

Moving Finger

Why We Are Conscious

Musings on philosophical zombies, epiphenomenalism, conscious inessentialism and the Hard Problem

Alex Christie, Doha, Qatar, May 13th, 2008

(published online at <http://www.moving-finger.com/papers/wwac.pdf>)

Abstract

It is perhaps just conceivable that we could have evolved as “phenomenally blind” automata without experiences, without feelings, just like the thermostat on the wall or the computer on the desk. But if this were the case, we would not possess the richness of the story that we tell ourselves about the world around us. By weaving this story for ourselves, we create higher-order models of what we perceive reality (the outside world) to be, and we use these models to develop higher-order concepts and ideas that relate to our interaction with that world. What p-consciousness “is”, then, is not just a blind emergence of qualia, of raw feeling and experience with no obvious purpose or function. Instead, what p-consciousness “is” is a dynamical system in the process of telling itself a story.

Possessing p-consciousness provides humans with one route to making higher order models of reality, and enables us to use these models to form higher-order concepts and ideas. It is not that it would be physically impossible to make such models without p-consciousness, but rather that p-consciousness provides the only accessible route for blind evolution to create such models.

On the evolutionary landscape, the zombie is on an evolutionarily inaccessible plateau, there is no bootstrap process whereby blind evolution could create such an agent (any more than a chess-playing computer could be created purely as the result of blind evolution). There are some places on the evolutionary landscape which you just cannot get to from here. The only accessible route for blind evolution to create an intelligent, volitional, intentional agent is via the tool of p-consciousness.

Definitions

A **philosophical zombie** (hereafter p-zombie, or just plain zombie) is a hypothetical being which is physically (and hence also functionally) just like a human being, but which has no conscious experiences, it possesses no phenomenal consciousness (p-consciousness). In the words of David Chalmers, there is “nothing it is like” to be a zombie (Chalmers).

Epiphenomenalism is the thesis that conscious experience is caused by (is supervenient upon) certain physical events in the brain, but such conscious experience in turn has no effect upon the physical.

Conscious inessentialism is the thesis that conscious experience is neither a necessary condition for, nor a necessary consequence of, any given behavior. In other words, consciousness does not supervene on the physical (hence physicalism is false).

The **Hard Problem** asks: If zombies are physically possible, why does the human brain give rise to p-consciousness? Why aren't we all zombies? Another way of asking this is: Why did evolution bother to give humans p-consciousness if zombies are physically possible?

Zombies, Sombies and Wombies

There often seems to be some confusion in the philosophical literature when zombies are discussed. I want to try and eradicate that confusion by distinguishing between two very different types of zombie. A strong zombie (sombie) I define as one which has properties (including

behavioural properties) that are in principle indistinguishable in all respects from the properties of a conscious human. In particular, a zombie would necessarily claim that it experiences phenomenal consciousness (just like its human counterpart), when in fact it does not. A weak zombie (wombie) on the other hand, is not so constrained. Although the behavior, language, character, personality etc of a wombie would render a wombie in most respects indistinguishable from a conscious human, unlike its human counterpart a wombie would not necessarily claim that it experienced phenomenal consciousness.

Now that we have set the scene, we can proceed to examine the consequences of these various concepts.

Considerations of Epiphenomenalism

Clearly, if epiphenomenalism is true then by definition there can be no such (naturalistic) benefit or competitive advantage to possessing consciousness (because epiphenomenalism says that consciousness has no effect on the physical). Thus, if zombies are physically possible and epiphenomenalism is true, then the Hard Problem is in principle unanswerable (at least from a naturalist perspective) – there can be no naturalist explanation as to why humans developed p-consciousness.

Note that the truth of epiphenomenalism on the one hand and the physical possibility of zombies on the other are mutually exclusive premises. If one is true, the other must be false. Why? Because a zombie is by definition both physically and behaviourally indistinguishable from a human, and if consciousness supervenes on the physical (the premise of epiphenomenalism) then either both zombies and humans are conscious, or neither zombies nor humans are conscious.

But what about wombies? Is it possible both for epiphenomenalism to be true, and for wombies to be physically possible? Yes, but only if there is also some physical, as well as a behavioural, difference between a wombie and a conscious human being. The answer to the Hard Problem, in this case, is then that humans possess consciousness by virtue of the fact that humans possess the right physical substrate on which that consciousness supervenes. Wombies do not possess the right physical substrate, hence wombies do not possess consciousness. We are still entitled to ask, however, if consciousness has no effect on the physical, then why would evolution bother to produce conscious humans rather than wombies? To this question, there seems to be no answer.

