I think that the, this whole question of, of sort of how,
The alignment of kind of things going on out there with the things we think about in our minds, the fact that science is sort of expanding the domain of things that we can think about in our minds, it's a fairly complicated interaction of those different kinds of ideas.
Now, you know, the next question is, well, let's say that there was a somewhat aligned kind of alien intelligence out there.
And let's say that it made technology.
What does that even mean? You know, what does technology mean for us? Technology for us is we take the world as it is, the natural world, whatever, the abstract computational world, and we say, how can we sample that world to get something that we think is useful for us?
That's essentially the operational definition of technology. Technology is, we take stuff from the world.
And we sample it in such a way that it's useful to us. We figure out that, well, yes, you can sort of cut down a tree that was there anyway, and use it to make timber that we can make a piece of a house out of, or something like this. Or we can take a thing that was, you know, a magnetic rock.
and we can suspend it in such a way that we make a compass with it. We're taking, we're sampling pieces of the natural world in those cases to make technology for us. Something I've spent a bunch of time doing is sampling the computational world, saying, look at all those possible programs that are out there. Which ones of those programs that are just out there
that you could abstractly construct which ones of those programs are useful for some particular purpose that we have. That's kind of a definition of technology for us, is how do we forage what's out there and find things that are useful for the purposes that we've defined for ourselves?
Okay, so now we ask, what does alien technology look like? Well, alien technology
is… Presumably, something that is useful for the purpose of the alien.
Well, what does that mean? I mean, you could say there is technology if you have, I don't know, a,
some flow of water, geologically. You could say, well, this thing that involves, you know, this particular rock that's porous or can dissolve in this way, that's technology that's useful for the alien intelligence that is that geological thing that's the flow of water, and so on. And I think that, that's kind of,
So, sort of purpose, is something that…

We can define for us humans, but to attribute purpose to other things is, at best, a very messy concept.
And, again, there's another sort of philosophical twist to this, which is any process that you look at, you can say, this is how the process works, step by step.
There's this mechanical thing that leads to that mechanical thing, and so on. It's an explanation of the process by the mechanism by which it works.
You can also look at that process, and you can say, well, you know what? That process is one that minimizes the action quantity in physics, or something like this, or gets the water from here to there as quickly as possible.
It's a process that you can define, describe in terms of a purpose that you attribute to that process.
And so, typically with any process, there's an explanation in terms of mechanism, and an explanation in terms of purpose.
sometimes the explanation in terms of mechanism will be much simpler, sometimes the explanation in terms of purpose will be much simpler. You know, if you say, why did the, I don't know, why did the cat reach out and grab the mouse?
Well, the purpose of… well, it wanted to, I don't know, eat the mouse, or whatever it is, that's a, That's an easy thing to describe. The mechanism of, oh, the electrophysiology of the… of the cat's brain led to this thing based on its eye, etc, etc, etc, etc, etc, that's a very complicated explanation in terms of mechanism.
So, these things where you say, well, let's look at the natural world, and let's say, do we explain this thing in the natural world in terms of purpose, or in terms of mechanism? That's a, that's a complicated thing, and you can't, sort of.
You can't… purposes,
when it comes to sort of defining things for technology, the sort of the base of that purpose ends up being, for us, human purposes, and it's not clear what another base would be. If we just see a thing out there in the world, we can say, well, gosh, we can use that
for technology, and, you know, you could say, well, that's a, you know, that's an alien technology. We're bringing in the alien technology. Let's take pulsar as an example. You say, well,
you know, are pulsars an example of an alien technology? Maybe we can think of the dynamics of a pulsar as being some sort of thing that we attribute some generalized alien form of intelligence to. Then the question is.
the thing it made, is it useful to us? Well, yes, in principle, we can use pulsars as a way of doing, sort of, GPS in interplanetary space. We can say, where are all the pulsars? We know we're measuring the timing of the pulses from the pulsars and so on. We are harnessing pulsars for human technology, and that's a sense in which we would be,
sort of taking something which is alien out there, but it's really all the same thing. It's really all taking the world as it is, the natural world, and finding a way to sample it for human purposes.
Now.
you know, I suppose it's interesting to sort of speculate, and much science fiction has done this, about what would happen if there was a thing close in rural space that was sort of comprehensible to us as being a human-like intelligence, but it didn't come from here. What would be… what would that be like? What would that… what would that experience be like? And so on. Well, obviously, in the history of human civilization.
There have been plenty of examples of that, of, you know, the uncontacted parts of human society that get contacted, and then there's all sorts of, well, what does this part do and that part

