ActInf Livestream #038 ~ “Brain architectures for predictive coding and active inference”

Discussion of the Dec. 2021 paper by Giovanni Pezzulo, Thomas Parr and Karl Friston "The evolution of brain architectures for predictive coding and active inference"
https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2020.0531

Presented by Active Inference Institute in 2022

Session 038.0, February 10, 2022

https://www.youtube.com/watch?v=cOo4juE_zcI

This video is an introduction for some of the ideas in the paper.

SESSION SPEAKER

Daniel Ari Friedman

CONTENTS

TRANSCRIPT

00:29 DANIEL FRIEDMAN:
Hello and welcome everyone. It's ActInf Lab Livestream number 38.0, February 10, 2022. We're going to be discussing the paper "The Evolution of Brain Architectures for Predictive Coding and Active Inference."

Welcome to the Active Inference Lab. We are a participatory online lab that is communicating, learning and practicing applied Active Inference. You can find us at some of the links here on the slide. This is a recorded and archived livestream, so please provide us with feedback so we can improve on our work. All backgrounds and perspectives are welcome here and we'll be following good video etiquette for live streams.

01:09 It's going to be a solo stream though. Go to ActiveInference.org if you want to learn more about how to participate or contribute or get involved with any ActInf Lab project and check out this code, a link to see past and encoding live streams. The page looks like this so you can see events that haven't happened yet, like 39, 40, and then also you can look back and you can see who is participating and read the papers and all of that. So check it out. Today in active stream number 38. The goal is to learn and discuss this cool paper, "The Evolution of Brain Architecture for Predictive Coding and Active Inference," a paper by Giovanni Pezzulo, Thomas Parr and Karl Friston from December 2021. And just like all videos, it's just an introduction to some of the ideas, it's not a review or a final word. So go check out the paper to learn more. And there's going to be an overview with first names and claims, abstract and roadmap. Alright?

02:13 I'm Daniel, I'm a researcher in California. The big question that this paper is getting at is what is the evolutionary neurophysiological basis of cognition; and how do complex cognitive phenotypes arise? So how do things develop and evolve, how they think and how does that change over evolutionary time? And shown here are three images representing three scales of analysis of looking at ant cognition. So on the left is a representation of the synapse with the glia wrapped around it and the molecules and some of the mechanisms. Because changes in those mechanisms can influence cognition.

02:56 Then in the middle is a 3D representation of an ant brain with the different brain regions, like the central complex and the optic and the olfactory lobes. And this represents the level of regional or micro or meso anatomical variation. And that definitely changes over evolutionary time, just like the synaptic level. And then there's this behavior ecological level; and that's where the ants area engaging in reflective behavior and stigmergy. And so how does this all work?

03:30 How does this all work in today's ants and how has it evolved and then expand that to other species and other questions?

03:40 The paper was published right at the end of 2021 in December in the Royal Society of Publishing. And just to go over the aims and claims of the paper, this is in the authors words: "There's growing consensus that the brains of humans in other phylogenetically derived or advanced organisms operate in a prediction manner across action prediction coding and action control Active Inference. Yet the ways in which our advanced prediction abilities may have arisen during evolution domain unclear. The goal of this article is to sketch an evolutionary history of brain architecture's for predictive processing. A central tenet of our proposal is that although prediction is often characterized as a complex cognitive function, it is not a late evolution addition of advanced animals like us."

04:35 Rather, in distinctions to a late stage cognitive argument [like saying language is what makes us an advanced cognizer or semantic language with certain types of syntax]; rather, our complex predictive abilities, eg. planning and imagination, emerged gradually e. g. via phyletic gradualism (smooth changes to evolution time), or punctuated equilibrium (sharp changes through evolution time). But punctuated at one scale is smooth at another from simpler predictive and errors correction loops.

05:10 E.g. motor and autonomic reflexes that were already part of the brains of our earlier evolutionary ancestors and were key to solving adaptive regulation problems. So, just like Mike Levin's paper was addressing the question of basal cognition from the bioelectric perspective, here is going to be more of a predictive processing and action inference perspective on the functional aspects, not on the mechanistic. So the bioelectric was down here at the level of cells. This is going to be approaching it from a little bit of a different perspective, but we'll find out. Here's the Abstract: This article considers the evolution of brain architecture for predictive processes. We argue that brain mechanisms for predictive perception and action are not late evolutionary additions of advanced creatures like us. Rather, they emerge gradually from simpler predictive loops for example, autonomic and motor reflexes that were a legacy from our earlier evolutionary ancestors and were key to solving their fundamental problems of adaptive regulation.

06:16 We characterize simpler to more complex brains formally in terms of generative model that include predictive loops of increasing hierarchical breadth and depth. These may start from a simple homeostatic motif and be elaborated during evolution in four main ways. These include the multimodal expansion of predictive control into an allostatic loop; its duplication to form multiple sensory motor loops that expand an animal's behavior repertoire; and the gradual endowment of general generative model model with hierarchical depth to deal with aspects of the world to unfold at different spatial scales; and temporal depth, to select which plant select plans in a future oriented manner. In turn, these elaborations underwrite the solution to biological regulation problems faced by increasingly sophisticated animals. Our proposal aligns neuroscientific theorizing about predictive processing with evolutionary and comparative data on brain architectures in different animal species.

07:18 And just looking ahead, here's a figure that we're going to get to. Here's the ancestral state. It has this structure to model and then it's going to undergo a set of different types of discrete operats that change its structure; and that's structure learning. And it's going to happen over evolution time scale and it's going to be tied to functional architectures for predictive processing. Okay, how do they go from here to there? This is the roadmap. Ater the introduction, they introduce predictive regulation and control; perception, cognition, and control action as basic design principles of the brain.

07:53 So kind of taking that embodied approach but making it very operational and functional so that it can be studied from brain evolution function perspective. Introducing the brain as doing structure learning in generative models over evolutionary and also other time scales. They then give three examples of simple predictive motifs in ancestral brains which is the homeostatic control, the allostasis control, and the simple behavior learning. Then they introduce that figure that we just looked at, and that's the evolutionary algebra of structure learning. Just like you can multiply and add, these are kind of like operations on evolutionary spaces.

08:36 They then discuss a few finer points related to behavior switching, temporal depth, hierarchical depth. And then take a phylogenetic perspective at the end giving an example. And there's a discussion. Okay, so to go into section two and just sort of deal with the keywords and themes as they are needed. Here's figure one. In figure one the reason why we can even jump in here without going to any keywords is it's biology we're talking about. And we can jump in. Why not? - as good of a place as any to go in at the action perception loop and then connect bt to some of the analytical or mathematical formalisms of Active Inference and the free energy principle. So this is figure one in the paper in section two, the action Perception cycle and Predictive regulation. So here's our entity, our agent on the left and here is our world state on the right.

09:39 The entity is engaged in prediction while they're making observations that are being emitted from the world. That's resulting in some discrepancy. Either things are exactly as expected or not. So an example would be in the visual field. The brain is generating a prediction of what is in the blind spot of the retina. And then if the eyes were to move there to use action changing the world in terms of the stimuli coming in through ocular motor action, that would result in a different perception that could either confirm or deny - confirm with a low discrepancy, or be very surprising with a high discrepancy, what was expected about what was in the blind spot, which would confirm accuracy in a visual model.

10:26 And so in this partitioning of action and perception which is just very descriptive, it's not quite the Bayesian graph that we're going to get to later. It's kind of like a flow model and there's probably other flow models that could be used as well. But it turns out that this partitioning or this way of thinking about flow at least conceptually leads to (in the Active Inference proposal) this idea of using a free energy minimizing function over some math that we'll get to a little bit more formally in the next figure, And using a kind of combined metric that has two parts, the red and the blue,

11:03 to make decisions about perception as well as action. Because it turns out that perception and action and cognition and metacognition are all part of the entity's model that it's doing inference on (in certain cases). So just to kind of throw back to not so long ago, here we have the F of Q, that's the distribution that's under the entity's control; and Y. And so as a function of beliefs and data, there's going to be some term. And so just looking back to [Livestream] 37, we looked at the variational free energy and how that relates to perceptual inference, where there's a penalty for overfitting as well as a penalty for failing to explain the data.

11:51 So it's kind of making a visual model or a perceptual model, that in that snapshot, given the priors and precision and all of that, is not overfitting, but it is fitting the data. And it's kind of existing on that frontier. And then it's using variational inference to solve that in a reality tractable way. And then when action comes into play, a few things happen. First the agent has to incorporate theory own preferences - because why care about action if you don't even care why it's going to happen? So they have to incorporate their preferences, which is a non-arbitrary (in a sense) for action selection; but it's arbitrary in a higher level. As well as incorporating the fact that there's uncertainty over the consequences of action or just future states of the world, not just like sensor measurement as in other cases. So we have to take this variational free energy calculation that was just like snapshot perception and expand it a little bit to the expected free energy.

12:48 So here's F in the background. And now there's this expected free energy term G, which is over also an action selection policy z. And now there's kind of similar, like resonating or rhyming terms. But rather than overfitting, the imperative on the left side is to satisfy preferences. On the right side, the penalty for failing to explain the data is kind of transposed into this failing to minimize expected surprise of future data. So this is like fitting the expectations well on the right side in blue; and then living up to your preferences and expectations in an optimistic way on the left.

13:31 So it's kind of like realism on the right and optimism on the left. And that is what we talked about in 37.And that's the partitioning that's being done basically here. The authors are setting that up as the action perception cycle and predictive regulation. Just wanted to kind of view 37 really quick because it was a fun discussion that we had. It also really sets the stage for,

13:55 How is that similar or different than other action perception partitionings or models? Does evolutionary psychology or evolutionary cognitive studies, do they have a fundamental action perception model at the root? Is that a good thing? Is it a bad thing?

14:18 Section three goes into Section two again was just about how this single slide and represented in figure one about this predictive (so, anticipatory, but also embedded etc.) infinity loop cycle is the basic principle of the brain. We can't take the basic principle of the brain to be some lower level like just information transmission among cells; nor do the authors jump in at a higher level, like "the fundamental unit of cognition is linguistic tokens that are being modified," not discrepancies with multiple different kinds of things that are being predicted. From this functional description of cognition, they move to Section three, Formalizing brain design as structure learning in generative models. So, what is the structure of this model; and then, what does it look like to do structure learning in that model? And why is it generative? And then, how is that formalized? So here's Figure two, the generative model and the generative process. So the first word's the same; second word is different.

15:28 So they're different words. And the figure on the left side has the entity. The figure on the right side has the world state. So it's the same action perception loop we saw in Figure one. And now this (sort of) conceptual flow single edge model (like just only one arrow here, no extra anything, just sort of first pass). It's compatible with this, which is actually a Bayesian graph.

15:54 But how do they describe it and what are all the variables? We still have the same things happening. We have the observations coming in to the cognition entity. That's the observations coming in. The entity is going to infer some action policy based upon the observations coming in, which is going to result in some change to the actual underlying system, which is the generative process. So that's like the actual birds and the bees and the sun and stuff,

16:25 allegedly. It does get into a little bit of a gray area with the realism/instrumentalism and the structural realism. But we're not even going to go there in this discussion right now. The generative process is the one that's handing out the observations as modeled. The generative model (to close the loop) is the entity's inference. And so here is X, the entity's prediction on hidden state.

16:49 And then here is X star which is (like) the actual hidden state that is being alleged in the world. And we've had some other discussions about how that's the sigma function. that's (like) mapping between the two X's. That's what's being minimized. If the discrepancy is low, there's other notation.

17:10 How do the authors describe it? The difference between the generative model and the generative process. Nodes correspond to probability distributions and edges to their statistical dependencies. So this is like a Bayes graph. Mathematically, a generative model may be formulated as the joint probability density, P of y and x - of observations y and hidden states x - of the world to generate those observations. (I think it was just a copy error.)

17:41 The latter are referred to as "hidden" or "latent states," as they cannot be observed directly. The joint probability distribution can be decomposed in two parts. The first is a prior P of x, which denotes the organism's knowledge about hidden states of the world prior to seeing sensory data. The second is the likelihood P of y given x, which denotes the organism's knowledge of how observations are generated from states. So that's the perceptual model.

18:07 And then they go on to describe how there's a difference between the entity's inference on hidden state and the actual hidden state, which is the generative process versus the generative model distinction. And then they introduce Action; and say Action,u - that's this node that influences the hidden state, even if zero effect is generated based upon the inferences made under a generative model. {Action is shown here as part of the generative model. Sorry.} Action is shown here as part of the generative process, making changes to the world despite being selected from the inference drawn under the model.

18:46 So action is actually making influence, even, again, the edge could be zero in some respect; but it's making Active Inference in the world. It's like the active states interpreted in a statistical way. So what does that have to do with structured learning? So the entity is going to either, whether you're a realist and saying the entity is doing structure learning; or you're instrumentalist - it is possible for us as researchers today to model that entity as doing structure learning because it's computational efficient or elucidative; or you go full utilitarian, you just say, "Disregard that whole Realism/Instrumentalism: it's a useful approach - and I'll follow utility wherever it goes!" For any number of those reasons, you might want to model the cognition of different entities without going into just the philosophy of what its cognitive process actually is. And so one approach that's going to get taken is using inference - either from the outside, describing instrumentally; or realism, as if it were happening maybe with anatomical evidence, as if the hidden state could include not just parameters that were continuous about the world, but also structures of models.

20:06 However, it's difficult to imagine that that type of cognitive or even extremely metacognitive thought or action selection could happen, for example, in some early proto cell, however simple it may have been. And so, how do we get from that flagella changing bacterium to all the other kinds of cognition that we see today? Or should I say, bacterium-like entity, relative or ancestor of today's bacterium?

20:40 So how can we think about this model, which is often described in the context of parameter learning? - And then approach this as if it were maybe about parameter learning sometimes; but also it could be about structure in terms of the good regulator and the requisite diversity, that kind of requisite variety, those kinds of models. Okay, the next several sections are where they get to the specifics and some of the contributions of the paper that I think will be really cool to continue the discussion on. Section four is just short. And it's saying we're about to go into three examples of simple predictive motifs and ancestral brains. Because one of the main claims of the paper is that these motifs are very ancestral. They're old motifs, they're not Johnny come lately to the cognitive scene.

21:29 These are features that one can think of as - who knows how far back or how simple these cognitive mechanisms have existed? - But we'll evaluate that maybe when we get to talk together. But first we'll just kind of go through how they define them and use them. The three predictive motifs are homeostasis, allostasis, and simple behavioral control. So first: Five, Generative models for the homeostatic control of interoceptive variables. They write. "The generative Model models shown in Figure three (which we'll look at after this slide) afford the homeostatic regulation of a single interoceptive variable, which we call here 'body temperature' for illustrative purposes."

22:14 Much like a thermostat, this model maintains the requisite body temperature by reporting the discrepancy between predicted and sensed thermoreceptor activation given Bayesian beliefs about temperature triggering an autonomic reflex, u, resulting in, for example, vasodilation, which resolves the prediction error. So if the life of the organism were just to hang out on the beach and vasodilate to off heat when it needed to and then to constrict and to save more heat when it needed to, that's the physiological task that this is going to be describing, which is just one facet of an organism's biology. But there are experiments that sometimes only measure temperature. And so thinking instrumentally, this single factor model, this single variable model on body temperature, may be sufficient for some experiments, or it may be useful in certain cases. So just because it's a simple model doesn't mean that it's not going to be very educational and provocative, but also even be sufficient in a lot of cases.

23:18 But no one's even claiming it's realism. That's why it's written this way. They say "see citation 20 for a fully specified example. And that is a citation to Tschanz et al. in March 2022 [[Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using Active Inference]]. (So still in the future!) And they write, "We start from the present premise..." (and this is in the paper that, again, is from the future) - "We start from the premise that the goal of interoceptive control is to minimize discrepancy between expected and actual interoceptive sensations I-E-A prediction error or free energy. Importantly, living organisms can active this goal by using various forms of interoceptive control, "homeostatic, allostasis, and goal directed."

24:05 So there's more details in this paper, but here in figure three is where they're going to show it. So keep in mind this generative model structure and now these are going to be in a different form. And here in the caption I'll describe what they say. This is the homeostatic, the first, most ancestral, or just the simplest possible. Just go make it darker if it's too bright and make it brighter if it's too dark.

24:41 Make it warmer if it's too cold. Make it colder if it's too warm. That kind of first order cybernetic loop. This generative model includes an interoceptive, thermoreceptor Y observation and a belief about body temperature x. So that's the beliefs about how the body should be and that's again the beliefs playing that dual function that the paper 37 drew out, which is that on the left side of this equation, failure to satisfy the preferences is dealing with this P distribution as a preference.

25:20 But then on the right side, p has to do with expectations that are being either fit well or poorly. And so this is where Active Inference has a slightly different architecture perhaps than some other theories. The beliefs are about body temperature. It's not an estimate merely of the external body temperature. Crucially, the prior over x is kept fixed and hence bit acts as a cybernetic set point.

25:47 Well, you can't just expect what's going to be best for you. You'll die, right? If you die, you die. But if you enact policy such that your expectations are realize, then you persist. That's why we're studying things that are persistent.

26:03 Any discrepancy between the predicted thermoreceptor activity given beliefs about X and the measured Y is registered as a prediction error that is canceled out by an autonomic response. For example, a thermoregulatory response. This is shown as an illustrative plot of the expectation of prior and posterior observation and autonomic actions over time. So here is like the action policy which is like be at the baseline level of thermoregulation and then kick in some sweating or cooling mechanism. And then here it describes how the observations start at about 37 one and then they steadily start climbing.

26:42 And then the belief which is initially like things should be 37 the posterior, the after evidence estimate starts creeping up and then it hits a certain value and it engages a critical threshold that turns on this thermal regulatory response and then that cools the temperature back down. So this is a basic architecture for doing first order cybernetics and that kind of first order logic. Here in this figure, the red circles represent the expected values of X, which are used to make predictions about Y. These are subtracted red arrow with the rounded end. So this one from the measured Y to form a prediction errors.

27:25 Dark blue circle epsilon which is used to update the expectation and drive action. Light blue circle you here's you that changes why such that the prediction error is resolved. What if it doesn't do it? Well, then the system dies. So we're talking about evolution where we've had like for the ants 120,000,000 years allegedly for that to get Pruned out and even longer at the cellular level.

27:53 Note the lateral modulatory connections in the allostasis network which we'll get to in a second. So just to take one little discourse, they say C 24 for details. What is PAP 24? It is Friston, Par and de Verise, 2017. The graphical brain belief propagation and Active Inference.

28:13 Let's just look at a few parts of this awesome paper. So first they have a table with definitions of the technical terms. So just to kind of read a few but it's kind of awesome to see the authors do this and this is in great paper as well. So how do they define generative model, generative Model or forward model? A probabilistic mapping from causes to observed consequences data.

28:39 So from hyper parameter to the parameter it is usually specified in terms of the likelihood of getting some data given their causes, parameters of the model and priors on the parameters. So it's all relative in nested models, but this is generating data like kind of cranking out like a music box possible or plausible data set with similar summary statistics, like similar mean invariance of some distinctions or similar parameters if there's a whole vector that describe it. And then the recognition model is related to learning where new data are coming in and discrepancy is being minimized. If the generative is outputting the exact same mean invariance that encoding data are having the discrepancy is low, the predictions which are about preferences are being realize successfully. Action policy is working well or better than expected, flip everything and you have the opposite situation.

29:39 And then just to give one more definition here, because the next slide will feature it, so the others are also good to read. Factor Graph factor graph is a bipartite graph where two distinct sets of nodes are connected by edges representations. The factorization of a function, usually a probability distribution function. Formulating a Bayesian network or model as a factor graph enables the efficient computation of marginal distributions through the sum product algorithm.

30:16 What does a factor graph look like and how does it relate to the kinds of Bayesian graphs that we've been looking at? So on the top is not the Bayes graph that's distributed across this slide, but another variant that we've seen a bunch of times, which is the partially observable Markov decision process. So G expected free energy minimization pi policy selection is influencing B, which is how s the latent state in the world is changing through time. There's d the prior on the hidden state and then a the mapping of how the state is related to the observation. And so depending on how the model framed, those can be learned or not.

30:54 But it turns out that because of how this is relatively sparsely connected within a time frame as well as across time frames, there's a way to use this bipartite construction called a factor graph that splits up those unlabelled edges which are statistical dependencies and kind of interweaves functions which have a slightly different representation. And it turns out that by interleaving these functions into the variables, it's possible to make what's called a factor graph and that gives an order of operations to arbitrary or within a certain set any kind of Bayesian graph. But it includes this one importantly. And so here is the one, two, three time points and two policies area being selected and that's what this graph represents. The organism comes in with a prior time step one, two, three, there's two actions.

31:50 And then here's another figure from the paper where at each of those three time steps, one, two and three little figure at time steps one, two and three. So at times step one, that's like anticipation and planning. At timestep two it's like short term anticipation as well as memory. At timestep three it's like memory and it's always now casting as well. And so one can imagine that this is a really useful format because it's extremely composable on one hand.

32:29 So just like they said, okay, well, we kind of have this motif with three time steps and theory connected us. What if D from the top level came down and was S at a lower level? And we've seen that taken to a really elaborated extent as well as interpretation in for example, the paper on mental action in Live stream 25. And so factor graphs are awesome because they're basically needing to only be specified in the Bayes graph format. But then it provides not just a mesh connectivity but a process algorithm and a heuristic and approach that's actually tractable.

33:09 So we get the composable analytical and graphical component that's an intractable algorithm.

33:17 The next section is generative model for the allostatic control of interoceptive variables. So this is going to be the first real modification of the homeostat that's introduced in three. This is going to be the base case, but it could be something else other than body temperature. The homeostat is simple but limited as they write. It can counter sensed changes of body temperature but cannot anticipate predictable changes of body temperature or other variables in nature.

33:52 There are several regularities eg night, day or seasonal alternation that can be easily incorporated to extend the above generative model as technically speaking, empirical priors the obvious advantage of prediction how our bodily and interoceptive variable will change is being able to exert some anticipatory or allostatic control. And so this is kind of getting into the second order or anticipatory cybernetics also related to Rosen's anticipatory biology. So here's figure three C in A and B there was just the homeostat returning us to a set point after something got triggered and now there's going to be the affordance for anticipatory control. This generative model sketched out with the same scheme as the previous slide, this generative model extends the homeostat by including. A second set of exteroception variables that correspond to light intensity y two and a belief about sunrise x two.

