Prof. Karl Friston

1st Applied Active Inference Symposium

Active Inference Lab

June 21, 2021

Abstract:

On June 21st, 2021, Active Inference Lab (activeinference.org/) hosted its first Applied Active Inference Symposium, featuring Professor Karl Friston. The Symposium was structured in three sections, corresponding to the Organizational Units of the Active Inference Lab: Education, Communication, and Tools. This publication reflects an edited transcript of the proceedings of the Symposium.

Authors (Participants)

Karl Friston [1], David Standish Douglass [2,3], Maria Luiza Iennaco de Vasconcelos [2,4], Stephen Sillett [2,5], Lorena Sganzerla [6], Dean Tickles [2], Ivan Metelkin [2,7], Alex Vyatkin [2,7], Daniel Friedman [2,8]

v1.0 (December 2021)

Author Affiliations

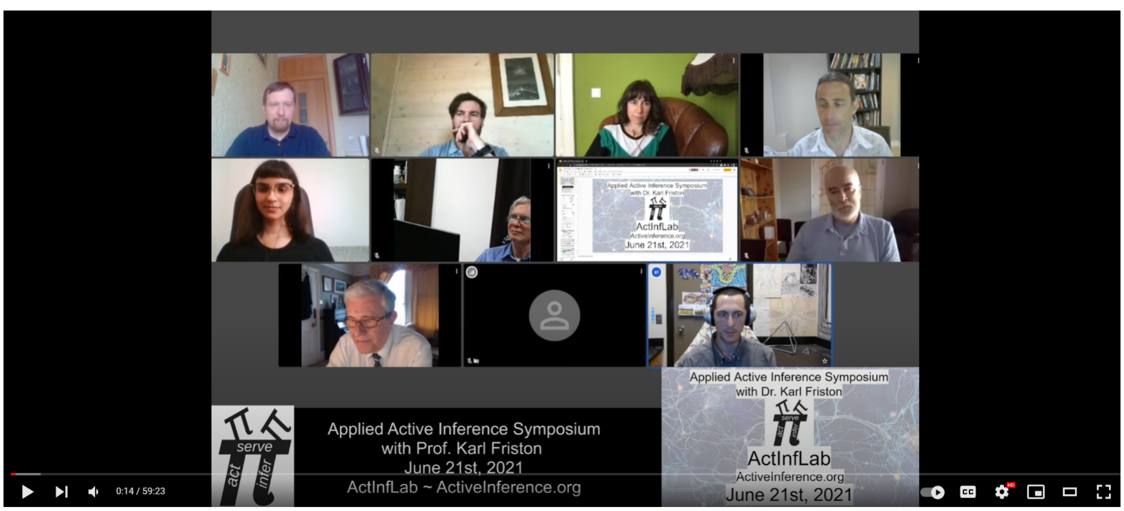
1. Wellcome Centre for Human Neuroimaging, Institute of Neurology, University College London, London WC1N 3AR, UK
2. Active Inference Lab, activeinference.org
3. American Society for Cybernetics
4. University of São Paulo, São Paulo, Brazil. Department of Philosophy
5. Canterbury Christ Church University, Salomons Institute for Applied Psychology
6. Berlin School of Mind and Brain, Humboldt-Universität zu Berlin
7. Systems Management School (Moscow)
8. University of California, Davis, USA, Department of Entomology & Nematology

Video Sections

Part 1: Education (.edu), page 3.
youtube.com/watch?v=INRaCBikpso

Part 2: Communication (.comms), page 18.
youtube.com/watch?v=X2GwqUVLlcs

Part 3: Tools (.tools), page 30.
youtube.com/watch?v=hW9IiOujS1E

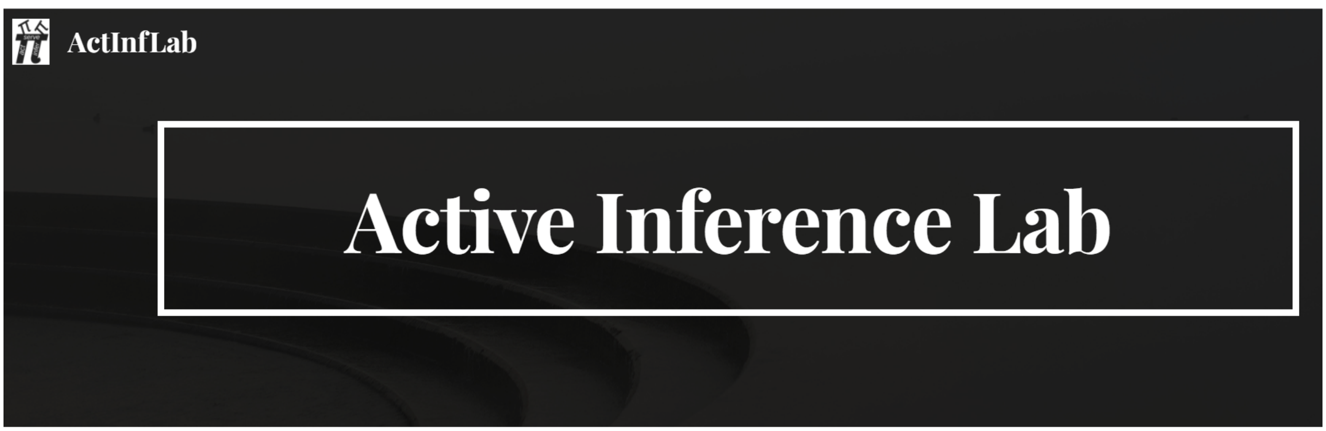


Part 1: Education (.edu)

00:17 Friedman:

Hello, and welcome to the Active Inference Lab, to our first ever applied Active Inference symposium. Today it's June 21st, 2021; and we're very honored to be here with Professor Karl Friston, and many of our lab participants.

Just as a way of quick introduction, the Active Inference Lab is a non-profit organization, that is a participatory open science laboratory, and we're working to curate and develop applications related to the Active Inference Framework – something that, hopefully, we'll be going into a lot more in detail today. And this is a screenshot of our website.



Active Inference Lab (ActInfLab) is a non-profit organization that is a participatory open-science laboratory curating and developing applications related to the Active Inference framework.

The goal of ActInfLab is to produce cutting-edge research and enable real-world applications of Active Inference.

ActInfLab also seeks to scaffold the Active Inference community: increasing the competency of participants & raising broader awareness of these topics.

As far as the overview of this symposium, there are three organizational units in the lab: .edu (education), .comms (communication), and .tools. And each of these units are going to facilitate a 45 minute or so session, and we'll have a short break in between sessions. So, in our weekly meetings over the past weeks, for each organizational unit, we've been developing questions and getting excited about things that we wanted to talk to you about.

As far as a few overarching themes that were kind of spoken to really through the whole journey of our lab, but also across organizational units.

The first theme is applying Active Inference across systems (again something that will come up probably in all sections);

The second theme is that of research debt, the idea that we don't want to be developing research frameworks that have a huge burden on those who are learning and applying; and that, especially early in the formalization of frameworks it's extremely valuable to increase the accessibility, so that we don't end up with major headaches and incompatibilities later on;

The third theme is collective intelligence and the ways in which it is manifest across different systems; transdisciplinary teams, projects and communities, which are kind of like nested levels of organization. But transdisciplinarity is something that is necessary for the type of work that we're all interested in; and also just modern challenges and opportunities for research and all that that means related to online and everything else; and, of course, anything else that you have tumbling around, and wanted to bring to the table, thematically.

So there we are, with our sort of lab overview and introduction.

Let's go to our first organizational unit, .edu. The goal of .edu is to scaffold and create a participatory and dynamic Active Inference body of knowledge, which we'll talk more about in a second.

And our progress and actions this year have been to release a terms list, V1, which benefited greatly from your feedback. And also we're now updating the terms list to version 2, which now includes five complete language translations and many references and citations for the terms. The way that we're approaching the development of the terms is by using approaches that place ontology, and progressively more formalized versions of ontologies, as kind of the backbone of an educational body of knowledge.

So we started on the left side here, with a terms list in the first quarter of 2021, and the ontology working group is like a train that's pushing to the right, as they're learning ontology by doing, and developing progressively stronger and stronger ways of relating the terms and the concepts that are essential for understanding Active Inference. And this will help us develop principled educational material that's also able to be translated rapidly.

Alex, do you want to give a quick thought on where knowledge engineering comes into play?

05:28 Vyatkin:

Yeah, thanks. At this slide, we are showing this work with ontology with systems engineering approach, which we are also using in the lab; and considering possible deliverables of working on educational materials and creating them. We should have at some point of time textbooks and educational courses, and actually maybe this lab is started from the idea that a textbook for Active Inference should be created.

Also, we see some connections that can be applied to organizational management for creating translations and to make it multi-language from the beginning. And also we should see for some domain specific use cases that we can understand in terms of that ontology that we are going to create.

05:42 Friedman:

Thanks, Alex. So on to the questions section. We're going to start off pretty broad here in the .edu:

How do we go about determining the core ideas and terms for Active Inference? This will be the format of the question slides, Karl, so, feel free to jump in!

07:37 Friston:

Right, I guess it will be structured around the key ideas, and essentially ingredients that underwrite the Free Energy Principle, and how that translates into Active Inference. So, you know, without thinking about it too deeply, my mind just goes to what are the things, what are the basic ingredients you need to explain to somebody, what Active Inference is, and why it works. And it normally starts off with the notion of a generative model, and then from that, you spin off all the appropriate mathematical ideas and constructs and descriptions that would attend that. I mean it may be best to reflect the question back to you.

So, this is a really neat idea – having an ontology! And it's certainly my experience that people are entertained by, sometimes the poetic use of phrases and descriptions, such as epistemic affordance, when trying to grapple with "What are the fundamental ideas behind Active Inference?" Some of them are fundamental and some of them are not. So, it's certainly an interesting idea to try and tie down the ontology.

But let me ask you: this ontology just means what it says, in the sense that you're trying to define the essential concepts and how they relate to each other? Is that the basic idea?

08:37 Friedman:

Yep. Going back to this slide here, we want to have a continuum from a list of terms, potentially, that could be developed into coherent and, again, principled course material and competencies; but also develop a logic. And we're developing within the SUMO ontology development framework, which defines not just relational edges, but actually, an actual logic. And so we hope to be able to ask, like: "Is this a complete Active Inference model? Have we really checked off all the boxes and used those kinds of logical tools that are accessible to the well-developed ontological frameworks?"

08:41 Friston:

Okay. Well that's very compelling and very clear. It strikes me then that you know, it would be useful to link that operational ontology to the underlying maths. So – you're much of the... much of the conceptual steps, both in understanding and implementing Active Inference, usually in terms of simulating your interesting behavior or using it as an observation model to explain some empirical data from a study.

Much of it can be developed in terms of a series of moves that usually (or, in fact, almost universally) inherit from – are framed in terms of – either information theory or linear algebra or differential equations; and you can just build the story from that. So, if you're looking for that degree of formal and useful detail, then... it would be... one principle you might refer to is basically: what your... "Where does one equality assertion or description or variable or object – where does it come from in terms of inheriting from the more basic formalism?"

So, what I'm thinking of here is: "Where does Active Inference *start*? And how do you get to the calculus and the Bayesian mechanics that you'd associate with Active Inference?" And my guess is: given the structure or the way that you have approached the ontology, you've probably actually done that already or are in the process of doing that.

Are you going to go through some examples that would sort of highlight the, you know, the strategy and the problems, which are usually more illuminating than the solutions that you've encountered so far?

10:52 Friedman:

Sure! I'll switch here to this screenshot of the current state of what it looks like. And we're starting just in tabular form by compiling up to five references and citable definitions. First just looking for exact cases where a term is used. And then we'll go from how the term *has* been used, towards synthetic definitions that capture different senses of the term. And then along with the concise narrative of the field, and also ontology experts who are here with us, we're going to then be working to make the actual logical underpinnings, elucidated in terms of specifiable code, rather than just concise English definitions. And then from that sort of generator of the formal relationships we'll be able to descend into mathematical formalisms, or other natural human languages.

11:48 Friston:

Yeah.

11:51 Friedman:

We'll keep you posted on this project though, for sure. Let's go to this next question and imagine that we had that set of terms in development (it's going to be a work in progress our whole lives):

How would we go from core terms and ideas to an interactive and enlivening education that speaks to people from many different backgrounds?

12:45 Friston:

So, I'm going to answer this question from the point of view of my experience as a supervisor, which is probably a little bit of a narrow remit from your more general ambition. I imagine that this is related to this notion of – (was it sort of "research debt" – I can't remember now) – but this notion that you don't want to put too much pressure on people, when becoming acquainted with the utility and application of either the code or the ideas.

In my experience, in an academic setting, just having toy simulations is usually the best way to give people a feel for what this approach *does* and how it can be used. So, it's enormously potent in terms of demystifying and also illustrating the functionality at hand, or that can be accessed. Having a sort of a working, or at least a *toy* model sort of provides a proof of principle, and that can strip away the magic as well.

And, I think your ambition to try and make this accessible to people who are not necessarily fluent in the underlying information theory or dynamical systems, is very laudable and perfectly feasible. So, again, in my experience, some of the most creative applications of Active Inference can be by people who don't really necessarily wonder too much, "What's underneath the hood?".

It all comes back again to the design of the generative model. So, if you get the generative model right, and it's apt to describe the thing that you want to understand or to simulate, then usually everything else follows suit.

And I mean that in the sense that you can just take off-the-shelf software, which I presume that your ultimate ambition is to make available, and make it work in the service of sort of saying, well... "What would this agent (or this synthetic creature or person) *do* in exchange with her environment, *if* this was the generative model and this was the generative process?"

So, a lot of this really, I imagine – in terms of answering your question – "how do we go from core terms to interactive and enlivening education?" – is just establishing a language, a lexicon, that allows you to talk through somebody in constructing their own simulations that speak to the issues, that engage them either academically, or beyond academia.

