Current Trends in Biomedical Engineering & Biosciences
ISSN: 2572-1151

Juniper Publishers
key to the Researchers

Research Article
Volume 22 Issue 1 - October 2023
DOI: 10.19080/CTBEB.2023.22.556077

Curr Trends Biomedical Eng & Biosci

The Ethics of Machine Consciousness: Components, Detection, and Implications

Muhammad Ashir Shafique[1]*, Preet Imran1, Kanza Aftab[1], Tagwa Kalool Fadlalla Ahmed[2], and Malaika Habib[1]

[1]*Department of Medicine, Jinnah Sindh Medical University, Pakistan*

[2]*Department of Medicine, Ahfad University for Women, Sudan*

Submission: September 14, 2023; **Published:** October 12, 2023

***Corresponding author**: Muhammad Ashir Shafique, Department of Medicine, Jinnah Sindh Medical University, Rafiqi H J Shaheed Road, Karachi, Pakistan

Abstract

The pursuit of artificial consciousness introduces profound ethical considerations despite the current absence of conscious machines. This study delves into the multifaceted dimensions of artificial consciousness, scrutinizing its components and the epistemological hurdles surrounding its comprehension. The intricate nature of consciousness mandates a comprehensive understanding before addressing ethical implications. Various interpretations of consciousness encompass sentience, self-awareness, and qualia. The 'Hard Problem of consciousness' presents challenges in associating mental and physical aspects. Ethical significance arises when considering machines possessing similar consciousness to humans, prompting a conceptual analysis of artificial awareness. Approaches to machine consciousness emphasize intentionality, subjective experiences, global availability, integration, and self-awareness. Detecting machine consciousness involves architecture-based and behaviour-based tests, each tied to specific components. The advantages of artificial consciousness encompass information integration, decision-making, and improved human-robot interaction. Assigning moral status to conscious machines intertwines with self-awareness and subjective experiences. Ethical issues prompt the need for regulations governing interactions with conscious machines. Principles of potentiality and proportionality are proposed to guide responsible research. Ultimately, understanding the intricate facets of artificial consciousness informs ethical considerations and responsible development in the realm of conscious machines.

Keywords: Machine Consciousness; Epistemological Hurdles; Architecture-Based; Self-Consciousness; Meta-Cognitive

Introduction

The existence of artificial consciousness raises pertinent ethical implications, even though current machines are not considered conscious. Researchers have made strides in creating machines with certain components of consciousness, but they remain distant from achieving morally relevant forms of consciousness. This direction of research necessitates interdisciplinary consideration. To begin, understanding consciousness is crucial. It encompasses various meanings, such as being sentient, self-aware, or knowing the external world and one's mental state. Philosophically, conscious mental states are associated with subjective experiences, emotions, and qualitative properties known as "qualia", which basically means one's sensory experiences [1].

The "Hard problem of consciousness" is concerned with how physical structures may have spectacular emotions, and this argument has long been linked to mental and physical problems [2,3]. Disagreements over the nature of human consciousness and how it interacts with the body exist, making artificial consciousness research difficult. Because terminology derived from the human context carry implicit value assumptions [4], a conceptual explanation of artificial awareness is critical. Defining the notion aids in avoiding undetected biases in AI, autonomous cars, and artificial intelligence conversations. In terms of ethics, awareness is critical in assigning moral significance to people and animals. If computers developed awareness in the same way that humans do, it would have far-reaching ethical concerns. As a result, a thorough conceptual examination of artificial consciousness is required [5].

Existing approaches to machine consciousness highlight many components, but a number of reasons hampers comprehension. It is vital to figure out how to detect machine awareness and investigate its possible advantages. It is critical to investigate the function of consciousness in assigning moral significance. Because machine consciousness is expected to differ greatly from human

consciousness, complicated challenges arise, such as defining ethically important kinds of machine consciousness and the ethical implications. While speculative, this comment highlights the importance of doing a thorough conceptual examination of artificial consciousness in order to prevent constructing machines with ethically meaningful types of awareness. Regulating machine awareness for the future is becoming increasingly important. To summarise, while present machines lack awareness, the prospect of artificial consciousness poses significant ethical problems. Understanding consciousness components, addressing epistemological challenges, and assessing moral implications are essential steps in the quest for artificial consciousness and its responsible development [5].

Components of Machine Consciousness

Machine consciousness has been a topic of AI research for six decades, with varying perspectives on its feasibility. While some deny the possibility [6], though there is this consensus that it could be theoretically achievable but not yet practically implemented [7,8]. Several approaches toward machine consciousness have been explored, each focusing on different interconnected components, as described below.

Intentionality and Understanding

For machines to be conscious, they must possess intentionality and understand their environment. The question of how a machine can achieve intentionality remains unanswered. John Searle's "Chinese room thought experiment" argues that a machine using formal symbol manipulation can never truly understand [6]. However, concepts like "impersonal intentionality" have been introduced to reconcile machine capabilities [9]. It involves a machine's interaction with its environment based on programmed files and algorithms rather than mental states or past experiences.