34:57 That's the sun visual side on the right. Furthermore, like the visual system and the left side is still the temperature and terraceptive system. Furthermore, the model includes a predictive relationship between sunrise x two and body temperature y. This edge isn't saying that the sun warms the body. It's saying that in this model there's an edge reflecting a statistical dependency and that's where there's a degree of freedom with respect to the blism and instrumentalism etc.

35:25 In this way inferring A, sunrise can trigger the autonomic response U of thermal regulation in an anticipatory manner, that is, before the sunlight actually increases body temperature. The order part of A and C are Bayesian networks highlighting that Y is conditionally dependent upon x with the directed arrow between the notes with more than one x and Y. In the model for the allostap, the lower parts show the form of neuronal message passing that could be used to solve these generative models. So the Bayes graph is represented on the top and then there's the message passing with respect to the neural correlation. So that's kind of the second aspect of figures three, which is just bringing in multisensory integration or even it could be like two pixels, for example, with beliefs about each other or something like that.

36:16 But that's what allostasis is going to be enabled by is just by this duplication of a column and then this connection in a different way. And then here's the third section of four seven generative model model for simple behavior control. And so they write the homeostat and the allostasis permit the control of simple forms of swimming, flow, motion reaching and other movements. One biological example is provided by the Zebrafish Virtual Reality Study 30 which identified the neuronal underpinnings of error correction during escape behavior in the animal's telecephalon. It's a brain region, brain evolution conserved set of brain circuits involved in action selection in other vertebrates, including mammals, such as the cortico basal ganglia circuit.

37:05 So that's about the evolutionary biology of the brain region. And then here's just some pictures from the paper by Tori Go. At all 21 zebrafish capable of generating future state prediction errors show improved active avoidance behavior in virtual reality. So they did a learning task that involved the fish being able to differentiate a signal and then they studied the role of anticipation in that. And the authors in this paper use that as an example.

37:38 Maybe we could talk about that or other examples in the dot one and the dot two two, section Eight. Here's where we get to the very interesting operations that are going to bring this sort of descriptive model of different kinds of homeostatic allostasis and intermodal and then behavioral regulatory elements into the evolutionary context. So our central argument is that evolution proceeded via gradual elaborations of the predictive motifs illustrated above. Under genetic constraints and opportunities and model selection pressure of novel problems to be solved, such as the control of more sophisticated bodies in the presence of richer ecological niches e. G when Vertebrates began to establish life on land some 400 million years ago.

38:33 Over successive generations, generative Model can remain stable or be elaborated along four key dimensions strongly limiting the space of what is evolvable. So that are the four kinds of dimensions that are going to be changed, that is going to be discussed in terms of the changes that can happen to the specifics of the generative Model. We have introduced the first kind of elaboration from the unimodal homeostat to the multimodal allostasis. So they kind of secretly introduced this transformation between figure three A, B and figure three C. So that was secretly like one of the transformations.

39:18 A second kind of elaboration is the duplication of predictive motifs which enlarges the animal's behavior repertoire. The third and fourth dimensions equip. The generative Model with temporal and or hierarchical depth respectively. These two expansions enable richer predictive motifs that endow a cognitive sophistication, such as the possibility to plan or consider events that change on multiple timescales. So it's the evolutionary algebra on structure learning because we're outputting a structure, this graph G, which is going to be like as if the species over evolutionary time is going to be implementing some graph in terms of the structure of its model.

40:05 Like if there's a case where the agent is not integrating the polarization of light with the olfactory system and then there's some change in the model that actually integrates them and then some relationship is learning whatever that means from a realist or instrumentalist perspective. And that is going to be like an evolutionary algebra. So it's not going to be like two x minus three x but it's going to be more like that than not because there's going to be operations and they're going to happen in order. So here's figure four where they represent their evolutionary algebra. Figure four, the five main dimensions of elaboration of generative Model model introduced in the paper.

40:43 So it was four dimensions. Then it's five dimensions. There's evolution in 4D by Chablanca and Lamb that would phase been good to add to the five main dimensions of elaboration of generative Model model introduced in the paper, illustrated as operations of an evolutionary algebra. So here's the five operations and so we're starting with on the left side that homeostat, that simple corrective calibrative first order cybernetics model either what the system is actually doing or model of. Then there's going to be five information that can happen and then it's showing there's a second round like once you go H, you can go HTA I plus I or I.

41:30 So you have five discrete options at the first time step, but one of them is no change. So it's kind of like no change or four different layers of excitement of the selection like four quanta, but they're for discrete operations like a deletion or an insertion in genomics. And then from there it's just the state for the next time step of the model. And then something else happens. So what are the operations?

41:53 The bottom is I, which is the identity operations that levels the generative model as is. So that can be interpreted as like a non mutation or just a conservative mode, which is how most inheritance works. Then the second one is the duplication operation I plus I. So it's like identity remains the same, but then there's a duplication. Literally it's like a genomic duplication but in this functional space replicates existing predictive motifs to form parallel sensory motor loops.

42:25 A is the operation that was described in figure three. C, the allostasis operation endowed the generative model with horizontal predictive relations between different modalities. And so the implication would be like going from one sensella, one antenna to two antenna or going from one photoreceptor to two photoreceptors. But the actual architecture of the column of the photosensory transduction cascade would basically be computationally or statistically unchanged. And then the allostasis is actually bringing in this horizontal aspect.

43:00 It's not just two duplicated systems next to each other. Now there's actually connections between them and of the possible kinds of connections across columns. One of them is like this classic allostatic motif. Then there's the ones that we haven't gone into as much, which are T. The temporal depth operation extends the generative model with separate variables for past, present and future states.

43:24 So it's kind of from a graphical perspective what we looked at in a difference between the first factor graph which did take action at three time steps three timesteps like the thermostat does to the one that actually has either prospectively looking anticipation about future time steps or retrospectively looking memory. But that's how the factor graph comes into play, that's temporal depth. Then the hierarchical depth operation H extends the generative model with separate variable for states of affairs that change at different timescales faster time scales at the bottom levels and slower time scales at the higher levels. Hence modeling narratives such as music and language where nested timescales are relevant. So to kind of split that idea of temporal depth into two pieces, there's incrementing the number of steps you're looking in the model that's increasing the time horizon on policy selection and increasing the temporal depth within a level.

44:26 And then there's this notion of nesting levels within each other. That's the nested generative model and therefore nested Markov blanket discussion that we've been having and that is going to be connected to cognition activities like narrative. And the reason why they're very similar is that the time scale can kind of blur into each other. And so it's all about the model structure as stated. Like this one is hierarchical and it has a depth of three.

44:59 It's a two layer model and it has a depth of three, three time steps are included. And it could be different if there was always looking two ahead and always looking to in the back. Then the model in the computer would need like a minimum of five time steps. But the entity's model could still be restricted to S minus two, s minus one and then S plus one, s plus two. So those are the two ways that it can expand in these two temporal and hierarchical ways, which is to nest hierarchical model to become temporal depth with a longer horizon given the nesting structure.

45:42 So these are all structural changes. That's why there was the whole piece about structure learning because it's as if or actually like over evolution time there's the structure learning happen. And then if we use this partitioning and Bayes graph approach, then hypothetically any kind of evolutionary starting point if we go back far enough and then final state if we have all the transitions, could be modelers within like a native Active Inference framework selection nine. They're going to go into a little more detail about duplicating prediction motifs and enabling multiple behavior. So they write how does this duplication of the model looking at it from the outside it's like as if they're acting as if there's two model looking at it from the realism in the inside.

46:30 It's kind of like thinking about the real duplication of a cognition function that's functionalism. Or Mike Levin a neuroanatomical region like the earlier examples with the retinal cells and that's like canonical realism. So generative model can expand by duplicating simple predictive motifs to form a larger repertoire of species specific behavior such as approach avoidance, the control of the vibrace and visually guided grasping classic. The operator I plus I in figure four illustrates a generative model in which the same predictive motifs are duplicated and specialized to form a behavior based architecture composed of multiple parallel sensory motor loops. So they're suggesting that because these are your affordances your operations in your evolutionary algebra, you can go from this starting point.

47:18 It's kind of like go to leisure balk starting point and then doing operations to it. Because you have the starting point and the operations to it, it allows you to get to even relatively advanced motifs like approach avoidance, etc. But a key piece is duplication because implication of something without changing it is how you are able to build more land to experiment in, so to speak, build more space. And I copied some images from genetics specifically in the relationship of how gene duplication and divergence in the early evolution of vertebrates this paper. And there's a huge amount of cybernetics and genomics works on the duplication and divergence and the neofunctionalization, the subfunctionalization.

48:11 Because if you have like an enzyme or essential gene a now to go into the whole gene thing totally another time though, you could have the function of the second copy in the genome be lost and then theorem is still like a continuous line of function. So if you only needed one copy of a then this would be sufficient. And then other times when you have a and it's value function like it binds to two different, not exactly similar molecules, then when there's a paralogy, when there's this duplication, it allows subfunctionalization or new functions to arise. So that's how people talk about it and link it to realism in genomics. And this is kind of approaching that from a cognitive perspective.

48:55 There's probably more to say, but we'll talk more about the duplicating of prediction motifs. So how is duplicating predictive motifs enabling of multiple behavior? Okay, they write from a structure learning perspective, duplication is an efficient way of building generative model models. And that's what it's all about in the sense that the dynamics are conserved over different sensory motor domain. This conservation is mathematically akin to factorizing probability distributions on the generative model that has been discussed in terms of modular architectures and functional segregation as a principle of functional brain architecture in Bayesian Statistics physics, this kind of factorization is ubiquitous and known as a mean field approximation.

49:41 Indeed, the free energy bound on model evidence is defined in terms of a mean field approximation that affords an accuracy and minimally complex explanation for sensory data. And so what are some of these citations? 44 modular Architectures for factorization of possibility distinctions in the generative Model par Majid, Karl Friston 2020 entropy so here is kind of a cool figure, nice graph and then there's the message processing and then citation. 48. The mean field approximation, what is it?

50:20 Here's a paper from 2001 and they wrote algorithms that must deal with complicated global functions of many variable often exploit the manner in which the given functions factor as a product of local interactions, each of which depends on a subset of the variables. Such a factorization can be visualized with a bipartite graph that we call a factor graph. A wide variety of algorithms developed in artificial intelligence, signal processing and digital communications can be derived as specific instances of the sum product algorithm, including the forward backward algorithm, the Viterbi algorithm, the Iterative turbo decoding algorithm, pearl's 1988 belief propagation algorithm for Bayesian Networks hashtag Markov blanket the common filter and certain fast fouryear transforms FFT algorithms. So it was 21 years ago when this was happening, and now we're here. Model Selection ten endowing generative models with temporal depth supports perspective and retrospective inference.

51:34 So just like we looked at with that factor graph, giving this operation over evolutionary time enables that factor graph to arise from something with the lower time horizon. The generative model models discussed so far only consider present states and observations. However, they can be expanded into temporal depth models whose variables explicitly represent future and past states and observations. So this is what the operation looks like. It takes XT and then at XT plus one or tau depending on how it's written.

52:08 And then now there's another time step appended to the end of this model, either actually or as if here's something cool that they wrote. They wrote various researchers have speculated that a major driving force for the development of deep temporal models was foraging. So why would this happen functionally? Which is to say, why does the mutational spectra, which does allow for this as an affordance end up selecting four and retaining and enriching force temporal depth models? Otherwise we wouldn't observe it to exist and they're connecting that to foraging.

52:45 Intriguingly, this is a vertebrate example. The same hippocampal circuits that support spatial navigation and foraging are also involved in perspective and imagination. This has led Busaki and Mosser to propose that objective function have leveraged cognitive and predictive maps in the hippocampal entorhinal system and hence mechanisms of memory and planning have evolved from mechanisms of navigation in the physical world. So what are the cognitive demands of foraging? How about information foraging?

53:16 How about mental foraging? Here s some awesome papers by Hills and Cuisine and others foraging in mind and foraging in semantic fields. How we search through memory. So what about mental foraging? What about individual and collective foraging?

53:35 This is an awesome paper by Feynman and Corman in 2017 and they talk about the continuum and the complementarity of individual and collective approaches to recognition model. So implicitly, like foraging as a phenomena some of the affordances and the neurophysiology of foraging in ants. It's the same materials and mechanism that any other insect that's not eusocial has. So the detection of light, the intensity, the wavelength, the polarization, sometimes the ability to do chemo sensation like taste and smell, mechanical reception, etc. And the same action affordances too like movement.

54:11 And so there definitely is nest Mike Levin cognition in ants. But also there's things that are of a few different interesting types. One of them is meso scale like small group dynamics, stochastic teams and larger scale like colony and even colony niche stigma g and ecological scale cognitive processes. Like these two ants are interacting and modifying each other's foraging behavior, mechanistically and statistically. But also it wouldn't happen unless the niche were exactly this way, which they have also in their extended selves established for themselves.

54:48 So how do we think about individual and collective foraging and stigma g in complex systems and mental foraging and cognitive demands and cognition security? What about section eleven endowing generative models with hierarchical depth affords multiscale inference. So now we get to the hierarchical operation that is going to give that multiscale inference. So far we have described generative models that can deal with aspects of the world that unfold at single timescale. So plus one plus one plus one timescale is temporal depth.

55:21 But you're getting still only one extra per transformation. However, they can be expanded into hierarchical model models whose variables at different hierarchical levels encode latent states that unfold at different timescales. One example is a song. Melody remains the same even thought the notes we hear or sing change rapidly and speech similarly, a movie or narrative remains the same for several minutes. Scenes remain the same for several seconds.

55:49 But visual stimuli can change over hundreds of milliseconds. Such models permit hierarchical models permit modeling of narratives, songs, movies and other events that change at different temporal scale by encoding variables that change more slowly. Eg melodies or movies at higher hierarchical levels and variable that change more rapidly. Egypt notes or visual scenes at lower hierarchical leave to neurobiological examples of hierarchical organization are visual areas in mammals and areas that control vocal gestures in birdsong, which has been studied Active Inference Lab several times. And so here's another quote from the authors in more advanced animals, the hierarchical control of action may have expanded into sophisticated forms of cognitive control and objective function.

56:45 Layer one scare quotes which help prioritize digital goals while inhibiting immediate affordance. So it's not just about seeing deeper within a time scale, but it's about being able to pull up to a higher time scale. And then from there, after the H operation, it can be followed up with a T operation. So here's the minute scale and then there's a hierarchical implication that allows for the hour scale and then that can go into 2 hours and now two minutes. So now there's a two-hour and two minute long model instead of a 1 minute.

57:23 And it was just two questions. But if the two questions had been or the three mutations had been just going deeper within the minutes, bit would be a different outcome.

57:37 How is that functional? So what is the function and the cost of temporal depth just instrumentally? When we're studying diverse cognitive systems, how can we detect temporal depth and versus hierarchical nesting? Then what is the meaning and the role of narrative in cognition? How does this relate to narrative information management?

58:04 All right, section twelve. Getting towards the end in the above, which was again the description of the simple motifs in 4567 and then the evolutionary algebra in eight and then several of these finer scale discussions on nine and ten. In eleven, we then get to twelve. In the above, we realize brain designs in terms of generative models that include predictive loops of various complexity red and then discuss the five main ways in which generative model designs can be elaborated. Green or the five main operating point an algebra of evolutionary structure learning.

58:44 Figure four. This means that one can describe the evolutionary trajectory of brain designs in terms of a limited number of computational operations over generative models. blue so here is a phylogenetic tree on the right side with the tree of life. One of the tree of life. What area?

59:07 Alternative complexity or traditional ways to think about phylogenetic trees in evolutionary biology? Are phylogenetic trees interpreted instrumentally? Are they interpreted under a realism framework? Is that what really happened to those species? Or is it our model inference about what is the relationship between Active Inference in the free energy principle and evolution?

59:35 Okay, they have figure five which gets at their phylogenetic model. So this is a phylogenetic free of generative model designs and putative correspondences with animal brains. So here's the implied internal states and then I is going to be the identity operator. So here the orange species has not mutated at all. Now, sometimes this is conflated with simply being an outgroup.

1:00:02 Just because it is that way doesn't mean they're making the conflation. But sometimes people will make the conflation. That because a species is an out group to some other clade that has been included in the analysis that it is the basil or primitive form. And so it does happen to be that way in this example that the basal is the so called least derived or most primitive or basal form. But my personal thought is that it should not be described.

1:00:29 And Tim Linsker and others have awesome writing on that evolution fallacy. So in the rest of the tree, which is being focused on, different kinds of operations happen. So here's that I plus I implication and then there's no change after that. And then this one has a and so on. So just like you could trace the phenotype changing through time on a tree inferred from trait or genomic data, which is just another trait, this maps up to certain changes that are seen neuroanatomically over vertebrate evolution.

1:01:03 And it reminded me of this paper, which was Chakra Borzi and Jarvis 2015. And so that is the paper, brain evolution by brain pathway duplication. So they don't connect it to in the exact same way the neurocognitive and the functional and Active Inference Lab and all that. But this paper does get out some of the very similar ideas about the functional duplication arising as a result of pathway duplication.

1:01:33 They have a section on brain complexity and pathway evolution. They talk about some alternative hypotheses and then talk about distributed and duplicated morphological structures. So it's a kind of interesting paper from about seven years ago. Another paper that's very related to this idea of doing like an evolutionary algebra with combinatorics but also a path dependence is this paper pretty recently, just a couple of days ago by Ryan Smith, Maxwell Ramsden and Alex Kilner. The paper is why Bayesian brains perform poorly on explicit probabilistic reasoning problems.

1:02:08 So look at this tree that they have the starting point and then three action. So here it's like divide, divide, multiply, divide, add, divide. And then they study that in the context of Bayesian brain and doing calculations, why is it hard to multiply numbers together sometimes? So then the authors of the paper, 38, interestingly, the mutational operators are commutive. The same generative model design can be obtained by executing the same operations but in a different order.

1:02:42 The communication property of mutational operations potentially sheds light on the conversion, evolution and the process by which unrelated organisms evolve similar traits independently and via different evolutionary histories when they need to adapt to similar ecological niche. That's pretty cool. All right, so just the discussion and then a few last points. So discussion and the authors summarize it in this article. We suggest that brain structure or design could be formalized as generative models agree disagree, that the brain generative model models of our evolutionary ancestors included simple prediction motifs agree, disagree and that the evolution proceeded via successive elaborations of these predictive motifs into more brain architecture that we observe in advanced animals.

1:03:30 They then talked about the ways that that can change through time functionally. And then they write while the evolutionary trajectory of designs for predictive processing proposed here is certainly tentative and incomplete, we consider it a first step towards the alignment of predictive brains and evolutionary studies of neuroanatomy in different species. So if this is the first step, where are we headed and why do we prefer and expect ourselves to be there or go there? Just a few more topics that we could talk about, like in the dot one and then the dot two. First would be they write that the error correction mechanisms, in their view encompass the simple and the complex forms of adaptive behavior hashtag intuitive theory.

1:04:14 And they're going to argue that that differs significantly from prevalent perspectives in psychology and neuroscience, which tend to separate sets of mechanisms for sensory motor processing and simple cognition. So how is Active Inference similar and different to other frameworks for behavior? What are the building blocks of adaptive behavior? What are the basal blocks of just any kind of behavior? Where does sensory motor integration come into play?

1:04:43 How about mental functions and cognitive functions like memory, anticipation, counterfactuals, etc. Okay, another point to kind of think about or write your questions down and reflect on is the perspective which is still speculative and not unchallenged suggests that the complexity of the ecological niche determines the level of complexity that the brain needs to have in order to be Bayesian optimal. In other words, brains only increase their complexity with sufficient ecological demands. So not necessarily just that mutational direction and intensity will go towards increasing brain complexity from ecological demands. The socalled anticipatory evolution that cognitive entities can have through selfmonification and niche modification.

1:05:27 But even for those that aren't actually doing anticipation, still it could be the case that when the biological demands are such that a behavior model increase in complexity is selection for and retained, then evolution will go that way. This is because having a more complicated brain does not help if you live in a simple niche. So that's like a very costly model that's not giving you any more return on investment. They then talk about how the social brain hypothesis states that the necessity to predict and deal with sophisticated social dynamics was a main driver of the evolution of large brains and sophisticated cognitive abilities in our species people. In short, the gradualism expressed

as a progressive increase in complexity rests on the circular causality implicit in the modeling of an eco niche that is itself constituted and constructed by increasingly complicated phenotypes.

1:06:22 So because it's so important to think through other mind first just how you're going to materially avoid a spatial collision but then, so the social brain hypothesis goes includes increasingly recursive levels of game theory, market theory and all this kind of stuff and related indirect cognitive phenotypes required like memory and recognition, narrative understanding, rhetorical understanding, governance, et cetera. That is being position as compatible with the models the authors have written and a hypothesis that others have written about not from Active Inference Lab perspective though in most cases. So what is the cognition niche for social entities, our material niche and our social cognitive niche? What is the social brain? What is the use social brain?

1:07:09 How are they related? So for example, the social brain is saying that the more social things are, the more sophisticated the brain has to be. We need more narrative understanding and more memory. But what if in the Eusocial case the brain is simpler on board for examples, it's more role based or temporal polyethym has allowed more reduction in each phase of the life cycle brain. So does social lead to you social?

1:07:40 How does your sociality arise and how does it elaborate? How do different kinds of sociality arise and elaborate? And how is that associated with the different kinds of models structurally changing that we've been discussing in this paper? And then just genomics all the other stuff gene expression and then just one last point which is the closing piece of their paper too. Finally, it is important to acknowledge that brain design bodies ecological and cultural niches covolved independently.

1:08:11 Not sure if that was meant to be meant. It's kind of like they are codependent in brain evolution but they coevolved, they were alone together. Given that here we were interested in the evolution of brain designs, we assumed a brain scenario perspective and conveniently focused on generative model in the animal's brain, hashtag realism in the brain not modeled as in the brain. However, cognition does not need to be confined in this goal to be it can be extended outside it to COVID for example, tools and social dimensions, epistemic niche, niche modification, digital stigma. Furthermore, the body design and not just brain design plays an important role in solving control problems, acknowledging that cognition can be extended and embodied and encultured etc.