do? I think one of the things that's perhaps most striking about some of those encounters is the extent to which one
side of the encounter just looks at what the other side is doing and shrugs their shoulders. Like, we don't know what the point of that is. You know, people will come to some uncontacted tribe somewhere, and they'll see they're doing all kinds of things. And the anthropologists will describe that as, oh, it's ritual behavior.
But really, if you were in the mind of the person who's doing it, it's as meaningful behavior to them as lots of things we do.
I mean, lots of… lots of activities that we do, looked at by somebody who doesn't understand our internal point that we think we have for those activities, would just say, oh, it's ritual behavior that somebody is getting into their car every morning, driving to this place, you know, hanging out for a little while and driving back again.
We would say, it's, oh, that person is commuting to work, and we'd have a whole chain of, that's why that's happening, et cetera, et cetera, et cetera.
just as somebody who is, who is, doing something where they're, I don't know,
you know, making, arranging sticks in a particular way every morning and then rearranging them every evening. They'll give an explanation, too, about why they're doing that, and it's… it's, you know, we could have a whole discussion about who has the more sophisticated explanation and so on.
I think that's a, that's a different level of question, and there are some questions you can ask there, because there's sort of… there's questions about, sort of, what level of formal abstraction do you achieve on the different sides? And the truth is, most human activity, as such, does not achieve a great deal of formal abstraction. We have built tools, particularly computers, things like mathematics and so on, that allow us
to go many levels up in formal abstraction, and that's sort of a new thing that's different from kind of the raw behavior of us as humans. But it's sort of an interesting question what,
Kind of,
you know, what happens when you have very nearby, kind of, mines close in ruleial space? I think the answer tends to be that the things that aren't in common
are often things where one side or the other will just shrug and say, I don't see the point of that. Yes, you're doing that, but…
You know, so what, more or less.
I think, perhaps an interesting point to make in terms of,
Thinking about, kind of, the interaction of different, sort of, intelligences, computational elements, whatever, in a world where computational irreducibility
is this kind of limit to how quickly you can compute a particular kind of thing. For example, in our universe, there is, in some sense, a fundamental value to having done an irreducible computation.
I mean, ironically enough, that idea, which I, kind of first
started talking about it in the 1980s, may very well have ended up being an idea that turned into things like the proof-of-work, let me prove I did something idea for things on blockchains and cryptocurrencies and so on.
a kind of a let's just grind computation for the sake of showing that we did something. But one can imagine that sort of pulling in the results of irreducible computations from other parts of the universe kind of feels like trading with an alien intelligence in some sense. I'm not sure it's quite trading, it's just ingesting the results of that.

Anyway, I think,
a few… a few thoughts on, kind of the story of alien intelligence. I mean, I think one thing to say is that just scientific discovery is like the discovery of alien intelligence. Every time we discover some surprising thing that happens in the world.
It's kind of not far away from discovering some… something we can think of as an alien intelligence, and then trying to relate that alien intelligence, either through our science or through our technological purposes, to things that we do.
Let's see…
Jameson is asking, do you think it's more likely that alien life is more similar to humans, or is like bacteria or trilobites?
Well… The…
This relates to the question of what's life and what's special about it. I literally just wrote something this week about that very topic.
So… One question is, is life just chemistry?
In other words, if we see certain chemical processes, do we say, aha, that's life right there? I remember when I was a kid, the first spacecraft landed on Mars.
And they had experiments that tried to test, is there life on Mars? You scoop up a piece of soil and you say, is there life here?
Well…
some of their experiments were very chemical experiments, like you feed it sort of a sugar, and you see whether it has certain chemical reactions that look like metabolism and so on. I think that pure chemistry, and this has become clearer as we look at exoplanets and exoplanet atmospheres and things, or even the atmosphere of Venus, pure chemistry isn't really the story of life.
In other words, there's something different about life
than any kind of ordinary chemistry. What is ordinary chemistry? In ordinary chemistry, you expect that you have molecules that are just sort of bouncing around. In a gas, more often in a liquid. Liquids are the places where lots of kinds of chemistry, kind of happen, because molecules are just jiggling around and hitting each other at random, but yet there's a decent density of molecules, so they actually do hit each other. In a solid, the molecules are all locked in place.
So they don't tend to hit each other and be able to react with each other. But chemistry is mostly the story of things in particularly liquids, where molecules are just meeting strangers all the time, and as soon as there are molecules that happen to be such that if they hit in exactly this angle and so on, they'll react, they'll interact, the molecules will combine, or whatever else, then that's going to happen, just by chance, enough in a liquid
That you see things going on.
So the question is, if you look at us, living systems like us, are we made of liquid stuff?
The answer is no, not really. The main discovery, I suppose, meta-discovery of molecular biology in the last few decades has been, if you look down at the level of molecules, except for very tiny molecules like water and so on, by the time you're dealing with proteins and big, hunking molecules.
There… it is not the case that…
That those molecules are mostly just bouncing around randomly.
Okay.