So, clearly then, the core terms play the role of literally a language, in terms of communication, which brings us back again to the importance of the ontology, and having the terms linked in a formal way to the mathematical expressions and also procedures and processes. So, I guess that a precondition to use the core terms in an interactive and enlivening, educative sense will rest upon getting that ontology right.

In my experience, you know, the best way to get the ontology right, in the sense of it being enabling, is just to talk about the terms, until there's some consensus and everybody understands them, both in terms of their teleology, but also in terms of where they come from – from the point of view of the code and ultimately the maths that underwrites all this. Is that the sort of answer you're looking for here, or thinking along with? Are you thinking along the same lines?

16:48 Friedman:

It sounds great. There's so many dimensions there! Just to provide a summary, or just jump in at one place:

What is Active Inference, and what does Active Inference do?

17:31 Friston:

Right. That's perfect, because I was just thinking: it would be really useful just to go down the terms that you had in the previous-but-one slide highlighted in green, because all the heavy lifting here is really just shouting about, what are the core aspects and claims, or the core things that you're trying to communicate with any one of those terms.

So, for me, Active Inference would be a description of a process that can be seen as something that arises from the Free Energy Principle. So you can either tell that story from the point of view of a physicist, and say that Active Inference is a teleological description of processes that systems that self-organize must possess; or you could tell the story, or define Active Inference, from the point of view of neurobiology and ethology, from a point of view of, say, predictive processing, and describes what it entails.

And I've used the word "Bayesian mechanics" before, because from the point of view of the physics definition, it would be a teleological description of a Bayesian mechanics that necessarily arises, you know, (with certain assumptions) from any self-organizing system.

One key thing about Active Inference, which I think would be important to put in the definition in the ontology – (I'm not sure if it's already there but, you know, if you're in charge of sculpting the ontology, then you're in the position to make sure it's there) – is it's *beyond* predictive processing, it's beyond sentience; and it emphasizes, or reflects, the pragmatic turn at the beginning of this century: really sort of epitomized by the 4Es (you know, embodied, embedded, extended and the like), to make it clear that sentience is *active*, and that you are talking about, the circular causality of engagement of any particle, personal, or plant, with whatever is out there.

So, that would be certainly one thing to emphasize in terms of what Active Inference means. The "inference" is interesting, in the sense that it does imply a process, and a process with purpose, which is to infer, which is why I keep using the word. The other, teleological description of something – that's actually underneath the hood from the point of view of physics.

One final point here is: there's an easy confusion, I think, between, first of all, *Active* Inference and *passive* inference. So, that's certainly something which probably needs resolving, certainly in the philosophical literature. So, I often come across philosophers who say, "Well, there's *passive* inference, or perceptual inference (which is just basically inferring states of affairs in the world on the basis of some sensory evidence). And then there's the "extra" bit, which is the *active* bit which is: now you're in charge of *gathering* that sensory evidence upon which you are now going to prosecute your perceptual inference."

That's an interesting dichotomy, which I'm not sure is a *correct* dichotomy. If it's not right – I'm not sure that it is *not* right, in the sense that it is a useful distinction; but certainly is not what Active Inference was originally *termed* to mean.

You know, by conjoining "active" and "inference," there were a number of motivations. First it was a generalization of David MacKay's active learning, but probably more importantly, it was a nod to the notion of active sensing, and active , you know, active perception that perception is in and of itself, an *active* process, a *constructive* process – that you have to put policies, plans, and action into the game. So that I think would be one important aspect of Active Inference to define, and I don't know that it has been defined so far. So, you know, perhaps it's your job to define that.

The other thing which is important, I think, in terms of emphasizing what Active Inference entails actually comes from that *enactive* perspective, which is *inference about the consequences of action*.

And that has an important but really simple concomitant: that the consequences of action are in the *future*. And that means you now have to think: if you're thinking about Active Inference in terms of teleology or as a normative theory of behavior – of sentient behavior. And you have to now think about – I should qualify: When I say *normative*, I mean it can be operationally defined as an optimization process that, in turn, requires you to define the objective function or functional. And that's important practically, because if you're now thinking about sentient behavior, Active Inference, and its influence about things that haven't yet happened, because you haven't yet acted, then you're necessarily talking about objective functions or functionals that are about states of affairs in the future. And that is an important move and something that Active Inference embraces, which goes beyond predictive coding. So much of the literature in the 20th... You know in the 1990s, and subsequent, much of the literature that inspired that sort of enactive perception or active sensing; take on situated cognition, take on sentience originated in, you know, in things like predictive coding. But predictive coding is *not* what is meant by Active Inference: you can do predictive coding just by, if you're a statistician, just minimizing variational free energy. That's only half the game, once you move into the world of Active Inference.

From a teleological perspective..., all your you are... you have to do that, you have to form beliefs about hidden states of affairs in the world , using sort of the perceptual side of perceptual inference - but that is only in the service of rolling out into the future, and deciding what the best thing is to do next. And that running out into the future and deciding clearly calls for an objective function.

So in Active Inference, that would be the expected free energy, which may or may not be unfortunately named - but that's what it is. And therefore, Active Inference sort of implies that you are committed to optimizing an expected free energy and implicitly it's all about choosing the next thing to do.

So, for me those would be two... would be two... sort of *cardinal things* that should be embraced by a definition of Active Inference that, you know, transcend other normative approaches. So, for example, you know, reinforcement learning in behavioral psychology would be all about what the good things are to do, and you commit to a loss function, or a value function of states, if that was the kind of behavior that you're trying to describe.If, on the other hand, you were all about the psychophysics of perception, or just building optimal recognition systems, where you weren't in charge of gathering those data, then your objective functions would be very very different.

But what Active Inference says, well you can't... you can't... carve up the two problem domains, because they're just both sides of the same coin. And thereby... you're, you know,... you're now facing the problem of defining an objective function that is fit for purpose, that does both the belief updating about latent states or hidden states generating the data, and also the best way to solicit or *cause* those data or outcomes *under* some prior preferences or some goal-directed constraints. Is that a good long-winded answer?

26:50 Friedman:

Thank you for the comprehensive answer! It leads directly to our next questions, which are: What is the *Free Energy Principle*? And especially, what is the *relationship* between Active Inference and the Free Energy Principle?

27:06 Friston:

Right! Well, that's, I think, a slightly easier question to answer! The Free Energy Principle is just a variational principle of least action. Why is it special? – or *not* formally identical to all the other variational principles that we use?

If you look under the hood, right from quantum, through statistical and stochastic, to classical mechanics, well the only thing that differentiates, really the variational principle of least action that is the *Free Energy Principle*, is that you're paying careful attention to the separation of states to which you apply that principle the separation of states into the states *of* an agent, or a particle, or a part of a person – and the *outside* states. So technically, if you were in statistical thermodynamics, for example, you'd normally *assume* that separation in terms of some idealized gas that was contained within the container or heat reservoir, or a heat bath – without really worrying about where the heat bath or the heat reservoir *came from*. But the Free Energy Principle says: Well no, you can't really do that. You've really got to attend very carefully to... what *licenses* a separation of different *kinds* of states, so you can assign to the *inside* of something – and the *outside* of something – and the states that mediate the exchange between the inside and the outside. And then you get into the Markov blanket and Markov boundary literature.

So, just to summarize: A Free Energy Principle is just a principle of least action, by which I mean, that there is a description of dynamics in terms of the *most likely* paths any system will take. That is the special provenance of a partitioning, or a separation, of the states of some universe into the states that are owned by an agent (or a particle), and those that are not, and the states that mediate the exchange between them. So that would be the Free Energy Principle.

Active inference, as I say, is a sort of teleological spin-off from the Free Energy Principle, in the same sense that you have now at hand a principle of least action. It allows you to identify, simulate, define – the most likely paths, trajectories, or narratives – that a system will pursue under certain conditions. And those conditions are just that there is an attracting *set* of states which that system will converge to, or will look as if it's attracted to.

So, sorry, What I was working towards, was the notion of an attracting set, as a metaphor for equipping that physics with a teleology; and that teleology is nicely illustrated by the notion of *attraction*. So (you know) when mathematicians talk about attractors – in the particular case, in the Free Energy Principle, these are these sort of pullback attractors or the kind of attractors that you get in random dynamical systems.

There's a *proper* and *natural* tendency to think that these *particular* states of the attracting set *literally attract* in the sense of, you know, gravitational attraction, or any other kind of attraction – they *pull states towards* them. So that, to me, would be a teleological interpretation which, I think, is much closer to Active Inference – that you're saying, that *influence* is a process that has a *purpose* and the underlying Free Energy Principle allows you to say the way it *looks*, as if self-organizing systems, show these certain *properties*, they're *attracted* to certain *states*, they're attracted to certain *paths*, and we can describe those in terms of the teleological ontology. And that would be Active Inference.

One *practical* difference between Active Inference and the Free Energy Principle, is that the Free Energy Principle is *just* a principle. It's neither right or wrong; it's just like Beren Millidge has noted: it's like sort of Noether's Theorem or Hamilton's principle of least action. But as soon as you start to say, "Well, I think that this principle applies to this population or person or particle," that certainly commits you or requires you to define the attracting set of states - a pullback attractor (in another jargon, the equivalent would be a generative model); and as soon as you commit to a generative model to explain the teleology of *this* system, or *this* agent, or this person, then you've moved into the world of non-falsifiable principles into falsifiable hypotheses, because you could have chosen the wrong generative model, and thereby there will be evidence for choosing this generative model or that generative model.

So, the relationship between Active Inference and the Free Energy Principle is *operationally* quite simple, you know: Active Inference is the application of the Free Energy Principle to a particular system. But in that application you're bringing a lot of teleology to the table, and more specifically you're having to commit to a particular generative model. And as soon as you do that, that becomes your theory or your hypothesis about what is an apt description for this system. So – a number of (I think) sort of interesting distinctions, in terms of the relationship between Active Inference and Free Energy Principle that I imagine your ontology is already addressed, or it's certainly addressing.

33:50 Friedman:

Well, we'll get there! Thank you for that excellent answer! For the next question: Lorena, please read it out.

33:59 Sganzerla:

Oh, hi! So, still in the spirit of broad questions and broad terms, and that, I think it comes in line with what came before.

So, how and where does the idea of information play a role in the Free Energy Principle, and how does it relate with Active Inference – in the sense of, what is something to keep in mind when thinking about information dynamics in Active Inference?

34:36 Friston:

Right, well, these are great questions. I'm getting the hang of this now. You just want me to talk. I've presented a question [in response]! Which I'm very happy to do. Are you sure you want me to do that? Or should this be a conversation? Perhaps it'll turn into a conversation at some point.

So, information. So, it plays a dual role, in the sense that information theoretic formulations underpin most of the derivations behind that principle of least action; and it can be no other way – in the sense that all mechanics from physics, is really articulated in terms of probability densities or distributions. As soon as you have a mechanics, or a calculus, or probability distributions, you're effectively in the world of information theory. And you see that at many different levels. So one nice example of this is that the central quantity that we often use to score the likelihood of being in a particular state – if you're a statistician, that would be the marginal likelihood; if you're fluent with an FEP ontology, it would be surprisal (or more simply surprise) – and that is just basically the self-information. If you're a physicist, you look at this as a potential – it's a negative log probability.

So you *start*, really, when thinking about the physics, with this central concept of self-information, which can be read as a potential function, or a surprisal function, or surprise – Here, it is the thing that the variational free energy is a bound approximation to. So at that level – and then every other move you make mathematically, in terms of the expected self-information being the entropy – and why that is important as a characterization of various probability distributions in the setting of self-organization – would testify to the fact that information theory is absolutely central to all the maths that underlies the physics of the sentience that emerges from having a distinction between the states of the system and states that are not in the system, namely the Markov blanket.

Having said that, I think "*information*" to most people's minds usually means more. Certainly in the folk psychology context, it's really information *about* something. And the FEP Active Inference has I think something quite special to bring to the table here, that goes beyond information theoretic treatments that you get in communication and signal processing and rate distortion theorems. All of that kind of information is just your extensions of information theory that inherit from self-information or the implausibility of a particular event or message – or, in more abstract domains, such as sentience and consciousness, you would go to something like integrated information theory. But that is all about this "Shannon-esque" kind of information.

The opposite, the other kind of information, which is information *about* something – So, what I wanted to try and put on the table, is the very fact that you've got this Markov blanket or separation of states on the inside and states on the outside, means that now you can equip the states on the inside with the role of encoding posterior conditional Bayesian beliefs about states on the outside. And that introduces, technically, a different kind of information geometry, a different kind of information theory – where, crucially now, you can read the internal dynamics as containing or *having information about* what's going on, on the outside. And this is a really important move, equipping your neuronal dynamics or variational message passing or belief propagation in a computer, with an information geometry, that now allows you to read off the state of the computer or the state of the neural activity *in terms of* what it is encoding, or belief, or the information it contains *about* the outside. And so that sort of dual aspect information geometry has been celebrated to a minor extent in the philosophy literature by Wanja Wiese, asking the question, is this really the maths of sentience – where you now have information about things.

And in a sense that really is the heart of the Free Energy Principle – or Active Inference, anyway – in the sense that it equips that information geometry. I mean, technically, what you are saying is that any particular internal state of a computer, or a person, or a brain, now can be read as encoding a Bayesian or a posterior belief about other states, namely, hidden or latent causes outside the Markov blanket. And that defines, technically, something called a statistical manifold. And as soon as there is a statistical manifold, there's an information geometry. And any movement on that manifold necessarily implies a change in your Bayesian beliefs, namely Bayesian belief updating. Which means now there's an interpretation of neuronal dynamics, movements on a statistical manifold on the inside, in terms of belief updating. So the notion of Active Inference, as the process of belief updating, really , you know, rests upon this fundamental notion that there's information about stuff that is encoded or parameterized, by the internal machinations, and the mechanics, and the dynamics of the inside.

