Could machines ever experience emotions like we do?

'Emotions make us human.'

This claim has governed our beliefs about human nature for millennia. As technological advancement and humanoid robotics challenge the boundaries of artificial intelligence, modern thinkers have become increasingly concerned about whether machines can replicate emotional experiences.

Whilst machines may imitate emotional responses, they cannot truly *experience* emotions as humans do. The irreducibly subjective, relational, and corporeal nature of human emotions remains beyond the reach of mechanisation. From theological, dualist, and materialist perspectives, this essay will argue that emotional experience is rooted in qualities unique to human existence – qualities that such human-made automata, regardless of sophistication, fundamentally lack.

What does it mean to "experience emotions like we do"?

Philosophers have long debated the nature of human emotion. Biologically speaking, human emotions are cognitive, relational, and physiological experiences. For instance, sadness can lead to tears, and happiness releases dopamine. Scientist Michel Cabanac defines human emotions as "*any mental experience with high intensity and high hedonic content*",[1] which,

[1] Cabanac de Lafregeyre, Michel. (2002). What is emotion?. Behavioural processes. 60. 69-83. 10.1016/S0376-6357(02)00078-5. Such "hedonic content" refers to the pleasure or displeasure associated with an experience.

according to Richard Dawkins in '*The Selfish Gene*',[2] often serve evolutionary purposes, like fear prompting the fight-or-flight response or love promoting reproduction.

From a phenomenological perspective,[3] human emotions are embodied phenomena empirically inseparable from our relational existence. For example, sadness manifesting as grief is a direct reaction that results from the death of a loved one and the loss of a relationship. What is more, emotions can unpredictably influence our decision-making, often leading us to act impulsively or *with the heart* rather than *with the head*.

Why would one think that a machine could experience emotions like we do?

Machines may simulate emotional responses, using algorithms to mimic emotional outputs responding to various inputs or stimuli. 'Guanghua No. 1' is a case in point of how machines can perform empathy; this humanoid robot uses facial recognition and algorithms based on the dopamine reward system in the human brain.[4]

However, it is vital to understand the decisive distinction between genuine sentience and simple responsiveness in these machines.[5] While they may *appear* to feel, their responses are pre-determined by algorithms, lacking the emotional intelligence invoked by 'qualia' necessary for genuine emotions.[6]

---

[2] Dawkins, Richard. The Selfish Gene. 1976. 30th anniversary ed., United States, Oxford University Press, 2006.
[3] A philosophical approach that studies structures of experience and consciousness from a first-person viewpoint (Stanford Encyclopedia of Philosophy).
[4] Flores, I. (2024). Fudan University Unveils Revolutionary Emotional Humanoid Robot 'Guanghua No. 1' at AI Conference. Tech Times. [online] 8 Jul. Available at: https://www.techtimes.com/articles/306456/20240708/fudan-university-unveils-revolutionary-emotional-humanoid-robot-guanghua-1-ai.htm [Accessed 2 Jan. 2025]. Proclaimed the "first emotional robot."

[5] The quality of being able to experience feelings.
[6] The introspectively accessible, phenomenal aspects of our mental lives (Stanford Encyclopedia of Philosophy).

Theological perspectives on human emotions and artificial replication

Judaeo-Christian theologians generally hold that the divine provenance of the Creator, beyond rational understanding, shapes human emotions. Thomas Aquinas defined emotion as central to our identity, tying inextricably to our souls rather than our mortal physicality.[7] He viewed such "*passions*" as spiritual connections coexisting with the Judaeo-Christian belief in humanity being created in "*imago Dei*".[8] In this light, emotions as human phenomena evidence divine guidance and connection to the noumena:[9] God.

It is inarguable, therefore, from a theistic perspective for man to "play God" and create a machine capable of such complex emotions, however intelligent and rational it may be.

Materialist and dualist perspectives on the mind-body problem

Materialist perspective

Materialism argues that humans are entirely physical beings, with emotions as purely biological processes, reducible to chemical reactions. Dawkins characterises humans as "*survival machines*" programmed to preserve their genes,[10] explaining all aspects of humanity through matter. He states emotions "*emerge from* […] *physical entities within the brain.*"[11] From this viewpoint, since little distinguishes between machines and humans, the

[7] Summa Theologiæ, 1a2æ 24,1 reply
[8] A theological concept meaning "image of God," referring to the belief that humans are created in God's likeness.
[9] In Kantian philosophy, the reality that exists independently of human perception; often associated with the divine or ultimate truth.
[10] Dawkins, 2006
[11] Dawkins, 2006

former could theoretically experience human-like emotions by emulating neural activity and mimicking the associated physiological reactions, such as releasing neurotransmitters. Dawkins suggests:[12]

> "*I see no reason why in the future we shouldn't reach the point where a human-made robot is capable of consciousness and of feeling pain*."

