Narrowing the Explanatory Gap with Bridge Metaphors

Jeff Yoshimi (jyoshimi@ucmerced.edu)

Cognitive and Information Sciences, UC Merced

Merced, CA, 95343 USA

Abstract

A central problem in philosophy of mind concerns the relationship between subjective experiences and the physical processes that subserve them. There seems to be an unbridgeable "explanatory gap" between the two. Whereas other scientific explanations (e.g. the explanation of temperature in terms of kinetic energy) involve determinate relationships between two kinds of phenomena, correlations between patterns of neural activity and conscious experiences seem to be arbitrary. I argue that by developing computer models of embodied agents, and interpreting them using the tools of philosophical phenomenology, the relationship between neural and conscious processes can be seen to be systematic, and non-arbitrary. Visualizations of these models serve as "bridge metaphors" that further emphasize how systematic neuro-phenomenological relations are. By showing that links between brain states and conscious states are non-arbitrary, the explanatory gap is narrowed.

Keywords: Explanatory Gap; Hard Problem of Consciousness, Dynamical Systems Theory; Neural Networks; Phenomenology; Reinforcement Learning; Husserl

The Explanatory Gap

The familiar mind-body problem is rooted in the difficulty we have understanding how subjective feelings could be reducible to (or identical with) objective physical facts. How could my rush of excitement at the birth of a child "just be" a distributed pattern of spiking neurons? This problem is associated with a kind of explanatory challenge. When it comes to consciousness, there seems to be an explanatory gap (Levine, 1983), or a "hard problem" (Chalmers, 1995), that does not arise when inter-theoretic reductions are pursued in other areas of science. The problem is a matter of ongoing discussion, and there is a fairly well-developed taxonomy of responses to it; for review see (Tye, 2013).[1]

The explanatory gap argument can be presented as a *modus tollens*:

1. To explain a phenomenon *x* requires making *x* intelligible.

2. The relationship between consciousness and neural activity cannot be made intelligible.

3. The relationship between consciousness and neural activity cannot be explained.

Let us consider the two premises and conclusion of this argument in more detail.

The first premise describes an intelligibility condition on explanation.[2] According to Levine, when something is made intelligible, "There doesn't seem to be anything 'left out'" (Levine, 1983, p. 358). In the explanatory gap literature this has also been characterized as a feeling of necessity or non-arbitrariness. If an explanation makes a phenomenon intelligible, we feel that the *explanandum* is necessitated by the *explanans*. To use Levine's example, that water boils at 100 degrees Celsius at sea level, seems necessary, given the laws of chemistry. Or, to use an example from Tye, "Once one learns that in solid things, the molecules are not free to move around as they are in liquids, one immediately grasps that solid things do not pour easily, that they tend to retain their shape and volume. Having been told the physical story, to ask: 'Yes, but why are things with molecules that are fixed in place solid? Why shouldn't such things not be solid?' is to show a conceptual confusion" (Tye, 1999, p. 705). In these cases it does not seem like the phenomenon explained (boiling water, solidity) could be any other way, given the explanation. We could not just change the mapping from low level physical states of a thing to temperatures, or to states of matter (solid, liquid, gas, etc.) in any arbitrary way.

The intelligibility condition can be summarized as follows:

> If an explanation of some phenomenon *p* in terms of physical state *s* and laws *L* is intelligible, then the relationship between *s* and *p* is felt to be necessary, rather than arbitrary.