So I think it might be quite important to – if you're trying to describe or educate people in terms of how they should understand information. I think it'd be important to differentiate between the mathematical notions of Shannon information or self-information, and the information implicit in an information geometry; namely, the information about something. This second kind of information is implicit in an information geometry – the sentience that is afforded by Active Inference – when now you're understanding *neuronal* dynamics or message passing in the computer on some Forney – on some sort of factor graph – because in *this* instance, each of those messages, or those neuronal dynamics, can now be read as belief updating – namely changing your *mind* about other things – so that the stuff on the inside has information about stuff on the outside.

	Associated Topics
Information Theory (Information of the first kind)	Shannon information Self information Surprisal Surprise Log evidence Log marginal likelihood Mutual information
Information Geometry Information of the second kind	Bayesian beliefs Posterior beliefs Conditional beliefs Information length Information geometry

42:38 Friedman:

Thanks for this important answer. And we're going to pass over a few questions and go to slide 18, with Stephen reading the question to continue on this theme about the separation of the inside, and the outside. So thank you, Stephen, and please read off slide 18.

42:56 Sillett:

Thank you. Just gonna ask how can the integrity of the Active Inference process theory be maintained when blankets, blanket states, and generative models are being interpreted in novel ways? We were thinking about what do you think of discussions around Markov, or Pearl, Friston blankets, etc.?

43:13 Friston:

All right, that's an excellent question. I have quite a technical answer; so if it's getting too technical tell me, now. I'll try and get back to what you were really trying to unearth. So this is not a fast moving field; but certainly it has been a delicate and important area of discussion over the past few years.

So, in the original introduction of Markov blankets there was an explicit nod to Pearl's construction of Markov blankets, and how Markov blankets are used *practically* in terms of simplifying message passing in computer science. However, that may have been something of an oversimplification. Because from the point of view of the Free Energy Principle, the kind of causality that the Free Energy Principle deals with is *not* the kind of causality that people, particularly people like Pearl, but also people dealing with things like Granger Causality deal with.

So from the point of view of the Free Energy Principle: that starts with a stochastic differential equation or a random dynamical system written as a random dynamical equation, and OU (Ornstein-Uhlenbeck) processes being simple examples – in physics these would be Langevin-like equations. Common to all of these starting points is time, and evolution, and dynamics.

Now, there is *nothing* in Pearl's formulations (well, certainly there's nothing in Pearl's book, on causality, that deals with time; and I know that, because I – before the days of PDFs and being able to go and search particular words – I had to go through (laughs) and find out – there's one paragraph that mentions dynamics!) If you were in statistics, computer science, you know – this will be the world of dynamic Bayesian nets – this is their take on something which is actually much more universal, which is basically the universe as a Markovian dynamical process.

So, just stepping back, the challenge now, is to articulate *independences* – that underwrite Markov blankets in the sense of Pearl, in terms of dynamics. So you've now got to *link* two quite distinct fields, which is basically the fields of dynamics and Langevin processes, and things that have paths of least action – to the world of statistics and Pearl-esque independences and causality cast as interventions that have observable consequences. The problem in doing that linking is that you have to really abandon the notion of causality in the world of Granger Causality and Pearl, because causality is baked into, and is inherent in, writing down any differential equation (be it stochastic or random or deterministic), in the sense that *states cause motion*.

So the causality in this context would be a more control- theoretic causality. So, that means that you *can't* then use the causality concept later on; but it does mean that you've now got to derive from a dynamical Markovian calculus the necessary conditions that would lead *to* the conditional independencies that are necessary to define Markov boundaries.

Just to slip in here: The Markov blanket is composed of *minimal* blankets: namely boundaries, in the sense of Pearl; and on most recent analyses, it looks as if the blanket is actually *two* Markov boundaries, in the sense of Pearl. But to get to the sense of Pearl, you've got to think very carefully about "what are the constraints that lead to the conditional independencies?" – where those constraints are specified in terms of equations of motion and things like the amplitude of random fluctuations.

So, once you've seen that that is the link that needs to be made, that actually simplifies the thing! It simplifies things, in the sense there's no real latitude for interpretation. So, I'm going back to the part of your question: "blankets and generative models are being interpreted in novel ways." And I don't think there's any latitude of... any novel interpretation other than the… Sorry! If "in novel ways" you mean "the best way" or "the correct way" – and we just haven't got that yet – then I would, you know, I'd concur entirely with that! (Laughs.)

If you think that there is some latitude, there's some library of insightful reinterpretations and redefinitions, all of which have equal veracity, then I would suggest that's *not* the case. There's only one way, there's only one Markov blanket, or there's only one particular *partition* that can be articulated in terms of Markov blankets; and the only novelness there is really in tying down very precisely and defensibly how you get from a Langevin formulation to a Markov blanket.

At the moment, the novel way of doing that looks as if it's that the conditional independencies arise from sparse dynamical coupling, or causal coupling. So, if you read the causality as the influence that a state has on the motion of itself or any other state, in this sort of minimal Langevin-like description of the universe, then it is the sparsity of influence, or the sparsity of coupling, that *leads* to conditional independencies. And if the system has a sufficiently *rich sparsity* of conditional independences and implicit coupling, then it will have a particular partition ; and if it has that particular partition, then the Free Energy Principle holds.

So, I think the discussions around Markov, Pearl, Friston and blankets are essential – they're fascinating – and the conclusions of those discussions that I think I'm gonna probably have to refer back to the underlying maths - and that maths is all about connecting Langevin formulations of physics to the kind of calculus that Pearl has established, in a more statistical sense.

51:08 Friedman:

Thank you for the educational answer. This brings us almost to the end of the .edu section. So, I will pass to the final question to be read by Dean, who had several excellent points and questions. So, Dean, feel free to ask however you would like.

51:26 Tickles

The question is: What's the difference between a subject matter expert and a prediction matter expert – and how does this relate to your mode of interaction?

51:40 Friston:

You're going to have to unpack what "subject" and "prediction matter experts" means, for me.

52:38 Tickles:

Yeah! So, for me, interesting, you become a subject matter expert by gaining a certain amount of concentration in a particular field or area, and you become a prediction matter expert when you are able to think more distributively, more dispersively. And so, I think what when I read some of the things and listen to some of the stuff that I've heard you talk about... You brought these two worlds together. And so, I'm kind of interested in hearing what you think - in terms of introducing some of the ideas and principles that you brought into a world where, traditionally, we focused on concentrating – whether it's materializing something from an engineering perspective, or deciding what's in and what's out. You've brought in another aspect to look at, and I'm curious what you think of that.

52:49 Friston:

Okay, that's a fascinating distinction! I'm not sure it's terribly important what I think about it, because clearly you're the expert on this. But it certainly would be fascinating to consider the conditions under which you were able to *simulate* the emergence of a subject matter versus a prediction matter expert *in silico*, for example. This is a proof of principle that these are both effectively Bayes optimal ways of responding to a particular environment.

And my guess is that you would be able to do that relatively easily, by appealing to the ideas that you find in applying some of Active Inference notions to structure learning and development, where the basic idea is: if you've got a very volatile environment, by which I mean that there's lots of uncertainty in the contingencies; or possibly there are lots of random fluctuations, that are irreducible, in terms of your ability to predict the outcome of the trajectory of latent states of the world in which you are becoming an expert.

Then when you parameterize your uncertainty, you're usually formally – in terms of the precision of various likelihood mappings or probability transition matrices in a sort of discrete state space generative models – when you parameterize your beliefs about that irreducible uncertainty and volatility, then agents that believe or have inferred that they are in a very volatile, changeable, capricious world usually become better at the prediction side of things, in the sense that they rely less upon deep past experience, and assign more precision or more potency to the more recent evidence. So, they have a different style of evidence accumulation that enables – and also, they have the right level of uncertainty about what will happen next. So, it looks as if in their predictive engagement and epistemic foraging in that world, it looks as if they are better at predicting changes – because they're not committed to a particular explanation or understanding of how their world works.

On the other hand, if you create a world which is incredibly predictable and learnable, then, over time, the natural pressure to minimize free energy translates into a pressure to minimize complexity, namely a way of modeling your world and your exchange with it, in the simplest way possible, in accordance with Ockham's principle. And what that leads to is somebody - it sounds as if – it's somebody who becomes a subject matter expert. But the subject matter is their lived world, that has now become so predictable that they do not entertain all possible other outcomes, because they have precise beliefs about the way that things will unfold, and they can make very wise, very parsimonious – or using parsimonious degrees of freedom, they can make moves and become very expert in the way that this particular non-volatile, predictable (i.e precise) world works.

And the link with aging here is that: If you allow for the fact that we create our own environments, and your many levels of Active Inference will permit – or is a way of framing – our eco-niche construction. The story people tend to tell is that: as you get older, you basically make your world more predictable and you become a subject matter expert in your own lived world. So, I like the example: I no longer go bungee jumping nor go to discos, because my world is very, very predictable. And I'm, you know, very much an expert – because my world is basically my conservatory, my study and my bedroom. So, I'm a complete subject expert on that! (Laughs.) You take me out on, you know, to disco and I will not be able to predict what's going to happen next. Because I'm old.

Whereas, you know, adolescents and children and certainly newborn infants or newborn artifacts discovering their world, which is full of uncertainty, and they are not yet subject experts. And the epistemic pressures or motivation for them to learn about, you know, what happens if I do that, and what can I control? What can't I control? That will make them very quickly into prediction experts, until they become sufficiently fluent that they can now engineer their world to make it non-volatile. And then they presumably will become subject matter experts. So, yeah. I'm sure that would be fairly simple to simulate, using all the toy Active Inference schemes that we currently use. And it would be really interesting if these two different kinds of synthetic agents did develop some cognitive styles and confidence in what they were doing that looked exactly like the distinction you're talking about! I'm not sure it would work, but if it does that would be, I think, an illuminating proof of principle.

59:03 Friedman:

Thanks for this answer – and for this session from the lab and .edu. That last answer really spoke to the importance, also, of intergenerational learning. At this point, we will take a five minute break and we will return for .comms. So, thanks again and we'll see everybody in five minutes.

Part 2: Communication (.comms)

00:01 Friedman

Hello, welcome back. This is the second session of the ActInfLab symposium on June 21st, 2021. We're here in the .comms, or communication, organizational unit. The goal of .comms is to organize the lab's internal projects and activities, as well as to carry out all forms of communication with external entities. So it's like our connective tissue and our neuroectoderm, in a way. What has been done so far, is that we've had about 75 live streams, variously on presentations and participatory discussions, since July 28th 2020.

And we're taking an Active Inference approach to communication, and learning by doing. Our very first live stream was on the paper "Narrative as Active Inference," and shortly after "A World unto itself: Human Communication as Active Inference", so that's something that we like thinking about here and that we want to explore. And, also some of our lab members framed online communication and team collaboration in terms of Active Inference in the 2020 paper "Active Inference and behavior engineering for teams". Our aim here is to make Active Inference accessible, well known, and well understood. So let's get right to the questions, Karl.

The first question is, "How can we show – not tell – the idea of Active Inference, for example through embodied experience, experiments, or what other mechanisms? How can we best communicate in a way that makes it resonate?

02:36 Friston:

Well... Pursuing that very interesting notion – of using the principles of Active Inference to optimize the role of the .comms team. At its simplest, Active Inference means that the imperatives for all behavior, and it's likely that most behavior is of an epistemic sort, is to resolve uncertainty. So if you want to engage people and be a service to the people you want to engage – which may be internal members of your own team – and then you've got to know, what they *don't* know. Because, that will define the epistemic affordances that will get them engaged with you, and you with them, that will incur the best kind of belief updating.

So, you know, practically, what that might mean, is that, it may well be that... you one has to identify didactic or informative illustrations that are tailored or specialized to the person or people that you're talking about.

You know, I have never thought about using embodied experiences before! But that's a brilliant idea!. You're just illustrating to people who want to know "how *my* body and *my* mind works?" – to illustrate to them. the mechanics of it working, using the language of Active Inference. And it can be very, very powerful!

So as I'm talking, you know, what one example of this would be using saccadic suppression as an illustration of the potency of getting your beliefs about the predictability of sensory evidence right. If you're a physiologist,you would... you would know this as sensory attenuation. If you were in machine learning, you're working with transformers, this would be, I think, attentional selection, basically deploying the gain, or switching on the right channels, in order to select those data that are going to resolve the most uncertainty or maximize... the information gain, subsequent on this sort of covert action; or, you know, often sold as a sort of "mental action" in the philosophy literature. So that *mental action* is really endemic, and a vital part of our sensory engagement with the world –, and beautifully illustrated by saccadic suppression. So… This... speaks to the notion of *attenuating* the sensory consequences of your own action; so that any evidence that you're not actually acting is precluded from your belief updating.

So a clinical example of this would be Parkinson's disease, for example: If I'm sitting still; and I wasn't able to ignore all the messages from my muscles that tell me, "at the moment I am sitting still," then I am never going to be able to realize an *a priori* intention that I'm going to initiate a movement; because as soon as I initiate a movement, I have… I put in place a plausible hypothesis that I'm lifting, that I'm going to stand up. Immediately, all the evidence at hand suggests that I am not standing up, so I'm going to revise my belief: "No I'm not standing up, I'm not in the process of standing up.", And it becomes impossible to move. So that would be an example of what would happen if you didn't have this capacity to attend to, or to select or apply the principles of optimal Bayesian design, in terms of selecting those data for your own belief updating.

But a really pragmatic and easy example of that is saccadic suppression –, when we do the simplest of movements, epistemic foraging –, which is moving our eyes, making saccadic eye movements. Because, when we do that, we actually induce *masses* of visual information on the photoreceptors in the retina, sometimes referred to as retinal slips. So when I look from the left to the right, there's a *flood* of information that *I have caused*. And yet it is not *useful* information, because it doesn't tell me anything I didn't already know. So... what the... what the brain does,... it suppresses that information by transiently suspending the precision... or the Kalman gain, if you are taking a Kalman filter-like perspective on predictive coding as one kind of variational filtering.