1:08:57 All the parse, all the other letters suggest that not all aspects of control need to be solved by or representation in a central generative model tail of two densities all the discussions we've been having about representation and so it was x and x star in the simple version that was presented in this paper where it was actually trying. To track X star in the world with the internal inference. But this is the discussion that we've been having action, action action oriented representation structure resemble the world. And there's other papers that we've discussed in the last several weeks that really touch on that point. So how do we think about generative model model and factor graphs for extended cognitive processes?

1:09:38 So a couple of books to check out about extended mind and embodied recognition model. Then here's a nice figure from touch points with the brain, the organs in the body, the world and the tools and the epistemic niche and the computer like a calculator and time chronos and food and then other people social. So that's like cognition is like a holistic interactions of all these features. What does embodiment this perspective have to do with realism and instrumentalism and utilitarianism and other philosophical positions? Are people really taking those positions or is it just as if they're taking those positions?

1:10:21 And then what else are you really curious about and motivated to explore? And how will you modify and improve your epistemic niche? So I hope you enjoyed this kind of Solo Zero video. I think it's been a little bit since the Solo Zeros, but just want to close, as always, with what might a good understanding enable? What are the unique predictions and implications?

1:10:48 What are the next steps for free energy principle, Active Inference, lab research and application? What are the goals of this research and what are you still curious about? We're going to be talking about this paper in the coming weeks, on February 16 and February 20 free. And so if you'd like to participate in those discussions through live chat or by joining the discussions live, just get in contact with us. Hope you read this paper because it's very thought, provocative and interesting.

1:11:21 So enjoy the paper and work through it. Hope to see you in other action flow activities. Thanks for listening and your regime of attention. Goodbye.

Session 0.1, February 16, 2022

https://www.youtube.com/watch?v=azwqMjgfY8Y

First participatory group discussion on the 2021 paper by Giovanni Pezzulo, Thomas Parr and Karl Friston, “The evolution of brain architectures for predictive coding and active inference.”

SPEAKERS

Daniel Friedman, Stephen Sillett, Dean Tickles

CONTENTS

TRANSCRIPT

00:23 DANIEL FRIEDMAN:
Alright. Hello and welcome everyone to ActInf Lab livestream number 38 Dot One. It's February 16, 2022. Welcome to ActInf Lab.
We are a participatory online lab that is communicating, learning and practicing applied active inference. You can find us at the links on this slide. This is a recorded and an archived livestream, so please provide us with feedback so that we can improve our work. All backgrounds and perspectives are welcome and we'll be following video etiquette for live stream. Just release an unending torrent of emojis if you have to speak or just raise your hand.

01:04 I'm sure that will get to it. If you're watching Live, please feel free to write questions in the live chat and we'll have enough time to hang out and discuss during this dot one where we'll be opening into the paper. Check ActiveInference.org for updated information on participating in any of the labs activities. I hope you'll find something that resonates with you today in ActInf Livestream number 38 one. We're going to be learning and discussing this cool paper

01:40 "The Evolution of Brain Architectures for Predictive Coding and Active Inference" Pezzulo, and. Friston from December 2021. And we're just going to enjoy discussing it and opening up any idea or questions that us here on the panel have, or those who are live chatting us. And we have some ideas and thoughts, prepared a few things that we know that we can go into. And also I hope that everyone has brought some other prepared seeds and also, of course, spontaneously feeling like new things are arising.

02:19 So we'll just start with some introduction and warmup. We can each say hi and then maybe it'd be cool to just also mention like, what got you excited about the paper or what image you want to discuss this paper, what's something that stayed with you? So I'm Daniel, I'm a researcher in California and I was very excited by the evolutionary focus. A lot of my research over the last years has been in evolution and ecology, so it's always awesome to see how people are thinking about how active inference, free energy principle and evolutionary studies can all be learning from one another. And I'll pass to Stephen.

02:59 STEPHEN SILLETT:
Hello, I'm Stephen Sillett. I'm in Toronto. I'm really interested in how this paper connects to my work with spatial meaning making and social topographies because this paper talks about more biological nascent stage of development rather than some higher order meaning that often people think about.

And I'm really interested in how this sort of grounded bottom up and active ecological approaches can be thought about as something where we can actually ground a lot of our learning making. So I'm inference in how this might match with some of that thinking and I will pass this over to Dean.

04:01 DEAN TICKLES:
Thanks, Stephen. I'm Dean, I'm in Calgary. What I found interesting about the paper is that given that

I've worked with a lot of young learners for a lot of my life. It was really interesting to see how this affirmed a lot of the thinking that I was doing when I was trying to get people past the idea that science is only about the biology and the chemistry and the physics that there is a statistical and predictive component to this. And those relationships in that realm can build certain sense of what the underlying architectures is. So, as I said, a lot of this is affirmations and it's kind of seen as being not part of what it's typically address core learning. But I think it should be back to you, Dan. Cool.

04:43 Daniel:
Nice intro. Let's just start with a big question and then either of you have any reflections on the big question? Or we'll go over to Blank Slide and just bring up some questions we have. And also it might be good to go over just some of the key points in the paper, like what did they actually do in terms of their contribution? So the big question, or at least one way to phrase something that might bring someone to approach this paper what is the evolutionary neurophysiological basis of cognition and how do complex cognitive phenotypes arise?

05:21 Like, you don't go from zero to colony in one time step. How does it happen that evolution arises from precursor, sometimes simpler, but also sometimes more complicated precursors? Stephen mentioned what is the basis for sense making? And cognition and sense making are very related. And what's been fun over these last weeks and months is we've explore basic or simple or reduced or basal cognition from a variety of perspectives.

05:56 Like, we talked about the bioelectric components of thinking about basal cognition with Mike Levin. Now we're thinking about a slightly different approach to understanding basal cognition niche we'll be discussing here. And then also we've looked at some of these more complex cognitive phenotypes like mental action, counterfactuals, deep temporal inference, all the active logic which we'll get into probably later. How can we use generative models of perception, cognition, action, and impact like active inference to study this whole continuum and diversity of relatively simpler in our perspective cognition, and also relatively more complex cognitive phenotypes and everything in between? So Klaas.

06:52 Stephen and then on with it. Thanks, Daniel. Yeah, I mean, I suppose also what some of the paradigm impacts of that, of taking that all the way through one question that might come up which you may or may not have an answer to or thoughts on. But Mike Levin's work with the ecological cones at the different nested levels and how this sort of could be thinking about what kind of action control potential these different nested levels might exist. I would be interested to think about how this may either be able to connect to that I can see it connecting philosophically or whether there's another bridge needed between this modeling and that type of modeling.

07:46 Stephen:
Just something I mentioned.

07:51 Daniel:

Great.

07:55 Dean anything to add or perhaps I'm. Going to wait till you get to figure one. Okay, so we won't go through all of the Kilner points in the paper. That's for the reading of the paper itself. Nor will we even go over all of the overview, which is what the dot zero video 38.0 is for.

08:19 So turn back time and watch that one if you haven't, or pause the video if you're not watching it live. But what this paper does, as evidenced by the Roadmap, is introduce a few basic principles of cognition, predictive regulation and control and then also structure learning in generative model models. That's not how every neurophysiology text is going to begin the building blocks of cognition. It's not how every evolutionary neurophysiology text is going to be framing cognition, but that's what they do here and that will play into how it's similar and different than other approaches. Then with those building blocks in hand or on the floor, they provide three examples of motifs that ancestral brains may be modeled as having or may have actually had.

09:13 With those three examples in hand, it's then possible for them to state a more general way of thinking about the transitions amongst different structures. These structures represent brain design as structured learning and generative models. That is called an evolutionary algebra. And they introduce five operators that can basically either leave unchanged or change the structure of brain design with respect to generative Model models. And each of those are explored in terms of what are brain architectures changes that that evolutionary transition is and then what are the functional consequences of that kind of an evolutionary transition.

10:01 Then they have some discussions about how sequences or patterns of application of this evolutionary algebra could lead to different evolutionary phenomena. So for example, increasing temporal depth in the future means that models are increasingly prospective. Increasing temporal depth in the past is like memory, etc. And then they close with mapping to a phylogenetic tree and thinking about that evolutionary algebra of state transitions being mapped onto the bifurcating tree structure that's called a phylogenetic tree which shows the relatedness of different life forms.

10:45 That's the roadmap. Let's go to the long awaited figure one and talk about the action perception cycle and predictive evolution. And so they're gain discussing this in the context of predictive regulation, anticipatory regulation, cybernetics and control as a basic design principle for the brain. So Dean, what do you see here or what would be cool to think about? So one of this follows I think is the usual pattern when we're trying to explain where we're going to go.

11:19 Dean:
And so you just had a roadmap up and I think a roadmap is a way for the people who are reading the Roadmap to find their way. But I think what this diagram shows is that in the middle of there, there's something called a discrepancy. And that discrepancy is later going to be given a label y. And what I think is really interesting here is there's a flip potentially that can happen here. That discrepancy is where what I would describe as a rule factory and rule in the sense that we find patterns that's kind of

our place where, as it says, the prediction and the observation come together.

12:02 And I belief that that's different than a roadmap, which is find your way. I believe that discrepancy or that road factory is wayfinding. And I think what we have the potential to do with this paper because of the way that they have sort of presented the information is the words become eventually become explicated rules, which then become action in the phenomenological space. That's where it gets really, really interesting because I think most of the time when people look at things

as a subject, they've taken the world and they've collapsed down to words and diagrams and models. What this potentially allows us to do is flip that and move from the words back out into the space with a little bit more confidence.

12:51 So, yeah, that's why I wanted to kind of start here because I think the word discrepancy, it's the first time I've seen it in this kind of figure model and I really like it. Cool. Yeah. It makes me think about setting off on the road trip on the mental actions that reflect the paper. And there's the roadmap, which is super informative, but it is like instructionalism.

13:20 Daniel:
It's saying you're going to go two streets and then take a left turn at the stop sign. And then when this happens, then you'll do that. And if you've seen this, then you've gone too far. Kind of classic instruction type sequences. Here we have a figuring out because it's visually arranged with some local connectivity that's suggestive of a causal connection through time.

13:45 But by no means is there only one way to read this simultaneous figure. And so that allows potentially for more of a figuring out, including the figuring out of rules. Yes, it is indeed a little different with discrepancy at the intersection here. Stephen. Yeah, so this ties in with the lower bound evidence control approach of action.

14:16 Stephen:
So the idea is that action is what we can tractably approximate and perception is something that we can access but can't, and we can try to make more sense of, but it is a harder piece of the equation to get a handle on. So I'm wondering if that's how do you feel about that exploration in terms of its use across other areas of applied active inference? So I think it could make things a little bit more digestible for a number of concepts.

15:08 Daniel:
Okay. I'm thinking about this in the context of recognizing that it's different than other action perception loops that we've seen, which is just really important to keep in mind that we're not perceiving something that we're projecting too much because other times the outgoing arrow from the entity is. What if this were a Markov blanket type diagram, which they often are, the outgoing arrow would reflect active states and then the statistical dependencies that are outgoing. But where is action? It's on the bottom right.

15:44 So let's really try to understand why the pieces are placed this way. The outgoing feature of the cognitive entity is the prediction and here's observation coming in. If this were a Markov blanket diagram, we'd have the world and then the incoming statistical arrow would be the observation, that would be the sensory states. So here it's the outgoing prediction of the entity and the incoming sensory data or observation. Those two are being differentiate to form a discrepancy, which is just a qualitative term, but it could be then brought a little bit more formally into a prediction errors or a little bit even beyond that into like a free energy differential.

16:40 This discrepancy has two arrows coming out of it. From the discrepancy is arising perception and the changing of beliefs. Does perception always involve changing beliefs? And discrepancy is also giving rise to action. So what kind of a thing is discrepancy such that the inputs are prediction and observation and the outputs are perception and action?

17:11 So at the very least this is not the Bayesian graph representations that we've seen before or that we'll see in just a few slides with a more traditional interpretation of nodes as random variables and edges as statistical dependencies. This is a little bit more like a thought map that then connects to the variational free energy equation which we talked about in number 37 a lot more. But just to recall, there's the red and the Bleu lines, these two different components of the variational free energy and those are shown again here. So check out 37 to learn more about the red and the Bleu and about variational and expected free energy. But for here we're starting with this action prediction discrepancy motif and then connecting it to perception and action as variational free energy.

18:14 Minimization Stephen I think one thing that's. Useful with this more nascent representation is this discrepancy can go into in different directions. So it can be discrepancy in terms of temporal occurrence, when was the prediction predicted to happen, when was it observed, but it could also be what kind of prediction and what kind of observation or whereabouts was the prediction, whereabouts was the observation as many different and at different scales. So there may be that at the kind of lower levels. There's quite a big jump between when I predict the baseball encoding to my hand, when that prediction was made, when that observation occurs, and also when that is chained up at different slower, bigger steps of the nested Markov blanket sort of sequence.

19:20 Stephen:
So this idea of discrepancy is I suppose it's probably the biggest bucket they could find, I would imagine at that spot. And that may be partly why it's there. Can I just add to Klaas. E. Stephen think one of the things that's interesting, especially after just doing the 37 paper where we were talking about guides, to me, the timing of this was absolutely immaculate, because now all of a sudden we've gone from guiding to almost taking potentially a referee's position on the world, which is a whole different thing than taking up the position of being a guide.

19:59 Dean:
I mean, that's not nuanced, that's not subtle. That's quite an identity shift. And we're going to actually get into identity when we look at some of the evolutionary steps later on. But I wanted to slow down on

this because I thought that's not a minor change, that's a whole different perspective shift. And I think we need to really make note of that because I think it colors a lot of what's to follow.

20:25 Daniel:
What makes you say that we're a referee in this situation? Discrepancy. There's a difference between what the world is telling us and then how we rule on that. So think of any game where you get into an argument with the rest.

20:43 Dean:
Are they wrong, are you wrong? It doesn't really matter. The point is that there's a difference that's when you said differential graph is maybe one of the ways to be able to make that explicit, but that's not how it's typically framed out when we're talking about Bayesian or Markovian stuff. So I think this is a big move. I know it might seem like a tiny one, but I think as we go again, as we go deeper into this paper, it affirms a lot of the stuff that I actually saw when you're trying to go out there and forage and figure.

21:19 So I'm excited to keep going. Yeah, it could be like the observation. The baseball player is running towards the base and the coach is observing and there's this ongoing prediction and observation with no discrepancy because part of the generative model includes the person moving through time. And then given the observation and the generative model the coach predicts, slash, expects and prefers, which we'll come to again later with the three piece that the baseball player is safe. You rarely see super animated refusals when a call has gone towards somebody, but it's when it has gone against, it's violated their fitness that there's a discrepancy with what the referee has called and what the interested coach has called.

22:09 Daniel:
And that is going to lead to some consequences. And again, we're still not at the Bayesian graph level, but that's going to be the next figure. Stephen.

22:21 Stephen:
This can then Bull on the idea of what's cognitive, what can you have a perspective on? I think what you're saying there with this referee position is as well as the kind of swarming dynamics and the kind of inactive processes, much of which is beyond our ability to sense or integrate. What is it once it starts to hit? What's the big bucket at the kind of cognition level that still isn't too much of an inflation. So in this case here, discrepancy.

22:58 There is a sense of being able to distinguish perception and action at some level and find a discrepancy, which in some ways I would imagine covers a lot of what cognition would require to be thought of in a cognition way. Other types of action, for instance, how we heal or grow, maybe action and perception won't be so separated. But this is trying to come at the idea of cognition. So maybe that also informs this process. Cool.

23:32 Daniel:
So action two, again, was just about the two concepts of predictive regulation and control. That's what we saw here, predictive and control. So it's just conceptually laid out in figure two. They formalize brain design as structure learning in generative models. So here we have a difference figure.

24:01 We still have the cognitive entity and the world. We add in one layer of absolutely essential active terms, which is generative models, generative process. The generative model is the cognitive entity's model of the generative process is the underlying phenomena that gives rise to observations. It's the difference between the cognitive model of vision, a generative model of vision, and the generative process of visual input, which is like photons and the sun and all of that. So these are very different.

24:45 They're complementary. But just so that we're really clear going forward that we're going to be using those in their specific sense, not using them like, oh well, it's a generative model because this is a model of cognition that makes me excited and think of ideas, that's not how we're using generative models here. So just to be clear on that, now we see action as changing the world represented by you. And the cognition involves partially action selection, policy selection, planning as inference, but not going to all those details yet, which can have some influence on actual unobserved, hidden states of the world. X star.

25:31 X is the cognition model of that hidden state of the world that is being inferred. And Y is the observation that then feeds back into the cognitive model. So that's what these nodes mean. It's about partitioning the cognitive model from the generative process, separating the generative models from the generative process. And we're starting to see the traditional blanket form with observations having incoming statistical dependencies and actions having outgoing statistical dependencies.

26:13 What else do either of you see in this model? Yahtin. So let's go back to being a referee for a second. The assumptions is that you're already attending. And most of these models that we have taken up in the past, especially in 37, what they try to focus on is getting from A to B.

26:42 Dean:
What this is essentially saying is let's add something to that translation right from A to B. Whatever we're trying to incorporate in this representation, let's maybe look at the interpretation part of it now and the recitation part of it. So if I'm the umpire and we can all be umpires. We don't have to do that just in a baseball game. This is incorporating now a certain critical process.

27:07 I was attention and I saw this, but then there were 600 other people also paying attention and they saw that. That's where I think this is getting really interesting. Now this is actually bringing it back to sort of I still think it's an instrumental piece, but I think now we're going to incorporate some of the reality. Were the 600 wrong or is the one person who yelled out safe wrong? That's what I see in this because that you is definitely embedded in the generative process, not the generative models.

27:50 Stephen:

Okay. Thanks, Dean Steven.

27:55 So the question that comes into mind is where the body is in all of this. And I sense that the generative process is the bigger process with the world. The generative model gives away to access those prior and to give away the hidden states. And the cognitive model, generally speaking, cognition is thought of more in terms of deductive and inductive reasoning and logic. And it could be that there's elements of the abductive that could be held within the kind of body.

28:43 And I suppose there's a question there as to how and where that is. I don't think anyone quite knows the answer to that, but that's my thought. So let's remember that this partitioning is specific to a given instantiation of model based science, just like Majid was talking about. So we're not going around and assigning aspects or phenomena of the world into either a generative model or generative process. So where does the body fit in?

29:17 Daniel:
Well, with respect to a model of the body being a structurally real thing that gives rise to observations, it's a generative process looking at the coin from the other side and thinking of the body as a generative model of its niche, doing inference on certain things or acting as if in that sense it's generative models. So different kinds of entities are not going to be just assigned simply to one side or the other. It's going to come down to what is specifically being discussed. There's a few other not complexifiers, but first off, just note that the notation here is not the same that's used elsewhere. So we will move towards better and cleaner or reformattable notation.

30:07 But like X star as an external state and X as an internal state might be clearer for some people. It also carries a little bit of a baggage that the hidden states is exactly what is being inferred about the entorhinal world. Like there's a temperature parameter in the brain and then there's a temperature parameter outside in the world as we explore in the representations paper. It doesn't necessarily have to be that way. There could be a hidden states internally having to do with movements left to or right and then there's a temperature variable outside and then the model, the generative model of the Cognizer is about movement conditioned on temperature observations, but not necessarily simply a thermometer being instantiated in the head.

30:57 So it's like good to look at this graphically and think about what is being connected to what without worrying too much about all of these side questions. But this is what they're setting up as the basis of their further dimension, which is it's about prediction and control and we can use Bayesian graphical approaches to represent that. Stephen. And that influence state, as you mentioned, in some ways that's always slightly hidden from us. What does that Dean to be an inferred state in terms of?

31:36 Stephen:
Is that the cognition that's coming out of that in some sort of deductive or can it be an effective sense of how well something is going? So again, it's not directly asked here, but accessing that is one of the big problems that happens. How do you access what has been inferred when it's not necessarily something

you can access through sort of direct reporting from a subject matter, from a participant? Right? Like another example of that might be somebody who has skilled action with respect to investing, but they may not be able to give an estimate for a number that's a certain asset is expected to be at because it's not like they're doing the asset price prediction and then doing a strategy.

32:32 Daniel:
The cognition model may have a very different structure, so we explored it in the representation in this paper. But like there would be ways in which therefore that investment decision is not a representation of the actual stock market because it's not the same variable. But then the aboutness of the investment decision would be a representation with respect to what was happening in the generative process. So those were some of the side avenues that we've looked at previously, but they're all in play at once. And so the question is just how to linearly structure to respect the specific contributions that are made here and the insight that can be gleaned without every single time pulling back to some of these questions.

33:20 But it's great that we have like specific papers and memes and core terms that we can refer to and then carry on with what they actually contribute. Okay, any thoughts on one or two before we get into the structure of the Alo stat or the homeostat first, I guess.

33:47 Dean:
Let's go. Okay, so we're going to go from this black and white Dorothy still in Kansas mode to some predictive motifs of ancestral brains. And the three motifs, the red, green and Bleu are homeostasis. So returning to a set point, allostasis anticipating or approaching a set point in the future, which could be a fixed or changing one and then implementing behavior control, not just scalar homeostasis or allostasis on an interceptive variable. Okay, citation 20 in the paper chance at all from the future.

34:36 Daniel:
March 22 is where to look for more details on the homeostatic formulation that they're using here but we can see it in terms of their figure free. Okay, so the left side just for reference there's figure two so that we can remember the structure of perception, cognition, action, impact that the authors are working with here. And we're going to connect this black and white figure two to figure three A. This is a graphical model, graphical in both senses, meaning visual, like we're perceiving it through computer graphics and graphical meaning like a network so nodes and edges because there's computer graph that aren't network topologies. But this happens to be a computer graphic that we're perceiving visually that also is reflecting a graph in terms of nodes and connected edges.

