



Memory: A Breakthrough Point in The Construction of Ethical Identity for Social Robots [†]

Ruofan Li ¹ , Zhaolong Peng ^{2,*} and Dazhou Wang ³

¹ School of Marxism, Fudan University, Shanghai 200433, China; liruofan2020@163.com

² School of Marxism, Hohai University, Nanjing 210098, China

³ School of Humanities, University of Chinese Academy of Sciences, Beijing 100049, China; dzwang@ucas.ac.cn

* Correspondence: rfli22@m.fudan.edu.cn; Tel.: +86-135-4300-9370

[†] Presented at Forum on Information Philosophy—The 6th International Conference of Philosophy of Information, IS4SI Summit 2023, Beijing, China, 14 August 2023.

Abstract: Social robots are tangible autonomous robots capable of interacting and communicating with humans. The contradiction arises between their rejection of moral nihilism and the current lack of ethical identity, leading to a dual identity crisis of “agent-receiver”. Constructing an ethical identity is advantageous in enhancing the socialization of social robots and facilitating their integration into society. Through a three-dimensional analysis of the “institution-literature-experience”, the key to constructing the ethical identity of social robots lies in “memory”: social robots need to possess “memory morality” software that enables them to adapt to diverse moral contexts using multiple moral behaviors. Memory assists in establishing the ethical identity of social robots, emphasizing the importance of considering the subject status of social robots and harnessing the mediating role of various subjects to achieve a transition from “moral memory” to “memory morality”.

Keywords: social robots; ethical identity; memory



Citation: Li, R.; Peng, Z.; Wang, D. Memory: A Breakthrough Point in The Construction of Ethical Identity for Social Robots. *Comput. Sci. Math. Forum* **2023**, *8*, 11. <https://doi.org/10.3390/cmsf2023008011>

Academic Editors: Zhongzhi Shi and Wolfgang Hofkirchner

Published: 10 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

On 20 March 2022, the General Office of the Communist Party of China Central Committee and the General Office of the State Council issued the “Opinions on Strengthening the Governance of Science and Technology Ethics”, stating the importance of adhering to the principle of “ethics first” and promoting the coordinated and responsible development of scientific and technological activities with ethics, fostering a responsible innovation. Social robots represent the fusion of ethics and technology, serving as the embodiment of the interaction between “social beings” and “robots”. The significant ethical implications of this technological development have been widely recognized. However, without an ethical identity, “deep learning” cannot bring about “deep memory”. Social robots embody a profound dual identity conflict of “agent-receiver” and cannot find resolution within the existing academic framework. Therefore, this article aims to promote the socialization of social robots, using “memory” as the breakthrough point for philosophical and technological dialogue, and explores the specific path of memory’s assistance in constructing the ethical identity of social robots from the perspective of the “subject-object-process”. This approach not only contributes to the integration of philosophical research on social robots and innovative practices, but also helps reduce the irrationality in the process of social robot innovation and promotes responsible innovation in the field of social robots.

2. The Logical Framework for Constructing the Ethical Identity of Social Robots

Social robots are tangible autonomous robots capable of interacting and communicating with humans. They include categories such as educational robots, medical robots, and military robots. Unlike industrial robots, social robots have a wide range of interactions

and closer proximity, which entails higher ethical obligations. However, the lack of ethical identity places the “social” significance of social robots in an ambiguous position.

2.1. The Technological Advancement of Social Robots Requires Ethical Identity

The technological progress of social robots requires ethical identity. Scientists have tirelessly explored the goal of “humanoid” and pushed robots from the laboratory to the social scene. Moreover, social robots constantly surpass and simulate human thinking and behavior in practice [1]. The widespread infiltration of social robots into the social division of labor makes it inevitable for human–robot interaction scenes.

2.2. The Non-Neutrality of Social Robots Rejects Moral Nihilisms

The non-neutrality of social robots rejects moral emptiness. Social robots have sociality, deception, and ideology. Social robots are non-neutral and not morally innocent [2]. In the design, construction, and operation of social robots, they are not simply an extension of human will and ability, but a transcendent carrier form that can achieve goals completely different from the original intention of the operators in active or passive situations, becoming an incomplete moral agent.

