Ishaan Jhaveri
PHIL 2621: Minds and Machines
Professor: William Starr
Teaching Assistant: Chad McIntosh
May 14, 2015

I Propose that Consciousness is aimed at an Organism's Survival

To answer the question: "Can we build emotional machines?" we must first understand emotion. According to Antonio Damasio, there are three stages involved. First there is a physiological reaction to a stimulus – this is called emotion. Then there is an internal feeling of this emotion; this could be fleeting, or even subconscious. Finally, there is a conscious experience of this feeling, which allows the brain to rationally dictate some behavior to the organism. For example, consider the feeling of being proud of someone. It may begin because a loved one just achieved something; on hearing this one may find oneself involuntarily smiling – this is the emotion. The feeling of the emotion is that they are proud of their loved one. The conscious awareness of the feeling comes when they know they are proud of their loved one, when they reflect upon it and question why – and realize it is because of the loved one's achievement. This conscious awareness encourages them to congratulate the loved one, which is the behavior they finally exhibit in response to the initial emotion. He distinguishes between primary emotions, like sadness and happiness, and social or secondary emotions like jealousy and pride (like in the example above). Though all these emotions follow the same three stages, with the secondary emotions the feeling is more important – it is only after the feeling of pride is made conscious for example that we identify it as pride – whereas with happiness we can more easily identify it as happiness as soon as we break into that spontaneous smile – without spending too much time appraising the feeling before it is conscious.

Ishaan Jhaveri
PHIL 2621: Minds and Machines
Professor: William Starr
Teaching Assistant: Chad McIntosh
May 14, 2015

Damasio further argues that the primary way emotions achieve their "longer-lasting goal is through this [conscious rationalization of our inner feelings]" (which are the results, of course, of emotions) (Damasio 36-37). There are two important aspects to this claim: first the notion that emotions have "goals", second that conscious realization and rationalization of emotions allow them to achieve these long-lasting goals.

Consider the first aspect: Damasio talks about his patients who "were entirely rational [until, because of brain damage] they lost a certain class of emotions and, in a momentous parallel development, lost their ability to make rational decisions" (Damasio 41). According to him, these poor rational decisions displayed a "disturbance of [these peoples'] ability to decide advantageously in situations involving risk and conflict" (Damasio 41). A particular example of this is Damsio's patient S, who too trustworthy because of brain damage, and was thus often misled by people she trusted too easily. These examples bolster Damasio's claim that emotions have a goal – and that goal is to help people decide advantageously in situations involving risk and conflict. This is a very evolutionary view of emotion, as it suggests that the primary goal of emotions is to aid in survival, and homeostasis – the regulation of all the needs of the body.

Damasio's particular three stages are key to the understanding of emotion as evolutionary because he suggests that it is through emotions that we are guided to making correct rational decisions to promote homeostasis. "[He] suggests that certain levels of emotion processing probably point us to the sector of the decision-making space where our reason can operate most efficiently" (Damasio 42). The third of his three stages is this activation of the decision-making space.

Ishaan Jhaveri
PHIL 2621: Minds and Machines
Professor: William Starr
Teaching Assistant: Chad McIntosh
May 14, 2015

Further, if his three stages take an evolutionary approach to emotion, then surely the most important of the stages, the conscious processing, must also be in place to promote homeostasis in an organism. Indeed he "[proposes that] homeostasis is a key to the biology of consciousness" (Damasio 40). This shows us the importance of consciousness in humans. For lower organisms like the sea-snail Aplysia, homeostasis is fairly simple – it needs to reproduce, find food and find shelter and safety. When it is touched by what could be a predator, it propels itself in the opposite direction of the touch. Here the touch is the stimulus, for which the propulsion in the opposite reaction can be understood as the emotional reaction according to Damasio's three stages. For the Aplysia then, homeostasis is established without a need for conscious rationalization of the feeling of the emotion (fear of the predator presumably) – because it already fled. For humans on the other hand, we have evolved to have a much more complicated set of requirements for homeostasis, that needs consciousness to really be met.

We need more than food, sex and shelter to really feel like we have homeostasis. We need a sense of good government, peace, a sense of bettering the situations of people around us, to name but a few requirements of human homeostasis – so evolutionarily, we need consciousness to understand our negative emotions towards war to try and alleviate it, and so on.

Note that this view of consciousness suggests that emotions can "learn" to "connect homeostatic regulation and survival "values" to numerous events and objects in our autobiographical experience" – emotions can learn to associate war with discomfort, so that consciousness can be put to work to make an effort to stop war (Damasio 54-55).

Ishaan Jhaveri
PHIL 2621: Minds and Machines
Professor: William Starr
Teaching Assistant: Chad McIntosh
May 14, 2015

This is why, were I designing an organism, I would imbue it with emotion – because emotion could learn, adapt to new situations and guide consciousness in establishing homeostasis – no matter how the organism evolved. Emotions are trustworthy guides for survival and homeostasis – once they learn to associate some object with some feeling, they will always elicit that feeling in an organism – sometimes coming in crucially to aid "absentminded humans" (Damasio 54).