And that's really easy to to demonstrate to an audience, you know: just get them to either fixate on an essential stimulus, and then pay attention to something that's moving around; or the converse: they fixate on the thing that's moving around, and then ask them questions about the central stimulus. And with the right timing it's a very potent illustration of, beyond just gathering information, but actively selecting and triaging that information in accord with the principles of your optimal Bayesian design.

So I haven't thought about going beyond that; but I'm sure there are lots of lovely examples of embodied experiences that really do illustrate Active Inference in action, as it were. I'm just reminded, because Active Inference could be read as if you like a 21st century version of ideomotor theories, which were very popular in the 19th century. And, of course, that was demonstrated through embodied experiences in a very alluring way, through hypnotism and the like! So I can imagine somebody doing a sort of 21st century version of hypnotism and all those wonderful Victorian illusions about the way you use your sense organs or deploy them actively. But now in the service of just illustrating some basic phenomena that underwrite Active Inference.

In terms of experiments (or, you know, the classic ones that immediately come to mind) that are really engaging, are visual illusions. So on one reading, all visual illusions are just ways of getting out your perceptual priors in the context of Bayesian inference. that you know... If you can conjure a particular pattern of sensory information that *you* know was caused in one way, and yet you think your subject or your audience has sufficiently precise prior beliefs that it could only be caused in *another* way, which is not the way you caused it; and then you let them experience that, and then you reveal how you actually generated those data, then that's a very powerful way of demonstrating the innate priors, the sort of formal priors, in you know in terms of the connectome or the the sparse coupling on a factor graph. And that, again, is part of the lived experience. So I think visual illusions would be – and there are loads of beautiful illusions out there – and all that one would have to do, is to harness their beauty and allure; and use them as a vehicle to give people insight into the their own, usually sub-personal priors about the way that... the world is constructed.

And then, in *my* world, what you generally try to do, is to actually put this *in silico* by just you know creating little *in silico* creatures. And *because* you've now got Active Inference with this information geometry,... and you now have the opportunity, not *just* to simulate what these creatures *do*, but what they *perceive*. Because now you've got the quantitative estimates of their posterior beliefs. So you can actually show, you know, a subject who's *just* experienced a visual illusion, that this is perfectly Bayes optimal. And, indeed, when you write down this variational message passing scheme in this synthetic subject, this synthetic person *also* experiences *exactly* the same illusions; and... this is and this is and... this is Bayes optimal, for this... kind of... kind of work. So you can leverage Active Inference activities in that sense.

And deliberately referring back to the .edu discussions: You know... What Active Inference brings to the table because it's got *information* about, stuff *out there* in the numerics – you can go a lot further, than you can if you were doing - say deep learning or machine learning. Because you've *got* this dual aspect, You've got this information geometry at hand... The *state* of your variational autoencoder actually means something in relation to a belief about what generated the… those data, which – You can create lovely little movies, you know, showing what this *simulation* of you, was actually experiencing. So, you know... I'm interested: are there any other ways that you've thought about in terms of showing people?

12:38 Friedman:

Thanks for the answer! It's like "look left, look right, now you're an Active Inference agent!" And, as far as potential avenues for embodiment: some of the work with Ryan Smith and others, bringing people into the *somatosensory* dimensions and their own priors and expectations about their *body* and about *motion* could be very powerful, as well as auditory modalities. And, indeed Active Inference is a framework by which we can think about how our perceptions are related to our inference and our action. So, in various domains, I think they'll be excellent experiments

It brings us to our next question. You actually addressed several areas in your answer. You addressed machine learning, as well as neuroscience, as well as just everyday lived experience. So how can Active Inference engage in better dialogue with adjacent areas? For example: machine learning, systems engineering, psychiatry, and neuroscience, as well as any other fields that you think are relevant, too?

19:46 Friston:

The obvious answer here is in either academic or commercial collaboration, and what might you... What would license that? I think you know... The simplest answer is that the Free Energy Principle and its (if you like) teleological correlative, Active Inference, is not there and was never intended to replace extant theories. It was there to prove , to endorse them, and to reveal the interrelationships between them. So anything that's worked and survived into the 21st century has some veracity and a proven utility. And therefore it's just a question of reformulating or changing the words, so that people can see immediately how their particular formulation relates to somebody else's formulation, where both formulations are special cases of the most generic and simplest explanation – which would you know from my point of view would be the the Free Energy Principle and Active Inference in the case of sentience. So I think that's, you know, as an integrative framework, you're very well positioned to say: look, you know, can we understand the way that you think about this and can we now articulate this, either using simulations or you know mathematical analysis? Can we understand what you've been doing in this integrative framework? And if we can show how it relates to another discipline's formulation of this problem? And sometimes you can get synergistic or added value from doing that.

So you know there must be loads of examples. You've written machine learning systems, engineering, psychiatry, neuroscience here. So, machine learning, for example, you know: how would Active Inference help machine learning? So at the moment you know there seems to be two answers floating around. And we've already sort of discussed a couple of the key a couple of these issues in depth. So machine learning commits to usually a normative approach to good behavior that can be quantified by a loss or a value function. But we've just said, well if we now want these machines to learn to act then we have to go beyond state action value functions and consider the belief based calculus that is Active Inference, which is all about the reduction of uncertainty. So now you are in a position to say: well look if you consider your objective functions as a part of a more generic objective function do you think what you might be able to get from this? And of course what you might get from this is a deep learning scheme, that actually can now go and solicit the right kind of data to optimize its own learning, and you know people in Bayesian Reinforcement Learning might argue: "well that yeah that's what we're doing you know with a series of bright ideas and heuristics to try and augment classical value functions". But you know you can say well okay you’ve clearly put a lot of work into that, but there is there is actually you know a simple objective function already out there that is provably appropriate to describe systems that self-organize and maintain themselves, that actually has what you want why don't you try this for example? So that would be one example. You have to tread carefully because you know a lot of people have dedicated their lives to solving these problems and they're very reluctant to change their rhetoric or see their contributions as a special case, but in many instances certainly from my perspective mathematically they are special cases. And sometimes if you catch the entrepreneurs, the innovators, the creative academics at the right stage in their career before they have committed to a particular church, or ideology, or calculus, or group, or company, and you can actually point them in the right direction, they become extremely creative.

And I'm not so sure about systems engineering but certainly I always celebrate the expected free energy as with just taking away various bits and pieces, various sources of uncertainty as reducing to KL control and then what I say is what KL control is what grown-up engineers use in a control theoretic setting. And so that would be another example so you know you could also I don't know this because it's not my field but certainly in terms of introducing say a fault tolerance in in control theoretic approaches in engineering, where the fault tolerance required uncertainty about the operation of some external part, you could again motivate a more complete objective function that takes you beyond KL control and introduces the information gain into the mix. Because to get from the complete objective function to KL control, you have to ignore uncertainty about the latent states that are in the mapping from latent states of the plant you're controlling, to the sensors or the observables, and so you know you're moving from a partial observed Markov decision process for example, to an observable one, and then expected free energy becomes KL control, or risk sensitive control in economics. So you could say, "Well, look, you know, why don't you just augment your KL control, and then put this extra term in… And now what you've got is a kind of anticipatory fault tolerance, in the sense that if there's uncertainty about latent causes, that's automatically resolved, in the way that you go and switch on various sensors or switch off various sensors? – As you know, there's a principled way of doing that."

I think to have any influence, you need to be able to show or provide proof of principle, that this more integrated, more universal, normative approach to problems can offer speed ups or increased efficiency or do what the people actually in that field wanted to do. So for example you've got to be able to show that Active Inference can outperform sort of vanilla deep learning by an order of magnitude. Which is easy to do, because of course most benchmarks and machine learning are actually inference problems. So if you just recast it it's actually quite a trivial thing to do just by saying: "Well, actually, what you've been dealing with is an inference problem!" – which looks a lot like a one-shot learning from the point of view of somebody in machine learning. But I think there will be some pressure to get people's attention to make yourself attractive, in terms of you – they will now have some first of all find you interesting and also have the potential that they can place an epistemic trust in you. You've got to sort of, you know, give them a cue and a clue as to why they should engage with you; and very often there's a two-way or two-road exchange. So one simple example which I see emerging in the field, is the use of deep learning to amortize certain mappings when they can be amortized in Active Inference schemes to evaluate the expected energy for example or doing very deep tree searches. So that's the kind of innovation you've seen coming out of 20 year olds at the moment, who haven't yet decided whether they're going to do deep learning or Active Inference because they want to do both and do it very very effectively. So that's a nice example from my perspective on the sort of you know that integrative role that could be played or you could play.

22:28 Friedman:

Thanks! I really heard this "yes and" maxim from communication and improvisation. It's like yes there's been a disciplinary way of approaching it and we're going to be working together to come back to first principles or to make it more efficient so that's really powerful. How does Active Inference help us rethink the nature of online communication where so much of our communication nowadays does occur?

28:18 Friston:

That's a big question isn't it , and certainly in the context of social media, politics, fake news, and the like you could take that question in lots of lots of different directions, which I won't do because that's not my field of expertise. But just yeah off the cuff in terms of first principles and what is communication? it's the ability for me to infer what you meant; it's the hermeneutics problem. if it's a hermeneutics problem that's most efficiently resolved in terms of dyadic or multi-system interactions when we come back to first principles which is the generative model, when we share the same narrative or same generative model. So in terms of helping, actually how does it, how does active influence help us rethink the nature of online communication?

And I think just from a first principles point of view, it would be the importance of establishing who is talking to who, and if you want to optimize the efficiency of that exchange, literally from the point of view of this principle of least action, the speed with which you can resolve uncertainty and minimize your uncertainty or surprisal. And then it's ensuring that like-minded communicators are actually communicating, because it's only them that will understand each other. So everybody has to speak the same language, they have to commit to a shared narrative, and a shared generative model. And then just by things like rate distortion theorem or you know rewriting that in terms of Active Inference free energy – the joint free energy minimization between two interlocutors – and that's the most efficient sort of shared path of least action. You know, how does that help engineer or intervene on things? I'm not so sure, certainly just in reference to communications with people like Maxwell and the like and other colleagues, there is this interesting notion that if communication, if the real problem of communication is not really the messages that you send, but the inferring whether to send the messages to this person or not, that itself now becomes conditional upon inferring that's a member of my in-group or that's a creature or a person like me, and then the question is: you know how do how does self-organization say you know in terms of social media exchange, how is that underwritten by an inference about the kind of people who I am listening to or who I am talking to? And what are the basic principles of that?

And again in accord with the minimization of complexity in our generative models what we you know what then it may be a useful hypothesis to say that there's an inevitable coarse-graining, of the way that we conceive of the people that we generate information for say on social media, and reciprocally and the kinds of people that I will be able to solicit by listening to this Twitter feed or that Wikipedia page or this news channel. So understanding how people carve up whether they are like to the degree of similarity to them may be very useful in just getting an idea of the dynamics of message passing amongst communities that will be defined by on average how each member of that ensemble or individual coarse grains and has a generative model of the kinds of people in the communication grain.

And just to finish this which is something I've heard and I found it really interesting notion, that again would be great if one could simulate this and understand the maths behind it, is that the only evolutionarily stable from the point of view of the Free Energy Principle the only one that will be selected by a process of Bayesian model selection, the only partitioning into in-groups and out-groups is a 50-50 in-group out-group, and in the sense that anything that departs from that sort of dynamically unstable but evolutionary stable partition means that the smaller group, the out-group – the odd man out – will necessarily ultimately be absorbed into the larger group, so the only stable partitioning is 50-50. Which makes a lot of sense when you look at Trump versus Biden when you look at Brexit versus not Brexit, wherever you look all the important allegiances in terms of our political ideological and possibly even theological communication seems to be split right down the middle. And perhaps it can be no other way. So it'd be very very interesting to simulate that and see if that is a truism that inherits from all of these marginal likelihood or free energy minimizing processes, implemented at multiple levels of hierarchical multiple hierarchical levels, which is you know: communication is just message passing and message passing is just the way you articulate updating and believe updating just is the process of inference which just is the paths of least action according to the FEP.

28:49 Friedman:

The 50-50 politics, it's maximally confusing, something we all experience. And a few key points there about the nature of online communication, is that at the core it is dyadic even when you're broadcasting to many it's actually about that connection and the hermeneutic relationship of unpacking meaning. And then also you brought up the importance of context and identity and who's talking to who and our inferences about that which is essential to rhetoric, and something that often gets left off when people take big data approaches to online discourse.

The next question is: how does Active Inference help us think about science communication and participation? Specifically as we move into broader citizen science initiatives and as scientists are in the loop something you've been recently involved in as well with society and with decision making. So as science and the nature of science is changing: who is doing it, and how they communicate it, how does Active Inference help us navigate that?

30:52 Friston:

Right well I'm sure you've thought about this much more deeply than I have, it's just drawing upon my experience you know in terms of science communication during the coronavirus epidemic. Yeah I think you're absolutely right. You can, as with the previous questions, I think you can take the principles of Active Inference, and just think about what does that mean for optimal communication and and belief updating, and shared belief updating and shared narratives or not and use that as a point of reference for the way that you articulate your own science. And you've asked all the challenging and exactly right questions, you know about how you communicate, how you engage other scientists, or other, other partners within or beyond academia. And I think the same principles apply exactly to the public. And just to reinforce your, I think beautiful observation that all communication is dyadic from the point of view of the person communicating: you know, it's this kind of person as a unitary object I am talking to, or *this* population, or this mentality, or you know, on this discipline. So it is I think fundamentally you know dyadic from the point of view of the person generating the messages that will, may or may not incur belief updating in the recipients. And these kinds of principles, I'm sure, would be useful in terms of science communication. So at that level I don't know that there's much that I would have to do with what you already know and you know and will, and possibly already are implementing.