Those molecules are, instead, it seems, always being orchestrated in what they do. This molecule, this enzyme, is moving this molecule to this place and arranging it so it will do exactly this. This funky splicosome
Object is taking in little pieces of, of,
of proteins and splicing them together in complicated ways. It's like a lot of little pieces of machinery. It's not like a thing where, oh, the molecule's just bumping around randomly, and when they happen to geometrically fit, something will happen, or some such other thing.
And so there's this question of, well, is what's special about life?
this kind of bulk orchestration of molecular processes. I think the answer is, to some extent, yes. That that is a thing we do not see elsewhere. We don't tend to see this phenomenon where we have a lot of, sort of, components that are, in a sense, seem to be orchestrated in that way. It's just they're bumping around randomly and things are happening, rather than we can see a chain
Of, kind of, supply chain of what's going on
That is, well, more than just a supply chain, a kind of a mechanistic process that is leading to this or that thing happening.
So I think that's sort of a special feature of life. Where did it come from?
Well, I think it came from the fact that there is overall this kind of force of natural selection of biological evolution that is kind of a forcing function on the whole dynamics of the whole system.
So a thing that I discovered a few weeks ago, actually, that I've been thinking about for a long time, but it sort of came together very recently, is the idea that as soon as you know that a thing was the result, that the rules by which a thing operates were the result of adaptive evolution. Where the purpose, the objective in the adaptive evolution was something that is computationally simple to describe, as soon as you know that the rules of a system were made that way.
Even if you don't know what the objective was, it turns out that you can make certain statements about what rules like that will do. And the main statement seems to be that they show what I've been calling mechanoidal behavior, a behavior that, at least on a small scale, looks like little pieces of mechanism.
And I kind of think that that's sort of the special thing that we've ended up with in life, is this thing where there is an overall kind of
very macroscopic forcing function of, sort of, the fitness of the organism, and the surprise is that the presence of that sort of leads down in the trenches, so to speak, at a low level, leads to actual… that only certain kinds of rules
will be able to be used in systems which have that overall adaptive evolution, and you can say about all those kinds of rules that have… that are made that way, that they will have certain… certain properties. I've been calling this the Ruleiel Ensemble.
The collection of rules that are
that are determined by some overall constraint of achieving an adaptive purpose. That's kind of modeled on what one does in statistical mechanics, where you talk about an ensemble, like you might talk about the microcanonical ensemble.
Which means the collection of configurations of molecules in a gas that have a particular energy. So the overall constraint is they have a particular energy, but then you're looking at all those microscopic configurations, and the big discovery of statistical mechanics from 150 years ago or so now is that you can make statements just by knowing that you're looking at all these different configurations without knowing the details of which

what the configurations exactly are. And my basic claim is that with this idea of the rural ensemble, you can again make those kinds of statements in the space of possible rules and so on.
So I kind of think that the definition of life
I hadn't thought there was such a thing, but I kind of think it relates to this phenomenon of, of bulk orchestration.
that is the result of some kind of downward causation of this large-scale sort of phenomenon of achieving some fitness function that then has this consequence at a small scale. I kind of think that that's the… that's a feature of kind of, something that we should consider to be sort of a bit of a lifelike thing, so to speak.
And… and it's something where it has these different scales, these different dynamics of kind of, sort of, there's overall objective, and yet there's… there's… there's sort of microscopic consequences. Now, will we observe that?
in, you know, in the, I don't know, the oceans of Enceladus, or something like this.
I don't know. Good question, which actually I haven't thought about, and I should think about, is how would you make a detector?
for,
that kind of bulk orchestration. I'm not sure I know the answer. That's a good question. It's… it's, it's a good one for… I should have thought about that. I… I'm,
I goofed, I was, but, so, you know, how would you make a device that takes a scoop of stuff from the ocean of Enceladus and says, is there bulk orchestration happening here or not?
In a sense, the whole development of modern molecular biology is a story of validating that there is bulk orchestration in life on Earth. In other words, if you can fill a whole textbook talking about how the thing works, then you've probably got bulk orchestration.
If, instead, you just say, oh, it's a bunch of molecules randomly bouncing around, that's all you can say, and you're never able to say more than that, then you probably don't have this bulk orchestration phenomenon.
And if, in other words, if what you have is something that is more like pure thermodynamic behavior, where you can ultimately say something, but only by knowing, sort of, the configuration of every molecule.
that's something which isn't really bulk orchestration, but if you have something where there are little mechanisms that you can start talking about, that is something where I think one can… and little mechanisms, and many of them, that would sort of fill a biology-like textbook, then that's sort of an example of bulk orchestration. So I suppose what that suggests is that
When you kind of scoop up the ocean sample, the thing you should be looking for
is how many kind of pockets of reducibility, how many explainable mechanism things are there there that you can latch onto? If there are lots of them, then you have something that might be the result of some adaptive evolution process of the kind that exists for life on Earth.
If not, then maybe you don't. Now, another question, reasonable question, is might there be other things other than adaptive evolution that lead to this kind of bulk orchestration? I don't know the answer to that. Another good question.
But anyway, a few thoughts about that.
let's see… A reflect is asking, is probability a factor for life?
In other words, I suppose, is it useful or necessary that,
things are being sort of randomly jiggled around. Is that a thing? If everything was sort of perfect and crystalline and so on, would