Here, he challenges the assumption that emotions require organic embodiment and implies that perceived differences between human and machine emotional capabilities reflect a limitation in our understanding of the physical world, not an inherent impossibility. If neural activity is the sole fount of human emotions, machines could feasibly replicate this activity, raising the possibility of creating sentient machines; we just do not know how.

However, this physiological perspective is self-contradictorily flawed. Emotions are not only abstract cognitive states but link closely to our bodily sensations and organic integration of interactions with the external environment into our internal reality.[13] In contrast to our organicity, machines rely on sensors to gather information but lack this organic immersion and connection with the world around them.[14] Thus, even if future technology would allow machines to mimic cognitive embodiment more convincingly, they would still be deprived of the relational connections that make human emotions meaningful.

Dualist critique

An argument that delivers a devastating blow to materialism is dualism, which criticises the materialist reductionism of complex human emotions,[15] conflictingly asserting that such emotions arise from the interplay between physical bodies and immaterial souls. Descartes'

---

[12] Big Think. "Richard Dawkins: A.I. Might Run the World Better than Humans Do | Big Think." Www.youtube.com, YouTube, 2017, www.youtube.com/watch?v=SM__RSJXeHA. Accessed Dec. 2024.
[13] Merleau-Ponty, 1962
[14] Korsakova-Kreyn, M. (2021). Emotion, embodied cognition, and Artificial Intelligence. Academia Letters: Marina Korsakova-Kreyn.
[15] The belief that complex phenomena can be explained by reducing them to their component parts.

distinction between mind and body emphasises that one cannot reduce emotions merely to physical processes; even if a machine could replicate neural activity, it would still fundamentally lack the immaterial soul needed for genuine emotional experiences.

Emotions are inherently relational and often arise from how we interpret the actions of others. Machines, lacking physical embodiment and personal and social loyalties, cannot replicate these heavily human aspects of emotional experience.[16] Human emotions are not isolated phenomena but localise themselves in our social, moral, and spiritual contexts.

John Searle's Chinese Room thought experiment is an illustration of this concept.[17] In the scenario, a person manipulates Chinese symbols according to specific rules without understanding the language. Although the responses may appear fluent, the individual crucially lacks proper comprehension. Equivalently, machines can "know how" to simulate emotions through processing inputs and generating outputs (syntax: processing rules) without understanding their meaning (semantics: meaning). They lack the "know why" – the fathomless understanding and subjective experience of human emotions. Thus, Searle's contradistinction between knowledge and understanding presents in a different light the inherent inability of machines to grasp the idiosyncrasy of emotions experienced solely by humans.

Free Will, determinism, and emotional authenticity

Humans often paint themselves as epitomising the ideal, morally free agents with a remarkable ability to take the road not taken. In contrast are machines: deterministic,

[16] Armstrong & Malcom, 1984, 101.
[17] Cole, David, "The Chinese Room Argument", The Stanford Encyclopedia of Philosophy (Winter 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2024/entries/chinese-room/>. A philosophical argument by John Searle that illustrates the difference between rule-following and meaning, questioning whether machines truly "understand."

algorithm-driven entities following predefined binaries. These entirely predictable and constrained vehicles of externally influenced action lack our power to deviate from their predestined fork in the road.

Free will allows humans to reflect on, struggle with, and act unpredictably upon their emotions in ways that machines cannot replicate due to deterministic programming.[18] Accepting that humans have free will is important to this discussion, as our emotions have a genuine and unpredictable impact on behaviour. A human might feel guilt over an impulsive decision and choose to act differently in the future; thus, symbiosis exists between free will, introspection,[19] and emotionalism.

Friedrich Nietzsche might argue that AI has an 'unfree will' as its lack of intentionality and conscious reflection suggests that humans have comparably sophisticated agency. We can make conscious and moral decisions, while machines do not engage in "free" introspective or self-aware activities.

Despite seeing free will as illusory (a romantic notion we indulge in to delude ourselves of agency), Dawkins is content that "*we can override biology with free will.*"[20] He acknowledges its evolutionary basis as a function of human consciousness and thus reinforces the idea that free will – and, accordingly, the ability for genuine emotional experience – lies in human exceptionalism.

In affirming that machines can experience human emotions, cognitive psychologists would be guilty of machine reductionism, reducing complex mental processes to simple computations and thus ignoring the influence of such emotion and motivation on human cognition,[21] which

[18] The ability of agents to choose actions independently of external determinism.
[19] The examination or observation of one's own mental and emotional state
[20] BrainyQuote.com. (2024). Richards Dawkins Quotes. [online] Available at: https://www.brainyquote.com/quotes/richard_dawkins_447478 [Accessed 12 Dec. 2024].
[21] Ansari, M. (2020). Criticism of Cognitive Psychology. [online] apsmcollege.ac.in, p.7. Available at: https://www.apsmcollege.ac.in/glassimg/thumb_album/1599764013-26.pdf.

are crucial in understanding how people think and act. David Hume famously argued, "*reason is* […] *the slave of the passions,*"[22] suggesting that emotions are fundamental drivers of unpredictable human behaviour. By contrast, we can consider AI the "*slave*" of the algorithms: an *artificial* simulation of human intelligence. Thus, these mechanic shams are often fallibly equated to authentic human emotional experiences.