There is another level though which is using Active Inference not as a model for the way that we work and communicate and participate, but as more a statistical, or observation model of data. So in a sense you can use the principles of Active Inference really to make the most of data pertinent to a particular domain. So again I'm thinking here of the dynamic causal modeling of the epidemiological and behavioral data that has been generated by the coronavirus epidemic. You can certainly use the the perceptual inference side of it, if you like the Bayesian filtering side of it, and but also in principle the the data mining or the optimal Bayesian design to select which data are useful or not, in a very practical way when assimilating big data in the service of understanding the system at hand. So if the system at hand is how does a spike propagate from one neuron to another neuron in a neural network, or how does a virus propagate from one person to another person in a neural population, or a network, a population network, then you can certainly use the data to you can apply Active Inference to build generative models of how you think that occurs. And what immediately confronts you is you've got to put in all of the things that generate those data. So you can't miss out any factors that are important, be they psychological, be they behavioral, be they viral, be they transport-related – all of these things have to go into your generative model to best explain the data.

So you know when we do this in a practical way we both use the instance rates from PCR testing, and Google mobility data, and department of transportation data, anything that speaks to them and reduces uncertainty about all the factors necessary they're entailed by your generative model, in this instance not a discrete space it's written down – no this one is actually a discrete state space model. And the Active Inference is not explicitly part of it, in the sense that we're not trying to predict people's behavior; but it does serve as an indirect guide through the principles of Bayesian optimum design. And all that basically means is: do I invest computational resources and thereby incur computational and statistical complexity by including or attending to this kind of data or not? And then you can actually evaluate the information gain by including that data or that data. So for example you need Google retail estimates and or workplace activity or just one? If you include both, that means that your – the complexity increases and you literally have to wait another half hour before you get the results for your dashboard. Or do you or do you not and have a more parsimonious model in the sense that you have now in the same sense of that saccadic suppression of retinal slip, and you've actually said "no I don't need that I've got everything I need I got the right kind of data just by focusing on these data", and then once you've got that in mind you can now go foraging for different kinds of data, different collaborators from different disciplines, who've got different perspectives but also *crucially* different data, you know to try that will inform and shrink your uncertainty about the model parameters, and also very importantly about the structural form of the model: Do I need this node? How many, is this interaction important or not? Is this degree of nonlinearity justified by the data? So all of these questions affected your hypotheses about how this system is responding or would respond if you intervene on it – all of those questions now become amenable to an evidence-based analysis, because you've got a generative model underneath the hood. So that would be a more practical application of principles that underwrite Active Inference, even though your computer program is not actually doing Active Inference but it's certainly been deployed using the principles of Active Inference.

37:04 Friedman:

Awesome, and we heard that integrative approach: "yes we're going to include multiple data sets, potentially of unconventional type, and we're going to have a principled way of deciding how to include that data". And also as you brought up at the end, who to include in the conversation. And there was one piece you said in there about the dyadic nature of communication, where a speaker is always, I think you said something like speaking to a person, or to a group, or to a community, and it relates to our next question which is: how can we appropriately interact with shared and nested generative models potentially across scales be it person, team, or community? Do we think about these levels of analysis as Active Inference agents in their own right? Or how do we for example speak to a community or speak to a level of analysis that's broader than the personal?

42:57 Friston:

Yeah I think that's a great and challenging question. I mean, clearly there has been some provisional work in academia looking at Markov blankets, which is effectively, from a statistics point of view, from a physics point of view at least, what we're talking about here. As a physicist you'll be tackling this with things like the renormal- or the apparatus of the renormalization group, which tells you immediately something interesting – that the existence of this nested structure if underwritten by, or if it is a renormalization group, means that there are certain functional forms that are conserved. So what that means is from your practical point of view, that there will be certain kinds of behavior that are actually conserved at different scales. So what works in terms of talking to your children should also work as a president talking to your community, or a governor talking to your state, or a team leader talking to your assembled team, simply because in order for there to be a hierarchical nesting that supports that hierarchical structure, that has to be this conservation, usually mathematically written down as a the functional form of the Lagrangian, or it could be the sort of marginal likelihood, or the surprise that we're talking about that underwrites these sort of paths most likely paths or paths of least action. So that actually paradoxically, slightly makes the problem slightly simpler. Because what you're saying is: what works at one level will work at all levels - all you've got to do is find the coarse graining operator that takes you from one level to the next. So what that would look like I think would be very very application domain specific. So you know I think that there is a great challenge ahead which is taking the single particle FEP approach, now into a world where it matters, where the world is actually an ensemble of particles. And we've already discussed the importance of, you know, thinking about worlds where all the particles are identical, whereas all the particles actually vote for half the particles vote for Trump and the other half vote for Biden. And this is interesting to reflect upon pre-21st century physics that was so powerful in articulating this kind of dynamics, because it just dealt with the simplifying assumption that my idealized gas was an ensemble of identical particles. And then you can spin up from that equilibrium physics and everything that led from Carnot cycles and engines, through to current technology.

So it's a powerful assumption, if you just make some simplifying assumptions. But just because, you know we've already said well perhaps that's not the best kind of assumption to make when you're dealing with political mechanics with at least like one bi-partition in there, and so that would require a revisiting that kind of physics from our point of view or your point of view, basically simulating active agents, inference agents or ensembles of Active Inference agents, particles, but where now there's a heterogeneity in play, and then asking the questions: well what are at the next scale the free energy minimizing, or surprisal minimizing, or potential minimizing solutions at the next scale up? So we come back to our you know, why is it the case that people are all split 50-50? Which has an enormous impact on the interactions at the scale below. So I think to tease, I think all I'm just hand waving here, because

I don't think there are formal answers and I think those formal answers will probably have to come out of agent-based and possibly stochastic agent-based modeling initiatives, but with the twist you're making each agent itself an Active Inference agent. So while each individual member of the ensemble is trying to minimize their free energy, also the ensemble through cooperation and the shared narrative is minimizing the joint free energy, and what that means for when you move from one scale to the next scale. I, you know, this is – if you're in physics, I imagine that this is the problem of beyond non-equilibrium steady states because you're actually now dealing with the multi-scale aspect of non-equilibrium. So at best we have good models of turbulent flow and solenoidal dynamics in laser physics, that take us beyond equilibrium physics where all the particles are the same, into non-equilibrium physics. But I don't know that there's an equivalent maths or metaphor, in physics that would really speak to the hierarchical nesting. So I think this is a really open and important research area, that I can only recommend is dealt with by numerical analyses basically predicated on underlying principles.

43:18 Friedman:

Thanks for the answer there, and it made me wonder if "agree to disagree" is a narrative that can be shared even when there is a 50-50 split. And I think it brings us nicely to the final question of .comms which is: how can we move people and teams into a co-transformative space or as some of your recent work discussed, an interactionist space?

50:13 Friston:

Well I'm sure you know the answer to some of these. I don't know, I suspect now I'm now realizing you already know the answers because you're knowing smiles when I say something that you recognize. But so actually answering that from, on the basis of what you just said. So yeah I think that's another really useful insight that you know "agreeing to disagree" is a surprise minimizing, Bayes optimal, explanation for the exchange with others, and so, but it does rest upon committing to the hypothesis that you are not like me you are not like-minded, and that's okay. So I've now classified you as somebody who's not like-minded, and I've resolved the ambiguity in, among the hypotheses that you are either like-minded or you're not like-minded. Normally we resolve those within a few your first impressions within a few seconds based upon all these epistemic cues we offer each other to define the kind of, you know sort of person that we are, so we make that job as easy as possible for us, and signaling to make this so we know our place. And I use that phrase because of course you know there's a paper called "knowing your place" that exploited a shared generative model that allowed you to be in a particular position in some space, even if you should if you and I share the same understanding of political ideology, but I know my place because I'm right-wing and your place because you're left-wing, or vice versa – and then we can quite happily exchange but agreeing to disagree. So I think that that's a wonderful you know perspective to have, and to endorse it, and that is a Bayes optional perspective from both sides of the disagreement, that's resolving uncertainty in a bounded rational way.

So applying that notion to co-transformative space it reminds me of the problem where certain patients in psychiatry have committed to a particular inference about the, you know, whether they belong out there in that kind of environment or not, and they have decided that they do not belong out there and that those people are not like them and they start to avoid, so in a very simple-minded way and if they're in psychiatry, in this example, but I think it's illustrative and useful in that respect. So say take depressive depression or agoraphobia, you know, which is a completely Bayes optimal response, if I have committed to the hypothesis that out there is full of people who are not like me and potentially will upset, confuse, and render me uncertain, and possibly even injure me in some way. So withdrawing into that corner of your house or into that silo if you're working in you know in teams, is a perfectly Bayes optimal response that says that you've got a precise belief that this is where you belong, you know these are the people that you speak to, and not those people. So and, that's usually perfectly functional in psychiatry – that would be a neurotic defense; but it can become pathological when you become housebound. Or, say, if you've got a pathological hypothesis, like your body has dysmorphophobia, and you nearly die because of a failure to eat properly.

So when you say co-transformative I imagine what you mean is you want to transform two teams into one team, or at least enable them to work together. If that's right and you're nodding partially, so assume that it is , you're facing the same challenge that a psychiatrist faces in terms of enabling people to revise the precision of their precise be-, precise beliefs about who they can interact with and who they should interact with. That's not an easy thing but it's certainly doable. And it basically usually reduces to presenting evidence to a group or a person, that it can be another way, so that they start, you start to revise their prior beliefs, or at least the precision of their prior beliefs by, in a safe space where it's okay to explore other hypotheses, enabling them to think about other ways of interacting. So this would normally be the objective of psychotherapy, basically by illustration very much in the same way you were talking about sort of, you know illustrating or educating by embodied experience, very much psychotherapy is thought to work like this: You provide a psychologically embodied experience where you can try out different styles and different hypotheses, and in so doing you, paradoxically introduce the right kind of uncertainty that, about different styles of engagement and who you are and who you are talking to, and by relaxing that precision you enable you give the patient or the naughty team that's become too siloed, the latitude to explore other ways or other ways of behaving. So I would imagine that most of most of the tried and trusted procedures to get teams into a co- into a co-transformative space, and you know, use one of or more of those mechanisms.

What would Active Inference bring to the table? It would just bring the narrative that all everybody you're trying to get to talk to each other can come to share, so they can see through the process of becoming more collaborative or exchanging ideas more fluently or working to, working with the cert- with the same lexicon, or mechanics, or code, and it will enable them, it will enable them to to – having the same narrative will actually shape their prior space and understand the mechanics of actually enlarging the hypothesis space in terms of inter, interaction styles.

So that was an incredibly hand waving answer, but it was in part informed by, almost well, my understanding of the question from the point of view of the psychiatrist who wants to transform the way that a patient relates to her, to her world. And I mean that literally, those drugs that are responsible neurobiologically for setting the precision: if you can temporarily suspend the precision in order to reveal other latent *a priori* hypotheses in terms of the way that I am, or the way that I interact, and the way that I behave, or the way that I perceive, that can actually have long lasting effects, and that, on bringing those other hypotheses to the table in the moment in subsequent interactions. So perhaps the most compelling example of this which is trending at the moment is the use of psilocybin assisted therapy particularly in terminal care. So you know, if you know you're going to die of cancer in the next six months, there are certain hypotheses that are brought to bear in terms of how I would expect to feel and how I engage with the world, and how I engage with my loved ones, as a dying person who is near death, and the ultimate loss. Those are not necessarily the best or most functional hypotheses or ways of being. There are other ways of dying gracefully, and gloriously. But to get at them sometimes having a you know a managed challenge to your 5HT-2a receptors via things like psilocybin and other related drugs just allows you to suspend for a moment your very precise beliefs about the kind of thing I am, and allows you to experience other ways of being and perceiving, which can be very useful when it comes to just trying out other hypotheses, and you know, in your, this is your cancer journey, but you know one can also imagine similar scenarios when you get locked in to a particular way of interacting with either within a team or to between teams in a larger organization. So that would mean you have to go on a retreat and take lots of magic mushrooms.

53:03 Friedman:

Never thought I would hear from you Professor Friston, but there we have it: drugs for teams. Thank you for this excellent interval with the .comms unit. We're going to take another five-minute break and we'll return for the final session for .tools. Thanks again everyone, and we'll see you in five minutes.

Part 3: Tools (.tools)

00:48 Friedman:

Welcome back everyone. This is our third session of the Applied Active Inference Symposium with Professor Karl Friston, hosted by the Active Inference Lab, and it's June 21st 2021. We're here representing the .tools organizational unit of the lab - the third organizational unit in the lab - and the goal of .tools is to enable effective tool and instrument use for all Active Inference Lab processes - so that's just using the digital tools affordances that we have better. As well as exploring and designing affordances for our niche, modifying our niche, resulting in effective action as well as innovations in tool development. As with the other groups, we've been meeting weekly in tools and having a lot of awesome insights related to where Active Inference might come into play and that's what we're excited to talk to you about.

Some of the *core* insights from the work in this unit relates to learning by doing, the recognition that modern systems are cyberphysical - everything is really intercalated with the digital. And also we found it really refreshing, kind of like a two-stroke engine, to be sidestepping, or complementing, or augmenting some of these philosophical discussions with technical clarifications. And two ways in which we've seen that play out: on the left here is a quote from *you* during a 2019 Dropbox blog post when you wrote that "Technology is the natural extension of Active Inference beyond the single person", which of course brings technology far from being something artificial into the realm of extended and embedded *cognition* in our niche. And then on the right side is a slide from a very recent talk by Bert de Vries on "Beyond Deep Learning: natural AI systems", speaking to several applications in hardware and software of Active Inference - for example gesture recognition, robotic navigation, and also audiometry for hearing aids. And one effort that we're starting up *now* is a NetHack challenge. It's kind of a video game played in text characters and we're assembling a team with already multiple interested *participants* to get an Active Inference agent on the playing field so to speak, and have people maybe update their generative model when they see that it doesn't have to be a three billion parameter neural network trained for six months on the GPU but what if it's enough to just be curious and to want to succeed?! Those are the kinds of things that motivate us in .tools.