would sort of life be able to exist in the form it is? I think it's interesting, because I think that when it comes to adaptive evolution, a lot of what,
The… at least in the ways I've thought about adaptive evolution, you are rolling dice.
to decide what to do next. Now, you can look at the whole multi-way graph of all possible paths of adaptive evolution. That's an interesting thing to do at a theoretical level.
But in terms of, well, what happens in practice, I think the main thing to say is, it really doesn't matter much which path you take. And so, if you take a path by rolling a dice, that's… that's one thing. You could also take a path by having some level of systematic
kind of decision about which path to take. My guess is that most of the time, that doesn't matter.
Now, having said that, it's a little tricky.
And in fact, in our model of fundamental physics, we have this issue, because in our model of quantum mechanics, we have these many different paths of history, and you are essentially the things that you're… the universe is following all those paths of history. And…
we are sampling some subset of those paths of history. And there's a question of if you sufficiently pathologically sample
one particular kind of path of history, can you see a universe that is bizarrely different in its behavior from ours? And the answer is yes. And so then you have to start asking, well, does it matter that we're sampling this big block of paths of history? The answer is yes. Now, does it matter whether those… whether there's sort of randomness in which paths of history we sample? I think it's probably the case that any sort of patch of positive history will probably lead to very similar behavior, but I'm not sure about that.
Let's see…
Gosh. Alright.
Let's see…
ZAP is asking, we might infer from LLMs… it's a long question, so let me, I can't, I don't know what's going to happen at the end of the question, so I'm going to start reading it.
We might infer from LLMs that we can decode our world model, sum of knowledge, scientific paradigm, in a way that can be communicated with.
Assume interstellar travel never becomes viable. Imagine encapsulating something like human consciousness in an interstellar signal in the hope that someone somewhere would receive it and implement the LLM. A pulsar or quasar would be useful for that. What if a distant civilization has already done this?
In other words, you might say, what if the sum total of a civilization is sent out as a radio signal, and and we are hearing that radio signal, and we just say, oh, that's radio noise, we don't know what it is, you know, throw it in the trash type thing.
Well… Certainly, a number of things can be said about that. For example.
let's say, I don't know, you have a, a neutron star merger.
That two neutron stars spiral in, they collide with each other. Out comes a whole bunch of gravitational waves.
One of the things I know people are excited about is, in a neutron star merger like that, the neutron stars tear each other apart.
And as they tear each other apart, those gravitational waves are essentially sampling more and more layers in the neutron star. You're getting to do, sort of, some very bizarre form of tomography on a neutron star as it gets torn apart by another neutron star. So in some sense, the gravitational wave signal from those neutron stars is the last gasp

of the civilization, I put that in quotes, that was those neutron stars, comes out to us in a burst of gravitational wave energy.
And are we able to decode it? Well, who knows? I mean, are we able to detect it? It's pretty hard with gravitational waves, but that's an example of where, sort of, the whole story of the neutron stars, as they're torn apart, you know, we're seeing all those different layers of the neutron stars, and they all came to us in a pulse. And I can imagine the same thing happens in other astrophysical phenomena.
And, you know, one might say, well, well, let's use that pulse to explore the story of that… that thing, and that thing
If that thing was interpreted by us as an alien civilization, then, by all means, we will be getting that sort of, that last gasp of the alien civilization where it had packaged everything up.
you know, maybe in a supernova, something similar might be considered to happen of, sort of everything. It's a little different because, well, it's, you know, I don't know. In a neutron star, there might be all kinds of structure inside the neutron star. We don't know what kind of structure there is, whether it's a quark gluon plasma or what it is.
And whether it's a kind of a mushy, liquid-like thing, or whether it's more something which has all sorts of details.
Of the kind that, you know… So a reasonable question would be, let's say there is a bulk orchestrated thing.
of which a neutron star probably isn't an example, and let's say the thing is torn apart and sends out a radio signal that is a reflection of all that bulk orchestration. That would be sort of an example of, well, we'd be getting a signal from the cosmos which is
the result of endless amounts of computation. Like, for example, as I say, in life on Earth, 10 to the 40th organisms have lived. We are the beneficiaries of all those 10 to the 40th organisms that have lived.
We, in a sense, are a reflection of all the computational effort that's gone into the lives of all those organisms, and now there's the question of if you sort of take us and say, well, can we get the benefit of all that irreducible computation that's happened?
The sad thing is that, mostly, when we get that radio signal, it'll just look like noise to us. We won't have the… we won't have the kind of the context to be able to decode it. It's, again, a science problem, in the sense that it is a thing of the natural world
And there's a question of can we make a bridge from that thing of the natural world to something that we can narratively understand in our minds? And that's the question of, could we ever decode the signal? Or is the signal something that is too alien for us to decode? And we could say, well, there's this signal.
And, I mean, I see this all the time in studying just simple programs in the computational universe. The program does what it does. It looks kind of cool, but does it relate to anything that we directly have a way to think about? And the answer will often be no, and I think that's the same thing that will happen in that sort of,
burst of communication from the thing we might think of as an alien civilization. It'll be this thing that is perfectly meaningful to the alien civilization, but is so far away from the things that we normally think about that we don't really have
A way to kind of connect it to things we care about, to make a narrative that works for our minds.
Let's see… It's a question,