More formally, if we describe the relationship between physical states *s* and higher level phenomena *p* using a function *f*, then if an explanation using *f* is intelligible, we must feel that *f* is fixed by the relevant laws, and is not just arbitrarily stipulated.[3] For example, the mapping between low-level physical

---

[1]The relationship between the explanatory gap argument and the mind-body problem is subtle. The explanatory gap argument is epistemological: it concerns our ability to understand how brain states are related to conscious states. The mind-body problem is a metaphysical problem: it concerns how brain states and mental states are *actually* related to each other (in particular: are they identical or separate?). So they are distinct issues. Some have used the explanatory gap argument as a metaphysical argument against materialism (and in support of dualism), by adding a premise to the effect that what cannot be explained in terms of physics is not itself physical. Others defend the view that physicalism is true, and that the explanatory gap is just a problem we human cognizers have with understanding how physicalism is true. For an overview see (Chalmers, 2003). I am officially neutral on these issues. I argue that there is less of an explanatory gap than the authors of the original arguments suppose. This naturally supports physicalism, but is nonetheless strictly speaking compatible with dualism and even other forms of monism.

[2]I refer to explanations, the phenomena explained in an explanation, and the theoretical relationships involved in an explanation, as being "intelligible." These usages are related as follows: if a relationship is intelligible it is determinate and non-arbitrary (details are in the main text); if a phenomenon is made intelligible it is described in terms of an intelligible relationship; if an explanation is intelligible, it makes the phenomenon it explains intelligible.

[3]Notice the reference to "feeling": intelligibility is, at least in part, a phenomenological concept. Indeed, the phrase "felt contingency" occurs five times in Levine's paper.

states of a system and states of matter like solidity and liquidity is not arbitrary, but is fixed by the laws of physics. In a similar way, the mapping between configurations of particles in a volume and temperatures is (given a temperature scale) fixed, and not arbitrary.

The second premise states the main problem. The problem is that, when it comes to explanations of subjective conscious experiences in terms of physical states, it does not seem the intelligibility condition is fulfilled. Suppose it is shown that activity in the insula and orbito-frontal cortex is correlated with the taste of chocolate. Even if the correlational evidence is strong, the relationship between these kinds of brain states and taste experiences does not feel necessary. It seems like we could have just as easily been told that activity in these brain areas correlates with the taste of strawberries. We have this feeling that the brain states and conscious states "are stuck together in an arbitrary manner" (Levine, 1983, p. 359). This arbitrariness does not apply in the other cases: we could not just as well say that a crystal lattice of atoms in a diamond is "liquid" as we could say it is "solid". We could not just as well say a given gas is 20 degrees Celsius as we could say it is 30 degrees Celsius.

So whereas in the physical cases we have explanations that fulfill the intelligibility condition, in the case of consciousness we do not. So there is an explanatory gap. Or so the argument goes.

I will make a case that the argument is unsound, because the second premise is false. The relationship between activity in visual cortex and visual experience may seem arbitrary when states of the visual cortex are considered in isolation. A given brain state considered in isolation could just as well have given rise to an auditory experience as a visual experience. However, when we shift attention to the dynamics of the visual cortex (in a plausible environment) and consider what kinds of phenomenological processes "match" those dynamics, our choice of mapping from brain states to experiences will no longer seem arbitrary.

Narrowing the Gap with Bridge Metaphors

The explanatory gap problem was made prominent by Levine's paper as well as (Chalmers, 1995), which was subsequently responded to by numerous philosophers and cognitive scientists. Francisco Varela's response was called "Neurophenomenology–A Methodology Remedy to the Hard Problem" (Varela, 1996). Varela's strategy was to draw on the area of philosophical phenomenology to try to solve the hard problem and close the explanatory gap. Phenomenology originates in the work of the German philosopher Edmund Husserl (1859-1938). Husserl developed a complex technical apparatus for studying the structure and dynamics of consciousness. He did not merely describe the contents of his experience introspectively, using something like a literary stream of consciousness or autobiography of his own thoughts. Rather, he drew on his training in mathematics to develop a more rigorous, formalized theory of consciousness

(his Ph.D. was in mathematics, and he wrote a dissertation on the calculus of variation with Karl Weierstrass as his advisor). Rather than emphasizing individual conscious states he emphasized sets of possible conscious states, what he called "manifolds" of possible experience. He also described laws (which he thought of as invariant structures in these manifolds) constraining the kinds of processes that must occur if a person is to experience a stable world.[4]

Varela argued that by drawing on the methods and results of phenomenology and integrating them with neuroscience and dynamical systems theory (via a system of "mutual constraints"), we can make progress on the hard problem. In his response, Chalmers was enthusiastic: "the idea of 'neurophenomenology' sounds eminently sensible to me. The test will be whether it can be cashed out in the form of detailed results" (Chalmers, 1997). I think we can pass this test.