So we can start right off the bat with asking: How can we use Active Inference to structure the process of innovation and tool development. And, How can Active Inference concepts help us design for complex agents that are interacting in complex niches? For example, thinking about niche modification, extension of affordances, reduction of uncertainty, or structuring of communications.

04:05 Friston:

Again, great questions. So, the use of Active Inference to structure the process of innovation and tool development. That is, , in itself, an entertaining notion, in the sense that, you are a realization of Active Inference! And, you know, I'm mindful that your nice use of the emphasis on curiosity as the imperative that drives most of our behavior is *exactly* the imperative that, as a scientist, drives me and most of the people I know. And in a sense, I would imagine this also drives your initiative and your laboratory.

So, all the questions you are asking are really… you know… , how do I make the next move in order to resolve uncertainty about your particular *model* of how say Artificial Intelligence or human communication is going to evolve. So, yeah in that light, I think there are two levels to the answer: the first one is just to celebrate and acknowledge that you are engaging in the scientific process as formulated by Active Inference, that you are on a journey of trying to satisfy curiosity that will be never ending. And that speaks to one of your themes in the previous slide about "learning is doing". The only way you're going to resolve or sate that curiosity is to go out there and see what happens, and that is exactly the right thing to do.

A more practical level answer though, I think, speaks to the tool development, because one of the fundaments of Active Inference is the appreciation that if you just want to maximize the likelihood that *your* kind of world model or generative model that *entails* the way that you exchange with and interrogate and ping a world, is the right world that is articulated out there, in terms in the sense of extended cognition for example, in terms of the software tools, or the educational tools that you're making available, then all of this is still subject to the imperative to minimize complexity.

So, in maximizing the likelihood that these tools will be out there and in a sense you're saying this model this way of narrating the way that the world works, you provide an *accurate* description that is as simple as possible. So, you know, you cannot escape the complexity - I'm speaking like Jürgen Schmidhuber now – which is a good thing in this instance. So that means you've got to find the simplest tools, and it's interesting that you highlighted Bert de Vries' contribution. Because… you know… again, just practically thinking, what's the game here? The game here is to write down the… or find the best hypothesis, the best explanation for my lived world and my "me" could be Active Inference Lab's, and the lived world is everything that you have to engage with, in terms of educational, commercial, or academic partners. So, you've got to write down, you've got to explore the model space, in terms to find the right generative model of the way that your system or your organization works.

The first steps in writing down the generative model are basically to define its structure in terms of the sort of hidden factors or latent factors and their interactions, and all that good stuff, but it has to be done in the simplest way possible. So, what's the simplest way of writing down a generative model? Well, it's to write down a Bayesian, graphical model. What does that mean for the actual *coding*, practically, and the software schemes and implementation that you would either offer to people or pre-package in terms of user interfaces, then it's going to be message passing on those graphs. I'm trying to get back to Bert de Vries's factor lab, a FourneyLab formation. So, to my mind that's the simplest, most generic bit of computer science that you would come across in the service of finding the right software tools to build absolutely everything, because absolutely everything can be written down as a generative model. If there's a generative model, there’s a Bayesian dependency graph, if there's a

Bayesian dependency graph, you know there's a factor graph, if it's a factor graph then you know there's a message passing scheme. What is that message passing scheme? It's just a variational Free Energy minimizing message passing scheme. So , I would imagine that, as tool development increases, there will be a move towards a common language that will look very much like Bert's Forney-style message passing. And within that, you know, you've got very limited choices, which is a good thing because that again speaks to this minimization of complexity, and just coarse-graining the world, and your world at its coarsest level that will sustain an accurate account, or a precise account of what you want to achieve.

The tools just have to come in two flavors. They have to deal with continuous state space, generative models, to interface with, you know, of the kind you need for robotics. But also the other flavor will be in discrete state space, and your latent states, latent discrete states, models that you need to do for, say, computational linguistics, or, you know, modeling the climate in various states. And we know all the message passing schemes that would be entailed by a commitment to one of those two kinds of models in the sense of generalized Bayesian filtering for the continuous state space. And by generalized, you include generalized coordinates of motion which generalize things like Kalman filtering, and on the discrete state space side you're talking about either belief propagation or variational message passing.

So, when you just think about it, what you have to do in providing tools of a software kind or a simulation kind, you know, happily there aren't many choices you have to worry about. So, you know, in that sense, all you need to do is to make sure that your tools accommodate both generalized Bayesian filtering and belief propagation and/or variational message passing, and then you're using off-the-shelf technology. Which brings us back to, well what's the real problem then? Well, the real problem is writing down the generative model. What sort of problems, how would you unpack those problems in terms of innovation and tool development? Well, it's solving the model selection problem. So, sometimes I think you sort of when describing the space of problems that are faced say with generalized AI or AGI, you know, you can unpack them at different spatial temporal scales into the inference problem, into the learning problem, and into the selection problem - by which I mean using Bayesian model selection to get the right structure. You know, do I use six or twelve layers in my deep neural network? Do I use a convolutional model or do you use a transformer? These are basically problems, that are solved, if you have a mechanics that can score the structure enabling you to select the right form. So that, I think, is going to be a focus of innovation in the… yeah, it already is, but certainly in the near future, in terms of development. And in the sense that I think the inference and learning problems, they're solved problems that you can just go to Bert and get your favorite message passing scheme, or you can keep at the level of your educational or academic message passing user MATLAB schemes that we generate here in London for toy problems.

And what is not, I think, a solved problem, and will require an innovative solution, is the structure learning problem or the selection problem. Exploring not the right hypothesis, but, you know, we're in the principled way, exploring the space of generative models you might want to bring to the table. And that has many different issues. And, things that come to mind are, of course, that you could do it in a bottom-up way by trying out new hypotheses. Where you get those from, you get them from experts in the field because effectively they are bootstrapping themselves on the basis of our prior beliefs or your knowledge about how something works. You can do it in a top-down approach by having over-parameterized… over-expressive models but with very weak imprecise parameterizations and then use Bayesian model reduction to solve the selection problem - these are ways that people are thinking at the moment.

But this thinking is innovative because I don't think there are any clear answers. So how would you use Active Inference to solve the structure learning problem? Well, in a sense it's already being used in the sense of Bayesian model selection as natural selection, but you really want to speed that up and make it work within your commercial academic lifetimes, but I would imagine that exactly the same principles would be brought to bear there.

That almost answers the next one "how can active concepts… how to design complex agents interacting in complex niches?" You just have to build these things as a proof of principle and hypothesis testing, and the nice thing is, you know, all the machinery and the tools that would be requisite in building these things right from the variational message passing using, say, ForneyLab through to now, you've got the right fitness function when it comes to using say a genetic algorithm to explore a structure space. And what is that fitness function? It's the evidence lower bound or the variational free energy. So, you've got all the maths in place. This is a question that I think of simulating these things and providing proof of principle. How would you translate that into the real world? I don't know at this stage, I'm afraid. I think, you know, a challenging first step would be to actually use robotics or *in silico*, or sort of hardware, or possibly a lot of excitement at the moment using soft robotics, and actually, you know, in you design your niche and see what happens and then turn your attention to niche construction – where you now acknowledge that the niche itself is also succumbing to the principles not of Active Inference in and of itself, in the sense that niches don't plan, but certainly in the FEP sort of vanilla free energy minimizing approach.

So yeah. I haven't actually thought about that before that but that's an interesting asymmetry, when it comes to simulating multi-agent interactions in the context of niche construction – where often it is the case that the niche is just the other agents in an ensemble. But if you now actually include the environment as part of the niche that is playing host to all the denizens that are the ensemble of Active Inference agents. Then there is this distinction between the ability to plan the consequences of action that would entail optimization of the expected free energy versus simply reflexively minimizing surprisal by minimizing free energy as an evidence bound. And put that even more simply, more intuitively: you're either with generative models that support planning or not. You know so there's nothing, I think, fundamentally different between these approaches, it's just if you've got a generative model, that is a model of the paths into the future consequent upon how you act upon the world. That's a much *richer, deeper* generative model than the kinds of generative models that would be applicable for a thermostat or an environment. And it's likely that the environment that, you know, that I have in mind here - which is a warehouse where you've got a sentient robot going around trying to collect the right things, so the robot can plan but the environment, the niche can't. It will still conform to the principles of variational, you know the Free Energy Principle. There will still be particles and things that are conserved and they will still fall and behave in a predictable way, there may even be a thermostat controlling the temperature - but none of these things are planning. So, there's an interesting asymmetry that gets into the game when you're talking about complex agents interacting in complex niches. Part of that complexity has to be a specification of whether the complexity entails planning or not. And it just creates different problem spaces , certainly in the context of multi-agent simulations, it sort of carves up the problem spaces implicitly. Problem spaces that will only be explored by doing and by doing I just mean actually realizing physically these processes in the kind of situations that you think are going to be useful for the future.

19:09 Friedman:

Thanks for the answer, and it's really fascinating about using simulation so that selection can happen within the generation of, for example a startup, rather than between generations. Because of course we can let organisms or startups proliferate, and then let pruning occur at the generational scale. Or there could be ways to design so that selection occurs within a generation, more like learning and development rather than intergenerational selection. So awesome points there.

This could be a broad question but we're curious: What areas of applied Active Inference you think just might be exciting, promising, or important?

21:32 Friston:

So my personal/usual response to this comes in two flavors. The first is from the point of view of a theoretical biologist and a psychiatrist. So if you can understand how a normal sentient artifact or person behaves, then that creates a space in which you can think about false inference and false learning – or certainly suboptimal, from the point of view of minimizing surprisal or free energy. So that's a fancy way of saying understanding the computational basis of psychopathology. So, you know, there's a whole literature on using Active Inference as a, if you like, a *normative* framework within which to provide an ontology of false inference or failures or aberrant Active Inference. And why would you want to do that? Well, if it can all be reduced just to the good belief updating, and the good message passing, we actually have quite a comprehensive understanding of neuronal message passing, and all its physiology, and all the roles of various neurotransmitters, and microcircuits and your anatomy that underwrite that kind of neuronal message passing. And implicitly we also then have a fairly fine-grained understanding of the role of neurotransmitters and the consequences of pharmacological interventions in the context of experience dependent learning and an inference of the kind we've been talking about. So from a translational perspective, literally translating the formalism on offer from Active Inference into the clinical domain, that would be one motivation for developing this theoretical framework.

The other one is more in the line of technology and artificial general intelligence. So then the question is: well I now want to build sentient artifacts and not only build them, but build brothers and sisters so they are complex and interact and learn to love each other in a complex environment that could include me. And then, you've got a clear offer from Active Inference as to the design principles you might want to use to actually *build* these artifacts. And then there are interesting questions about what kind of artifact do you want to build? We've already discussed the difference between a thermostat and a sentient robot going around collecting your next sort of home delivery. There are different kinds of generative models. So now you ask the question, what are the exciting and promising kinds of artifacts, as defined by their generative models that one might expect to see in the future? And then we get into the world of generative models for support planning, so we're talking about deep generative models where they have a temporal depth.

What are the next stages that you might be looking at? Well there's also a sort of hierarchical depth that would at some point… first of all include the capacity to deploy precision, and why is that important? Well, as soon as you have the deploying the precision as a process of inference, you have now a normative theory for this kind of mental action or covert action. So, one example of this would be I don't know how the technology works but I can almost… I can be assured that I know what it's trying to do, but thinking about transformer networks, and the way that attentional selection operates in this context. What you're saying is you can actually optimize the attention selection as an inference process, using Active Inference or an evidence lower bound and where you're now predicting what things to attend to and what particular weights to switch on and which weights ways to switch off. And at that point you can understand that as mental action. So when the transformers or variational autoencoders start to now optimize their estimates of the posterior precision at lower layers in an auto encoder, it's now acquired the capacity for mental action, and it now will pay attention to various representations and possibly even various data sources. That's not magical! We do that every day, in the sort of MDP (Markov Decision Process) and use it to explain a lot of the attentional mechanisms implemented in the brain. If you can migrate that technology into deep learning, you would have taken one baby step towards true sentience, which is mental action.

The next step would be… how can I now minimize the complexity of my generative model, where my generative model now actually includes this "meta inference"? In the sense I am now providing predictions about my inference because I'm controlling the precision of hierarchically subordinate message passing. And at that point you start to think, well perhaps one way of simplifying the computational complexity of the complexity part of the inference would be to carve up different states of attentional deployment in exactly the same way we're talking about carving up people into Biden versus Trump voters. A simple, stable, complexity minimizing carving up, which suddenly suggests to you that you can now equip an artifact with states of mind. So that they can be in four states of mind – they can be happy, they can be sad, they can be confident, they can be unsure, and they will have to infer, given all the evidence at hand, including the message passing lower in the hierarchy, what state of mind it is in. And if you now include in terms of the sensory evidence, you know, the voltage on their batteries or some measurement of their interception, you now have something that's going very very close to, say, Ryan Smith's notion of emotions. So, now you've got a part of the generative model, is now inferring "what state of mind am I in?" as the best explanation for all these interoceptive, embodied sensations. Not just the proprioceptive state of my actuators, but also are they getting a bit sticky? is there some wear and tear? are my batteries charged? all of these things come together as part evidence in conjunction with all the usual visual, radar, acoustic inputs, to actually supply evidence for a posterior belief "I'm in this state of mind", "I'm anxious", "my battery's running out". This immediately creates different prior preferences, cost functions if you like, that would be applied to your policies because you've got a deep changing model that plans into the future. So now you've got an artifact that not only has the capacity for mental action, it's now got the capacity to be in different emotional states.