Yet placing such importance on consciousness leaves a major part of the theory unaddressed – the notion of having a feeling without being conscious of it. Damasio's theory allows room for behavior (be it as simple as smiling or as complex as having a subconscious preference for one person over another) without conscious awareness of the feeling motivating this behavior. Before discussing Damasio's evidence for this, I must note that Damasio argues that we can exhibit a behavior without *conscious awareness* of feeling that caused that behavior, but never without the *feeling* itself.

Damasio's key piece of evidence for the distinction between behavior without conscious awareness of it is his work with a patient known as David. David had a condition that didn't allow him to form any new memories whatsoever: thus he couldn't remember the names or faces of people he met. Yet Damasio observed that he still seemed to show a preference for, and aversion to certain people and decided to investigate how this was so without the ability to remember people. So he designed an experiment in which over a sustained period of time David was exposed to a person who treated him well, one who treated him neutrally and one, badly. David was then shown pictures of these people amidst three other random people and asked to pick the one who

he thought was his friend – he chose the good guy more often than random chance would dictate, the neutral guy about as often as chance would dictate, and the bad guy almost never – showing that without his conscious awareness – he did have some *feeling* towards these people that motivated his preference for the good over the bad.

This is astounding and suggests that there can be something in someone's mind that spurs them to act a certain way without their really understanding why they are acting in this way. Chalmers and Block believe this very phenomenon underscores the distinction they draw between what they call conscious *awareness* and conscious *experience*.

Chalmers calls zombies to mind – organisms that can behave exactly as humans do because of some proclivity to do so – without understanding why they are behaving as they are. He claims that the fact that such organisms are conceivable means there must be some difference between behaving in a certain way, and understanding what it is to behave in a certain way. Behaving then – actively performing some behavior as a response to an external stimulus is called conscious awareness – one still needs to be aware of the stimuli to process a response to it. This is distinguished from conscious experience – the understanding of what it is *like* to experience some stimulus, and behaving in a certain way in response that requires more than just knowledge of the stimulus – but this nebulous concept of knowledge of what is *like* to have this feeling and respond behaviorally to it.

This distinction is also illustrated by the neurophysiological phenomenon of tic douloureux, a condition where even the slightest touch to the face causes excruciating

pain. Damasio knew a patient with this condition who had to get an operation lesioning a part of his frontal lobe to reduce the pain. The patient later reported " "the pains were the same," but that he felt fine" (Damasio 75). It appears that the patient was now unable to *experience* the pain – to know what it was like to be in pain – yet he was still vividly aware that the stimuli inducing the pain were still present and that his body was "feeling" the pain. He went from conscious experience of pain to conscious awareness of pain, with the ability to respond to it without any conscious experience (excruciating pain) of it. This distinction frames our discussion of David. The experiments show that David clearly had a conscious awareness of the fact that he was asked to give a preference between some people and could clearly respond by choosing one of those people. He could respond yet he didn't have access to the information that caused him to respond – he didn't have an experience of the preference, he just knew how to exhibit it – by choosing the person his brain guided him to choose. As Damasio says, "David had not learned new knowledge of the type that can be deployed in one's mind in the form of an image. But something stayed in his brain and that something could produce results in nonimage form: in the form of actions and behavior" (Damasio 46).

Thus for consciousness to be truly fulfilling the homeostasis-seeking role Damasio so importantly bestows upon it, it seems that it must respond based on conscious experience of stimuli, as opposed to just conscious awareness. It appears that till date there are no machines that can exhibit conscious experience, though there are those that can experience conscious awareness. Yet, if what Damasio tells us about consciousness were true, that its main role is to find the homeostasis that emotions guide

it towards – there is no reason to think that there is anything uniquely biological about it. If there are already machines that can be programmed to seek certain types of homeostasis, there is nothing to suggest that eventually they will not get to a stage of "machine consciousness" where they can experience the feelings of needing some homeostasis and thus seek it as best they can – no one is suggesting that this homeostasis would be the same type that humans seek – thus no one can say that if consciousness is just a mechanism for most effectively responding to emotional needs (programmed or natural emotions) that machines will not eventually be able to be imbued with it – with true conscious experience – not just conscious awareness.

Adolphs similarly advocates for the perspective that should we design an organism, we must give it emotion. He agrees because "as [a neuroscientist acknowledges that] the processing of emotions [shares] mechanisms with the processing of thirst, hunger [and] pain" and thus must aid in the furthering of the organisms life (Adolphs 15). In his opinion the motivations for giving robots emotion is to understand more about emotion in general, as well as to understand human perception, in observing how human's interact with emotive robots.

References

1. Adolphs, Ralph. "Could a Robot Have Emotions?". *Who Needs Emotions*. Ed. Fellous, Jean-Marc & Arbib, Michael A. Oxford, 2005.

Ishaan Jhaveri
PHIL 2621: Minds and Machines
Professor: William Starr
Teaching Assistant: Chad McIntosh
May 14, 2015

2. Damasio, Antonio. *The Feeling of What Happens*. Orlando, 1999.