Recall that the key issue with the explanatory gap problem concerns the *intelligibility* of the mind-brain correlations discovered by cognitive neuroscience. According to the argument, any mapping discovered between neural states and conscious states will feel arbitrary and thus fail to meet the intelligibility condition. I argue that the intelligibility condition can be met.

Let B be the set of possible brain states for an agent and C be the set of possible conscious states for that agent. Consider a function $f : B \to C$ that maps the agent's brain states to their conscious states. Suppose a neural process occurs, which we characterize by a discrete approximation $b_1, \ldots, b_n$. Using f, we can consider the conscious process this neural process gives rise to: $f(b_1), \ldots, f(b_n)$. We can do other work with f as well. For example, given a way of saying how probable future brain states are given current brain states, we can use f to say how probable future conscious states are, given current conscious states.

In general, we can use the mapping f to take structures in B, and associate them with structures in C. Call the resulting structures "induced phenomenological structures." I will argue that we can test the arbitrariness and hence the intelligibility of our neural explanations of consciousness by asking how coherent induced phenomenological structures are, relative to different choices of f. If the degree to which induced phenomenological structures are coherent varies with our choice of f, then our choice of f is not arbitrary, and so our neural explanations of consciousness are intelligible.

A number of questions arise here. First, we don't know what the function f is. Second, it is not clear what the "coherence" of an induced phenomenological structure amounts to. Finally, it's not currently possible to measure brain activity in a manner that is precise enough to allow us to refer in any detailed way to specific neural processes $b_1, \ldots, b_n$.

We can address these questions by making use of simulations and theoretical constructs, which are sufficiently well motivated to make a case that our choice of f is not entirely

---

[4]For an overview of the ideas in phenomenology I draw on here, see (Yoshimi, 2012).

arbitrary. We can begin with a neural network simulation of an agent in an environment, using theories and methods of computational neuroscience. We can then posit a mapping from the states of this simulation to conscious states. As we will see, this can (at least for the cases I consider, which focus on sensory states), be done in a fairly natural way: for example, the states of an agent's neural network when it is near a flower stimulus can be thought of as experiences of smelling the flower. We do not need to assume this *must* be the way the mapping works (in fact, the key step in the argument will be to imagine variations on the function). It will be enough for our purposes just to start with something plausible. Finally, to determine the coherence of induced phenomenological structures, we can ask to what extent those structures are consistent with an independently developed theory of conscious experience, namely, Husserlian phenomenology.

The approach I am describing can be thought of as a stepwise "bridge metaphor" strategy:

1. Develop a neural network model of an agent in a virtual environment, using existing principles of computational neuroscience.

2. Project patterns of activity from the neural network's state space to a visualizable 2d or 3d space using dimensionality reduction techniques. Observe the sets of points or "point sets" that occur in the projected neural network state space.

3. Propose a mapping f from network states to conscious states.

4. Use f to induce phenomenological structures. Ask whether these are coherent relative to the principles of Husserlian phenomenology.

5. Consider alternative mappings f', and ask whether they induce phenomenological structures that are more or less coherent than f.

In the last step, if the alternative mappings f' induce phenomenological structures that are less coherent than f, we have evidence that f is not felt to be arbitrary, and thus that explanations of consciousness can be intelligible. To that extent we narrow the explanatory gap.