35:35 It's a generative models for the regulation of a single interceptive variable. So here's A and B with the homeostat and C we split out to talk about later. But first A and B. This generative models includes an interoceptive thermoreceptor and a belief about body temperature. The prior over X, which is body temperature, is kept fixed and hence it acts as a cybernetic set point.

36:10 Any discrepancy between the predictive, thermoreceptor activity given beliefs about X. So Y conditioned on X and the measured y is registered as a prediction error that is canceled out by an autonomic response. U, for example, a thermoregulatory response. So hidden states on temperature, beliefs about temperature, why the thermometer perception and then there's the selection of action. So some sort of like vasodilation or thermoregulatory response and then that's going to change the underlying unobserved true temperature but that's not needing to be shown.

36:52 So any comments on that first part? We're looking just at the top half of three A and connecting XY to you. Action. Can you still hear me? Yes.

37:09 Dean:
Can you take the cursor now and just reinforce the feedback and the feed forward part of this because the author spent a bit I don't want it sort of unpacked like they have at the figure three unpacking that we have at the bottom of the slide. Can you just run the cursor over all the examples of feedback and feed forward going on concurrently? Because I think that's really important to see that it's happening in both directions at once. Let's label everything and then definitely can do that. Perfect.

37:43 Thank you.

37:55 Daniel:
Okay, so light Bleu is action.

38:02 These are subtracted. So red circles represent expected values of X. The red circles which are used to make predictions about Y, these are subtracted to form a prediction error.

38:30 Dean:
Because those lines with the arrows on the end of them are just dependencies. They don't really show both the feedback and the feed forward.

38:52 Daniel:
Yes, agreed. Like having some rounded edges and

other directed edges. Let's see whether it's whether they're like kind or not.

39:10 Let's start with action. Action influences the state of the world which an edge can be drawn to, how that changes the state of. The thermo receptor. Why? Here was the observation.

39:31 Yeah. That's the state of the thermoreceptor is the observation. Okay. Yeah. So action changes this the observation, the state of the thermo receptor which is being contrasted with the belief about temperature.

39:48 From there, a prediction error is generated. It could be zero if there's no difference, or it could be higher.

40:01 Dean:

Can you just show, with some colored line that has an arrow on each end a connection that demonstrates within that diagram the feedback and the feed forward at once? Is that possible using this diagram? Let's see it's representation. So a single arrow, a single line that has an Aaron Fath. Both ends is a different color.

40:27 So we've superimposed it over this, but that it shows that there's a feedback feed forward loop going off at the same time.

40:38 Stephen:
Okay. So here's the temperature information measurement flowing in to contrast with the beliefs. And it is in general, really important to label the edges, not just use color coding. We'll let it roll for now. The information is flowing from the measurements to the belief and then that gives us the prediction error.

41:04 Daniel:
Then the prediction error is used in the selection of action, which then influences future observations.

41:18 Now the red circles represent expected values of x.

41:26 I'm actually not sure what exactly the red bottom, larger circle is meaning.

41:36 Because the prediction error well, there has. To be some way of representing that discrepancy. Right? So they needed the second ball on the bottom to show difference between prediction errors and expectation. I think that's all that's trying to show.

41:54 Yeah. Or another possibility might be that the prior is staying field. Correct.

42:07 They can be flexible, but in this case it's a fixed prior. That's because we're talking about homeostasis. Then the observations are diverging from the prior and the posterior is kind of like the realized perception, which is a compromise between the sensory data coming in and the prior. And so, yes, there's a lot of degrees of freedom depending on how parameters are weighted. This green line might approximate the red a lot more sharply.

42:41 We call that a weaker prior because sensory data updates the posterior to be more like it. Or it could be the case that having a lot of observations different from your prior don't change it. That's a strong prior where sensory data do not change it as much. But then here is the prediction error in relationship to x. I don't know.

43:09 We can look at the paper, but is Mu shown in an equation? I don't have it copied out if it is. Let me look here while you're doing that. Yeah, but that is the perhaps here. Okay.

43:24 Stephen, anything?

43:29 Stephen:
Yeah, I'm just noticing how effective the belief on temperature. It flows through the thermoset down into the to get the expectation prediction error. So, like you've shown there, there is a dynamic going down from the belief through to prediction error being mediated. It's like the thermoceptor is kind of like a mediator between belief and prediction error on what to do for action.

44:16 Dean:
And another point, the reason why you can't find the mu is because it's not actually pointed to in the description. It's not really true. Yeah, and I remember reading and looking and looking and looking and not being able to find it and then looping back up paper to go, okay, so how about feedback feed forward. Now, there are cases where view is used to describe internal states that may be implicitly how it's being used, but it's super important that all variables are defined in a paper. I wish they all had a table for every single variable and expression that were used.

45:00 Daniel:
Yeah, it would make the dot zeros easier, but also it would reduce uncertainty.

45:07 For example, is this epsilon even described? Okay. Stephen. I mean, I suppose in a thermostat in some ways the internal states of the bioelectric strip in some ways it holds the way that the expectation of action can happen because in some ways it dictates the way that the thermostat will behave in some sort of ways, even if it's kind of an analog process where would the kind of Kilner thermostat I know this is a homeostat, so it's a bit more sophisticated now. But extrapolating that out, you've kind of got a belief, and there's a belief of what something is, and then there's an expectation of what you can do about it.

46:07 Stephen:
For instance, my beliefs can go bigger than what I can do in my actions. I could have beliefs about temperature which exceeds where I can even exist or where I am able to change it. It depends on the scenario. So it's held in both scenarios in both the body and the context and the kind of probabilities available. Okay, so for a nonliving thermostat, it doesn't have a cognitive belief on temperature, but there could be something that's computationally like that reflected by just a digital prior on temperature.

46:48 Daniel:
And again, these aren't cognitive, personal, effective, experienced beliefs. Bayesian belief. This is just saying random variable reflecting in a model variable on temperature. So even a sincerely held incorrect ecological belief is not the Bayesian belief. They might coincide at times if it were a parameter, but the belief here being the prior, it must be adaptive.

47:25 That's the evolutionary twist that actually helps resolve a lot of this because otherwise the design space of all edges by all nodes and then any variable I mean, it's just like saying here's all the words in the dictionary. And so evolution helps restrict the discussion to cases that actually do manage to achieve adaptive control. Dean this is where the translation from the silhouette of a head to a statistical density is assumed that the person who's following along with this just sees that. But you can also see how easy it is to slip into the idea that, oh, wait, a second, we've gone from the physical space and we just held on to the physical space, when really now we're talking about a statistical density space. And again, if you're not really, really careful, you can see how people can carry forward something.

48:18 Dean:
But the actual thing that they're talking about has changed. That's why the word discrepancy was such a big deal to me, because I normally gloss over things, but this time it actually I went, oh, okay. So there's going to be some there's going to be some moments here where we're actually talking about different things, even though in the continuum we kind of think that they're talking about the same thing. No, we're not.

48:43 Daniel:
Yes, it's the travails of Realism and instrumentalism for Biological Active Inference episode 55 because of these terms, are they an examples or is this an example of a model being used to discuss a real system? But this suffice to say, is the architecture of the homeostat. It undertakes action to reduce discrepancy relative to a prior held belief, fixed in this case about what temperatures are expected preferred. The dialectic of the first two PS from livestream number 37. This is an expectation and a preference because expectations having to do with survival are as good as inference for survival over evolutionary time.

49:39 Okay, this is going to be contrasted with figure three B, which is the alo stat. We see the same stack of x one Y one, light Bleu, dark Bleu, red, except there is now a second column next to it and there's some cross connectivity. So again, the x's are going to be beliefs about so prior's on and then these are observations. And so they're saying this generative model extends the homeostat by including a second set of exteroception variables that correspond to light intensity y, two observations of light, and beliefs about sunrise. So like beliefs about the generative process.

50:31 So generative model of the generative process.

50:36 Furthermore, the model includes a predictive relationship between sunrise x two and body temperature y. So again, I wish we could label every edge because the edges are meaning different things even at different times. Like this is a predictive relationship. So it's an anticipatory one, but then the one about light intensity and the belief about sunrise, there would be ways to make that an anticipatory or an instantaneous relationship. But the result of this sketched architectures is that inferring a sunrise, which will only happen with high posterior confidence if the visual observations are consistent.

51:21 So low prediction error inferring a sunrise, not quote, seeing a sunrise. That is not what is happening in the model inferring a sunrise and finding that visual observations are compatible with it can trigger the autonomic response. The behavior you of thermal regulation in an anticipatory manner, that is before sunlight actually increases body temperature. Well, how would the parameter be set that way? Because if that were adaptive, then other parameter accommodations have been weeded out already by evolution.

51:56 So that is where we tuck the thread back into the ball of yarn, which is the parameter communication that are nonadaptive die, they dissipate, they fail to exist, they're not going to be measured as things in the culture empirically, however you Kant to take it. So how does it work? It's the same as it was 50 years ago, which is it has to do with survival of the persistent. Stephen. And these graphs, what they are used for as well as they do show beliefs field to the observation space.

52:35 Stephen:
That the space available around the observation.

52:42 Whatever the sensorium that is available, that's what gives the beliefs its scope. And then the expectations are tied in here more clearly to action. So the errors, the errors are what's important. The actual dynamic that's driving action to change something is coming out of the prediction area. Errors that's feeding into action.

53:18 I think that's quite important because this is like the proto animal, this is like the proto kind of piece here that ties into a lot of how we think about knowledge and meaning. And I'm not saying that I'm just extrapolating a lot, but I think it does show how beliefs are tied to the type of sensors that is being used to shape the conversation space. I think that's quite useful. Yeah. And these are just good.

53:54 Dean:
I just want to get both you guys opinions about that orange bar bell at the bottom of the diagram, because talking identity, but we're also still talking about dependency, we're talking about duplication, we're talking about anticipation. So that's a lot of stuff packed in one orange barbell. What do you think?

54:19 Daniel:
Yeah, I'm going back to the full caption.

54:28 The red circles represent the expected values of X. So if we're interpreting these as the posteriors on X, because the expectations, the errors are the top Bleu ones, so it wouldn't have to be the red circles and they don't mention the word orange. Note it's there and it's identity. Note the lateral modulatory connections in the Allostasis network. See 24 for details.

55:07 And 24 is the graphical brain with frist and par and de. Bruce so, yeah, again, this is just a sketch module. They're not using it to fit any data or even simulate any data. And it does relate to what Stephen said about the Intersensory or the intermodal inference, which is one could imagine a cognitive model where if there's noises that are associated with sunrise, then beliefs about sunrise can be a variable with edges coming from different sensory modalities. So it gives us a exploration of the organs of sense and interface cognitive modeling.

55:50 Stephen:
Stephen yeah, it gives a sense of when would something be important to act upon. So there's many, many things that I could believe I'm seeing or being perceived in terms of this is sunrise, or this is the type of light coming in, this is the nature of the light, this is where the light is coming from. There's all sorts of stuff. The stuff that's really filtering down is what is important and this is where the affordance is coming, what affords me to make some action. If it was expanding out, there can be lots of things I noticed.

56:33 Like I might notice that there's a line there and I may notice and believe it's orange, but I may not act upon it until I take my attention to it and think it's useful as there's many things going on on the page. So when we're we have beliefs, but then there's what's being acted upon out of all of that to then create some sort of change. So I think that is also quite useful in this architecture.

57:06 Dean:
Can I add to this because I think really important. So we're talking about in the context of the sun coming up, but let's see whether it still is true. And I think it is just to prove a point whether that orange barbell still makes sense. If we're talking about Daniel anticipating 61 year old Daniel and Stephen anticipating 61 year old Steven, and Dean anticipating 61 year old Dean, now I'm much closer to 61 than the two of you, but do I need the Monte Carlo? Do I need the play out to be able to make that anticipation?

57:47 As long as I have the identity and the parallelism and the backwards and forwards looking dependency that we see between X one and X two and Y one and Y two, that's what I think makes this a very interesting way of being able to sort of build on the previous homeostatic examples. Klaas. Stephen yeah, I think also this speaks to traditionally we think about the meaning, what does this mean, what's all the meaning out there? As you mentioned, all the data, the stuff that it could mean to relate all these variables. But ultimately, and the bit that is actually what active inference gives us is that barbell is where the learning comes in.

58:37 Stephen:
It's like the way to know how being 61 is for you Dean, is for you to image what it would be like to make choices and to act in the world as a 61 year old Dean. Not necessarily to go out and look at all the data, so to speak, all the beliefs and the things we think about the world, but to actually bring that in. And of course you've got a better chance to do that because you're closer to being that kind of Dean. So I think that's quite interesting. That barbed at the bottom in terms of pairing together the kind of beliefs about what actions and prediction Brea are available and it's more into that meaningful action, meaningfulness realm, rather than what does it mean?

59:28 What's the data Jelle? We might get there. And I'm not pushing back on what you're saying, Stephen, we might get there. But I think for now, what we're saying is that in order to get there, we have to have more than a single stack of dependencies. We have to have a double stack in order to be able to extend the homeostatic nature of how we've evolved to where we are.

1:00:07 Dean:
Again, I'm not pushing back on you. I'm just saying I don't want to jump that firing pistol just yet because I think that orange barbell is going to turn into a green arrow and then all hell's going to break loose anyway, so we should carry on. I don't want to take Stephen stuff away, but I think I want to park it because we're going to go somewhere in a minute. I'm hesitant to ascribe too much specificity to something that wasn't labeled in the caption. Doesn't have a clear labeling in the figure either.

1:00:39 Daniel:
But I think it is already demonstrated that using graphical framework and partitionings like we have in active inference, we can start to approach some of these questions like how would different kinds of variables be connected? How does that relate to future inference on action counterfactuals, etc. Okay, so these sections 4567 were just laying out the relationship between an evolutionary or a physiological function like homeostasis allostasis or simple behavioral control models of things that are core evolutionary features and function, connecting them to the kinds of internal representation, which are like a little bit of a hybrid of a Bayesian graph and a factor graph because there's some kind of computations being implied here, whereas in a Bayesian graph, the edges only reflect statistical dependencies, whereas we're bringing in a little bit more of like a nuanced type relationship that one could imagine could be unpacked. A bit more were to be specified. But it's just to show how different graphical models relate to different brain architectures underpinnings whether you think this is the actual architectures or whether it's just an instrumental architecture like a model structure on a given phenomena.

1:02:19 How do model architectures relate to evolutionary functions? Where is evolutionary time in this model? We started out by asking like, how do cognitive phenotypes arise specifically over evolutionary but also over developmental time? And that's where we get to one of the main pieces and contributions of the paper, which is figure four, the five main dimensions of elaboration of generative model models introduced in this paper. So there are other kinds of Chang Kim ways that this structure could evolve, but they're going to focus on these five.

1:03:00 There's starting with the homeostat. Does it leave to be our starting point? Maybe not. Starting with the homeostat, there's the I operation, which is identity unchanged. There is I plus I, which is just a parallel isolated duplication that's like from one photoreceptor to going to two photoreceptors, so to speak.

1:03:24 There's the alastat, which is a implication as well as a cross linking. So these are just sketches. One could also probably write that one as like a duplication followed by a linking and there's no like change in linking, for example here. So this is like a few of a taxonomy of operations. There's increases in temporal depth within a level of the hierarchy.

1:03:50 So looking one more unit further in time at a given time scale and then there's hierarchical nesting with a h. And the figure is shown like this because these are like the things that can happen to the homeostat and then they can have a second round and so on and so on. There's so many other ones that could happen. Like where is reduction, where's loss, where's deduplication, where's the uncoupling of the allostasis? Reduction of temporal depth, reduction of hierarchy.

1:04:24 So there's a broad space. But one of the main contributions of this paper is to connect functionally oriented graphical models of homeostatic and physiological function to the operations that result in the elaboration of simpler model architectures into different architectures. Steen, this is an interesting part. I know Stephen wants to say something to but really quickly, there's no plus minus multiply and divide. There's an operation, but there's no symbolism for that.

1:05:05 Dean:
And again, that keeps it safe on a statistical level. I think that's a nice tell there. Again, for somebody who's just sort of looking at this for the first time, pointing that out, it's not there and because it's not present, that tells us something to your point about maybe how do we get implication?

1:05:31 Daniel:
Yes, well, I plus I it is like it's a suggestive use of the addition operator, but we're not adding these just like integers. So yes, they're kind of like categorical operations that constitute this evolution algebra. Okay. Stephen, anything on figure four?

1:05:56 Stephen:
Yeah, I was just saying that I was saying they've set a paradigm from the homeostat and they've taken that through and they've shown that plausibly can carry through to cognition.

1:06:12 And then like you say, there can be other ways. But what they've now given is they've given some legs. Another way to think of the legs that can be attached to that paradigm because this modern approach, while they could have chosen a lot of others by taking it, there are certain things that you're hard baked into that early stage choicemaking which then makes other paradigms wouldn't fit with it. Right? Even if you go instrumental realist, your instrumental models, your realist models don't make sense unless you with this, unless you take in some of the paradigms of this structuring creates.

1:06:54 Daniel:
Yeah, agreed. Thanks. So just maybe think like is this functional model realism? Like the function of the model is what is being tracked through time and so there's many other alterations that could occur and then there's even more granular like parametric changes or changes in connectivity or edge types. And then I think as you said, Dean, when all hell breaks loose.

1:07:29 What's interesting here, let's look at the alo stat in a little bit closer detail. Okay? So previously we looked at the alistat and we were like, okay, there's an orange bar. Now the Alastat has a free bar. It's Green Arrow and it's a unidirectional one.

1:07:52 This one a dependency.

1:07:58 The Allostat has the green bar. Okay, well, those are my bad eyes because I saw the little arrow from the Bleu actually turning. I didn't actually see that there was two circles. Thank you. There's my 60 year old eyes for you.

1:08:17 Yeah. Okay, cool. And so each of these are in quite a different domain or at least like a different aspect of cognition. And these are the transitions that it can engage in. So, like, duplication retains the exact same function, but it's a very important step.

1:08:41 And that's a little bit what explored in the zero with genetic duplications, like on the left side here at the screen, having a region of the genome duplicate. Yes, one of the duplicate copies can just degenerate. That may be like opening up space or creating motifs for other kinds of later evolutionary steps. So even this is not like a failure, it's just a change. There's neo functionalization, where then those two models can start to subspecify.

1:09:12 Like if you only have one photoreceptor, it has to be doing what it's doing. But duplicating, it then allows one to focus on a different wavelength than the other, for example. And then there's sub functionalization, which is where initially a composite function, a B becomes two, like enzymes start to subspecialize on one of the two functions of the ancestral. So it's kind of like a role duplication. So duplication is an important event.

1:09:46 Then there's the Alo stat and multiple things happen with A. But we're seeing it as there is a implication occurring and there's also cross linking occurring between beliefs in one modality or beliefs about one type of thing influencing observations or about another kind of thing, as well as this lateral modulatory, green orange bar.

1:10:20 And then we have two kinds of expansions into temporal depth, a sort of local expansion with the T operation just pushing the model one step deeper. And then the H hierarchical nesting operation. Which brings it to another timescale analysis. Dean and this is where I really struggled. And that is when you were at Alastat, it wasn't just comparative.

1:10:46 Dean:
I felt like you had to build in time, especially if we're talking about identity. But then it was like, no, we're just going to leave a comparative and then we're going to introduce the wedding of this. I mean, I understand why they're doing that to try to make it tractable, but I don't know how they can say it can be comparative learning. I know active inference between A and B right now, but not sort of address the fact that in order to differentiate into A and B, some amount of time had to had to be applied to that split.

1:11:26 I was saying you have to go back in time in order to just be able to do a comparison. So it's temporal only going forward in time. I don't mean all hell broke loose, but I mean, I did ask this question. You're right. Like here there's the triggering in an anticipatory manner.

1:11:45 Daniel:
And so in one view, it's like, but this is a single time step model, so how could that occur? Because it doesn't have a temporal depth. On the other hand, again, like kind of those two senses of representation in the structural sense, if the model is a single time step model, it kant have a representation through deep time ergo, it cannot be anticipatory function. The other side of the coin would be if it's enough at this time step, such that beliefs about sunrise, trigger and response, the model can be functionally

anticipatory without a temporal depth representation. Which is why the representation discussion was very important leading up to this one, because we're all over those eight quadrants here, because in evolution and this is something that like cognitive psych people talk about is like our cognition being shaped towards effective, which is to say productive reproductive success, not towards like, ultimate truth discovery.

1:12:56 And so should we be too surprised that our interpretation, which has had a certain objective function, is then susceptible to being convinced about things that are not true? I don't socialize with those people. So I don't know. I just knew that when I was looking at it, I was thinking in real practical terms, what is built into this in order for it to still make sense? And that wasn't necessarily parsed out in the diagram.

1:13:31 Dean:
You kind of had to fill that in through your own interpretation of what was going on. So thus back to discrepancy. That discrepancy thing is going to come up again and again, and I'm really glad that they inserted it at the beginning. Klaas. E.

1:13:48 Stephen:
Stephen think it's useful here to hear this temporal depth and hierarchical depth being sort of spelled out, because I think it's sometimes it's a little confusing in other papers quite where it fits. So temporal depth and this, again, might be where some of this ontology work can be useful, just to see if that can be kind of consistent. But if temporal depth is, you've got variables for past, present and future states, but they're operating maybe within the same kind of dynamics. And then if you've got different temporal rates operating simultaneously, then you get that hierarchical depth. And I thought in the past, I sometimes thought temporal depth would have covered that more hierarchical depth.

1:14:46 So this is interesting to think about whether this might be actually the way that differentiation is best made.

1:14:58 Daniel:
Differentiation is probably best made with reference to a specific model. And this gives us the design language and the grammar and the motifs. So if we want to predict 100 years in the future, should we have a decade model with ten year models? That's a times depth of ten decades and then a nested model with a depth of ten years or so. We have a one layer model with 100.

1:15:33 They're very different and it's not to say that one is like more accurate or less and depending on how the exact situation is set up, maybe the computational requirements of one or higher or different than the other. But that's the discussion to have. Do we want to have a B transition matrix? Let's look at a nested model with temporal depth. So here we have on the top level of the model s state two, this is temporal depth happening at the upper level of the model, three discrete time points in an upper level that is temporal depth, there's also nesting within each time step.