2.3. The Dual Ethical Identity of Social Robots Creates a Crisis

The dual ethical identity of social robots creates a crisis. Independent and highly embedded social robots embody a dual ethical identity, mainly reflected in the position of a moral agent in moral action, but in moral evaluation, it transforms into a certain passive moral patient [3]. Even if an unethical concept becomes reality, social robots will not be subject to any effective moral condemnation or moral self-blame. The lack of self-discipline in moral behavior restricts the advancement of social robots, which means that any unethical behavior ultimately shifts from social robots to the designers, manufacturers, and users who are subject to moral norms and industry-specific regulations. Today, as deep learning continues to advance, more imaginative and futuristic suggestions point to a path that endows social robots with a more stable ethical identity using memory as an intermediary to alleviate potential moral crises.

3. The Memory Aspect of Constructing the Ethical Identity of Social Robots

In the context of the dual ethical identity conflict, the construction of the ethical identity of social robots is no longer a sensational issue. It requires a different approach to promote the development of the socialization level of social robots. “Memory” can serve as a breakthrough point for philosophical and technological dialogue, playing a crucial role in the construction of the ethical identity of social robots from the following perspectives.

3.1. Towards the Subject: From Subject Memory to Memory Subject

The ethical identity bestowed upon social robots is a product of the development of artificial intelligence technology to a certain stage. With the continuous evolution and advancement of artificial intelligence technology, highly autonomous decision-making and interactive perceptiveness have become prominent features of social robots. Although the word “robot” originates from the Czech word “robota”, meaning labor or hard work, with the deepening of human–machine interaction, robots have found broader applications. Social robots have gradually gained warmth in people’s perception, rather than being seen merely as cold and purely functional tools. Humans are also attempting to bestow ethical identity upon social robots [4]. When discussing the application of social robots from an ethical perspective, it is conceivable that social robots can develop moral capabilities through deep learning. Compared to humans, robots may exhibit more stable and enduring moral abilities, which makes it possible for social robots to participate in traditional ethical scenarios.

“Memory is an important category for understanding the relationship between artificial intelligence and humans” [5] (p. 187). For a long time, due to the absence of ethical roles, robots have been perceived as objects of human perception,

recording, and storage within memory. The limitations of memory capacity have led to a circular reasoning when assigning subject identity. Memory provides a reasonable entry logic for shaping the ethical identity of social robots and is an important condition for the existence of their ethical identity. Consequently, memory becomes a key factor in constructing the moral capabilities of robots."

3.2. *Towards the Process: From Moral Memory to Memory Ethics*

The transformation from "moral memory" firmware to "memory morality" software enables robots to adapt to different moral situations using diverse moral behaviors. Moral memory is a special form of memory that is the product of the memory ability of humans to record their unique moral life experiences. Similarly, as intelligent beings that integrate the mind and body, the prerequisite for social robots to obtain ethical identity is having memory ability. From the perspective of memory, the process of intelligent beings forming moral ability can also be understood as a process of deep learning, which is a process of continuous evolution based on experience. On this basis, the conclusion derived is that memory plays a crucial role in establishing an effective connection between evolution and intelligent beings. However, moral memory is always a low-level storage of moral experience, which does not include moral value judgment, selection, and evaluation. Memory morality, on the other hand, is an indicator of a higher level of moral value judgment, which is a moral form formed via value judgment, moral selection, and other processes of existing moral memory behaviors. It gives social robots a normativity that unifies their cognition and behavior: deliberating what kind of moral situation or moral life experience that should be remembered and applied to their own moral behavior.

4. The Specific Path of Memory Assisting in Constructing Ethical Identity for Social Robots

Through the analysis of existing institutional texts, academic literature, and practical experience, the construction of ethical identity for social robots with the assistance of memory needs to follow the following paths.

4.1. *Pay Sufficient Attention to the Subject Position of Social Robot Memory*

It is necessary to fully respect the subject status of social robots. The setting of the subject status of the moral memory of robots is the key to the problem and also a practical need. As objects created by humans, robots are increasingly exhibiting the trend of object-subjectification under data-driven algorithm programs. With the gradual improvement of their humanoid intelligence and behavioral abilities, the existence of intelligent machines that were once cold and ruthless has transformed into a certain form with a subject life structure.

4.2. *Effectively Utilize the Mediating and Regulatory Roles of the Two Subjects*

It is vital to play the intermediate regulatory role of the management subject and the task subject. The management subject is the relevant enterprise that develops social robots and should take responsibility while pursuing economic interests. The relationship between human-machine interaction has become increasingly complex, and the development of artificial intelligence is inseparable from the behavior of robots. This requires relevant enterprises to focus on infusing humanistic spirit into the program design of artificial intelligence technology during the development process and realize the value of guidance by human beings on intelligent technology. The task of the task subject, such as data annotators who directly participate in the development and operation of social robots, is to fit human cognition and provide learning materials for artificial intelligence. They need to produce standard data that meets the requirements of deep learning algorithms. This also requires attention to the moral level and moral judgment ability of data annotators and also allows data annotators to shift from passive to active. They play a crucial role in the field of artificial intelligence deep learning and cannot be viewed as mere tools while ignoring their value.