We can think about this process using the theory of conceptual metaphors (Lakoff, 1993). Conceptual metaphors are cognitive structures that allow one conceptual domain (the source domain) to be understood in terms of another (the target domain). Typically the source domain is more concrete and better understood, and in this way provides structure to the target domain. Examples are time is money, love is a journey, and argument is war. The structure of money, journeys, and wars are relatively concrete and embodied, and can thus be used to *provide structure* to the more abstract domains of time, love, and argument. So too here. We begin on the physical side, with a neural network model of an agent in an environment. Neural network models can be directly constrained by anatomical, physiological, and behavioral data.
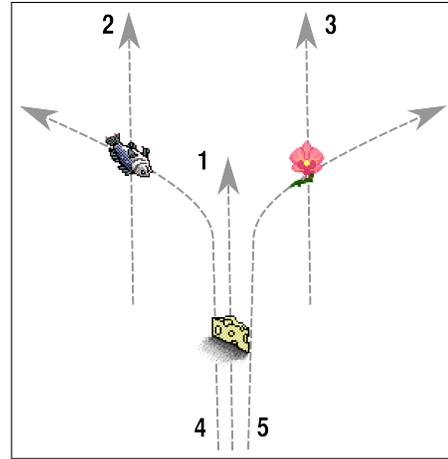

Figure 1: An agent in a 2d world with three objects, and the 5 excursions it takes. The dashed gray curves trace out the paths taken by the agent on the five excursions.

They are also relatively concrete and easy to interact with in computer simulations, and can be interpreted using a well developed body of theoretical work. The concrete structure of neural network dynamics is a kind of source domain that can be used to give structure to the domain of phenomenological dynamics, which is more abstract and thus stands to be benefit by this linkage. I call these "bridge metaphors" and the strategy above a "bridge metaphor strategy."

Whereas conceptual metaphors are mappings that happen *de facto* in the course of our everyday experience, I envision bridge metaphors as tools that are actively tuned in scientific practice, to maximally harmonize intuitions between disparate domains. We use the neural network dynamics to understand the phenomenological dynamics. We then use the phenomenological dynamics to interpret the neural dynamics. By proceeding back and forth in this way, we begin to develop a detailed understanding of neuro-phenomenological connections. We build a bridge between the two domains, and begin to narrow the gap of intelligibility between them.

Example

I now illustrate the 5 steps of the bridge metaphor strategy with a concrete example.

(Step 1). We begin with an idealized model of a network in a virtual environment. The model was developed using principles familiar from a wide range of models in cognitive science, including reinforcement learning models (Sutton & Barto, 1998), and Bayesian models of the brain as a prediction machine (Clark, 2013). It is also consistent with ideas from the earlier literature on animal learning going back to (Tolman, 1948), according to which behavioral data imply that agents' navigate environments using internal maps. A component of most models in these areas is a mathematical structure that learns to predict what future states will occur, often using an error based learning rule like the delta rule

(Bishop, 1996).

The agent and its virtual environment are shown in Figure 1. The network controlling the agent consists of several subnetworks: a sensory network consisting of neurons that respond to objects in the environment; a motor network consisting of neurons which, when activated, make the agent move straight or turn (or both); and a prediction network (trained using the delta rule) that predicts the next state of the sensory network based on the current activations of the sensory and motor neurons. Based on what the agent senses and how it moves, it predicts what it will sense next. Each sensory neuron responds to a specific object via a linear scaling function of the distance between the sensor and that object. For example, if a cheese-sensing neuron is on top of cheese, then it is maximally activated. As the agent moves away from the cheese that neuron's activation diminishes to 0.[5]

(Step 2). To get a feel for the way the agent represents its simple environment, five "excursions" were run, where the motor neurons were clamped to specific values at successive times according to one of five scripts (see Figure 1). At each iteration of the simulation, the network was updated, the agent was moved on the basis of the states of the motor neurons, and the current state of the sensory networkwas projected to a 2d plot using principal component analysis (Abdi & Williams, 2010). The points of the plot were then colored according to the predictions of the prediction subnetwork. The current state was colored green, and states that were predicted to occur next were colored red at varying degrees of saturation (and clipped to gray below some threshold), in proportion to the Euclidean distance between those states and the next predicted state.