Oh gosh, lots of questions here. Lots of interesting questions. Bob is asking, if you need time to do computations, then does time exist before computation, or does computation create time? I've written about that at some length, actually. The, just to say sort of a short answer to that.
Is, in our models of physics these days, comp… the…
The actual computation of the next state of the universe from the last one, the application of the rules of the universe to go from one state of the universe to the next, that process
is the progression of time. In other words, the fact that these rules are being applied, that's… that succession of application of rules is the progress of time.
And you say, well, well, from outside the universe, what happens if the rules just weren't applied? Well, the answer is then time does not progress. But as observers inside the universe, who are subject to the same laws as the rest of the universe.
time is progressing inside us as time is progressing in the rest of the universe, because both inside us and in the rest of the universe, there are these computations going on that update the state of the universe. It is a very non-trivial fact.
that time works the same way in our psychology, and in the natural world, and so on. But the reason for that, I think, is that we're all made of the same computational stuff. We're all… time for everything is the doing of those computations. And that's sort of why these different kinds of time align with each other.
let's see…
Question here from Upcycle. Could bulk orchestration be mathematically modeled in terms of compositions of actions of finite, simple groups?
I don't think so.
Let's see how to think about that.
Let's think… I mean, this question of How do you define mechanism?
is an interesting question that I don't think I've answered.
We can say some things about when there's mechanism, there's predictability. When there's predictability.
There's the ability to use some modeling method, whether it's statistical, data compression, whatever else. Something where we have a way of making models of things, we can apply that way of making models of things
To this thing that we say has a mechanism, and if it has a mechanism that kind of relates to those ways of making models of things.
we'll be able to compress what's going on. We'll be able to say, I understand that mechanism, it's just a, you know, a thing going in a circle, for example. I understand that kind of mechanism, it's just periodic, so I can predict what's going to happen.
So I think that, sort of, the ability to do compression, the ability to make models, is related to the presence of this kind of, you know, underlying mechanism.
Now, you know, what kinds of models might you make?
You know, the simple models that implicitly get made in data compression are things like, oh, there's a block of data that will occur repeatedly.
Or there's, you know, you have different, sections that are, like, things like words, like the word rhinoceros, will appear repeatedly, even though it's by just the probability of letters, the combination rhinoceros is pretty rare.
But it will occur, sort of more often than you expect, because it is a word in English, and then you can kind of compress by saying, well, there's a dictionary of words that appear in English,

and so on. So there are a variety of different methods for doing that, but I would love to have a more general way of identifying what
what is a mechanism? Actually, we're just working on that right now. It's… I think there may be some things that you can use from category theory and mathematics.
I don't know… I don't immediately see a way to think about things like group theory, although maybe… maybe there's a way of doing that by thinking about… yeah, I mean, I suppose to some extent that one could say there is.
In the sense that… okay, so things get a little tricky here, and this is going more into the science direction, but… but often in computation, it's kind of a… you have a thing, you apply a rule, you get another version of that thing. In mathematics.
It's much more, it's much less this progression through time of things. It's much more, you have two things, are those things equivalent? Is there some path that goes from here to there? Not thinking about walking down that path.
just thinking about, does the path exist? And so, for example, in something like group theory, that does tend to be much more of the, the, are these things equivalent, and so on.
Rather than, if I do this operation on this thing, will I land up at that thing?
But I think there's a way of thinking about… so, there's this question of.
For example, to say it a little bit technically for groups, so just for people's benefit, the, when you do something like, oh, I don't know,
when you do ordinary arithmetic.
you expect that A times B is the same as B times A.
and you expect various other properties of things like numbers. If A and B aren't numbers, but are some quite different kind of thing, and you say,
like, let's say A and B are operate… rotations you can make on a hexagon. So, you know, A is rotate by… by modern… by…
I don't know…
60 degrees or something like this, and then some other thing means is reflect the hexagon and so on.
you can ask, if I do A times A times A times A times A, I apply that… that rotate by 60 degrees, and I apply it, you know, 6 times or something, then I'll find out that I get back to the same thing again. So A to the 6th, in a sense, is equal to 1. And group theory is the story of…
Of making, kind of, setting up
Things that have various properties about operations with some kind of generalization or multiplication.
And so…
what you can do, you have these objects, like a thing like rotate through 60 degrees might be a generator of the group, of that group, and you can ask, well, I'll make this kind of composite thing by multiplying together generators. I do this, you know, I rotate by 60 degrees, I reflect it, I rotate by another 60 degrees, and so on. Each one of those sequences of generators
Is what we can call a word in the group.
And so you can build up these words.
But then one of the big questions is, well, which words are equivalent to which other words?
And often in group theory, as it's normally practiced, you would consider things which are…
which are written out differently, like it's A to the sixth, is written out differently from the identity, 1, but…