The next step is to say "hang on, so there are these different states… can I now equip it with minimal selfhood?" Can the hypothesis that I am actually an artifact provide empirical priors that reduce the complexity of my message passing at subordinate levels that is generating? That is, inferring the state of mind that I'm in that in turn optimizes the posteriors of the precisions of various likelihood mappings or preferences over policies. So at this point you're starting to get to artifacts that could have minimal self-awareness. The next stage would be, that's only going to be ever, useful when you consider dialectic interactions again because the only rationale for having self-awareness is to disambiguate self from other, which means that there must be some confusion or some uncertainty at hand, in order to justify the resolution of uncertainty, justify that complexity of the model, which means that you have to be interacting with or exchanging with things that are sufficiently like you, to license the inclusion in your generative model, of a self versus other, or that you are like me or not like me. So we actually come down back full circle to what we're talking about before, in terms of inferring *who* I am I talking to. So this structurally something quite fundamental about this inference problem: are you a creature like me or not? – or are you like one of those? – are you a pet? – are you a plant? Just being able to carve up this world in a way that is self-referential, necessarily entails a minimal selfhood in the inferences of these, that speaks to the importance of getting the necessary evidence from the environment. That would be, if you like, license that degree of complexity and the only kinds of environments that can supply that degree of complexity, are when that environment, that eco-niche is actually comprises other agents like me. That makes it, if you like, worthwhile me inferring "oh it's me, not you doing that."

So I would imagine then, the most promising applications of active influence in constructing sentient artifacts, pets and you know, carers or things that you can converse with, would be to grow them certainly with themselves, but more importantly with you there, so they can learn by their doing with you there, so they're curious about you, and you're curious about them. And at that point, one could argue that's the only scenario in which you're going to have any empathetic interaction with these artifacts. So I'm sure there are other applications in terms of climate change, or commerce, or whatever. But in terms of imagining what you could produce, what you could sell, I would imagine that you know, a mindful robot that actually is curious, genuinely curious about you, because that will teach you something about itself.

32:11 Friedman:

Thanks for that answer. The idea of tools for attention, and of design and engineering for regimes of attention, to use an Active Inference term, is really essential. And what you were talking about there with the phone: first off before the internet when there weren't other devices of similar kind, there was no need to communicate out, and what we've seen is that as there's more and more devices of similar or interoperable kinds, new levels of organization have to emerge. And then I thought about the anxiety that a person might feel when their phone is running low on battery. Right now that sensor reading is getting emotionally offloaded to the human. So we could have that anxiety on device, so let's have a more relaxing relationship with our phone and then as you pointed out it would be the incipient steps of selfhood, or perhaps what they could even call a self-phone, if I'm allowed one pun per symposium.

The next question is what kinds of tools have been most helpful in your work in research? Which includes many areas such as SPM, and DCM, that a lot of people who are just learning about Active Inference might not be very familiar with. And, what kinds of tools don't exist yet but might be helpful for Active Inference work?

36:14 Friston:

So, the mathematical tools so you know, and I'm often asked this question of students, "do I need to be able to do maths to contribute to this field? – and if so, what kind of maths?" I won't tell you what my answer is, but what I have found useful is certainly mathematics, but not necessarily very high end, this is always Wikipedia-level mathematics, and in particular, dynamical systems theory, information theory, and linear algebra are probably all you need to to do everything really. And indeed you could read most of quantum electrodynamics as basically linear algebra with a bit of probability theory underneath it. So that has been the mainstay, that is, if there is one tool that it would be the tool and the language of maths. And relatively simple maths. And the second thing is to the learning is doing you know a "see one do one teach one" ethos I think applies very pragmatically in this context, which means, it, you know, it's actually very useful if you can get students to actually build their own little simulated artifacts and even more useful when you know when they can actually code it out themselves. Which means you need access to a high level, at least third generation, programming language that you know a student can get fluent with should they want to. Not only to use the existing tools, but you know, try and write it down themselves without having to spend years training as a computer scientist. So I found MATLAB very useful in that respect, not because it's terribly efficient, although I have to say actually some of the matrix operators and under the hood tensor operators are actually much more efficient than people give them credit for, because it actually came from X-ray crystallography. However what's really useful about it, is that it uses the same syntax that you would find in a book on linear algebra, which didactically or educationally is really quite important when it comes to writing and reading the code. So we have deliberately stuck with MATLAB not because it's computationally efficient, or that it's open source , it should be I don't think it is, but simply because it's configured in a way that people reading standard texts, 101 texts and in linear algebra and the like would be able to see how it transcribes into a computer language, so that's been a really useful tool. And looking ahead I imagine that you know, one's gonna need open access and possibly more, I don't know could go out the way I'm just thinking about first of all people like Bert [de Vries] and Forney Lab in terms of very generic, very high-end specifications of message passing in computer science, that may be that that's the level you want people to actually compose their generative models and their artifacts and they don't even need to know about linear algebra, and even less information theory, what they need to know is that the language of relational, you know the object relations and how to specify just different classes of exponential probability distributions, and you know is it categorical, is it continuous, is it always positive, or can it be positive and negative, and that may be quite sufficient to write down a factor graph, or a generative model and then everything else is just off the shelf and it'll write itself.

So that would certainly be possibly helpful in the future and so I'm moving on to what kinds of tools don't exist at the moment. So I'm thinking of it, what I never used it but I would I imagine would be Bert's Forney Lab facilities, but offered as an application or a user interface, that allowed you to compose generative models, and then just hit compose a generative model, compose a generative process, the actual world that's going to be modeled, and then there's click run and see what happens. That, you know, that would be really useful I think. Having said that, the other side to future-scaping here, is I repeat this sort of leveraging more specialized or other fields, and the you know amortizing certain parts of inference or learning to infer, or indeed inferring to learn, or learning to plan, or learning to infer how you plan.

So, or starting to sort of see what parts of the inference process are so conserved that they could actually be amortized and learned? And certainly that looks as how that's what the brain has done. For example there are people who think that the cerebellum has basically learned how the motor cortex does its online KL control or Kalman filtering, and therefore lends a fluency and a computational efficiency to the message passing, which you know we which in its absence, it doesn't mean you can't do something it just means you can't do it as fluently, and as gracefully as quickly as you, your, as you know as you could with a cerebellum. Indeed when you have a cerebellar lesion all that really happens is you become a bit clumsy and slow. So, that, those kinds of tools, you know a quick and cheerful integration or importing various amortization and deep learning technology into a Forney-style message passing scheme that could support, you know, any kind of generative model, I think would be really really useful.

39:16 Friedman:

Awesome, thank you. Approaching this nexus from another angle, what kinds of tools and platforms could inform transdisciplinary, highly contextual, and engaged teams that are working with these approaches? ActInfLab, we hope to be working with others to be developing the Active Inference curriculum, and body of knowledge more broadly. But when teams are actually using these kinds of approaches what kinds of platforms might exist to enable their work?

41:02 Friston:

Yeah, okay, I have a strong suspicion that you know the answer to this , so I'm trying to guess at the answer that you know is the right answer, and I'm not doing very well here. So I, you know, I think that you, that we've already talked, and certainly implicitly in the way that you presented the ambitions, and implicitly send the questions, and you know all the answers are there. Whether that's trying to engage through education, whether it's trying to engage through insight, using say, you know embodied experience illustrations of the basic principles, whether it's supplying games or user interfaces, graphical user interfaces to facilitate the designing and enacting and the playing with generative models and Active Inference. I think these are all your obvious and laudable ways of leveraging what Active Inference has to offer. Participatory, yeah that's… . I mean the "learning is doing" thing and the "see one teach one do one" you know keeps coming back to mind, and the course completely licenses the participatory aspect. But what kind of participation did you have in mind, are you talking about sort of hackathons? Are you talking about sort of playing games with Active Inference? Computers that start to hate you or love you or what level of participation were you?

41:18 Friedman:

Yeah Stephen, do you want to give a quick thought on a few kinds of participation, or what does that mean to you?

41:49 Sillett:

Yeah, one area is quite interesting, is in psycho drama, they use action methods like action sociometry, or spatial activities, to look at how people relate to their experience in a dynamic way, so physically. So I've been looking at ways that spatial participatory approaches can unpack people's relationships to different niches or different workplaces or different types of embodied experience, and then that could be visible, to be put into Active Inference type geometries.

47:02 Friston:

I see. Well, there's a great example. So two things that I've come across before, are the architectural design and the importance of not just you know sort of pragmatic affordances that says, you know, can I walk up, can I sit there. But also the epistemic affordances you know if I look over there what would I learn about the space around me, if I go around that corner? So there is you know embryonic interest, in, you know, in my world, from this, the architectural sciences and you know and architecture in and of itself, that is, could in principle be motivated – it's an odd discipline, because it's half like art and half like science, but certainly some of their ideas are very much aligned with certain Gibsonian notions of affordance, and also the affordances, the dual aspect affordances brought by expected free energy under Active Inference. So it's not just you know, am I the kind of creature that can sit on this particular chair, but also what will I learn if I do so? And so things become epistemically attractive to engage with The other domain is in entertainment and in music and in particular the joy of synchronization and mutual predictability or minimizing free energy through mutual prediction, prediction when singing or dancing together, or indeed interacting with a slightly greater asymmetry in terms of being a member of an audience, watching a band for example.

So you know one of the key things that comes out of that kind of research is ways of measuring the implicit generalized synchrony that you get from having this information geometry that I was talking about before, that rests upon there being a synchronization manifold between the inside and the outside. But if the outside is another inside from another person's point of view, what you now have is something called a synchronization manifold. So there's a mathematical image or space, to actually talk about mutual inference and mutual Active Inference and engagement and communication, singing together for example, or diachronically exchanging messages, that does actually translate mathematically into movement and belief updating on a synchronization manifold. And that has real world correlates, you can measure that using kinematic measurements, you're putting LEDs on people who are dancing together for example, or measuring heart rate variability or galvanic skin responses or , doing eye tracking or indeed EEG, and start to – so there's quite a lot of work in , in things like hyper scanning, and in you know sort of ethology, and dance discipline, where in the arts, in the life sciences where they do use a lot of these techniques to quantify the degree of generalized synchrony. What it would be nice to do is actually try and model that synchrony, or understand that synchrony, in terms of movement on the synchronization manifold which is sort of the mutual belief updating. And one thing which comes out of that, just in discussion if not, if no if no further, is the reciprocal, the circular causality that is necessary to maintain that generalized synchrony, the particular synchronization manifold we're talking about for, from the point of view of Active Inference of course is mediated across the Markov blanket, as of the active and sensory states. And, but in general, you need to have reciprocal coupling in order to get synchronization, so directed coupling doesn't work. And if that's true, what that means is that engaging as an audience for example, or participating as a spectator, will only really work in terms of establishing that generalized synchrony that you are chasing, and while you're chasing it well as soon as you have a generalized synchrony, you've got predictability for free for all, and that's a good thing because that minimizes free energy. You know: the more predictable you can make the world better, the better it is, from the point of view of free energy. But you can only do that if as a member of the audience or a witness to something you can actually actively intervene on it.

So that brings to mind how could you get, for example discussion discussions with friends of Maxwell [Ramstead], if you wanted to promote virtual concerts, you know, online for example during the pandemic, what you don't have online which is what glues things together things like mosh pits in sort of carnivals and festivals, is you don't have the audience participation the applaud, the roars, the the lighter waving, or the light waving. So how would you get that back into a virtual virtual experience because that would be absolutely essential, you know, I think to actually engage people, otherwise it'll be just like a pop concert on television. So you know more than just if you like revealing the underlying correlates of that generalized synchrony in terms of the EEG traces of the dancers or doing some sensory mapping from their motion to auditory input. You know, just making it making the sensory evidence that supports the mutual inference, more precise and more available just by having it displayed, say, by putting motion in sound or sound in motion, or EEG, or electromagnetic measures of performance, or the audience, visualizing that – and that has been done by people like Paul Verschure in Barcelona. More than that to actually enable the audience to change what the performers are doing. You have to make, you know – or perhaps what other members of the audience are doing. So you have to empower them to close that circular causality to get that dynamical coupling place, so you get the right kind of generalized synchrony. So that, you know that that sort of dynamical systems perspective on synchronization and free energy minimization certainly speaks to a particular kind of participation and engagement, that does indeed rest upon action-oriented approaches. But crucially it's the action of the audience, on the performers, not the performer's action on the audience, that is usually what you need to pay more attention to. I don't know if that was that the kind of thing you were thinking about?

49:55 Sillett:

Yeah, that's really a useful answer actually. Yeah, we were thinking about that, and some participatory immersive theater type events, and other participation in collective meaning making, so that's the type of thing that we're looking at.

50:32 Friedman:

And it reminds me of the live stream affordance, which is relatively novel, but allows people to be asking questions. And it enables not just efficient production of material in a one-shot approach, but it allows the feedback, and I can't help but add that it's that affordance for participation, for example "speak now or forever hold your peace", that expands the wedding into the community, because there is the opportunity for feedback it's not just a breakaway clique, it's actually something that remains integrated through the affordance for participation.

So I'll turn to the last question for this section: How might future modeling involve large-scale patterns in social data sets, and working backwards to infer their hidden causes, for example in the case of pandemic modeling, governance, economic, other situations?

53:26 Friston:

Well this is a very practical and very prescient question, because of course a lot of people are asking themselves that now specifically with respect to pandemic models, but also the people who are exercised and have the interventional clout when it comes to COVID, are generally also the people who are invested in climate change problems as well. So there's a lot of , there's a lot of noise out there at the moment about you know, how we can harness the data assimilation and modeling advances made during the during COVID-19 and keep the momentum up to tackle climate change - and what you know not just climate, but the economic structures, that, and financial structures, and informational structures, that are deeply interwoven in terms of and climate change. And my answer is going to be somewhat deflationary. And I've had this kind of conversation before, again with Maxwell (Ramstead) and John Clippinger and friends, and due actually to have another conversation with him on certainly Open World, or (I can't remember), in the near future. There's a temptation to take all the high church of the Free Energy Principle and Active Inference, and epistemic foraging, and all of that good stuff we were just talking about, and say oh well now let's make it work in terms of understanding say the pandemic. And you don't need to do that. All you need to do is to apply the good scientific principles that things like Active Inference appeal to, to the problem at hand. And it all comes back to the generative model. So you know, all you're saying here is, how might future modeling involve large-scale patterns of social data to infer the hidden causes? Is just a statement of, we need the right generative models to make proper sense of the big data at hand. And in saying the right generative models, we need the equipment both to invert those models in the sense of inferring the parameter's interactions, using the simple tools we've just talked about. They will just be lifting it from the laboratory or you know continuing to use MATLAB.