On the first three excursions, the agent moves past each of the three objects individually. As the agent passes an object a single arc (in these cases line segments) unfolds in the state space. Intuitively the agent is representing (for example) a "faint flower smell, a little more flower smell.... maximal flower smell ... a little less flower smell..." etc. until it is back to smelling nothing. As it moves up to each of the objects a kind of "halo" of red (predicted next states) moves with the current point up to the end of the forming arc. As it passes the objects the halo moves back down the just-formed arc (you are encouraged to download the software, run the simulation, and see it yourself). After the agent has passed the three objects its state space contains the three-pronged structure shown at the core of Figure 2, what can be called a "bouquet of three arcs," where each arc corresponds to one of the represented objects.

In the fourth and fifth excursions, the agent visits two objects in succession, moving forward past the cheese and then turning either left or right. An example is shown in Figure 2. This creates new mixed representations that did not occur in the first three excursions, corresponding to two sensory nodes
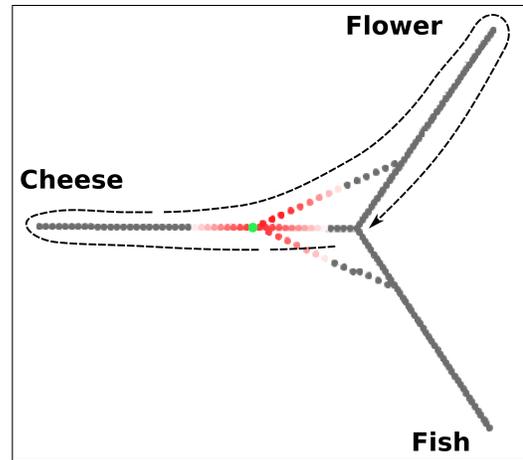

Figure 2: The "spandrelled bouquet" point set after all five excursions have occurred. Shown is the fifth excursion. The current point (shown in green) corresponds to the point when the agent has just passed the cheese and is turning towards the flower.

being active at once. In the fifth excursion, for example, we have a sequences of points that moves from the arc corresponding to the cheese alone, to a new connecting path or "spandrel" corresponding to mixed representations of flower and cheese, to the arc corresponding to the flower alone. As with the first three excursions, as the agent moves, a halo of red follows the current point in state space.

After the agent has gone on these five excursions it can be sent back on any of them again. In these cases the point set does not change (no new sensory states occur) but the structure of the halo moving through the point set does (the network's predictions improve). For example, after the first five excursions if the agent is sent on the fifth excursion again, then at the moment the agent begins turning towards the flower, it correctly predicts that it will soon be in a mixed cheese / flower state even though it occupies the same point in the point set as when it is about to have the mixed cheese / fish state of the fourth excursion. Figuratively, when it is at the "fork in the road" (shown in Figure 2), it correctly predicts which path (which spandrel) it will follow, based on what it is sensing in its sensory nodes and what action it is taking.[6]

(Step 3). Our provisional mapping f from points in the network's state space to associated conscious states was already suggested above. We can simply assume that states of the sensory network produce conscious perceptions of the represented objects. For example, the network state that occurs when the agent is right on top of Swiss cheese corresponds to a conscious perception of Swiss cheese as being near the agent, whereas the network states that occur when the agent is at various distances from the cheese correspond to percep-

---

[5]For a complete specification of the algorithm, the code can be requested from the author and run via Simbrain, a freely available simulation package (Yoshimi, 2008).

[6]This is hard to see in the figure, though it is the case that the output of the prediction network corresponds to the first point on the upper spandrel.

tions of the Swiss cheese at corresponding distances from the agent. When the agent is near multiple objects it has a mixed conscious state encompassing all the objects it perceives, at corresponding distances from it. When it is sufficiently far from the objects it consciously perceives no object.