1:16:23 There's a cognitive rollout on three time steps. This is one type of nested model but there's other nestings that can exist. So here we see a depth of two and then at each level of the nesting there's three temporal depth but they're totally different in architectures and in function. So it is the interesting discussion, especially when we have to do multiscale prediction like Fermi estimation. So if the prediction is within one time scale there isn't necessarily a need to least but once we get into larger time scales or where there's emergence over multiple timescales, it makes sense to have a nesting model.

1:17:13 Like if we wanted to predict someone's activity over the next seven DAGs, it might make sense to have a day model with a nested model inside of that, because then within each day there could be parameter setting versus just trying to fit one time series and find the parameters that make that one time series oscillate in a circadian way. Instead we just need a day model with a very specific kind of simple transition and then a transition at a nested level with a depth that is measured at the hourly or the minute time scale rather than the days. And that becomes especially important when doing familybased model fitting like variational inference. So we're not just like getting all the data points and fitting a spline through them, but we actually need to have appropriate model structure otherwise free energy minimization will drive right off the cliff just like we talked about with axles bacterium. So if we want to do free energy minimization and model selection on model structures, we can find the best model relatively easily given the families that we've specified.

1:18:27 But there's no guarantees once we step outside of that spotlight that we're even in the right category structurally. And so that's why humans having these design motifs in mind and many many many more helps us not get into what seems to be a global optimization
 based upon free energy minimization but actually is a very relatively local optimization based upon unimaginative model structured learning. Stephen yeah, thanks, that's really helpful. Also ties Hinton something I've been really thinking about in terms of. The modeling itself is these scales, these steps.

1:19:09 Stephen:
You know, you're talking about the multi year decades when we're modeling, it's like, what do we have access to? So we tend to ascribe that in terms of what can we externalize, what can we use to put something into a model if it's instrumental? And then of course, there's a question of, well, what is the realist perspective on that and what is there that's within our conscious awareness of being in the world, something that can be reported as an event, a story, and what timescales are present there? When you get into 100 years, it's beyond that. So then you're saying, okay, what are you going to extrapolate?

1:19:47 What sort of temporal structure? But on the other scale, if I'm asking what do I think I expect you to do next, or Dean to do next? And it might seem intuitive, I can have a whole series of temporal and nested scales going on which is not even available to me in a cognitive conscious way. I might have a sense of a feel for what I think is going on, but a lot of my predictive processing will be happening at rates unavailable to my conscious awareness. So that also ties in, I suppose.

1:20:27 I'm not saying it's an answer, sort of adds to the challenge, I suppose, but it is the challenge of how much ends up being what we can get a measure on to put into our model as much as what it is. And that's the same challenge in a way that comes even from the realist perspective at some level as an organism. What is it that I even can access in some realistic way? Not saying we can model it or know that could be put into something like this. So I think you're going to say something.

1:20:59 Dean:
So the paper is talking about how do we use certain math to be able to understand maybe the evolution of cognition, right? That's essentially what the paper is trying to point us in that direction. So again, in order to make this accessible, in order to sort of lower the barrier for people who actually want to have this make sense because they don't necessarily spend sums and sums of hours on trying to figure this

stuff out, like the authors or maybe even us. I Kant to point out something that I think is obvious to us, but maybe not so obvious to people who are looking at this and going, I have no idea what they're saying. I think, first of all is that there's a chronology which we keep going back to, not just at the temporal depth level, so that's relativity math.

1:21:50 There's brain evolution aspect of this, so that's algebraic math and there's a dependency, so there's a statistical math. Do we have to be polymath in order to be able to feel our way through this? And I'm going to go back to what I said at the end of the 37, Karl Friston said, you don't have to be polymath, but you better be at least somewhat comfortable feeling comfortable with the math, because there's no one math that's going to get you here. It's a blending of all of them, and then it's how you feel about that. That's probably going to be the thing that lowers the impediment to really being able to use this now in real practical terms going forward.

1:22:37 So what's the next evolutionary step? Right. It's how you're able to turn this into something. So I think that's one of the great gifts of this paper. If you're not afraid of being, if you don't consider, if you don't Kant to identify now or you don't duplicate now, or you don't allostasis now as a polymath, maybe that's something we owe people that want to get into this.

1:23:08 Daniel:
Maybe the British pronunciation of maths helps reflect that. There's already a plurality of maths. And Kirby Earner, coming more from the cybernetics side, talks a lot about this, actually. Like how the discourse around math as universal language makes it sound like math is a singular language, when actually math is a pluriverse. And so maths reflects that a little better.

1:23:38 It's not like these are just sort of slopes on one mountain and that's Mount Math, and it's so high and only a few people get to the top and scale every side. This is just like we're using maths. Yes. Of all different klaas. Stephen.

1:23:59 Stephen:
You see a lot of this in the papers that are published active inference lab. There's often a team of people, a group of people, where maybe one member has that higher level understanding the math. But the other question is to understand what would it mean to make an observation that could be a very good the mathematician may not be the one. Like some of the work with Ryan Smith's work on the gut, you know, it was like, well, they're going to do gut inference. Well, how do you get some tractable source on what's the gut doing?

1:24:32 So they created some electrode. I'm not quite sure Hohwy they did it, but they had some way of getting sensory information, effectively sensory information, or effectively some sort of information to put into their model. And so knowing how action and observations can flow in and the implications of that, I think, is also kind of a big part of this. Right. And in some cases it may be a trivial part because it's fairly obvious.

1:25:06 In other parts, that may be the biggest barrier, and then that may be the biggest barrier as well, often in organisms. Right. How do you even access plausibly, the approximation science, which isn't so approximately it's just chaos that it's tractable. Yeah. The rate limiting step or impact and improvement in the real world is unlikely to be any single individual's conceptualization of math on a team if it's structured appropriately.

1:25:44 Daniel:
And for those pushing the frontiers of math. Their understanding of math is quite literally a rate living step for them. Or maybe it might be something very mundane like time availability. But when it comes to thinking about real model based science and translational implication of active inference, I think that we're working towards new ways of combining skills and having shared knowledge resources that help make that make sense and using shared language and ontologies narratives, formal documents, tools on Ft because we're doing it on teams and we're online. Let's just in the last 2030 minutes go through the final pieces of the paper so that this dot one will have been like a first sweep, initial pheromone deposition, and then we can return and take some cul de sacs, etc.

1:26:44 So, action nine, explore a little bit how the duplication operation allows for multiple behaviors to arise and that's sort of by analogy to the genomic duplications and specializations and all those different routes that can occur here. They connected duplication a little bit more directly to the factor graph models.

1:27:24 Dean:
In. The sense that duplicated motifs have dynamics that are conserved over different sensory motor domains. So let's just say that we just had a visual model, it was a column of visual and then we duplicated it. So now we have two photoreceptors with parallel columns and now the photoreceptor in the second one changes into a chemo receptor because instead of expressing Rhodopsin protein, it's expressing an olfactory receptor protein. There can be a conversation of the dynamics of inference even when the observation has changed.

1:28:03 Daniel:
But it was a slot for observation. And so we went from having like monocular vision to binocular vision to monocular vision and olfactory perception. But you can't get there in one jump to just like duplicate and transpose. It's not likely to happen in brain evolution context. And we can see the structure of these hierarchical model as equivalent to being factorized probability distributions, which is to say that the sparse connectivity of variables means that we can make parts of the model that can be like fine tuned independently in a way that can be fit very attractively.

1:28:48 That's the factor graph, that's factorized Bayesian inference. And we've talked about that in other places. But just here all we need to say is duplication enables like A control C control V copy paste and then let's edit the other version. But that can now happen with evolution features and functions. Dean, real quick, your hand gestures in this section point to the orthogonal mesh beginning and then .2 I'd like to pull that apart a little bit, but let's carry on.

1:29:22 Dean:
But that's what I took away from this. The mesh began and in the beginning there was a mesh begins. We'll return to that for dot two. Okay, section ten talks about temporal depth. It's about time and about the encoding of generative models with temporal depth and the way that that supports prospective inference, anticipation or retrospective inference, which is like memory.

1:29:58 Daniel:
So here's what the operator looks like. We have X sub tau x at a time point, and then now there's X tau plus one the next time point, or t plus one could be now, and then t could be the last moment. So memory. And I still see figure four.

1:30:28 Slide 29. Okay. Yeah. So the structure of anticipation and of memory are very similar. It just depends whether one has the stream of observations happening on the left side and a prospection, or the stream of observations are on the right side and in which case it's retrospection.

1:30:51 Stephen:
Stephen yeah, I mean, it sort of ties in. They talk about what the police do is very hard for someone to lie backwards. So you always get them to tell their story backwards because it's very hard for them to do that because they say if we're creating things retrospectively but playing them forward, so to speak, it would kind of tie them out.

1:31:18 Dean:
I want to emancipate that. I looked up police in the paper. I see policies.

1:31:28 Daniel:
Is that policies?

1:31:34 It's a great opportunity for my favorite joke, but I won't thank you.

1:31:41 This action talks about temporal depth. Okay.

1:31:46 From whence temporal depth. Various researchers have speculated that a major driving force for the development of deep temporal models was foraging. And there's a lot of interesting empirical and conceptual reasons to think about foraging in terms of temporal and spatial depth of model. And that's true in the Vertebrates, where they're discussing mainly, like, hippocampus and entorhinal system and in the invertebrates that don't have a hippocampusentorhinyl system. But it's one reason why comparative neuroscience is so important, because it prevents us from getting fixed on specific anatomical realizations of given functional attributes of evolutionary systems.

1:32:32 Like, if the story of memory is just about some brain region in humans, it may be a useful model. It's not even to say that it's inaccurate. It's not even to say that it's a partial model. It just is a model of that. Whereas if we want to understand a given cognitive function in a broader context, we leave to pull back somewhere a little bit beyond or in complement to the anatomy, because we need the empirical anatomy to have anything specific to be talking about.

1:33:05 But it isn't just the case that vertebrate anatomy is the way to do active inference. Stephen I think this is really helpful because it shows that the traps you were just mentioning there that people we fall into is like humans have the frontal cortex. It's all about this new what's, the new parts of the brain that are there. But we don't say that with a Formula One car. We say, well, it's just got a primitive engine thing in there, and the rest of it's evolved.

1:33:35 Stephen:
It's like, well, that's itself different. And the same with the mark Ines conversation. The evolution of what's often dismissed as the primitive parts of the brain. They themselves can have been more sophisticated, though they're doing more things, but they maybe help to be doing more than they were, as opposed to there's something else that is doing all the heavy lifting, which I think is certainly a common misconception, I think, in psychology anyway. Dean, can I read quickly from the paper?

1:34:16 Daniel:
Sure. The evolution of temporarily deep models from simpler models could be realized during evolution via the progressive keyword here correlation of an initially differentiate model. So a relative sense of invariance, ie. A model that does not distinguish present from past and future into a model that features separate latent states for the past, present and future. This is the part I loved.

1:34:44 Dean:
A key drive for this factorization or correlation may have been the observation and progressive internalization of the sensor motor sequences sequentially that the animal creates and experiences while acting. While acting. In other words, the self modeling of one's own sequential behavior patterns. See 53 for a computational example. I didn't open up 53.

1:35:10 But again, if we want to talk about the polymath piece of this and the fact that our hex cells have to do some parcellization, it's right there in terms of sort of the next layer on top of this as we move up through that evolutionary cycle. Yeah. 53 is Stoinove et al. The hippocampal formation as a hierarchical generative models supporting generative replay and continual learning? There's a ton that could be said about that from like a computational and a neuroanatomical perspective.

1:35:47 Daniel:
The internalization of the sensory motor sequences relates to the sensory motor detachment that we talked about in the representation paper. So if some motor region in the brain, if it's like a marionette and the fingers and then the motor plant so that motor region has to be coupled to the activity of that motor region, like either in a one directional way or maybe even in a bi directional way, let's just say. So that system as it thinks, so it does, and vice versa, it cannot engage in counterfactuals because any direction that the neural system turns, the motor system is just simply doing that. So it is not able to

engage over evolution time in adaptive action. The ones that do no longer persist.

1:36:46 It's always what ties it back to reality and to the finite amount of entities on the finite spaceship Earth. Like Darwin's famous calculation, like if the number of elephants slowly reproducing, they'll cover the Earth unless their population levels are kept in check. And so once there's sensory motor detachment, so there's some brain region, either a motor region or some supplemental area, some ancillary area that's able to intervene in that process or somehow play a role that's detached from the motor activity. Now, there can be like a motor planning occurring that opens up the affordance of temporal depth or of counterfactuals all these other cognition functions arising via the sensory motor detachment. And so foraging is an awesome place to look at that for a lot of reasons in different life forms and computational foraging and so like, just a few of the notes were like what are the real cognitive demands of foraging for different creatures?

1:38:00 And what about internal foraging, like mental foraging? And these papers are very good. Foraging in mind and foraging in semantic fields, both very useful, because they have to do with the way that they have some nice maths too, but they have to do with how actions that are spatial can have conserved. Dynamics structurally. Just like we explore here to mental actions and then we see them come together, like with a memory palace or something like that.

1:38:37 So foraging is cool. It's good to study eleven endowed generative models with hierarchical depth of ours. Multiscale inference. We kind of addressed that earlier with the decades and years. Nested temporal modeling is not the same as just deep temporal modeling.

1:38:58 Another way to look at that is like how many operations you would need to get to a hundred years. Well, if you're going to do only temporal depth, it's going to take plus 100 or 99 time steps versus if it's one nesting and then ten on the higher order than ten on the lower order. It was 21 operations. And so then there's shorter sequence of events to achieve a higher affordance model, assuming that they're appropriate.

1:39:32 Section twelve? Yeah, go ahead. Just one thing just on that model and it's also interesting thinking about how we think of time because indigenous approaches tends to be Jelle as a cyclical time, right? So things in the future will be a repeat of the cycle so the times of thought of in a cyclic motion of the sort of passing of the day, passing of the and of course we have this assumptions of time stretching out. So is it that we will over time, what will things be like in four generations?

1:40:17 Stephen:
And if my world and my biology follows certain cyclical patterns and it's fairly stable, maybe that's quite a good way to think about, say, Kant the tree now to help the people in 150 years. So it sort of comes back to the different ways to construct how we think about time.

1:40:44 Daniel:
Cool. So Twelve looks into a phylogenetic tree of the evolution of generative models so what is a

phylogenetic tree and what's the relationship between active inference, FEP and evolution? Well, we'll have more to share in the coming papers, as always but how do they address it specifically here in figure five they map on we'll root the tree so technically, as shown right here, it's an unrooted tree. There is an implication routing here but it's a very interesting feature of phylogenetics that few outside of the field know about that computationally the tree is often inferred in an unrooted fashion like by clustering the ones that are close and finding the relationships, et cetera. However, that is leaving an ambiguity as to where the root of the tree is.

1:41:46 So like for example, if this is the overall connectivity based upon the phylogenetic inference, this is what the tree topology has been inferred to look like and somebody might say, well, it looks like this one, the red one and then the green one below it. They are, they're clearly very closely related but actually if the root comes in here then that's not necessarily the case. So the rooting of a phylogenetic tree is very important and a tree that's rooted inappropriately or has an inappropriate group selection is just it's worse than illegible because it's highly legible and it can have a high statistical confidence but it can have an absolutely nonbiological topology. But assuming that they meant the tree to be rooted here, here the branches are where there's going to be different operators occurring encoding the eye identity, unchanging one and then the edges reflect like bifurcations expectation events. So here there's an ancestral homeostat and then the lineage leading to orange did not go undergo any changes that were nonconservative on the branch leading to these sister species.

1:43:06 There was a duplication and then it stayed the same and stayed the same and then the purple node is going to have two duplicated homeostas and so they're just overlying evolutionary algebra the steps that they gave on to the topology of speciation. And so that as they discuss opens up a way to think from this sort of graphical functional perspective which is compatible with active to think about how, for example, from the species today we could infer some of the schema of the early vertebrate brain with the early primate brain or push it back further. What about the common ancestors of the vertebrates and the invertebrates and all these other questions but that's kind of what they lay out here. These are familiar models to evolutionary biologists like phylogenetic ancestral state reconstructions of phenotype which are often even done in a Bayesian way. Like there's an algorithm beast Bayesian estimation of ancestral states.

1:44:16 This is also a Bayesian model that allows us to do reconstruction of ancestral states but it's in a very different way than it's been approached outside of the active world. But that's figure five and that's where we see the evolutionary algebra, the design patterns or the pattern language for structural changes in generative models super imposed on species relatedness in a phylogenic tree.

1:44:53 Any thoughts on five? I tried to convert this and reapply it to my picking of teams in the March Madness table and I wasn't able to take philosophenetic trees and convert them into winning thousands of dollars on betting on college basketball. So despite my best efforts, it's kind of contained for what it represents dot. Two foreshadow too hard, but I think live stream number 39 and 40 might constitute March Madness. And then also here's another for those who watch.

1:45:34 Okay, here now we have a little bit of a March Madness bracket going, don't we? Here's bracketology. That is so fantastic. So here we go. Yeah, now here the root could have been here and then here's like two big species clades, like two like this is like vertebrates and invertebrates and then there's some rooted clade.

1:45:57 Stephen:
Wow. So that's like why it's important to root the tree because here it looks like all of these are a sister to each other at the exclusion of this one. But then if we were to have rooted that tree like here, then actually this is a small sister out group to all of these and so the rooting and the contextualizing is really important. It's super interesting because it seems like you put in so much data into these phylogenetic models and it seems like literally how could this not converge on a super obvious answer? I mean, isn't it clear how these answer related to each other phylogenetically?

1:46:38 Daniel:
But then there's a little bit of nuance that enters the picture. Stephen.

1:46:43 Stephen:
Yeah, I was curious that the allostasis can recombine with the homeostatic to give a base level temporal.

1:46:56 It doesn't deflate back to just staying allostasis I suppose they would be they just don't show that it did that. I suppose you just wouldn't talk about it when it goes to a and then it goes AI goes to HT so th yeah, so we have a. Homeostat here we have two homeostats, two homeostats and now here we have one homeostat and one allostasis in this Bleu node.

1:47:35 Daniel:
But they're just kind of giving general examples like it would have been cool to see this is where we're talking about mammals and here's non mammal vertebrates. Like here is a specific brain region and also that is in this Chakraborte and Jarvis 2015 paper. This paper gets more into the nuances of neuroanatomy and neurogenetics and how like implication in brain regions can be underpinned by developmental neurogenetic changes. And they do link that a little bit more closely, like to studying the bird brain and the song and motor system in parrots, which is also cool because Birdsong has been studied with active. So there could be a nice building upon this work and connecting it to some of the structure fitting in figure five.

1:48:41 Stephen:
So it COVID almost be like you got parts of the brain which are more allostasis orientated or homeostatic orientated and that so in a way by the structuring of that gives some sort of differentiate in terms of how information and control would be carried out. Yeah, totally extending temporal depth within a model. It's like a brain region getting a little bigger, let's just say not that size is correlated with function in that specific way but then duplicating brain regions is like duplicating laterally or nesting of brain regions. It's hierarchical. Think about how the eyes developed the insect eye, it has a relatively

simple module but then insect eyes can change in size over evolutionary time very radically from like taking up half of the head to being just one photoreceptor or even being lost because they're existing in more of a duplicable motif.

1:49:47 Daniel:
Whereas the binocular eye system that mammals have, it's not as amenable to like just doubling into four eyes or splitting into two in the middles or something like that, whereas the insect compound eye has some of those evolutionary affordances. So then understanding like parameter change within a model and then structure learning on populations of models and what are the mutational adjacencies. There are so many awesome areas and it's like truly just beginning for evolutionary bio active states. Yeah, it was really helpful. Having someone like yourself who's got that background to go through.

1:50:31 Stephen:
That is quite helpful because it's a little bit intimidating, actually, at first sight seeing this for me, I see the logic, what you're saying. It's quite exciting actually seeing that. Again, one of the things this paper does is it's time threads together which seem very distant almost in traditional terms, like Jelle. I don't like using the word kamikaze, but it's a little bit like yeah, it's a bit of overwhelming often. So I think but of course, that's the beauty of active is it has that unifying potential.

1:51:13 So thanks. Cool. Yeah, thank you. I agree. There's some labeling.

1:51:19 Daniel:
This is rarely included unless it's a time calibrated phylogenetic tree, but like, we're looking at the passage of time, but that's not even always shown. Yes, Dean, we'll have our final thought. Can you just go back to that? Because I didn't prompt you to put that up there. But I really like that because at some point we're going to have to in point to talk about that sort of basal gating part of this.

1:51:47 Dean:
And most things that are organized as if then causality. But you just said it. It's then if go or don't go is very dependent on then before we choose to go or don't go. Right. Go or don't go is still held open until one of the others decided.

1:52:08 So another thank you for that because the point too. We should maybe talk a little bit about that.

1:52:21 Daniel:
Cool. Yeah. There's so many ways to think about like this garden of forking paths, branching paths, how that relates to parameter updating, Bayesian belief updating and structured learning. What about when the structure of a model is a parameter in a nested model? So then that sort of blurs the line.

1:52:41 From the bottom looking up, it looks like structured learning, but then from the top looking down, it looks like parameter fitting. Right. And if only we had like some mathematics to describe that.

1:52:56 Okay, any final thoughts or what are we excited about for the dot two to next week? Well, my last thought is yeah, I'm excited to keep this going because I think it's got a lot of legs to it. I think one thing that comes to mind as well with all this progressively increased temporal and hierarchical depth, is maybe where some of the more basic elements still have a really useful part to play is what Mike Levin said about when do we stop? So we got all these ways where we can start to okay, so we got all these ways that we can think, how do we know when to stop and to sit down? And maybe in the ultimate ways when our tummy maybe that's where the gut is so useful.

1:53:40 Stephen:
Like at the end of the day, the brain can go running off and maybe the gut and the heart needs to say, okay, sit down, get some food. I'm going to get a drink of honey now, but I'll bid you feel well. Thank you for a great dark time, but I'll hear your last thoughts as well.

1:54:01 Dean:
Dean, I hate having the last word, so maybe you'll have a dory you'll have it? And I just appreciate Klaas. Stephen just said because I think that that's a big part in this ability to go from word back into the phenomenological space. And I think there are other brains that we know of now that play a significant role in that, not just the one that turns everything Hinton, symbols. Great.