they're written differently, but for purposes of, kind of, official group theory, you say, but they represent the same thing. Even though they're different words, they represent the same thing. I'm afraid this is getting slightly technical, but… but the whole question about, sort of, when you compute things, it's more like just making words
Rather than it's about saying, well, what's equivalent to what? Taking the whole word, sort of, in a big gulp, and just saying, is it equivalent to that other thing?
So I can imagine that this whole question about, is there a mechanism, is a little bit related to questions about, is there a way to talk about things, not in terms of the words, individual do this operation, this operation, this operation.
But instead, is there a way to aggregate that into sort of these big gulps of saying, they're just these elements that represent maybe infinite numbers of words and have certain properties that can be thought about at that level? So it's an interesting, interesting, perhaps, way to think about it, so thank you for that poke in that direction.
Gracie is asking, if aliens evolved in the universe with different fundamental rules, what might their version of math class look like?
Well, I don't think you have to go,
You know, you don't have to go that far, actually. I mean, you know, our math class is very much based on our sensory experience of the world. So, for example, the fact that we count things is
a feature of our sensory experience. We have objects, we pick up solid objects, it's like there are… there are five pebbles or something.
If we existed as, you know, in a gas, for example, where everything is just flowing around, we might not ever identify, kind of, discrete things that we could count.
We might be just saying, oh, there's this sort of deformation of the flow field of the gas, and that's what we're talking about. Even maybe for aquatic animals might have the same type of thing, although one kind of thinks that if they're picking up, you know, there's a fish, there's another fish, it's like, count the fish type thing. But, you know, you could imagine in terms of their experience of motion.
That it might be rather different, and not have,
You know, I, I am…
I was walking upstairs, and I went one step to the next, to the next, to the next, and so on. I'm not sure that's a very good analogy, but,
There are many things about the mathematics that we have that are pretty specific to our experience of the world.
And you can certainly imagine building, you know, when… what is mathematics? Well, you know, a common definition for the last 130 years or so has been, it's something that's based on sort of an axiomatic description of things. So you might say, what is arithmetic?
oh, it's the study of things where X plus Y is the same as Y plus X, where X plus 0 is X, where blah blah blah, a certain set of axioms that describe, this is the kind of thing we're talking about when we talk about arithmetic. That's one view of what mathematics is, is the thing you get by working through the consequences of certain axioms. Now, that approach
of just saying, given the axioms, we can just mechanically grind out all the facts of mathematics. That approach, that was kind of the Hilbert program from the beginning of the 20th century, where it was kind of the idea was, once we found the right axioms, we just grind out everything else. That turns out not to work.

Godel's theorem kind of showed that that doesn't work, in the sense that there are things that… there are plenty of things where the question of whether
you… where there's sort of an infinite path you need to go to get to that thing. It's a… it's a question you can reasonably ask about, for example, integers in arithmetic.
but which you just can't reach by sort of axiomatically developing from this kind of underlying description that you've got. Now, if you say, well, what kind of a math can you have?
If you're taking this axiomatic view of mathematics, well, you can just start generating axiom systems at random. You know, we, in doing human mathematics, there's maybe half a dozen axiom systems that we commonly use. The most, sort of, fanciest is, a version of set theory, the theory of infinite sets.
the, you could just start generating, enumerating all possible axiom systems. I've done this a bunch of times, actually, and asking, well, what are the consequences of each of those axiom systems? Each axiom system, in a sense, is a field of mathematics that you could imagine defining.
It's not one that we humans have ever looked at, but you can just generate those fields of mathematics by just writing down axiom systems, maybe generated at random, and seeing their consequences. I've done that a bunch.
you can work out the sort of rich structure of theorems that come from those axiom systems. If you say, what is the human significance? How do we tell a story about that? How do we say it's… we talk about it in terms of moving pebbles around and things like this, but we don't know how to do that. We haven't made the bridge
From that abstract thing out there in the world to something which is a sort of science-like narrative that we can deal with in our minds.
And in some sense, human mathematics is that
kind of human narrative way of talking about abstract kinds of things. And so, this view that mathematics is this thing that you get by generating things from axioms probably isn't really the right operational view of what mathematics is. Mathematics instead is something that is much more operating, not at the level, down at the level of these individual axioms, but is much more at sort of a collective level
Of kind of the sort of flow of mathematical knowledge
operating at the level of many axioms and many, many theorems, and so on. Very similar to what happens in fluid dynamics, where, in principle, the description of a fluid can be made in terms of every individual molecule colliding with every other individual molecule, but in fact, the more human-level description is in terms of flows of fluid from here to there, and so on. And I think the same is kind of true of mathematics.
But there are many different ways one can think about those kind of patterns of flow.
At least that they're ones that are natural for us humans with our particular cognitive and sensory capabilities. There will be different ones suitable for other kinds of, for… for… if you have other kinds of sensory apparatus, other kinds of cognitive capabilities.
Which raises the interesting question of when we're making sort of AI-like things that are somewhat aligned with humans, is there sort of emergent mathematics, the natural mathematics for them, the same as it is for humans? And the answer, to some extent, I think, is no. I think that, but…
It's also the case that the sort of mathematics that they would naturally do is, again, mathematics that just doesn't really connect to the way that we humans think about things. So it's, again, one of these kind of, we shrug when we're… when we're confronted with that.