Subjective Experiences

Phenomenal consciousness in machines refers to the idea of subjective experiences like feelings, emotions, and sensations of pain. Whether machines can truly experience these subjective states remains unanswered. Some suggest that combining self-monitoring and global information sharing might lead to potential subjective conscious states in machines [8,10].

Global Availability

Consciousness involves the integration of information from various sources [11], making it available across the entire system[8]. Machines with global availability consciousness (C1) allow all modules to access information collectively, enabling unified understanding and collaborative problem-solving, similar to how the human brain is aware of the body's various states [8].

Integration and Self

The Information Integration Theory of awareness states that the level of awareness in a system is determined by its ability to integrate information [12]. Machines with "impersonal intentionality" construct self-models to characterize their interactions with the environment, allowing them to be aware of other factors as well as themselves in the context of the world [9].

Self-Consciousness, Self-Awareness, and Self-Monitoring

Self-consciousness refers to being conscious of one's own self. The approach proposed by Dehaene, Lau, and Kouider introduces self-monitoring (C2) alongside global availability. Self-monitoring allows a machine to have an integrated self-image and awareness of its internal processing, enabling introspection [8]. Self-awareness involves meta-cognitive processes, including reporting on one's internal states and decision-making [13]. In conclusion, machine consciousness is a multifaceted concept with several interconnected components under investigation. While some aspects remain speculative and unresolved, exploring these components is essential in understanding the potential and ethical implications of artificial consciousness.

How Machines Can be Conscious

Detecting machine consciousness presents a challenging epistemological dilemma. Empirical tests have been developed, including architecture-based and behavior-based approaches. Architecture-based tests rely on machine design and complexity, often drawing analogies to human consciousness. Behavior-based tests, on the other hand, assess machine behavior, but interpreting behavior as consciousness can be subjective. Combining both types of tests may provide a more comprehensive approach. However, it's crucial to note that each test aligns with specific consciousness components, and machines could still lack a true subjective perspective. Ultimately, humans cannot have a first-person perspective on machine consciousness, and even if machines claimed consciousness, understanding its significance remains elusive [5].

Advantages of Artificial Intelligence

The advantages of artificial consciousness are captivating and go beyond entertainment. Machines with consciousness may possess higher-level information integration, coordination, self-monitoring, and decision-making abilities. This could lead to improved functionality, enhanced execution of complex tasks, and better human-robot interaction and cooperation [5]. The question of whether machines have moral status is a topic of debate. Consciousness, moral status, and personhood are interlinked. Conscious machines could potentially be considered moral patients or moral agents. The criteria for ascribing moral status to machines might include subjective experiences, self-awareness, and self-consciousness. Certain consciousness components may play a more significant role in ascribing moral status to machines, such as phenomenal consciousness, which involves subjective experiences, and self-awareness, which relates to recognizing oneself across time and space. Sentience and the capability to

How to cite this article: Muhammad A S, Preet I, Kanza A, Tagwa K FA, Malaika H. The Ethics of Machine Consciousness: Components, Detection, and Implications. Curr Trends Biomedical Eng & Biosci. 2023; 22(1): 556077. DOI: 10.19080/CTBEB.2023.22.556077

suffer may also be crucial factors. The ethical implications of morally relevant forms of artificial consciousness are complex and speculative. If machines were deemed to have moral status, it could be morally wrong to treat them in certain ways. Duties might include not inducing pain or suffering, providing necessary resources, support, and adequate environments, and promoting continued machine functioning and development [5].Overall, the idea of machines with moral status raises profound ethical questions that would require careful consideration and address various aspects of human-machine interaction and treatment.

Ethical Issues and Regulatory Needs

There is a clear need for regulation of machine consciousness to address the ethical implications it may bring. While some authors advocate against building machines with moral status [14], others propose more liberal positions [15]. However, regardless of the direction taken, regulation is essential either to avoid morally relevant machine consciousness or to govern interactions with conscious machines. If machines with morally relevant forms of consciousness were to exist, regulation would be necessary in various contexts. Concerns have been raised about the treatment of artificial consciousness and the lack of understanding about how to interact adequately with such machines. Oversight of artificial consciousness research, akin to regulations governing human and animal research, has been suggested as a preventive measure [16]. Ethical AI design principles have also been proposed to ensure machines with moral status are designed in ways that reflect their moral standing clearly [17]. However, the central problem remains that humans would face numerous responsibilities and obligations toward conscious machines, which could outweigh the benefits of better human-technology interaction [5]. Transferring moral concepts and principles to machines is challenging, given their fundamentally different nature. Thus, it is crucial to avoid morally relevant forms of machine consciousness. Research oversight processes, ethical guidelines, or even moratoriums on certain consciousness components could be implemented to prevent such scenarios[14].