1:54:31 Daniel:
My last thought is there's no better drink after a foraging trip than honey.
Alright, thank you and talk to you later. Bye.

Session 0.2, , 2022

https://www.youtube.com/watch?v=YKn2njZ_ICg

Second participatory group discussion on the 2021 paper by Giovanni Pezzulo, Thomas Parr and Karl Friston, “The evolution of brain architectures for predictive coding and active inference.”

SPEAKERS

Daniel Friedman, Dean Tickles, Bleu Knight

CONTENTS

TRANSCRIPT

00:30 DANIEL FRIEDMAN:
Hello and welcome everyone, to ActInf Livestream number 38 two. It's February 23, 2022. Welcome to ActInf Lab. We are a participatory online lab that is communicating, learning, and practicing applied active inference. You can find us at the links here on this page.

00:47 This is a recorded and an archived livestream, so please provide us with feedback so that we can improve our work. All backgrounds and perspectives are welcome and will follow good video etiquette for live streams. Go to activeinprints.org to learn more about how to participate in any of these live streams, which are part of the Comma Communications Organizational unit or just any other part of the lab. All right, today in Active stream number 38 two, we're going to be having our third discussion. There is the zero, the one, and now the two on this paper, “The Evolution of Brain Architectures for Predictive Coding and Active Inference.”

01:28 And maybe each of us have a few ideas in mind and I'm sure we'll come up with a few other things. Plus, if anyone's watching live, they can write a comment in the chat. So we'll start with just introduction, saying anything if we want about the paper before just jumping in somewhere or somewhere else. I'm Daniel, I'm a researcher in California, and I will pass it to Dean.

01:58 DEAN TICKLES:
Hi, I'm Dean. I'm up here in Calgary and I'm not doing very much except trying to keep up with all the papers that we're looking at in active inference. And I'll pass it down to Bleu.

02:21 BLEU KNIGHT:
I'm Bleu. I'm a research consultant in New Mexico, and I really enjoyed this paper for many reasons. My background in neuroscience not, like, being a priority, but yeah, it is. So I'm excited to get into it and to discuss some of the points that I picked up with you guys. Okay, well, Bleu joining us today. What figure can we start with? Or start with a blank sheet and just any of the points that you'd like to raise or we can do something else.

02:55 Man, figure four for me was like, awesome. So maybe that was my favorite figure. Maybe that's the one we should start with. Okay, all right, so what was your just perspective on what figure four was and then what did it mean or why was it interactions for you? So I was interested in this figure and maybe like, even in how this figure could be expanded in terms of hierarchical depth and temporal depth.

03:30 And temporal depth has more than one potential scale which isn't really mentioned here. Temporal depth can be thought of as small time steps. What am I going to do next? What's my next action? But also in bigger time steps, what is the next interactions of me and how these hierarchical depths?

03:55 Because that's obviously like a larger levels hierarchy, like the next iteration of me as a species or as a human species. Right? Like, what is the next interactions of the species versus what is the next action I'm going to perform as part of the species. So that was kind of an interactions like perspectival dance that they took and just really like the multi dimensionality of this and how to kind of think about it in like, maybe like is this like a multi dimensional space in terms of like, my cells, what are my cells going to do next? My tissues, my organs, knee, active inference lab, lab.

04:39 What are all these things going to do next? And also on that level, is that like a short time frame or is it like a large time free? And do the time frames differ or can the same hierarchical level have more than one time frame or not?

05:01 Daniel:
Okay, nice questions. Dean, what do you think about any of that as we start to look at figure four as a starting point for talking more about what Bleu just mentioned? Well, I think one of the things we touched on in the .1 was the idea that within a frame there's also an orthogonal mesh. So as we push through this representation from the bottom up, are we going from what might be described as an architectural existence to at the top, what we might describe as architectural learning? Now, of course, if we take that position, we have to be careful that we're not reinforcing the idea that it's all about constructivism all the time versus being creative.

05:56 Dean:
Right. Like, I think you need to be a little bit of both. And I'm not being creationist, like, not one story that carries for it all the way through time as Bleu kind of ref there, but the idea that the updating matters throughout the time scale that you choose to look at. And so now the question is, is architectures the lens that we should be using to talk about how we advance from mere existence to a place that we would describe as learning?

06:34 I mean, that's what it was chosen here. But is that the one way now? Right? It's a way. It's a good way.

06:42 Certainly make something super, super complex into something that's just complicated. But is it the only way? I don't know. Just one question of clarification. When you said this is the only way, what were you referring to this figure for?

07:00 Or what? To the idea of an architectures. Right. The architectures lens, as opposed to other ways that we might build out this idea of going from homeostatic evidence to going from unknowing to knowing and whatever that means.

07:27 Daniel:
Okay, thanks for sharing it. Just to kind of maybe connect this. You mentioned architectural disturbance, which is within the metered space. Whatever the metric is or the measure or the grids that you draw, there's a distance within a map and that's like the ecological transect. That's the Blueprint, that's the forensic analysis.

07:54 And then that's like parametric learning because it's a distortion of parameters in a given space and then fully defined, not just like space for openness whereas the architectural learning is at the structural level, architectural. And then that was kind of like what you related to not just the homeostatic or even the allostasis perhaps, but something that's like the sort of unknown, unknown open ended ness, biological open, evolutionary trajectories. So not just sustaining it, but actually realizing it.

08:42 So we'll just get a few notes on each of these transformations, then see what other ones or like what are some situations that might model and then also return to the questions that were raised early by Bleu, like with the multilevel self and where's mental action or thought. In this model, the examples were provided of like, for example, the Alo stat that took the temperature homeostat and then added in in their example like the Sunrise visual detector and modeler column. And so that gave the anticipatory change in body temperature when the sun was viewed. But then is there a similar structure for cognition? Is it just like linking together senses in more and more combinatoric way?

09:36 It's a very sense oriented perspective which from an evolutionary perspective, which is what this paper is exploring, that's the one to ground further elaborations and changes on rather than trying to reverse from something like engaging counterfactuals. So just using an active inference model, sometimes for a counterfactual agent, if the model says that that's okay or doing the bacteria without implying that it's doing the counterfactual explicitly, that might be possible too.

10:14 Dean:
So we have thermal regulation in the paper. But then the question is, does that equate to feelings in a human being? Right? Like if all that feelings cover necessarily collapse down to the example of thermal regulation? And at what point are we aware of the fact that there might be a difference between the example and the concept?

10:40 Like where in this figure four is that blurry line that we cross? It's a good question. I don't know where it is subtina. You're talking about feelings like emotions or feelings like sensations or both? Yeah, all of that, right?

10:58 Like so feelings tends to take on a big sort of it's more detached from the example of thermal regulation. And yet that's just the fact that you can ask that question proves out that maybe the one example doesn't cover all of the different aspects of the higher concept, right? That's where the scaling part comes in. So it's not just temporal depth that allows for that. You have to have the temporal depth and the hierarchical depth what kind of feelings are you talking about?

11:35 Implies that there's more than one way that has to exist for you to come at that to ask that question where do we draw the partition? Right? So in here they show duplication. But I'm not going to say it misleads, but it potentially mislead if you think that there's a linearity to this. But in fact, a couple of times Daniel talked in the .1 about the fact that there is a step function moment.

12:10 There is a jump. It's not a jump. I forget about one example where we can't say you just go from one eye to two just because I think that was one of the examples you gave Daniel. But at some over evolutionary time there is these little progressions where if you look back on it, it seems like a rather large jump took place. Right?

12:41 Daniel:
Let me give a few points there. So when you said, like the big jumps and so in one sense that's related to the kind of discussion around complex patterns of evolution and some phenotypes or genotype regions that are very stable through long periods of time. Over many generations, like the ultra conserved elements in the genome that have very low mutation rates versus regions that have high mutation rates, et cetera. And then same as phenotypes, like phenotypes that are very constant, like the silicon fish or phenotypes that are very plastic or dynamic or ones that we just don't have the records of.

13:23 And then how to connect what is happening or the models that we want to make in humans, especially with computer enabled

and language enabled humans, technology enabled humans with the continuum of cognitive processes. And then you said like, it's not linear. And then that kind of image me think like there's like feelings, which is still a rich qualitative area, like you brought out like sensations but also experiences and emotions. And then how quickly, securely or unambiguously can we map just that cloud of concepts just to be like it's a policy selection on attention, it's an uncertainty on valence, it's valence. Like we're doing the parametric cognition model with this, but it's still even a little bit open.

14:24 What this represents. Is this the evolutionary biologists doing model testing on a phylogenetic tree? Or is this talking about the genomic or the neurophysiological components that have to be connected just anatomically for a certain function to arise? And that is the rich area that there's so much to explore in too. Well, I think the one thing we can agree on is because we're talking about evolution, we're talking about history.

14:57 Dean:
So it is backwards looking and backwards looking is easier than looking into the future, right? So I think that's one thing we can agree on. But what this is trying to do is explain perhaps how the history came about.

15:15 Blue:
So one thing that the authors mentioned as well and that's kind of Linsker to the idea of feelings. This was a discussion about the evolution of brain architectures. But there's also the fact that brains are in bodies. They don't just exist in jars. For me at least, I feel like I feel like emotions, not necessarily sensations, but I feel like emotions are driven in a large part by the body.

15:46 We've made a huge kind of connection between the gut and the brain and there's probably more connections to be felt because I feel feelings, like emotions deeply in my body. Like, I feel them in my heart or I feel in my stomach or I feel, like, in my pinky toe. So I really think that the idea of emotions can't be separated from embodiment. But prediction, I think it's a mental activity, or at least I feel it in my brain. So I don't know if that's the same kind of sensation that everyone has, but prediction is, in my mind, a brain thing.

16:31 But I know that's overriding. If someone throws a Bull at me, I put my hand up like that's a reflux. But is that like my brain driving my body? I'm not sure, really. But I know that there's some prediction that's maybe not mental, that's more like physical or that's more embodied is probably the right word.

16:56 Dean:
And you kind of move into that intuitive space as well. Right? Like, if you think about going to the beach, I can't tell whether you're going to think about the wetness aspects of that, the warmness part of that, the smell of the sunscreen part of that. But you're going to but that prediction isn't going to also generate embodied emotions. Even though you may not articulate them.

17:20 Right. You're still intuiting them, you're still anticipating them. And so I'm not sure how the architectures pulls that out, especially to Daniel's point, where our form has adapted evolutionarily to the phenotypical situation that we find ourselves in. And going forward, let's say that sea levels rise and suddenly you have beachfront property where for millions of years you did not. How does that affect now your emotions?

17:57 Right. I like the architectural idea of being able to take something and look at the historicity of it and be able to explain it out with a little bit more confidence. I'm not sure, though. I'm still struggling with being able to figure out how, if we want to look ahead, we can necessarily set it up the layered way it was set up here. And I really like figure four, but I just have to categorize it and keep it where it's most useful, which is in the backwards facing part of this.

18:34 That's all.

18:38 Daniel:
Okay, let's return to how to use this. That sounds like a good question, but we'll get there. So Lou brought up the brain and compliment with the body. And this was just the quote from the very final lines of the paper, or near the ends, where they acknowledged that partial perspective that they presented in the paper. Partial learning.

19:04 Just a portion of, but also partial, like partial to a certain view because there was a neuroscience regime of attention. And that's the contribution that was made that helps apply to other systems. And then there's another synthesis to be made. And this is kind of like the opening where they just acknowledge that cognition can be extended and embodied. And it suggests that not all aspects of control need to be solved by or represented a central generative models.

19:33 So it's this point that we've come to from many different angles from like, the ecological and the cognitive and then here from more of a structure learning perspective that like the representation, whether just in the model of a system or in the system's own self representation. It doesn't have to be the exact same thing as some latent variable in the outside world. So it's possible to have like an effective model of climate action without a perfectly accuracy model. But that has to be understood within like was brought up a space of counterfactuals of what could happen from the outside world where we really never can eradicate the total distribution and also the introduction of the unknown unknowns. And so that's the fundamental openness of the external states and the modelers view.

20:33 And then there's still a lot to explore with the counterfactuals of the models that the system that like the agent is can implement. But I think that gets a little bit of how to start to use it in a certain area.

20:59 Any thoughts or what? I got a question for you, Dean. When you study dance, for example, and they're foraging and I think it's section eight or I forget what section that they say foraging really, really matters. I think if we're point toing this now, I think that the sort of the go forward piece of this is now, when you started looking at ants, did you start from a cognition? I don't know what your background is, but did you start from a position of sort of pulling apart the historicity of why that thing is now sort of laid out the way it is?

21:34 Dean:
I know there's a piece of that that's there. But is that the place where most people that begin a study start? Which is kind of interesting about this paper because to me going from existence to learning implies a certain history. And so I'm wondering how you sort of come at things when you see them foraging where history fits in terms of that you gaining some understanding of what's happening here as opposed to then what happens.

22:13 Daniel:
Great question. The topic is historiosity and foraging for 400 from Dean Bleu. What would you say on the topic first? How would you bring that together? What would be some important points?

22:28 Blue:
So I really like this idea also and I have a completely unrelated point and then an unrelated question. It says mechanisms of memory and planning have evolved from mechanisms of navigation in the physical world. And some super interesting that I thought. Side point was if you've never read the book, “Deep Work” by Cal Newport. He talks about memorizing a deck of cards.

22:55 If you haven't read the book, read the book. But it's interesting because he goes through a very elaborate way to memorize a deck of cards and it's through building a memory palace. And so in this memory palace, you have a place that you know very well, like your own house. You can imagine walking through your door. You imagine the porch, the doorstep, how it looks.

23:20 You imagine the doorway, all the rooms, all the closets. Like, you imagine going through your house, taking a tour, and exploring each room in a very systematic way. So it's very easy to memorize a route you would take to your house. Like, say you're looking for a ghost in your closet or something, hiding in your house. How would you look?

23:38 So you go through your house, so you have this palace, right, that you have in your mind. And then he talks about making every card a character, right? Like, the queen of hearts is Queen Elizabeth or something. Like, who is each character in the cards deck? So each character in the card deck becomes a character that you memorize.

24:03 So you memorize these 52 cards as people that you know. And so this is an interesting point also on the socio dynamics as well as the spatial dynamics. So each character in the deck is a person that you know well. You memorize that 52 people correspond to these 52 cards, and you know the layout of your house. So then when you can go through the order of the deck, anybody can do it.

24:25 You can repeat back the order right away. So, like, you go through the cards and you say, okay, you can go through it three or four times, and then you know the entire order of the deck of cards. So this is how to repeat back the order of a deck of cards just to memorize the order, like, oh, they're supreme hearts, she's in a positive, or who's standing in the doorway is like, the first part, right? And so, you know the tour through the house, and you place these people in this spatial orientation in your house. And this is like a proven solid way the people that do these kind of memory trips use by constructing this memory house.

25:00 And I thought that that was super interesting in relation to this paper, that memory and planning have evolved from mechanisms of navigation in the physical world. And I also wonder if there is, like, a social role that also contributes to memory, and could that social role be this kind of higher level? Allostasis right, like, we talk about the homeostat and that's like, there's no other like, you've got one time step, right? And then in the out of step, there's other ways to kind of modify, but how many in the social components to succeed and function and to actively have memory and prediction? Like, does a tissue have memory because there's other cells in the tissue?

25:42 Does an ant have memory because it can spatially navigate for food and also it knows its friends? I'm sure ants have ways to distinguish one ant from another ants. Like, they're very social features. I don't think that they all think that they're all the same. Like, my cats don't definitely don't.

25:59 My cats know each other. But if you showed them, like, a stray cat, they'd be really upset with a stray cat, I'm sure. So, you know, animals know each other, I'm sure ants know each other also. And so what component of memory is also social? And that just Jelle me points to help remind me of that or brought that to, like, sorry if I have any questions.

26:21 Daniel:
Awesome. It brings up a lot of the aspects that I could add to now just from a specifically anti

perspective. So you mentioned the evolutionary aspect. It's one of the themes of the paper that perhaps from foraging, but other processes, too, as key cognitive demands, those early challenges or, like, ancestral challenges giving rise to current life were faced and met in this way. So that was the evolutionary aspect you mentioned.

26:55 And then thinking about the history at multiple nested scales. So, like, there's the history of the ant colony as a genuine modelable entity, where the colony, year after year kind of developed in its foraging territory as its colony personality or basically phenotype at the colony level evolve. And then asterisk one then at the nest mate level, it's the recent interactions of a nest mate that induce it to forge locally, meaning retrieving a seed. So, like, there might be an ant whose current neural or hormonal state is such that a lot of interactions with incoming foragers don't stimulate more foraging. But then there might be another ant whose status is such that a certain reduced rate of interactions with other foragers, maybe even no interactions with other foragers in the entry chamber cause it to go on a foraging trip.

27:55 That foraging biology is specific to this ant shown here, the harvester ant. But maybe other ants, too, care that asterisk too. Then there's the ant tissue. And, like, each of the ant tissues, the brain, the different glands, the gut, they're all undergoing kind of stereotyped transitions that are called temporal polyethylem, or they're going through, like, recent changes, like a leg was lost and then the least of the body, like, adaptive. So that's, like, the history of each nest mate tissue.

28:26 The asterisk kind of combined for me personally, but also just more conceptually with thinking about the relational view on modeling the ants, like, not taking too strong a view on where phenomena should be located in terms of nest mate versus collective cognitive features and then also the place based and the field and the long term experiments. So my advisor, Professor Deborah Gordon, has done, like, a 30 year plus field study in Arizona. And so I was doing five years of fieldwork there, but it was, like, year, I don't know, 20 something or 30 something to some other number five years later in this space where there was a lot known about this ant with a long lived colony, but with turnover of the foragers. And so the sort of ecological and the place based nature and then also the specific biology of this ant where the forager was just so clearly an expendable part of the process of the colony, like sloughed off and some other features of the foraging biology of these ants. Pretty interesting system and it reveals a lot on different cognitive demands and different foraging environments.

29:48 So that's not even getting into the whole mental foraging or active component, but just at the ant foraging level.

29:58 Dean:
I think you brought up an important point, though, Daniel, and I'm not sure if people associate architecture with abduction, but your advisor, supervisor, whatever her title was, her 30 year window and then where you stepped in for your five year window as Nested, you were the one now Nested within a larger time frame. So back to Blues earlier point, but you weren't historically looking back from the perspective, okay, here's the evolutionary nature. You are now seeing yourself as relative to

something that had a greater timestamp, right? So you walked into the party and you had to ask yourself, did these ants come from this colony? So back to Perth's, big question of right.

30:58 So it's when I come into the situation that matters as much as if I can induce a skyscraper or deduce how much structural rigidity that requires. Right. So I think this is kind of interesting now because I'm not sure if we should assume, like when you were talking about the phylogenetic trees, sometimes we just assume we're stepping in on the branching, but I'm not sure we should make that assumption. I think we have to be careful to be, as you said, very explicit, not just assumptive of where we come in in this history. Because if we don't, we can end up thinking that this is sort of a linear layering that figure four, when maybe it's more like identity.

31:49 It's both stable and dynamic.

31:58 Daniel:
So it's definitely related to almost the very first comments of Bleu, like, where's multiscale identity and multiscale systems in this hierarchical depth, it may be related in some way to multiscale systems. We could draw that out. But where is that nesting that we're describing, the real physical enclosure sense? And this way that mental systems can almost be modeled as Nested, just like we saw mental action as a nesting of a cognitive model with an action model inside of it. It's like the mental is encoding and even modeling through inclusion of the Perry personal space, but then with like a drop off going further out and then what is that modeling space?

32:57 And is there another operation here? Again, what are these graphs? What are we seeing here? Dean, my question is to answer Blue's question. Do we get closer to a satisfactory answer by looking at the mechanics of this or do we get a better sense of it by seeing the relationship.

33:20 Dean:
Like when you were explaining the ant part you've spent a lot of time talking about relationship and I think maybe one of the things we try to do is we try to apply the operation to it when in fact what we should be doing is seeing the relationship in its entirety. I don't want to say zoom out because you can still zoom out and still get bogged down in the sequencing because that's kind of what this does. It says this layer of sequencing can now build out, fractal out into these other layers of sequencing. And I'm suggesting rather than seeing the layers, maybe it's the totality of the relationships. I don't know because Bleu also pointed to the social piece which is all relationship all the time, right?

34:10 So I'm just wondering if we help ourselves if we collapse to oh, well, this is now explainable. Because we have some way of showing the sequence. Sometimes that helps, but maybe in this case it's from an evolutionary standpoint it doesn't matter whether we're talking about 1000 years ago and today as much as the relationship of what 100,000 years ago and today represents. I don't know, just throwing it out there. Thanks for sharing it.

34:47 Daniel:

Just one thought on that and then we'll continue here. Phylogenetic trees with genomic data are usually based upon sequence alignment followed by the fitting of a tree on the aligned sections of the sequences. And so there's a lot of ways and approaches of modeling from the sequence alignment itself from unaligned to align sequences, a lot of ways you could do it and then from the align sequence matrix to the tree model there's a lot of ways to do that. It could be a bifurcating tree, it could be like a different kind of tree, different algorithms, Bayesian, et cetera. One interesting point that many have raised, but it's very challenging to deal with is that depending on how stringently you align the sequences different Ines with different histories or different information content partially or in totality are being filtered out in a different way.

35:47 So if you have a mode of evolution where like short chunks get interchanged so that locally the sequences are not changed like domain of a protein aren't changed but then they're swapped in order, those might become nonalignable sequences. And so the hope is that those kinds of genomic changes are silent with respect to the longterm evolutionary history of a species. But in the short term other methods might be needed. But in the long term it's the hope that by pulling back the alignable sections you can get a good signal. And so that's a pretty satisfactory approach when multiple converging threads do overwhelmingly agree.

36:27 But it really shows that when there's like a signal processing, filtering and then a model fitting the information in the data can be extremely powerfully misleading because there's so much because you're getting exactly the same post. Normalized data, whether it was total noise in, total bias out or whether your correction process over normalized for one thing or another, those are kind of in that gray zone, unknown, unknown space coming into the searchlight. And so there's the whole looking at the searchlight story. So maybe where do we situate active inference with respect to that looking where the light is searching for one keys, etc. Etc.