Let's see… Let's see, there's a long… comment here from Brandon.
So we might just be able to use or create technology within the environment that we are currently living in.
most likely, then, if there's an intelligent life form out there, would have only its available resources to create technology. If it was possible.
in a way that they discovered their information to be used as technology. Would it be possible that we could anticipate what technology
or what environment the other possible alien intelligence would have in order to create a similar technology. In other words, we might have to just try to assume what type of environments they would be in. So, I mean.
This question of, given environment.
what's useful technology for that environment is another slippery slope kind of thing. Because if you say,
were we… I mean, again.
We say, in that environment, what would be useful technology? Well, it depends on what you're trying to achieve. If you're trying to achieve very human-like things, if you want that little human who operates at, you know, one atmosphere of pressure and, you know, an atmosphere that has a bunch of oxygen in it, and this and this and this and this and this, then you can say, well, yes, we have a purpose of making a spacesuit or something that will, make us able to operate the way we normally operate.
in this… in this alien environment, but, you know, at a more abstract level, it's much harder, I think, to say, you know, to… to… okay, so what we're doing, again.
Throughout technology, we're taking what is out there in the world.
and sampling it for human purposes. A lot of what I've done of sampling the computational universe, we can say, we can sample this computational universe, how do we make a bridge to some useful human purpose? And actually making that bridge is often the most difficult thing to do. So, for example, if we look at modern technology.
We now have LLMs.
They do all kinds of things. They're kind of wild animals. They're alien intelligences. They do things, they do certain kinds of things. The real question of whether we can find them useful is can we put harnesses on them that connect what they're doing to things we humans want to do? So, you know, in the analogy of the LLM is like the wild horse or something, if we can domesticate it and put a plow on it, we can use it to, you know, to plow a field. Well, similarly, if we have this LLM, and we're trying to sort of put a harness on it to make it useful for…
I don't know… Generating comedy acts or something, then…
that would be one type of way to harness it. We might say, well, we'd really like to put a harness on it that lets it solve math problems in a, you know, in a direct way, or lets it do, let's say, computational processing of things.
well, that harness is not going to work with this type of wild horse, so to speak. It's, you know, the LLM doesn't have, the, kind of… it just doesn't have that kind of thing
in it to do. It's something where… but again, this question of, you know, how do we take
The thing as it is, and connect it to human purposes, that's all about
what is the kind of very high-value thing as a practical level in sort of technology around LLMs, is how do you connect LLMs to, for example, computational tools like the ones we make, or to particular use cases in such a way that these things will meet human purposes, as opposed to just be off on their own, kind of having fun as LLMs

Rollicking in the computational universe, so to speak, doing things which are completely disconnected from what we humans are interested in.
Okay, another slightly different kind of question here. Do you think social science research will begin to revolve around large-scale multi-agentic simulations?
I have to say, I've thought that's a pretty interesting direction, and I know people have tried it a bit. I keep on hearing people say, we're gonna do this, it's gonna be amazing, and then I don't hear from them, and that's always a bit disappointing.
But, you know, kind of the idea of what Do, sort of.
what do a bunch of humans do when you put them on the desert island and sort of, you know, what do they do with each other? How do they form a society? Do they all eat each other, or what happens? You know, you can imagine getting some level of simulation of that done with LLMs. You know, perhaps with LLMs that have been prompted to say, I'm going to act like a… this type of personality, I'm going to act like that type of personality. How will these things interact? I think it's a really good direction, and, you know, I suppose I've seen some amount of work on this, but not as much as I think could be done.
And, you know, one of the problems… okay, every kind of experimental science takes a whole bunch of judgment.
If you're doing experiments on
physical systems. It's like, well, how much isolation do you need in that experiment? You need to be doing it in space. Is it enough to just keep it at constant temperature? You know, what do you need to do? It's much worse in biology, where it's like, does it matter whether the mouse You know, had, had breakfast before it did the maze.
Or, or, does it matter at a more molecular level, does it matter that the thing was exposed to electromagnetic radiation before this happened, or whatever it is?
Even more difficult in psychology experiments, where, you know, it's like, well, you do this experiment on somebody, and does that experiment reflect back to something they learned when they were 5 years old, of which you have no control?
You know, when you have synesthesia, and you associate colors with letters, do you associate those colors? Because when you learnt your alphabet, you happened to have, you know, a quilt that had those letters in those different colors on the quilt?
Those are things that are very hard to sort of deal with at an experimental level. And when you do computer experiments, as I do a lot.
there's a question of how do you do a clean computer experiment. Often, you'll do a computer experiment, you'll search a million objects, and lots and lots of different things will happen. How do you tease out of that a narratively useful thing for human purposes? Just saying a million different things happened isn't useful. What you have to do is get something out of there that is sort of simple enough
that it becomes a human narrative. I remember back in the 80s, when people were kind of studying artificial life and making kind of simulations of different kinds of ecosystems and so on. It was not uncommon for people to say, I've got this model of a rainforest, and it's got a hundred different species, and they each have this property, and that property, and that property, and look, this very, you know, this basic thing comes out of that simulation. And it's like, well, that's not surprising, because, you know, you put all this stuff in that we know works more or less the way rainforests work, you're getting a very simple thing out. The real value
In experiments, is when you get out