(Step 4). Thus, under f, in each of the first three excursions a sequence of conscious perceptual states occurs beginning with faint perception, rising to maximal perception of one object, then falling back to faint perception. Each of the arcs corresponds to a perceptual processes of moving past one of the three objects. The origin of the bouquet, where the three arcs intersect, corresponds to perceiving no objects. In the fourth and fifth excursions, the agent first perceives one object, then perceives two objects (while it is on the spandrels), then perceives the second object alone. This is the phenomenological structure induced by the spandrelled bouquet of arcs under the mapping f.

We can also consider the phenomenological structure induced by the prediction subnetwork under f. After some training, the states of the prediction subnetwork—visible as a moving halo of red—accurately predict what the agent willsense next. These predictions correspond to phenomenological predictions about what the agent will consciously perceive at successive times. We can also associate the mean error at each time step (the discrepancy between the states predicted to occur and the states that actually occur, which is used by the delta rule to update the weights), with feelings of rightness or wrongness.

All of this is consistent with Husserlian phenomenology. According to Husserl, we consciously perceive objects as "identities in manifolds", such that when we move our bodies in a smooth, continuous way relative to an object, we perceive it as changing in a correspondingly smooth, continuous way. Husserl also develops a complex phenomenology of expectation (or as he also says, "protention"), whereby as we move with respect to a familiar object we constantly have expectations about how the object will look relative to our movements. These expectations are "fulfilled" or "frustrated" in varying degrees depending on what we actually perceive in successive moments. As we learn more about an environment these expectations become more specific, in the sense that there are fewer subsequent perceptions that would fulfill our expectations. This corresponds to the idea that the halo of predictions surrounding the current point narrows as the agent repeatedly goes on the same excursions.

(Step 5) We can now ask how coherent these induced phenomenological structures are, relative to other choices of f. If f were different, we would get different induced phenomenological structures. Would these different structures still make sense relative to Husserlian phenomenology? For many such differences, the answer is obviously "no". Suppose, for example, that we chose an f' such that on the first excursion the agent experienced a haphazard sequence of perceptions of the fish, cheese, and flower, in rapid succession as it passes the cheese. This would not be a sensible or coherent experi-

ential process. The agent could not, in Husserl's terms, have any kind of experience of stable things or "posited unities." A subjective process would occur, but it would be a kind of "maelstrom", a chaos of unrelated perceptual states. Husserl admits that something like this is imaginable, but says that in such a case all sense of reality would be lost:

> Could it not be that, from one temporal moment on... the series of appearances would run into one another in such a way that no posited unity could ultimately be maintained... Thus we arrive at the possibility of a phenomenological maelstrom... it would be a maelstrom so meaningless that there would be no I and no thou, as well as no physical world—in short, no reality (Husserl, 1997, 249-250)

Though I think this is enough to show that our choice of f is not arbitrary, let us briefly consider the prediction network and the phenomenological structures it induces. If we consider an expanded f that not only maps sensory states to conscious perceptions, but also maps from error states of the network to experiences of fulfillment and frustration, then the only way this map would make sense of the phenomenology would be if low error corresponded to fulfillment and high error corresponded to frustration. For example, when the agent begins to turn right, having just sensed the cheese (excursion 5), it predicts it will sense the flower. When it actually does sense the flower, the prediction is confirmed and error is low. If this mapped to a feeling of surprise under an expanded f', that would make no phenomenological sense. We are not surprised when our subsequent perceptions match our previous expectations: we experience something like the kind of smooth feeling of fulfillment Husserl describes.