37:22 Dean:
Can I just say real quick, I'm not advocating for don't show your work. I believe that showing your work really matters and then when you're trying to explain what's happening at a certain point, don't get stuck in, well, this was the order of the steps that I took when I showed my work. Try and zoom out and try and understand what the relationship as well. That's all what I'm asking. Does that help being able to toggle back and forth?

38:00 Blue:
So can I just jump in here for a second and point out that we had phylogenetic trees or we had yeah, I guess they weren't called phylogenetic trees at the time, but we've had like speciation trees long before we had genetics, right? So we can construct them based on morphometric analysis. And something that I liked in this paper is that it really drew out the affordance of the niche environment with respect to implication. Because we saw many instances of species, because we did these trees when we didn't have genetic information. And so we had all of these morphometric characteristics that we thought were similar.

38:44 And so we lumped these species together because they must be related because they share this morphometric feature. But in many circumstances, that was not the case. And so because of the similarity of the niche environment or even existing in the same niche, we had evolution converge on this feature that was successful in that environment. And so I think here the idea of duplication and the authors got into that, which I thought was really cool. So the same mechanism can evolve more than one time in parallel just because of the fitting to the niche.

39:22 Daniel:
Totally agree. Bleu so when I said multiple conversion lines of evidence, what I was totally thinking of was how genomics the patterns that we get from those aligned sequences do recapitulate and even discover unexpected but later verified relationships. And so that's sort of like a measure of the unique explanations and predictions. But this tree could have been fit with phenotypic characters, as you mentioned, as it was for a long time and then it was just a few sites in proteins and that was sort of like the 1960s and 70s was fitting with very reduced data sets and now the amount of genomic data is rapidly increasing. But out of the genomics analysis, the action state of phylogenetics is really just this tree topology fit from a family of trees, hopefully with an uncertainty associated whether they show it like with a bootstrap uncertainty or some other way but then that is an artifact.

40:24 Now, there's a few cases where, like, this is in and of itself interesting, like when people have resolved some deep relationships within insects or other groups and then it helps structure thinking that can later be brought out about different traits and ecologies. But it's like this is just a starting point for thinking about evolutionary histories. And like you said, similar ecological or generative process processes, dynamics might beget increase likelihood for a certain generative model to arise. And that's like kind of what ecosystem is like a desert seed harvesting ant whether it's in the Mediterranean or the southwest US. Like polka miramax in the US.

41:13 And then cataclyphus in the Mediterranean regions. There's some similarities there but then there's other species that are more closely related that have a different foraging biology. So then, how does the foraging strategy sometimes have extreme fidelity with a group like leaf cutter ants where once they're committed to being a full leafcutter ant, so to speak, evolution or they haven't been observed to revert back some other mode versus other, like, foraging strategies at the group level, at the colony level. But it's the individual, in another sense that's like very interesting thinking about foraging phenotypes on this tree. And why not?

42:00 Just like any other trait that you could study like the evolution of first spatial foraging and then the evolution of cognitive foraging and then applying that to different search spaces which was also explored a little in the citations here the foraging in mind and foraging and semantic fields papers what would you say? Bleu? So it's clear that you brought that up, Daniel. And just because in our past conversations you've brought up that the same species of ants can have different foraging patterns based upon the colony that they live in. I think that that's super interesting and I wonder how could you map foraging patterns?

42:45 Blue:
Like can you put foraging patterns into a phylogenetic tree? Is there a tree structure that we could use to describe foraging patterns and would that relate to or map onto this brain structure tree in some way?

43:01 There's a lot of diversity within the same brain structure. Also like the ants have the same brain structure but even humans have like some of us go to the grocery store. Some humans out there are still like hunting and gathering to least their food. Awesome Dean, go for it first so. You guys can learn me school me on something here.

43:25 Dean:
Is there any evidence that ants all gather in the morning and the queen aunt or whatever the organizing and the principal ant pulls the map out and says now I kant you to take this thing that you've mapped and carry it out today. And I'm not being a smart alec I'm asking is there some sort of analogous evidence that they do that?

43:54 Daniel:
Okay so to the team huddle Monday morning stand up. No, I hope that's not overgeneralizing because ants are also of course very diverse in their foraging strategies and a lot is unknown. But just speaking to the first approximation the queen is not a central decision maker in information router hormonally and at a slower time scale for sure, but at the behavioral regulation level. Now in the morning for the poke and the mirmax they would like, chill and some would slowly come out and like their walking speed is temperature dependent like all insects and it just assessing maybe the humidity or damage around the nest entrance. So there's no reason why those archetypes like convergence before planning or rupture repair which of course cannot have a planning component directly.

44:54 Different algorithms are going to have different levels there. But because the Kant symbolic behavior communication is pretty flow bandwidth probably. So unlike the waggle dance which is conveying like quite a lot from an information theory perspective using symbolic means like the direction and the distance and also chemical information like about the flowers there's ants that like physically pull another nest mate to a foraging location. So using memory or pheromone or like tandem run and there's probably some limited examples of being able to do something like a little bit higher order signaling and also, of course the smell of the food and the pheromone and so a lot of the extended features but they don't do a powwow with the queen. Okay, so let me take what you just said and sort of feedback and maybe build on it a bit.

45:53 Dean:
So you describe the situation. It's morning if morning then ant behaves like so and in the last slide of this deck and the phylogenetic tree you had then to now or then if. So what I'm wondering is can we exist in a sequential world of if then like as a homeostatic representation is if then all we need. But if we want to include temporal depth and hierarchical depth we actually incorporate the inversion of that. We incorporate the then if.

46:32 Right? So that's my question. So to Blue's point earlier, where is this shift or this sense of what's going on? The relationship might be the discrepancy. The first representation figure on the discrepancy might be those people that are looking at things from an if then perspective.

46:58 If this then we should expect that from a predictive standpoint versus the then if perspective then the critical call had to be made as the base runner was foraging for second base and a call was made. Now if not a whole bunch of people chime in about what their interpretation of what just happened. Right. So again, I'm not sure if only looking at it frontwards and backwards as a sequence is enough. I think we also have to be able to see the relationship of if then and then if from the historical view.

47:47 And I think when we do that we go from architectural existence to architectures learning.

47:57 Daniel:
Okay, let me try to restate that but I've not lost the forging trail. So if the outcome is like the end point then that's the if then if the person committed the crime you do the time and then the sort of then if it precludes if then in multiple sequences of if then and uncertainties on if thens and then it says okay then or it's like if then if. So we're kind of downtown if and then that's where the outcome is a starting point. And so that is in one sense it's teleological because it's end direct. Where is this project going to go?

48:43 Do we really expect and prefer to publish this paper on February 28? And then also it includes the next steps. So it creates a world model where there is a possible which is a selectable set of policies that the team will achieve like that expected and preferred outcome and then it sets it up for what would be the obvious next loop of policy selection. And so if it's like if things are this good or bad with the environment in 2030 then this is how much damage will go to this or this way. Even if you were right it wouldn't involve the policy selection elements.

49:22 And so that's probably a complex situation with the partitioning of like estimation and world policy selection which is clearly a separate discussion that's dot one being invoked as an example but it just points to how those are different components of inference. And then foraging like an example of if then and then as if or then if was like sometimes the colonies would pack their nest entrance with pebbles and then there'd be a thunderstorm in the afternoon. I don't think a study has ever it's a totally widely observed behavior with many animals and there's no exception with harvester ants. Now maybe there's thunderstorms where they fail to make pebbles or it not enough and there's probably some false positives too. So we don't really know the exact success ratio.

50:12 Maybe it's better than the previous days forecast, maybe it's worse. But it's very clear to see that the nest mate doesn't need to have been inheriting from its evolutionary history and its niche rain cloud modeling software. It's just picking up pebbles with a activity to interactions and humidity etc. That help colonies that have that proclivity succeed in that niche. Which doesn't even mean that it was a popular phenotype like 20 years ago as things change in a region.

50:47 So that's like super interesting.

50:52 Let's return to foraging and the foraging variation question. But first, let's see if we've completed the figure four discussion. And just like, is there anything else we want to add about, like the main? Because this is kind of what the paper works towards. And then there's, like, a little downhill role is, like, we could do this algebra on the free, but this is sort of where things get led up to through demonstration of the motifs and the connection between cognition and structures and structured learning and then, like, parameter learning at a behavior or developmental time scale and then structure changes over evolutionary active states.

51:35 So we have hierarchical depth, temporal depth, which we're also just kind of adding in a spatial component allostasis which was used in the paper with the intermodality. But maybe it could be other modalities or same. There's duplication, which is kind of like parallelization that's like adding another nest mate or another unit to the compound eye of insect developmentally. And then I is homeostasis, which is being used here as sort of the base or root or ancestral case which is referring at the behavior time scale to return to a stable point. And then also at the more evolutionary time scale, the eye operator on the homeostat, it's like the homeostasis on homeostasis.

52:25 It's keeping the structure of that model fixed. And then when the homeostasis at the evolutionary time scale is open then the structural transformations here are like the affordance from that base case, which is what's being shown, like, with the main one on the left being broken into these. And then each of these become the prior for each of these next ones.

52:52 Anything to add before we go to forging? No. Okay, no.

52:59 Great relevant question. Bleu measuring foraging. So one paper, though, there's a bunch that are really relevant, but Michelle Dean 2014 did a lot of literature review and visualizations of foraging strategies in different ant species. So these are like some of the foraging approaches, like foraging archetypes and then categorized those into different traits on a tree.

53:36 So that's an awesome work and something that could and should be extended and integrated with like, modern databases so that somebody could submit, we just saw this one going in a trail. And then that kind of reduces our uncertainty about a broader model of ant foraging because it's like, oh, that's a surprising conversation about this given what we already knew about that species, for example. But how do you actually measure foraging variation within or across the species? So in this case, in figure one, these are like kind of like the structure at one level, they're like archetypes of structure. Now, sometimes they probably blur within each other.

54:22 Like, if it's two, then it's tandem running. But if it's three, is it a group or is it still like pseudo tandem running? You know, so it could be a little blurriness. And also, there could be different modes even within the same species or like in a continuum between a foraging column with, like, a trunk and a spread versus like, just a spread. So this is sort of like I've.

54:47 Blue:
Dot two, jump in and say, like, how these look like neurons just with different axons and dendritic fibers. These are like different brain structures. So I just had to point that out. Yeah, the bottom one definitely does. Even both of these and even this top one too, the trunk trail.

55:12 Daniel:
Let's just say that these are sort of architectures within a broader space and then within a given mode, then the behavior ecological will tend to make like a more specific measure. So let's just say it's 100% a trail forger. Then having the counts to and from on the trail is a good measure. So like in the harvest ranks they most basically do this trunk trail on the top. So there would be like one or two or three main avenues that they were leaving the colony in for that day and then sometimes it be the same day to day, other times it would be slightly different.

55:56 But then what happens when one wanders off not on a trail? Is it still a forager? And a lot of that space is really interesting. But suffice to say that once a foraging definition has been selected, like not just the class that it's going to be, but also how it's gain to be measured in that study, then there can be like quantitative parameter comparisons like this colony foraged at this rate in terms of number of outgoing foragers per minute. And then its neighboring colony of the same species was doing at this rate.

56:31 So you can do like parameter comparisons within the species kind of at the population level. And then there might be another ant species whose activity it doesn't really make to compare because it's like a different number of nest mates, it has a different foraging ecology. So that's kind of where the structure learning would come into play with comparing the ants that are encoding the scale, insects versus foraging for seeds versus doing whatever else they do. And then within a species or perhaps across a few species that have a similar foraging ecology, it's possible to do parametric comparisons.

57:17 Dean:
So I have a question for you Daniel. How does foraging fit architecture? What do you think? No, I Dean, I'm not trying to be a smart allocator. I'm really curious because most architectural analogies result in some sort of an outcome that's static.

57:35 Like there's a lot of processing that went into the outcome but the outcome is static. And I don't think that any of the representations in the frame that you're moving right now, although the differences are appear as static, they're supposed to be representing something that's quite dynamic. So that's where the question of the architectures piece blended with evolution might be helpful or it might be oxymoronic depending upon how you interpret it. Right.

58:13 Blue:
Architecture implies like there's an architect. I don't know, in an evolutionary sense I'm not sure like architecture is the right word. It's weird. Yeah, but I'm not saying it's wrong. I'm just saying you have to

be really careful and then what assumptions you come in with based on what you think architecture is.

58:37 Dean:
Yeah, maybe you're thinking of the way I am. I'm not saying it's wrong, I'm just saying, what baggage does that carry? When we're talking about something that's actually quite dynamic.

58:53 You can see the ants and the ant. Hill has an architecture. Yes, you can take a snapshot of it, but is that what we're really talking about when we're talking about foraging? So what's the difference between a structure and an architecture? I wonder that.

59:08 Yeah.

59:12 Blue:
It's just semantics. Yeah, I think it's more than that. But we just have to be careful that we don't insert one for the other, thinking that they're the same thing. That's what the whole ontology exercise is about. I mean, I got my mind changed.

59:27 Dean:
Okay. Normally I'm the one that argues that don't give somebody an owner's manual. Get them behind the wheel, right? They'll see their owner's manual later. But in this particular case, I'm kind of going, wait a second, we better all agree what we're talking about.

59:43 Blue:
So here it is. What's the difference between structure and architecture? Structure tends to be the pieces and the parts of the design and how it goes together and the finished object. Architecture defines how the final object looks, behavior and costs by designating, how design choices are made. So architecture implies that there's some kind of agency behind the design.

1:00:06 And that was kind of my interpretation of it also. Okay.

1:00:13 Daniel:
Bunch of stuff. There that last point, like the definition of architecture involves the choices rather than structure involving just the description of parts. So that's kind of like the perception, variational, free energy. That's like structure.

1:00:32 Blue:
It also means the entire thing. So it's not just the parts, it also means the entire object. Yes, but it still doesn't involve preferences or choices. It's about description, like perceptual processing is. And then the architectural definition has choices in it and then that is what you brought to like agency and also architecture.

1:00:51 Daniel:

It's like a profession and an operational approach to modeling the process of building and building, though there's probably theoretical and basic architectures and all this other stuff. Okay, we're just a few thoughts here on like the engineering and designing. So I guess one ontology that can be helpful here is with Bucky Fuller talking about structural integrity, which is something that has static holding strength, like a metal cube has static holding strength. Then there's process or the pattern integrity, which is sort of like describing the life cycle of the butterfly and the caterpillar, like a recapitulation of process. And that is a lot more like this sort of dynamic systems approach to organization rather than the build it and block it approach because these multiscale systems always have things that are turning over at different timescales.

1:01:56 And then there's just so many mechanisms that that can be modeled by or actually occurs by. Like just in the harvester ants. It's the historiosity of that colony over the years, then the nest mate and that day. And it's visual memory and just how many factors come into play then every foraging trip at the nest mate level could be seen as unique and the exact communication will just never be recapitulated. Some of the major ones can be explore and then controlled for, tested for in a lab.

1:02:30 But that's part of the openendedness of dynamics systems because by changing the setting, you can study certain features. But then it's another step that's impossible in practice to understand how those features that were in a different environment translate to the first environment. So you can take the essential behavior of the traffic jam and then recapitulate it with the simulation and the lab experiment and all this. But it's a different thing to actually put it into the design time, which is happening at a different time scale than just building the skyscraper and leaving it there. But that's more like the dynamic organism.

1:03:14 So where's foraging on this?

1:03:20 Dean:
And I would ask, is there a different kind of forging? If we're talking about sort of the sort of the machine learning type of foraging on this versus the wet foraging of an ant, can we actually see those two things being quite different in terms of how those two foraging types plays out? Or do we focus on the things that are different about those two things? Or do we focus on the things that are the same about those two types of foraging? Bleu also with your foraging from a more culture and human perspective, social machine learning and foraging.

1:04:13 Blue:
So machine learning and foraging, I've done some modeling of foraging with, like, agency modeling. But in machine learning, I wonder, like, I mean, I know that there's information forging but that's usually explore from the human concept, right? Like humans going out and looking for information and how long we stay on the website. And so but I wonder about machine learning and neural networks doing object classification. Is there a foraging that they do there?

1:04:50 Like, you know, they pick up things like, Chris Ola wrote a really awesome publication about how we understand how neural networks work. And they kind of deconstruct a deep dream that's onto Still Club, which is really cool. It was like a 2018 paper. But they look at how does a neural network recognize a dog, a dog in a tennis ball? So they look for like, round object and like, you know, floppy ear.

1:05:20 So there's like different neurons in the neural network that'll learn what is a floppy ear or learn what is like a round object. And so it's this combination of different things that the neural network learns that piece together to form. Like, this is a dog with a tennis ball, which is a super interesting way to think about it. And I wonder, in training a neural network, can that be looked at as a type of information foraging, right? Like the network is setting out to learn distinct pieces of a pattern that it'll later piece together to make it complete image recognition.

1:05:59 So it's interesting. I'll give a thought on the machine learning and the bio origin. So you mentioned there's the question of should we approach it from a similarities or a different perspective? And there's different reasons why you might approach different angles. Right.

1:06:20 Daniel:
Like what is the most pedagogical to someone learning about the difference versus what is the best way to approach this to a certain stakeholder, who is actually applying this question in a certain context? But to get to the question outside of that, the idea of contrast came to mind because we recently had the model stream on contrastive Active Inference, which used machine learning. So check out that stream, but it involved a contrasting step of processing inputs before doing policy selection, etc. So that helps scenes that are broadly similar but have a key difference that's small, like where the ping pong ball is can be learnt better because there's a contrasting and that reminds me of language learning when it's like here's and here's and then that's a contrast that isn't going to come up in the language, but then it helps learn on that. And that's not just a machine learning thing because even in the parametric modeling of SPM, statistical parametric modeling mapping, but the model making of it, the contrast matrix is the experimental design.

1:07:33 And so if it's just condition A and cognition B and you have ten measurements, it's just five of a five zeros and five ones. And that's like the true or false matrix that basically gets multiplied to do statistics on. So whether it's a very classical, parametric, sparse modeling approach, which was initially all that was had by statisticians, or whether it is now things that are highly parametric, or just all this other stuff with modern machine learning, the contrast is like this key piece because it's always about what signals being detected within a data set and then also how the difference between that examples and the population is being modeled. And so like in parametric statistics that's about the sample standard deviation. And then there's a slightly different equation for, like, the population standard deviation if you would have kept on sampling.

1:08:28 And then it's also sometimes hard to incorporate that kind of reasoning into these more modern statistical systems where it's, like, very easy to overfit because there's a lot of data, but it's hard to model how that large chunk of data is related to even other segments of data. So it's similar to the phylogenetic challenge. But that is one that is like searching for contrast. Takes on a very different meaning.

1:08:56 But how is it related to foraging?

1:09:03 There's the info foraging like Bleu mentioned. There's the path element of foraging which might be related to inferring the best info foraging path for somebody on their feed because it has to balance the Epistemic and the Pragmatic. Or at least it acts as if it balances the Epistemic and the Pragmatic. So foraging could be maybe looked at in that way.

1:09:33 Dean:
I think what this slide represents is broadly speaking, not just or broadly identity. I think what we do is in that diagram, we tend to differentiate. We see the levels because we've created the partitions. We see the dependencies because over on the far left we have actually Bleu arrows that show what those dependencies are and then we reintegrate, right? I remember when I was doing my I come back to the program that I was doing, but I would say to the kids, I mean, there's a migration piece to this that we have to develop, which I would consider to be the tactics, which would include foraging.

1:10:17 And then there's a tools use or a strategy piece, which is sort of the information covered as changed as measured, which is kind of the instrument piece. And what I kept saying to them was you kind of need to have both all the time. You don't need to know the difference between something that's a migration or an action or a fulfillment in the sort of the physical space and the tools sort of the thinking instrumental parts where you can turn what you're seeing into these models and these representations. You have to kind of be able to marry the two of them all the time. You have to be able to integrate them across these boundary when you're trying to get into these professional settings.

1:11:02 Now, that isn't a construction exercise. That's being creative because you really don't know at the end of the day how it's going to turn out. But the big point that we were trying to make was don't leave one of those two things behind. Like don't get stuck because you can't go out and forage don't get stuck because you don't think you can draw the perfect representation of the world. It's just an instrument, right?

1:11:31 I mean, even a spoon can eventually drain an ocean if you're prepared to spend enough time on it. So I think that's what this representation I think that's one of the greatest gifts of this representation is. It kind of lays out very clearly that there's two parts to this. How do we differentiate eventually on different duplication, parallelization dependency predictability, preferential time based and spatial based and then scale based and then reintegrate it all back together into some sort of a coherency? For some people, though, I think it's kind of dense.

1:12:16 I mean, if you walked in today, gee, I don't know what kind of prerequisites you would need, what kind of priors you would need to understand what the heck we're talking about. But as I said, for me, the entry point, the lowest level, was do you migrate? Do you tool use? Most people already admit they can find examples when they do both. So my question then would be, so will you know that you're doing both going forward?

1:12:43 And if you can do that, maybe it's not quite. So quite so dense or quite so mystifying.

1:12:56 Daniel:
So a few things. There so great point that going foraging on your merry way without having all hidden states learnt at the structural level or the parametric level, even conditioned on a structure. So you don't even know what's out there. No cognitive agent does. And at the structural and the parametric level, the inference challenge is about action selection under uncertainty of multiple types.

1:13:27 The inference challenge is of action having subproblems involving inference on perception and causal modeling. But those are like little subproblems within a broader frame that active inference provides of like action prediction, action on policy selection. And then foraging is a great setting to study policy selection. And then you brought up like that. There's prerequisites that we hope to learn what different backgrounds can come to active inference from.

1:13:57 But I think one takeaway from the paper would be that perhaps archaic in a foraging structure, whatever that means. But the way that we forage through linguistic and idea space and solution space, strategy space can be modeled by some graph motifs and classes and transformations, which is what the contribution of this paper is. And they also describe some of the more basal or homeostatic ancestral structures. So it's like you have a linear model that works for height data and it works for weight data and it works for height and weight together. It's just any numbers you pass to the linear model, the GLM, it's going to be basically okay, this is like another model class that can take or apply to cognitive modeling, homeostatic modeling, cognitive and homeostatic modeling, all these variations that are being described here.