But the bigger problem is what we talked about, which is the selection of the structure learning problem. So this goes beyond just higher you know how many layers do I have in my deep network? Much more important, I think it's a factorization - it's knowing how many conditionally independent factors do I need to minimize the complexity to get the right granularity, the right way of carving up the latent causes behind all the data that is available to me? So I think the pandemic modeling is a beautiful example of this, because you know the factors that determine whether I infect you can certainly be written down in terms of virology and the ACE receptors, ACE-2 receptors, and basically reproduction numbers, and transmission strengths, and transmission risks, and you know, the spike proteins - but that's only half the story. The other half of the story is, how likely are you to be at work or at home, when I'm at work? Are you likely to be wearing a face mask? Are you going to, or are we going to be one or two meters apart? So all these behavioral aspects start to become really important factors. And even beyond that when it comes to making sense of the model, the likelihood part of the model that actually generates the data, it can become extremely difficult to optimize when you start to think about what kind of data is at hand? For example just notification rates of new cases per day in the, of SARS, of coronavirus. Now you might think oh that's really great data. It's really difficult data to handle because the different kinds of tests not only have differential false positive and false negative rates, but the different ways in which they are deployed, really compounds that in terms of the selection bias. So are you testing people who are symptomatic, what's the probability of being affected if you're symptomatic, are you not are you doing survey testing, are you doing the same amount of testing this week as you were doing last week, all of these what would be if you like from an epidemiological or a behavioral science perspective really un-interesting factors, suddenly now become the most important factors in making sense of those data. But you only know that when you start to do the model comparison, the structure learning, when you actually commit to writing down the congenital models And that's certainly what I've learned over the past year, now coming up for a year and a half.

You know the future of modeling is .... First of all it's obvious what the future is, it's just basically writing down the right kind of dynamical state-space models that account for data. But the future is really dealing with the problems of structure learning and model selection for any data, but in particular from the big data at hand in terms of pandemics or trafficking on the web, you know or climate change. So it's a really exciting opportunity. Why do people want to do it? Well once you've got the most evidenced i.e the, you know, the minimum free energy model at hand and you've got posteriors over all the model parameters and all the right interactions, then you can do all sorts of stuff in terms of reducing people's uncertainty about the future, because you've quantified the uncertainty and explained to them things that were once uncertain about and what isn't uncertain about. That has enormous implications for mental health and well-being, and possibly even feeding back into finance, because you always hear well, the biggest determinant in terms of the markets is the market confidence. It's all about the uncertainty, so if you can do uncertainty quantification in a principled way, using this kind of modeling you've done a big thing already. But then you come to monitoring putative interventions you've now got a direct handle, posterior estimate on the latent states you actually want to make decisions on. So it's not the notification rates or the number of new cases in California today, it's a number of new people that have become infected today. And that's a very difficult thing to infer given all of these complicated aspects of the generative model. And then of course once you've established the validity of this model, in terms of its construct and predictability, then you can intervene on it. Then you can say well, what would happen if I changed this? Or what happened if I changed that? And what would happen now, what would happen in the future? So that, you know, you're suddenly in a world of quantitative modeling, where you can start to ask some very powerful questions, and also share with everybody who matters, and the products of your inference. So you can now start to think about supplementing the weather forecast with an epidemic forecast, you know, the virus in your area, and tomorrow we expect you know… You can also do that for the markets. And these kinds of things I think are going to be more important when people, or when the current generation get to – your generation I guess, start to wrestle more with climate change, because they're going to want to not just know what whether it's going to rain tomorrow, they're going to want to know you know, at the level, not just the weather but the climate, what are the indicators? Because those indicators really contextualize, and inform their generative models about their place in the world, and that global, that global scale. But to provide that kind of weather forecasting, that meteorology beyond the weather, you're going to need to have these state space models probably optimized, and you know, in a first principle way in relation to their marginal likelihood on their evidence bounds. And we've just read governance here, because governance is just policy decision making based upon counterfactual outcomes, so that is always underwritten by these Bayesian beliefs. But you can't get the basic beliefs, unless you've got a generative model and has the consequences of action in the future there would be also interventions either politically, or financially, or or otherwise.

60:38 Friedman:

Thank you so much again for joining this symposium. It was really a special moment for the Lab, and we look forward to continued interaction. So, much appreciated, and we will see you all soon. Thanks for everyone who's watching, and we hope that you participate in ActInfLab. So, thanks everyone! Bye.

Supplemental Lists:

1. Subjects
2. References
3. Mentioned Names & People
4. Other Resources

1. Subjects

4E Cognition (Embedded, Extended, Embodied, Enactive Cognition)
Serotonin 5HT-2a receptors
a priori hypothesis
Attract
Attracting set
Active Inference
Active Inference framework
Active Inference Lab (ActInfLab)
Active Learning
Active Perception
Active Sensing
Agent
Ambiguity
Attentional Selection
Bayesian Belief Updating
Bayes Optimal
Bayesian Mechanics
Bayesian Reinforcement Learning
Belief-based Schemes
Bounded Rationality
Circular Causality
Co-transformative space
Death
Deep learning
Dyadic
Dysmorphophobia
Enactive perception
Epistemic foraging
Epistemic trust
Expected free energy
Factor graph
ForneyLab (https://biaslab.Github.Io/project/forneylab/)
Gain
Generative Model
Hierarchical Nesting

Hermeneutics
Ideomotor
in silico
Infer
Information Gain
Information Geometry
Integrative Framework
Interdisciplinarity
Kullback-Leibler (KL) control
Language
Least Action
Lexicon
Linking
Loss Function
Machine Learning
Marginal Likelihood
Markov Blanket
Markov Decision Process (MDP)
Mental Action
Metaphor
Neural Network
Neuroectoderm
Normative Theory of Sentient Behavior
Objective Function
Objective Functional
Optimal Bayesian Design
Optimization
Partition
Pathological Hypothesis
Perceptual priors
Partially Observed Markov Decision Process (POMDP)
Passive Inference
Plausible Hypothesis
Posterior Belief
Probability Transition Matrix
Process
Psilocybin

Psychotherapy
Principle of Least Action
Process Theory
Pullback Attractors
Purpose
Reinforcement Learning
Renormalization
Retinal Slips
Risk-sensitive Control
Saccadic Eye Movement
Saccadic Suppression
Self-information
Self-organization
Sensory Modality
Sentience
Separation of States
Shannon Information
Situated Cognition
Somatosensory Dimension
Sparse Coupling
State-action Value Function
Statistical Manifold
Sub-personal Prior
Surprisal
Surprise
Synergy
Teleology
Trajectory
Transformer
Variational Autoencoder
Variational Free Energy
Variational Message Passing
Variational Principle
Variational Principle of Least Action
Visual illusion

2. References

Bouizegarene, Nabil ; Maxwell J. D. Ramstead; Axel Constant; Karl J. Friston; Laurence J. Kirmayer (2020) "Narrative as Active Inference: an integrative account of the functions of narratives." Accessed at https://www.researchgate.net/publication/342828597_Narrative_as_active_inference

Constant, Axel; Andy Clark; Michael Kirchhoff; Karl J. Friston (2020) "Extended Active Inference: Constructing predictive cognition beyond skulls." https://doi.org/10.1111/mila.12330. Accessed 2021-12-14 at https://onlinelibrary.wiley.com/doi/10.1111/mila.12330

Engel, A.K.; Friston, K.J.; Kragic, D. (2016). The pragmatic turn: Toward action-oriented views in cognitive science. LCCN 2015048280. ISBN 9780262034326.

Engel, A.; Maye, A.; Kurthen, M.; & Konig, P. (2013). Where's the action? The pragmatic turn in cognitive science. Trends in Cognitive Sciences, 17(5), 202-209. doi:10.1016/j.tics.2013.03.006.

Forney, G. David (2001) 'Codes on graphs: Normal realizations.'IEEE Transactions on Information Theory 47.2 (2001): 520-548.

Forney, G. David (2012) 'Codes on Graphs: Fundamentals.' Accessed 2021-11-28 at https://arxiv.org/abs/1306.6264

Smith, Ryan; Karl J. Friston2; Christopher J. Whyte (2020) "A Step-by-Step Tutorial on Active Inference and its Application to Empirical Data."

Vasil, Jared; Paul B. Badcock; Axel Constant; Karl Friston; Maxwell J. D. Ramstead (2020) "A World unto Itself: Human Communication as Active Inference." Accessed at https://arxiv.org/abs/1906.10538

3. Mentioned Names & People

- John Henry Clippinger
    - MIT Media Lab
    - https://spatialwebfoundation.org/
    - https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=824432
- Bert de Vries
    - https://biaslab.github.io/
    - https://kdvsfoundation.org/medical__board/dr-bert-de-vries-md-phd/
- G. David Forney
    - https://ieeexplore.ieee.org/author/38558627400
- David MacKay
    - http://www.inference.phy.cam.ac.uk/mackay/
    - https://www.inference.org.uk/is/
- A. A. Markov (1856-1922) (Andrei Andreyevich Markov "senior")
- Beren Millidge
    - https://github.com/BerenMillidge
- Judea Pearl (b. 1936)
- Maxwell J. D. Ramstead
    - https://www.researchgate.net/profile/Maxwell-Ramstead
    - https://scholar.google.com/citations?user=ILpGOMkAAAAJ&hl=th
- Claude E. Shannon (1916-2001)
- Ryan Smith, PhD
    - https://www.laureateinstitute.org/ryan-smith.html
- Paul Verschure
    - https://specs-lab.com/portfolio-items/staff-paul-verschure/
- Wanja Wiese
    - https://predictive-mind.net/ppp-contributors/Wanja_Wiese

4. Other Resources

- Blake, A. and Yuille, A. (Eds.) (1992) Active vision, Cambridge, Mass.: MIT Press.
- Churchland, P., V. Ramachandran, and T. Sejnowski. 1994. A critique of pure vision. (In Large-scale neuronal theories of the brain, ed. C. Koch and J. Davis. Cambridge, MA: MIT Press.) Accessed 2021-12-12 at https://cogsci.ucsd.edu/~nunez/COGS1_F07/Sejnowski_rdg.pdf
- Clippinger, John Henry. (1999) ed. The Biology Of Business: De-Coding The Natural Laws Of Enterprise. https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=824432
- ForneyLab https://biaslab.github.io/project/forneylab/
- Friston, Karl; Conor Heins; Kai Ueltzhöffer; Lancelot Da Costa; Thomas Parr (2021) "Stochastic Chaos and Markov Blankets." https://doi.org/10.3390/e23091220.
- Karl Friston, Thomas FitzGerald, Francesco Rigoli, Philipp Schwartenbeck, and Giovanni Pezzulo (2020) Active inference: A process theory. Neural Computation, 29(1):1-49, 2017. fil.ion.ucl.ac.uk/~karl/Active%20Inference%20A%20Process%20Theory.pdf
- "Cybernetics is the science of defensible metaphors." – Gordon Pask (1975), Conversation Theory.
- "David MacKay's work" and “CSC412/2506 Probabilistic Learning and Reasoning”. Accessed 2021-12-12 http://www.cs.toronto.edu/~jessebett/CSC412/content/week1/lecture1.pdf
- "Research Debt" https://80000hours.org/podcast/episodes/chris-olah-unconventional-career-path/
- Hohwy, J. (2018). The predictive processing hypothesis. The Oxford handbook of 4E cognition, 129-146.
- Hohwy, Jakob (2014) The Predictive Mind.
- Leibniz, G.F.W. (1678) "Nature, proposing some end to itself, chooses the optimal means." in "Definitiones Cogitationesque Metaphysicae"
- Lilly, John C; Antonietta Lilly (1978) The dyadic cyclone : the autobiography of a couple. Accessed online 2021-12-15 at https://archive.org/details/dyadiccycloneaut00lillrich
- Millidge, Beren (2021) Applications of the Free Energy Principle to Machine Learning and Neuroscience. arXiv:2107.00140v1 [cs.AI] 30 Jun 2021. Accessed 2021-12-14 at https://arxiv.org/abs/2107.00140

- Newen, Albert; Leon De Bruin; Shaun Gallagher (eds.) (2018) The Oxford handbook of 4E cognition
- Noë, A. (2004). Action in Perception. Cambridge, Mass.: MIT Press
- Papo, David et al (2014) Complex network theory and the brain. Accessed 2021-12-15 at https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.2013.0520
- Pezzulo, G.; Donnarumma, F.; Iodice, P.; Maisto, D.; & Stoianov, I. (2017). Model-based approaches to active perception and control. Entropy, 19(6), 266. https://doi.org/10.3390/. Accessed 2021-12-12 at https://www.mdpi.com/1099-4300/19/6/266
- Schwartenbeck, Philipp; Johannes Passecker; Tobias U Hauser; Thomas HB FitzGerald; Martin Kronbichler; Karl J Friston (2018) Computational mechanisms of curiosity and goal-directed exploration. DOI: https://doi.org/10.7554/eLife.41703
- Sejnowski, Terrence J. 1994. A critique of pure vision. Accessed 2021-12-12 at https://cogsci.ucsd.edu/~nunez/COGS1_F07/Sejnowski_rdg.pdf. Also see https://cogsci.ucsd.edu/~nunez/COGS1_F07/Sejnowski_slds.pdf
- Varnes, Erich W. (2004) Noether's Theorem. Accessed 2021-12-14 at http://www.physics.arizona.edu/~varnes/Teaching/321Fall2004/Notes/Lecture14.pdf
- Wesolowski, Lukasz et al (2014) TRAM - Optimizing Fine-grained Communication with Topological Routing and Aggregation of Messages. Accessed 2021-12-15 at http://charm.cs.illinois.edu/newPapers/14-18/paper.pdf