Thus, our choice of f is not arbitrary. We cannot simply associate brain states with conscious states however we like, and expect the result to make phenomenological sense. Some choices of f produce phenomenological structures that are consistent with Husserlian phenomenology; other choices of f produce phenomenological structures that are inconsistent with Husserlian phenomenology. So premise 2 is false, the explanatory gap argument is unsound, and there is hope yet for an intelligible explanation of consciousness in terms of neural structures.[7]

Conclusion

I have used the bridge metaphor strategy to address a specific philosophical argument, the explanatory gap argument. Like Varela, I think neuro-phenomenology can be used to narrow

---

[7]The upshot of all this is that the explanatory gap can be narrowed. I do not say it can be "closed", because I believe that there is some latitude in how we specify f. For some range of f's I think that induced phenomenological structures can remain coherent. In fact some fairly radical permutations of f, akin to color inversion cases (Byrne, 2014), may preserve phenomenological coherence. So, I am not sure a unique mapping from brain states to conscious states can be achieved. But I do think I have shown that there are substantial constraints on how f is specified, and to that extent I believe I have narrowed the explanatory gap.

the explanatory gap (though I do not think it can close the gap altogether), and make progress on the hard problem.

In developing this argument, I have been motivated by a broader research agenda: to show how these two extensive bodies of research—computational and cognitive neuroscience on the one hand, and Husserlian phenomenology on the other—can be integrated using the tools of computer simulation and dynamical systems theory. We have seen that computer simulations can be used to visualize an agent's model of its environment. The point-set that develops as an agent explores its environment has a specific and meaningful *shape*. In the simple case considered above, where three objects were repeatedly explored in five ways, the point set was a spandrelled bouquet of three arcs.[8] As the agent explores its environment it traverses this structure, and a halo of predictions, which narrows over time, moves with it. Using a mapping from brain states to conscious states, we can interpret this structure as being simultaneously a picture of the agent's neural representation of its environment, and as a phenomenological structure, which shows how well the agent knows its way around its environment, what it perceives at each moment, what it expects at each moment, and the degree to which it is surprised by what it perceives.

Acknowledgments

References

Abdi, H., & Williams, L. J. (2010, July). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459.

Bishop, C. M. (1996). *Neural networks for pattern recognition*. Oxford : New York: Oxford University Press.

Byrne, A. (2014). Inverted qualia. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Summer 2014 ed.).

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of consciousness studies*, *2*(3), 200–219.

Chalmers, D. (1997). Moving forward on the problem of consciousness. *Journal of Consciousness studies*, *4*(1), 3–46.

Chalmers, D. (2003). Consciousness and its place in nature. *Blackwell guide to the philosophy of mind*, 102-142.

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 1–73.

Husserl, E. (1997). *Thing and space: Lectures of 1907* (Vol. 7; R. Rojcewicz, Trans.). Springer.

Lakoff, G. (1993). The contemporary theory of metaphor. *Metaphor and thought*, *2*, 202–251.

Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific philosophical quarterly*, *64*(4), 354–361.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction* (Vol. 1) (No. 1). Cambridge Univ Press.

Tolman, E. (1948). Cognitive maps in rats and men. *Psychological review*, *55*(4), 189.

Tye, M. (1999). Phenomenal consciousness: The explanatory gap as a cognitive illusion. *Mind*, *108*(432), 705–725.

Tye, M. (2013). Qualia. In E. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Fall 2013 ed.).

Varela, F. (1996). Neuro-phenomenology: a methodological remedy to the hard problem. *Journal of consciousness studies*, *3*(4), 330-349.

Yoshimi, J. (2008). Simbrain: A visual framework for neural network analysis and education. *Brains, Minds, and Media*, *3*.

Yoshimi, J. (2012). Two dynamical themes in husserl. In *Being in time: Dynamical models of phenomenal experience* (Vol. 88, pp. 165–184).

---

[8]This approach can be extended in various ways. Adding more objects to the environment adds more arcs to the bouquet. Adding more excursions fills in the region between the arcs in different ways. Adding decay to the neurons creates more curved paths in the state space, some of which look like curlicues. These kinds of additions can be understood in terms of a more complex representations of the environment and a correspondingly more complex phenomenology.