Dispensing with Zombies, and with Conscious Inessentialism

Since the properties of zombies are identical to the properties of humans in all respects (including all behavioural aspects), but such zombies are devoid of p-consciousness, it follows that if zombies are physically possible then conscious inessentialism is true and physicalism is false (ie consciousness does not supervene on the physical). This is the “anti-physicalist” argument based on the zombie premise. It is relatively easy to show, however, that the properties of zombies will not necessarily be precisely identical to the properties of humans in all respects (ie a zombie is an incoherent concept), hence the support for conscious inessentialism crumbles and the argument fails. (See for example Moody)

But what of the Hard Problem itself? For the purpose of the Hard Problem, it is not necessary to assume either that zombies are physically possible, or that conscious inessentialism is true. Rather, it is only necessary that wombies be physically possible, ie a zombie must be capable of displaying similar (but not necessarily completely identical) complexity of behavior, character, personality and intelligence to that of a p-conscious human.

To address the Hard Problem, we need to ask ourselves why did evolution bother to give humans p-consciousness if zombies are physically possible? What possible benefit or competitive advantage accrues to an agent from the possession of p-consciousness?

Examining Phenomenal Consciousness

Analytically, any machine equipped with some form of colour-sensitive receptor could “see red”, no phenomenal content, no consciousness, is needed. And from an analytical viewpoint, when a machine registers that it “sees red” all it is doing is producing an output based on simple numerical inputs and comparing these against a pre-determined “colour scale”. There is no higher-order processing, no phenomenal experience, no consciousness, going on.

But for humans, there is no pre-determined “colour scale” in our heads that we can use. We must assess all of what we see by brute comparison of our visual field at any one moment with the memories of previous visual experiences. In practice, we must make a massive store of visual memories, which serve as our databank for assessing new visual experiences. There is no micro-analytical reductive processing going on – all the more complex processing within the human brain is necessarily of a higher order. Thus, for us humans we cannot say “oh, that is red because it registers a reading of between 1.7 and 1.85 on my visual receptor” (which is what a machine would presumably do). Instead, we must say “oh that is red because it looks like the stored memories of other types of red images that I already have in my brain”. And here immediately we see the emergence of the phenomenal “it looks like” account.

We register what something looks like, what it feels like, what it tastes like, because THAT is exactly how we humans identify, categorise and assess new objects in the world about us. Most of the time we do not indulge in a reductive micro-analysis of these objects – we do not have the time and the wherewithal to do so. Instead we must make short-cuts, quick judgements based not on (what for humans would be) painstakingly slow reductive micro-analysis, but instead on the qualities of a much broader and coarser phenomenal perception.

The How and Why of p-Consciousness

There is an account (Humphrey) of a certain visually impaired primate, Helen, whose capacity for normal vision was destroyed by surgical removal of her entire visual cortex. Although Helen originally gave up even looking at things, she regained certain visual capacities. She improved so greatly over the next few years that eventually she could move deftly through a room full of obstacles and pick up tiny currants from the floor. She could even reach out and catch a passing fly. Her 3-D spatial vision and her ability to discriminate between objects that differed in size or brightness became almost perfect. Nonetheless, after six years she remained unable to identify even those things most familiar to her (e.g., a carrot). She did not recover the ability to recognize shapes or colors.

There is obviously something lacking in Helen’s cognitive functions. She can no longer register what a carrot looks like, and since she also has no access to any micro-analytical means of determining whether something has carrot-like properties or not, she is simply unable to identify carrots.

Pose this problem to an agent: How would one go about identifying an object as a carrot? I suggest there are at least two main ways that this could be done.

- 1) The reductive micro-analytical route: Define the necessary and sufficient qualities of “carrotiness”, then test for the presence or absence of these qualities. This is the way that we might naively design and program a machine to detect carrots.
- 2) The higher-order phenomenal route: Build up a store of many different phenomenal representations of carrots, and then compare and contrast the phenomenal experience of the unknown object with this databank of representations. This is the way that humans do it. But doing it this way requires that we possess phenomenal consciousness – we must know “what it is like” to see/taste/smell/feel a carrot. This “what it is likeness” necessarily has an essentially ineffable “feel” to it, simply because it involves no fine-grained or reductive analysis. We make decisions about phenomenal content based purely on their qualitative and analogue-like “feel” to us, comparing and contrasting with other similar remembered phenomenal experiences. No wonder they are ineffable.