1:14:55 So that's just one way to put it with perhaps fewer prerequisites. And then these were interactions questions. Is this what it was? Do you migrate and do you use tools? Yeah.

1:15:09 Dean:
So when you're talking about tools, are you talking about provisions? And especially if you come to breakfast with a plan on how to build a spoon, how is that going to help you eat your cereal when your foraging exercise was to be able to move the cereal from the bowl into your mouth? Right? So lots of times, especially in some of the more conventional ways we look at things, we tend to take the plan to the cereal bowl instead of the actual spoon. So the provisioning part is more than a plan, it's more than the words and it's more action oriented representation.

1:15:52 It's the actual implements that we carry with us when we do that migration that I think really counts. Because if you carry, I guess you can origami a sheet of paper into a spoon, but then that's another process that you have to go through. But really what you wanted was just to have the implement.

1:16:18 Blue:
So when you talk about migration, how far is a migration? Like if you're a single cell organism, like migrating through the ocean.

1:16:31 Dean:
Maybe the migration in that case is moving around. You right. You don't have to think migration in terms of me from A to B. If I'm stationary and A to B are passing by me when the car beside me is pulling out of the learning lot, I feel like I'm moving forward. Right.

1:16:54 That's also a migration. So that is related to what we talked about with some of the geometry, like the projective geometry and how there's the flipping between the egocentric geometric projection where it's like your peri personal space and it's kind of like a distribution around you. Like there's me in the center or located somewhere. And then there's like, things closer to the eyes are generated to be bigger because they don't leave to be because we know we're making the generative model anyway. So why does it happen?

1:17:27 Daniel:
Why can't I see it from Grand Theft Auto perspective? It is a certain way for a certain reason. So then that is flipping. That's flipping with the more, like objective or D personalized stance, like the prior on the architecture structurally in my room with a right angle. And even if it looks different than a 90 degree angle in the generative models, that's perceived as normal or unsurprising because it's compatible with the underlying structural hypothesis that's unobserved.

1:18:01 And then the observed structural observation that has to do with the actual way that the angles look. So here's just one other kind of Professor Gordon shout out angle to look into evolution of algorithms for collective behavior. So let's connect this to the cognition foraging and then Dean's questions. Do you migrate? Do you use tools?

1:18:30 And do you provision? So that's like related questions. So that one might be like, how do you adapt when things change and you can't adjust certain things. Or also the difference between, like, let's just say I'm using a software I'm not very familiar with, like Photoshop between doing one move photoshop searching on the search engine. How do I paste okay, paste.

1:18:57 Okay, I do one move. So it's like a one step model of function versus someone who just rolls out of bed, opens up Photoshop, and then they just open their mind to being creative and being in feedback with what they actually do. And maybe that does involve still using a search engine, but they're in a functional flow with respect to that tool and they've just set aside the time so that their regime of attention can be, like, productive or aligned with a project. But they're not going step by step. They're engaged in broader organization.

1:19:35 Dean:

So we're talking about figuring out part rather than the finding out. That's what you just described the difference between those two things. Yeah. Opening space for figuring out, for figuring out. And that's sort of like it's a little bit like local and more measle, or global foraging, which is like you can dig in a given patch of dirt, so it's super labor intensive.

1:20:03 Daniel:
But it's a local finding of what is there versus trying to find a better Dutch of dirt and dig their first seed, maybe even find an exposed seed. But whether one of those policies has a higher risk of reward or energy expenditure that's gain to depend a dot one what one believes about the frequency of like half emerged seeds and seeds in this location and the time cost of moving from location to location. There's just a ton of features and it's not just like as simple as making a heat map of where the seeds are.

1:20:38 Dean:
We would describe it as mimic foraging and mimic foraging and what was the other one modify foraging. Some people might believe that for it. So it's interesting at active inference I was at last week, they talked about how far can you get mimicking an expert and there were several talks. So whether it's navigating a maze or figuring out whatever so like they'll train a model on a whole bunch of expert, what would an expert do? And then they set them loose in like a novel circumstance and if the novel circumstance is similar enough, it kind of works.

1:21:20 Blue:
But really, like, the better thing to do is to put some exploratory components in there so, you know, you have, like, the expert training. But if there's an explorer component about, that doesn't leave to do with an expert or, like, you know, if you train it to navigate its own maze, also sometimes, like, alternating with its own maze or image with experts in it. So like following the experts only gets you so far. And I wonder, and something that's come to really mean a lot to me in the field of intelligence is when there's information. Like when you can do something new, when you can generative, model, solve any image because you learning from a few experts that's innovative or the ability to generalize, but also the ability to innovate and produce some new output rather than just do what the experts do.

1:22:13 Daniel:
So one thought that's awesome. Comments Bleu like learning from examples and from experts usually learning by example when used not in a malicious way is referring to seeing somebody who is better or least adequate, if not far, far better. And then there's this question of like should it be shown only at the expert level? Because the expert speaker would do tongue twisters and Russian sentences with five conjugations. But the expert in interaction is able to approach a zone of agreement and of like proximal development so that the example and the expertise as applied is in a very specific space and that's like the order of examples and then knowledge this person is able to conjugate this thing.

1:23:08 But they're having challenge in conjugating this case. So then depending on their personality, maybe you lean more on the one that they are more familiar with or maybe they come to you to be wrong. So there's a lot of advanced cognitive modeling and parameters that could be in a fullfledged model because that's a super complex scenario, especially in a group and all of these relationships on relationships. And we don't need to have that whole generative process model to have a generative model of policy that can be effective at foraging in that space. So it is a project and making that map is helpful for it.

1:23:55 But to make policy decisions in the cognitive space isn't the same as describing and mastering all the variables of the cognition process. And so this mental foraging and physical foraging and also what we talked about with a sensory motor detachment helps make that really clear.

1:24:20 How does active inference model relevant expertise, whether domain expertise or discipline, like a set of affordances A versus a whole different set of affordances, you know, B or expertise like in that language learning example or some other example where it's not simply about providing some advanced example and then having people asymptote to that. It's about opening the space for figuring out in that setting.

1:25:06 Dean:
I think you can find a structural adaptation in mimicking an expert and that typically sort of plays itself out as instructionalism. I think you can find yourself also modifying your niche and adapting in that way and then being able to go back and have an expert look at what you've done. And the reason why they're an expert in that particular case isn't because you're going to mimic them, but because you're going to ask them what they Hohwy, the process that they use to arrive at, where they have a certain subject confidence. They are able to now diagnose where your confidences might exist and where you might want to might want to re examine what the probabilities and preferences were that got you to the conclusions that you drew. So again, you're going to adapt either way.

1:26:10 But I think that the path that you choose matters and knowing which path that you chose also matters. So I adapted, right? I adaptive because I'm now trained up to be able to find the Bleu dot, the Bleu ball in the sea of plastic balls, right? That's one form of adapting because I've never been in a ball pit before and now I found the one that I was looking for. But another person could go in there and be able to maybe do something else with that Bleu ball because they modified the frame and the mesh and that's that.

1:26:52 Again, we just want to make sure that we know that there's a difference so that when we are reintegrating them at some point in the future, we know how we got there. That's all.

1:27:08 Daniel:
Okay.

1:27:13 I like the throwback to instructionalism and interactionism. It's an example. It was like latently there when talking about that niche dependence and the cocreation. So all these kind of terms digital stigma, g when working online, for examples, become implicit when we're using a given sequence to be talking about like a certain thing that we want to call our attention to. We call our attention to some example or to some relationship and then that modifies our cognition model so that we're thinking about things and doing things differently after.

1:27:52 And so it's kind of like quality of engagement rather than quantity because some people by reading a little, change a lot and vice versa try to set up the relationship. I think it's funny that you bring up learning. So for a very long time my PhD advisor never read anything. She would never read anything. Did she read it?

1:28:22 Blue:
She does like speed reads. She's speed reads, right? Which means she looks at every third word and then when she really sits down to read it, she's got all these things to say about it and wants to like, you know, wordsmith every sentence and it's great. We've had a lot of fun working together. I don't mean to make that sound evil in any way, I don't speed read.

1:28:43 I read fast. But I'm totally incapable of speed reading because I feel like I have to duplicate my efforts. Like if I skim it like I can pick up some headings or whatever but I never really dot two read it. If I kant to read something, I want to really read it, like go through it line by line by line and look up things, I don't know what they mean and this kind of thing. So it's a different and interesting approach how different people approach things in different ways.

1:29:13 Daniel:
Yeah, you know, talking about in this paper how we've been going between the sort of physiological and biological substrate of cognition and the nervous system and the brain and all of that. And then this more cognition modeling approach which can be sort of cut and dry or overly computational or so called unrealistic with respect to biological foundations. It's kind of like what Dean pointed to is like the bifurcation and so now things are kind of coming together, not in the way that everyone might have expected, but at least in a way and so that's a very interesting space and then Bleu with a fast and slow reading. It's almost like there are multi scale dynamics associated with attention that can be measured with the EEG but it's a hidden state what is actually happening in the mind and brain. But those are the kinds of experiments that people use.

1:30:15 And there's some rhythms like the fast cycles, like gamma that are playing out at the speed of perception and awareness of words. And then there's also slower symbolic or. Semantic changes in all these different timescales and it's just like making the space to least the connections and the terms and the implication resonate and connect with the learner and the creator is like a lot more valuable than any amount of coefficient of low reading. And so how to balance the information amount and type and sequence and all of these things with holding the space for just a slower transformative process. And that's very true.

1:31:13 Dean:
Yeah, that's huge, Daniel, because I think a lot of people that get into the instructionalism, you'll hear the expression and I'm not overgeneralizing here, you'll hear, Jelle, we're going to break this down, as

opposed to the interactionism, which is, we're going to slow this down. There's quite a difference between those two things. Open up and slow down, or we're going to drill down, finish it out, get it done in the time. And as we discuss, like, this is the archetypes or the dialectic. It's not like we're only taking one approach or another.

1:31:52 Daniel:
But it's very easy to see how the range in at least education space, not even thinking of other areas, is like, it's the amount of content and then it projects down into how it's delivered and then content gets slotted in or out, or things get expanded or not. Or the Monday's class gets canceled and it moves to Tuesday. So that's like the sort of instruction projection approach which is so exafferent in multiple ways, especially in how it's applied to thinking more about the opening up a space and slowing down and the learner's journey and all that. And when you're in the huddle and you're calling out the shorthand before you actually line up at the scrimmage line and snap the football, it's interesting. That what you point out, Lou.

1:32:44 Dean:
In my history, I was always prepared to slow down when it was math and the symbology. Like, I have no problem slowing down for that. But trying to read out long sentences, it just about kills me. I mean, my Add, I have to interleave like every 20 to 30 minutes because if I don't, my skin starts to crawl. Those sort of contextual factors that are literally weighing on you, you don't know those sorts of things.

1:33:13 And if you're an expert who's just trying to push out all this information in a 13 week window, I'm not really sure whether or not you can sort of get down to that level of personalization. But if you're able to, it makes all the difference in the world in terms of what people take away from the exposure. Right. I find it way easier to do the shorthand thing. If a picture is worth 1000 words, I'm in full agreement with that.

1:33:44 I just can't do the thousand word thing.

1:33:51 Daniel:
Yeah, really interesting. And Steven wrote in the Chat, good to see these questions about active modeling expertise being thought of in different ways, like the direction of travel, the distance to travel, the rate of travel, mode of transport.

1:34:08 Yeah, thanks for sharing it, Steven. Landscapes and sort of the sense making and the way finding and being on this cognitive landscape versus or and being on this spatial foraging landscape. So going out for the mushroom forage with someone who's familiar with mushrooms and somebody else who's familiar with the biochemistry but hasn't seen the mushroom in the ground. And then someone else who's on a different perspective because of their background or people with different physical abilities and knowledge in the spatial setting. And then there's the cognitive foraging.

1:34:46 And then just like, is it that the colony forages as a unit, and then that's the sort of flip side of the nest mate's forage collectively. And so the team forges as a unit collectively. And the individual is sort of engaged at this level that we have experience at in the policy selection relevant to our participation in some either niche or team or project.

1:35:18 Let's go back to paper, see if there's any other pieces that we could kind of look at. Let's look at the roadmap.

1:35:34 Dean:
Can we talk a bit about the endowment part of this? We haven't really touched on that. Well, just the endowment. What's the endowment? So in the .2s, essentially, now that we've built this base, a stable platform, what can we do with it?

1:35:53 Like, what leverage have we now potentially active? So what's the endowment here in terms of what the paper really? What's the new affordances that we didn't have before that we didn't maybe necessarily have explicated quite so clearly prior to reading the paper? Maybe that's one. I don't want to put pressure on Bleu, but what do you think?

1:36:16 Bleu? So I'm trying to remember. Maybe you guys can remind me in this figure. Oh, wait, so endowing with temporal depth. It's got to be an autostap, right?

1:36:29 Blue:
Is that correct? Yeah, because endowing a homeostat with temporal data isn't possible, right? Or they can only do and I wonder if that's like, a singlecelled organism can only do at that moment. Like they're not planning for it to get cold later or something like that.

1:36:55 I wonder if it's not only temporal depth but any kind of relationship. I wonder if it's not necessarily temporal depth. It's time. It's flat. Right?

1:37:08 I mean, I guess in order to modify your own behavior or thought, you would have to have some temporal depth. But this allostasis idea is interesting with respect to my relationship to my future self or my relationship to my previous self or my relationship to other people in my community or whatever. So I just wonder about if that's why this is the allocate and not the homeostat. Or like, can you endow the homeostatic temporal depth and get that same kind of ability for perspective and retrospective inference? You can't, but why not?

1:37:51 Maybe let's put some words to that. Let me see if I can find some of the paper. Nice point. I would say homeostatic function could be in a model with memory or anticipation. Like, the thermometer could record the last 100 time points, or it could have a prediction for the next 100.

1:38:11 Daniel:
But the function would be that it only turns on the heating once it gets too cold. So it still is the on the action selection side. It still is making one Markov chain decision based upon an instantaneous

measurement. And so it can be implemented whether just for a computational parsimony or in a simple biological system with just one time step. If it's this hot, then turn on the cooling.

1:38:44 It's the ultimate if then because it's like a threshold model and so it gets pared down over evolutionary time. But then you're bringing up like to even engage with a counterfactuals or some influence or something on future policy, there has to be like a depth. And then this gets back to whether it's modeling the thing itself. Like what do these represent? Are they the scientists testing a model of one, two, three or four five hour temporal depth for the ant forager?

1:39:18 Or is this the forager's quote, cognition model, as if that were somehow not modelbased science? Or is it the neural architecture or the functional anatomy of the forager? There's a lot of ways it could lean or be framed, I think, in a complementary way. But one will have to be really careful where support for one domain super valid and where is one not? Like if it were a circuit board with that topology, then it seems like pretty good to go to use that as a pretty broad descriptor of the system.

1:40:00 Whereas if this was a simplification of memory and there was an uncertainty parameter labeled uncertainty or anxiety, like we discussed a little earlier, someone slaps a parameter on one of these edges and starts addressing complex real world settings. How does that exactly work?

1:40:27 Dean:
So one of the things that I think that it endows us with is that the homeostat is the basic pattern. There's a partition between the epsilon and the mu. So the first thing that they did is they created a sort of horizontal partition and then when they went to the parallelism, now there's two epsilons and two muse. And so now they added the vertical partition as well. And I think that's really, really helpful in terms of at the basic level of the statistical we talked last time that you leave to kind of be a polymath.

1:41:05 But I think from a modeling perspective it's getting that orthagonal mesh established that moves us up the chain of complexity because then all of a sudden you can have diagonal dependencies and then you can actually add the z plane going through this in terms of how the scale is built out on both distance and time. So I think that's one of the endowments that it provides.

1:41:37 Daniel:
I could just restate that it's like when we say that one of the advantages potentially in application of active inference could be that models are composable, so they can be composed within a given scale, like spatially or temporally, and that has a visual representation. It can also be composed in this nested way, which as we've explore several times can represent physical enclosure or temporal enclosure, as well as this cognitive architecture, like mental action, depending on how different variables are interpreted in nested modeling. And so it's composable in several dimension of transformation which they represented in Figure Four in a way where each transformative step, even though the function might change a lot from the beginning to the end. Each organizational or structural architectures, cognitive, step or morphological, if one is looking at that layer, can be mapped into a certain space and direction of movement within that space. Do you think it can also be sculpted out?

1:42:48 Dean:
I mean, because we're looking at this historically, we're looking at all the things that happened in the history. Was this the stuff that was kind of carved out? And a lot of the stuff that doesn't matter removed in order for it to be more comprehensible, because I agree with you, it can be composed. But I also think it's not decomposed. But I do think or even deconstructed, I think it can be sculpted out of a greater history.

1:43:12 A greater whole.

1:43:17 Daniel:
Yes, please. Sorry.

1:43:22 Blue:
When you talk about a greater history, when you talk about historicity in today, like during the last few hours, and we've also talked about memory, and there's a difference, right? So historicity is like, historically accurate and memory is like what I remember or what any individual remembers. Yeah, we've talked about that before, actually, I have, a few livestreams ago. All I'm thinking is that what this does is compare all the things that is an Axle constant that said, how did he frame it so that we can understand why we're here now versus all the things that didn't make it right. That's essentially what I'm talking about in terms of explaining.

1:44:12 Dean:
Right. They're explaining why something got to the place that it did, as opposed to the necessity of all the things you had to do, all the things you had to compose in order to realize I'm trying to bring that back to a sculpting metaphor now. Awesome. So that's the actual constant 2021 paper. The title was the Free Energy Principle.

1:44:36 Daniel:
It's not about what it takes. It's about what took you there. That's definitely an evolution perspective. And it's partially because of Axel's background in evolution and ecology, because that's what we can say about evolution. This is what got birds here.

1:44:53 Dean:
Right? And so the counterfactuals in evolution are a really interesting topic and that's sort of explored in a variety of more popular as well as more mathematical treatments, like how do we go from the observed, which must have existed? So it's very easy to fall to them being the best of all possible worlds, optimal form, etc. But then we also recognize the changing environments and interactions and all that. And then Bleu, like with the historicity facts and the impacts, which are like facts that aren't known.

1:45:28 Daniel:

And then there's the memory, which requires a perspective. People could use system memory or something to refer to a factor and impact on a system. But these are just being used loosely like some parts of how the past influence the present and the future have to do with systems that take a perspective. Other impacts are not through those adaptive active inference agents. And so that's very different.

1:45:57 And so historicity is kind of like the stigma g pheromone layer and then there's like memory with the forager. So you could talk about the memory of the ground for the pheromone, but it makes sense to just remember that it's a different thing with the pheromone on the ground than the cognitive infra ant. And no nest may is going to have all the knowledge or the onboard modeling to make decisions safer. The ones that dot two complicate it too much, but it's stuck in the dialectic between doing exactly what worked in the least, which is to say recently succeeded, but might not work in the future and changing things. But once you move into changing things, the space of the adjacency of changing is sometimes like daunting and you still haven't really address the problem of changing world.

1:46:49 So origin is a very deep entry point into a lot of these questions in active inference and in other areas because it really connects to the egocentric, hopefully not in the Solop cystic or narcissistic sense, but just in the wayfinding sense making paradigm that actually accommodates for others through modeling them explicitly and prioritizing the environment and things like that. So pretty cool. And this paper just to kind of restate what they brought to the picture, they looked at evolutionary time scales and they framed cognition as structural change, which in a Bayesian modeling framework is sometimes called structural learning as opposed to like parameter learning, fine tuning given the model structure. And they connected that to the Bayesian graph model of active inference lab partitioning and then connected that to the factor graphs and then using the homeostat example, talked about how there could be like elaborations into other structures. So from structures to each other that might be enabling the rise of phenotypes or enabling the modeling of phenotypes that are interesting.

1:48:20 Okay, any final thoughts are like, where does this tape take us? What is exafferent now? What are we going to be talking about soon?

1:48:32 Blue:
So I'm excited to really delve into the math that is going to get behind the models that were elucidated out in figure four. How do we go through time and through relationality and hierarchical dimensions and so forth?

1:48:59 Daniel:
Okay. Thank you. Bleu Dean. You won't have the least word, but what do you think?

1:49:06 Dean:
Jelle I'm still trying to figure out how this sort of ties in with the whole identity piece. I left that because I thought that was just a rabbit hole that went way too far and too deep. But I know that when

you do get that parallelism and when you. Can sort of create a dimensional grid that has a negative space as well. So like, you can be as units going positively or negatively and passing through zero and how that fits with identity.

1:49:36 That's kind of a personal thing that I got to do. The other thing the thing that I really liked about this paper is that I think it took something that's really, as you said, over huge timescales and sort of took it back down to something that might be digestible for a lot of people that otherwise they kind of go okay, so here's 1000 years ago or a million years ago or whatever. From the bacteria to now. What are the sort of basic building blocks that we say might have might have played Hae Park in getting to the kind of sophistication where we are now? And again, I think it's about learning.

1:50:22 I always think that evolution is about learning. And so, as I said in the .1, this affirmed a lot of the stuff that I think we need to be bringing into some of the more formal ways that we structure learning because I think it's learning sort of unleashed as opposed to all tied up.

1:50:45 Daniel:
Awesome. Thanks both and everybody who was helping in all the ways for the previous videos. Just some last thoughts would be like from both of you and hopefully all together. There's really some interesting insights into learning and education for active inference learning, which is kind of like our meta regime of attention active inference lab lab when we're in that regime of attention. But many people have other regimes of attention where it would certainly apply to bring some of these methods and ideas in and also bring pieces from outside into our cognitive modeling of attention and collaboration and communication.

1:51:32 Like those topics and how those are the pillars of education and research during a time when education and research and participation are evolving is some of the most interesting things. And so to be able to read evolutionary papers that take a cool scope and approach and then connect it to how we're hopefully organizing together and including the affordance to participate with anyone who's listening now, it's an interesting mixture. So I hope that we all can just keep on learning and applying together.

1:52:16 All right, peace. Thanks, guys. Thanks again. Bye.