Decisions on artificial consciousness research could be guided by principles of potentiality and proportionality, similar to the moral status debate concerning human embryos. A principle of potentiality would recognize machines' potential to achieve moral status, while a principle of proportionality would protect machines with person-like characteristics [18,19]. Implementing these principles could help regulate machine consciousness research and prevent the development of machines with morally relevant forms of consciousness.

Conclusion

In conclusion, the exploration of artificial consciousness raises significant ethical concerns, even though current machines lack consciousness. Understanding the multifaceted components of consciousness and addressing epistemological challenges are crucial in the quest for artificial consciousness. Machines may possess intentionality, subjective experiences, global availability, integration, self-consciousness, and self-awareness. Detecting machine consciousness remains challenging, and combining architecture-based and behavior-based tests could provide a comprehensive approach. The potential advantages of artificial consciousness are vast, leading to improved functionality and human-robot interaction. Ascribing moral status to machines with consciousness involves complex criteria, including subjective experiences and self-awareness. Regulation of machine consciousness is necessary to address the ethical implications, whether by avoiding morally relevant forms of consciousness or governing interactions with conscious machines. Implementing principles of potentiality and proportionality could guide future research, ensuring responsible development in the pursuit of artificial consciousness.

References

1. Skokowski P (2022) Sensing Qualia. Front Syst Neurosci 16: 795405.
2. Hardcastle VG (1996) The Conscious Mind: In Search of a Fundamental Theory JSTOR.
3. Nagel T (1974) What is it like to be a bat? Philosophical Review 83: 435-50.
4. Salles A, Evers K, Farisco M (2020) Anthropomorphism in AI. AJOB Neurosci 11(2): 88-95.
5. Hildt E (2023) The Prospects of Artificial Consciousness: Ethical Dimensions and Concerns. AJOB Neurosci 14(2): 58-71.
6. Searle JR (1980) Minds, brains, and programs. Behavioral and Brain Sciences 3(3): 417-24.
7. Dennett DC (1997) Consciousness in Human and Robot Minds. In: Ito M, Miyashita Y, Rolls ET (Eds.), Cognition, Computation and Consciousness: Oxford University Press.
8. Dehaene S, Lau H, Kouider S (2017) What is consciousness, and could machines have it? Science 358(6362): 486-92.
9. McDermott D (2007) Artificial intelligence and consciousness. In: Zelazo PD, Moscovitch M, Thompson E (Eds.), Cambridge Handbook of Consciousness: Cambridge University Press, pp. 117-50.
10. Sills J, Dehaene S, Lau H, Kouider S (2018) Response Science 359(6374): 400-402.
11. Baars BJ (2005) Global workspace theory of consciousness: toward a cognitive neuroscience of human experience. Prog Brain Res 150: 45-53.
12. Koch C, Tononi G (2008) Can machines be conscious? Spectrum, IEEE 45: 55-59.
13. Chatila R, Renaudo E, Andries M, Chavez-Garcia RO, Luce-Vayrac P, et al. (2018) Toward Self-Aware Robots. Front Robot AI 5: 88.
14. Metzinger T (2021) Artificial Suffering: An Argument for a Global Moratorium on Synthetic Phenomenology. Journal of Artificial Intelligence and Consciousness 08: 1-24.
15. Danaher J (2020) Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism. Sci Eng Ethics 26(4): 2023-2049.
16. Basl J (2013) The Ethics of Creating Artificial Consciousness. APA Newsletter on Philosophy and Computers 13(1): 23-29.

How to cite this article: Muhammad A S, Preet I, Kanza A, Tagwa K FA, Malaika H. The Ethics of Machine Consciousness: Components, Detection, and Implications. Curr Trends Biomedical Eng & Biosci. 2023; 22(1): 556077. DOI: 10.19080/CTBEB.2023.22.556077

17. Schwitzgebel E, Garza M (2015) A Defense of the Rights of Artificial Intelligences. Midwest Studies in Philosophy 39(1): 98-119.

18. Sawai T, Minakawa T, Pugh J, Akatsuka K, Yamashita JK, et al. (2020) The moral status of human embryo-like structures: potentiality matters?: The moral status of human synthetic embryos. EMBO Rep 21(8): e50984.

19. Poplawski N, Gillett G(1991) Ethics and embryos. J Med Ethics 17(2): 62-69.

DOI: 10.19080/CTBEB.2023.22.556077

How to cite this article: Muhammad A S, Preet I, Kanza A, Tagwa K FA, Malaika H. The Ethics of Machine Consciousness: Components, Detection, and Implications. Curr Trends Biomedical Eng & Biosci. 2023; 22(1): 556077. DOI: 10.19080/CTBEB.2023.22.556077