You put in a very simple experiment, you get out some very surprising, rich thing. Now, how do you do that for social science experiments with agentic, kind of players in the experiments? I don't know the answer to that. It sort of relates to what are the questions that you want to ask in sociology, anthropology, political science, kind of economics, and so on. What are some big, robust questions that you can start to ask? I would say one of the obvious questions is.
How does coherent society get formed?
Is society coherent? I mean, or is it the case that you have something where you set things up, and it's just like every person is hanging out doing their own thing.
Versus there's a coherent direction to what society does. You know, do the minds flock in rule space, so to speak? Is it the case that, you know, when it's a society made of human-like agents, do they behave like a flock of birds that go in a certain direction, and then when things change, they suddenly go in another direction, or do they all just sort of randomly
wander around doing their own thing. You know, I think it's an interesting phenomenon
That, we, you know, we look at human society, and we all feel we have a certain amount of free will about what we do. We all feel that we have certain constraints and norms from the way that society operates, but there's sort of a question of to what extent, when do things work?
in this kind of herd way, and when is it everybody's just doing their free thing? And I think that's a thing you could imagine studying by doing sort of agentic simulations of societies, but you have to be kind of careful about that experiment, because, sort of, if everybody is identically the same, it's kind of…
a little different than if you have a certain distribution of different personality types and so on. You know, you could ask, you know, imagine you're doing,
you know, one of those simple kind of personality analyses, and you ask, you know, what fraction of the population is ENTJ, or something like this? And then, let's prompt the LLMs to be, you know, 20% ENTJ, and I don't know all these, all these acronyms, but, you know, 10% some other thing.
or let's, you know, I know people have thought about, well, what if you take, sort of, census data, and you know that X percent of the population
is… does this, and another percent does that, and they have different occupations. Or maybe you take surveys that people do. I know dating sites have a lot of data on kind of funky questions and how people answered them.
And I think you can imagine sort of trying to segment your agentic population by some of those kinds of things. I suspect that's a very slippery slope.
to somewhat meaningless kinds of results. Because I suspect that, insofar as you just say, we're going to make something that is a faithful model of the way, kind of, humans work and are distributed, and then we say, well, we get this kind of effect that is sort of like what we get with with humans and so on, it's like, well, that's not very surprising. You kind of put in the answer. The question is a more, you know, small stuff in, big stuff out.
You know, is there a way, for example, we know for flocking of birds and fishes and things like this, that there are very simple models where each organism just looks at its neighbors to decide what to do that lead to flocking behavior that's at least somewhat realistic.
And I think that,
That question is, you know, where, we could, we could ask questions about,
The, the, the, you know, what…
what amount of agentic diversity can you have while still getting herd behavior? My guess is rather a lot, actually.

But then, you know, we know things from studies of animal behavior, like, you know, the… the herd of goats or something, where, you know, there's a question of, is there one leader goat that's always the leader goat? Well, no, probably not. There may be a group of them, and on different days, there's a different leader, like in ancient Athens or something, where they would elect different, you know, on different days, different people would be in charge. Not a system that I think particularly worked well in human society.

But, apparently in goat society, it's not disastrous. But, you know, a question is, what, you know, under what circumstances… like, an example of a somewhat robust question would be, under what circumstances do you get what levels of coherence in society, and under what circumstances do you have a sort of a schism of that? You know, what are the basic phenomena that lead to that? Is there kind of a phase change when you sort of adjust the fraction of of awkward agents that all want to do their own thing. You know, is there a phase change when you have, you know, 12% of agents that want to do that? Suddenly society falls apart or something? Or how does that all work? I can imagine that there are real experiments to be done along those lines, but I also think those experiments will need a lot of judgment to not end up being very mushy, because it's just kind of like you've thrown all this all this stuff in. I mean, the worst case in science is when you have… you have a model where the model has tons and tons of stuff that go into the model, and then out of it comes this one morsel. of results. That's the bad situation. The good situation, which I'm happy to say is happening, for example, with our physics project, is the model that goes in is tiny. The model that goes in has very little that you have to say to describe it, and yet what comes out is incredibly rich.

All right, I think I have to run and go back to my day job, and I'm sorry I didn't get to lots of other very interesting questions, but,

Thanks for, thanks for these comments and questions, and, actually you've now given me some homework to think about, so… so thanks for that, and, look forward to chatting with you all another time.

Bye for now.