Conclusion

In our age of AI, it is tempting to draw moral lessons from melodramatic Sci-Fi that warns us against the impending doom of a race of sentient machines set on bringing about human destruction. However, in any such imagination, machines would be unable to achieve emotional parity with us, as emotions are not mere algorithmic outputs; they are rooted in the nuanced interplay of reason, body, and will, which we cannot simplify as trivial practicalities.

Judaeo-Christian theology emphasises the role of the soul and divine purpose in shaping emotional experience, which materialist doctrines contend with by stressing emotions as biological mechanisms driven by our evolutionary survival instincts. Furthermore, phenomenological and dualist approaches argue that emotions are indissoluble from the human consciousness, free will, and embodiment, asserting that emotions are fundamental to what it means to be human and makes it clear that machines, no matter how advanced, will never be able to replicate the depth and authenticity of human emotion. These diverse philosophical approaches ultimately lead one to believe in the impossibility of disentangling the intricacies of emotional experience from the benevolent existence of human beings. Even arguments that mistakenly support the possibility of emotional AI succumb to their own inescapable pitfalls which essentially reveal the limitations of trying to replicate or understand human emotions through artificial means.

[22] (T II.3.3 415)

Despite the potential for future technological advancements, it is consequently irrefutable that machines cannot experience the ineffable depths, relational richness, and transcendence of human emotion – our irreplaceable paragon that no algorithm can replicate or relegate. Thus, the futile pursuit of creating emotional machines is not just a mechanic technicality but a phenomenal impossibility. Our humanity is an indisputable manifestation of our nonpareil capacity for emotional introspection, and in recognising this, we attest that the true marvel lies not in robotic performance but in the indefinite cognisance of our sentient existence.

Reference list

Ansari, M. (2020). Criticism of Cognitive Psychology. In *apsmcollege.ac.in* (p. 7). https://www.apsmcollege.ac.in/glassimg/thumb_album/1599764013-26.pdf

Aquinas, T. (1920). *Summa Theologiæ* (2nd ed.). Kevin Knight. https://www.newadvent.org/summa/. Original work published 1265-1274.

Armstrong, D. M., & Malcolm, N. (1984). *Consciousness and Causality: A Debate on the Nature of Mind*. Basil Blackwell Publisher Limited.

Big Think. (2017). *Richard Dawkins: A.I. Might Run the World Better Than Humans Do | Big Think*. Www.youtube.com; YouTube. https://www.youtube.com/watch?v=SM__RSJXeHA

Cabanac M. What is emotion? Behav Processes. 2002 Nov;60(2):69-83. doi: 10.1016/s0376-6357(02)00078-5. PMID: 12426062.

Cole, David, "The Chinese Room Argument", The Stanford Encyclopedia of Philosophy (Winter 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <https://plato.stanford.edu/archives/win2024/entries/chinese-room/>.

Dawkins, R. (2006). *The Selfish Gene* (30th anniversary). Oxford University Press. (Original work published 1976)

Flores, I. (2024, July 8). Fudan University Unveils Revolutionary Emotional Humanoid Robot “Guanghua No. 1” at AI Conference. *Tech Times*.

https://www.techtimes.com/articles/306456/20240708/fudan-university-unveils-revolutionary-emotional-humanoid-robot-guanghua-1-ai.htm

Goodyear, D. (2012, November 27). *11. The Soul as Software?* Philosophicaleggs.com. https://philosophicaleggs.com/11-the-soul-as-software/

Korsakova-Kreyn, M. (2021). *Emotion, embodied cognition, and Artificial Intelligence*. Marina Korsakova-Kreyn.

Krista, H. D. (2012). *Thomas Aquinas: Soul-Body Connection and the Afterlife* (p. 11). http://irl.umsl.edu/thesis261

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* (p. 3). University of Chicago Press.

McCarthy, J. (2007, November 12). *What is artificial intelligence?* Formal Reasoning Group; Computer Science Department, Stanford University. https://www-formal.stanford.edu/jmc/whatisai.pdf

McNerney, S. (2011, November 4). *A Brief Guide to Embodied Cognition: Why You Are Not Your Brain*. Scientific American Blog Network. https://blogs.scientificamerican.com/guest-blog/a-brief-guide-to-embodied-cognition-why-you-are-not-your-brain/

Merleau-Ponty, M. (1962). *Phenomenology of Perception*. Routledge.

Richards Dawkins Quotes. (2024). BrainyQuote.com; BrainyMedia Inc. https://www.brainyquote.com/quotes/richard_dawkins_447478

Tye, Michael, "Qualia", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/qualia/>.