4.3. Effectively Utilize the Mediating and Regulatory Roles of the Two Subjects

To enable social robots to transition from moral memory to memory morality, the following approaches can be employed: First, supplement moral knowledge to enhance moral memory. Social robots need to acquire a foundational and systematic understanding of existing moral norms, principles, and basic categories. This forms the basis for constructing their ethical identity. Only when social robots possess the relevant moral knowledge and actively engage in moral life, forming moral experiences, can they have the prerequisite for moral memory. This lays a solid foundation for the subsequent transition to memory morality.

Second, create moral scenarios to shape moral memory. By engaging in specific moral practices, robots can develop situational memories that contribute to their moral capabilities. Although humans have innate moral instincts, moral life is complex, and moral contexts vary. People's moral cognition and judgment of certain events differ across different situations. Therefore, when constructing moral scenario memories for social robots, it is important to expose them to diverse moral contexts, allowing them to imitate and learn from humans with higher moral standards. This facilitates the formation of situational memories and, more importantly, the construction of memory morality. Memory morality should serve as a guiding principle, directing moral memory. In future real-world applications, when social robots are placed in such moral contexts, they can utilize stored moral memories to engage in appropriate moral behavior and judgment.

Third, foster individual interactions to reinforce memory morality. The essence of morality lies in the behavioral norms that regulate the relationships between individuals and between individuals and society. To truly discuss the construction of ethical identity in social robots, it is necessary to analyze the trust and obedience relationship between humans and machines within the framework of Edmund Husserl's phenomenology. Current research focuses more on human trust in social robots, while Matthew E. Gladden examines the obedience relationship between humans and social robots from a more specific perspective, particularly in the context of robot leadership. For a robot to assume a leadership role, it requires not only a physical presence but also an emotional "heart" [5]. Professor Yang Qingfeng proposed the concepts of "rational intelligent agent", "moral intelligent agent", and "emotional intelligent agent" [6] (p. 210), which provide deeper possibilities for the discussion of leadership robots. A leadership robot should possess the ability to think rationally and make moral judgments. Although humans are the creators of leadership robots, in practical operations, these robots should exhibit an intelligent agency. Can a leadership robot exist as a stable moral entity? Yuval Noah Harari suggests that artificial intelligence is more stable than humans in executing a predetermined political and moral position. However, it lacks its own position since computer algorithms are not shaped by natural selection and do not possess emotions or intuitions. Therefore, in critical moments, their ability to adhere to ethical morals surpasses that of humans. By employing deep learning, social robots can establish memory morality, which may result in a moral capacity that is more enduring and stable than what humans possess. When facing ethical dilemmas, social robots are likely to exhibit a stronger moral will in their decision-making processes.

Author Contributions: Conceptualization, R.L., Z.P. and D.W.; methodology, D.W.; writing—original draft preparation, R.L.; writing—review and editing, Z.P.; supervision, D.W.; project administration, D.W.; funding acquisition, D.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by University of Chinese Academy of Sciences; grant number: 20204002340. And The APC was funded by University of Chinese Academy of Sciences.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study did not create new data.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lake, B.M.; Salakhutdinov, R.; Tenenbaum, J.B. Human-level concept learning through probabilistic program induction. *Science* **2015**, *350*, 1332–1338. [[CrossRef](#)] [[PubMed](#)]
2. Freenberg, A. *Critical Theory of Technology*; Oxford University Press: New York, NY, USA, 1991; p. 5.
3. Verbeek, P.P. *Moralizing Technology: Understanding and Designing the Morality of Things*; Shanghai Jiaotong University Press: Shanghai, China, 2016; p. 9.
4. Frey, C.B.; Osborne, M.A. The future of employment: How susceptible are jobs to computerisation. *Technol. Forecast. Soc. Chang.* **2017**, *114*, 254–280. [[CrossRef](#)]
5. Matthew, E.G. The Social Robot as ‘Charismatic Leader’: A phenomenology of Human Submission to Nonhuman Power. In *Proceedings of the Robot-Philosophy*, Copenhagen, Denmark, 1 December 2014; pp. 329–339.
6. Qingfeng, Y. *Memory Research and Artificial Intelligence*; Shanghai University Press: Shanghai, China, 2020; p. 187.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.