Now, we may very well be able to design and develop machines which can accomplish feats of reductive micro-analysis relatively quickly. But just because we can design a machine that can accomplish this does not mean that such a solution is open to blind evolution. Just as the probability of creating a Rolex watch by blind evolution would be vanishingly small, so too would be the creation of an agent which interprets the world about it in purely reductive micro-analytical terms rather than in higher-order phenomenological terms.

Bootstrapping Intelligence

Possessing p-consciousness provides us with one route to making higher order models of reality, and enables us to use these models to form higher-order concepts and ideas. It is not that it would be physically impossible to make such models without p-consciousness, but rather that p-consciousness provides the only accessible route for blind evolution to create such models. It's a bit like the difference between a chess-playing computer, designed by humans, and a chess-playing human. The chess-playing computer contains no higher-order concepts and ideas, it hasn't developed its chess-playing abilities in a bootstrap fashion based on those ideas. Instead its ability to play chess is "hardwired" through the intelligent design of its hardware and software. Humans cannot learn how to play chess via such hardwired intelligent design, instead they must acquire their intelligent abilities in a bootstrap fashion, based initially on the internal development of basic models and concepts, and subsequent higher-order models and concepts, of the world around them. Although it is possible in principle to create a non-conscious, intelligent chess-playing agent by design, it is my thesis is that such an agent cannot evolve (ie create itself) in the "bootstrap fashion" required by blind evolution, and that the useful tool of phenomenal consciousness provides the only bootstrap-accessible "short-cut" to achieving such emergent intelligence which is available to blind evolution.

It is perhaps just conceivable that we could have evolved as phenomenally "blind" automata without experiences, without feelings, just like the thermostat on the wall or the computer on the desk. But if this were the case, we would not possess the richness of the story that we tell ourselves about the world around us. By weaving this story for ourselves, we create higher-order models of what we perceive reality (the outside world) to be, and we use these models to develop higher-order concepts and ideas that relate to our interaction with that world. One of the first things that we do is to create a centre of phenomenal gravity (we create the concept of "self"), which forms a source of subjective perspective on which these higher-order models, concepts and ideas are anchored. The "self" is created when consciousness tells itself a story. It is not essential that we tell ourselves such a story, but perhaps evolution has found that the easiest way to create a volitional, intentional, intelligent being is to allow that being to tell itself a story. What phenomenal consciousness "is", then, is not just a blind emergence of qualia, of raw feeling and experience with no obvious purpose or function. Instead, what phenomenal consciousness "is" is a dynamical system in the process of telling itself a story.

Conclusion: Climbing Mount Improbable

It is not that it would be physically impossible for an agent to develop complex models, concepts and ideas, and intelligence, without p-consciousness, but rather that p-consciousness provides the only accessible route for blind evolution to create such things. So what I'm saying is that yes, zombies are perhaps conceivable and even physically possible – but from an evolutionary standpoint zombies are so very implausible or so unlikely that they simply do not exist. In other words, it's a question of "not being able to get there from here". On the evolutionary landscape, the zombie is on an evolutionarily inaccessible plateau, there is no bootstrap process whereby blind evolution could create such an agent (any more than a chess-playing computer could be created purely as the result of blind evolution). The only accessible route for blind evolution to create an intelligent, volitional, intentional agent is via the tool of p-consciousness.

Climbing mount improbable is possible in blind evolutionary terms if and only if the slopes of mount improbable provide a reasonably continuous ascent. Blind evolution is incapable of

jumping across gaping chasms and gorges. It stands to reason therefore that there will always be some things which though physically possible are simply inaccessible to blind evolution alone. There are some places on the evolutionary landscape which you just cannot get to from here.

References

Chalmers, D; (1996) *The Conscious Mind: In Search of a Fundamental Theory*: Oxford University Press

Humphrey, N. (1992). *A History of the Mind: Evolution and the Birth of Consciousness*. New York: Simon and Schuster

Moody, T; "Conversations with Zombies" In *Journal of Consciousness Studies*, 1 (2), 1994, pp. 196-200 (online at http://www.imprint.co.uk/Moody_zombies.html)