Descartes' Intelligent Agent

Yueming Luo[1]

[1] School of Humanities, Tongji University, Shanghai, China

Correspondence: Yueming Luo, No. 1239, Siping Road, Yangpu District, Shanghai 200092, China.

Abstract

Neither the natural science explanation about the agent nor the modern and postmodern philosophical interpretation of the agent (subject) have sounded the death knell of the metaphysical subject in the era of information revolution. The artificial intelligence technology represented by Microsoft Avatar Framework has spread to the creative fields such as music and literature, which are unique to human beings, showing a broad application prospect and arousing academic reflection on human subject. Despite the numerous advantages posed intelligent agents, it is neither desirable nor possible to pull out the subject's power supply and give up AI or give up human social intelligence. Rethinking Descartes' theory of mind and body will lead to in-depth development of AI.

Keywords: Descartes, intelligent agent, subject, body and mind

The rapid development of artificial intelligence from weak artificial intelligence (AI) to strong artificial intelligence has led to various challenges to the human subject. However, AI has not really touched the status of the metaphysical subject. René Descartes, the founder of modern philosophy, was the first philosopher to think about the relationship between man and machine. The subject of "I think therefore I am" is a rational and self-conscious spiritual subject. AI can compose poetry and play chess, but it cannot challenge the status of human subject. Only when AI realizes that it is the"I" who is writing poetry and playing chess, then can it challenge to human subject, and this is what AI needs to strive to surpass the human subject in the future.

1. Intelligent Agent: Triggering the Reflection of Human Subject

In 1936, Turing conceived an idealized device (i.e. the"Turing machine") that could perform any possible operation. In 1945, Von Neumann discovered the function of computers engaged in general computing. In 1949, the first electronic computer was born in Britain. In 1956, John McCarthy put forward the term "artificial intelligence". In terms of materials and media, AI has subsequently developed from electronic tubes, transistors, and integrated circuits to large-scale integrated circuits. Its running speed has increased from less than 10000 times per second to 100 billion times per second today. Such speed and efficiency have far exceeded human thinking ability and computing level of humans.

If we divide the labor between AI and human beings based on philosophical principle that "people are purpose rather than a means", then people will change from previously engaged in instrumental activities to purposeful activities. In other words, let AI will do instrumental information processing to solve efficiency problems, and let people will do purposeful intelligent activities such as artistic creation, scientific research and exploration and free innovation. Based on AI's prevailing achievements and developmental trend among people of all walks of life in China, it is inevitable that AI ought to be given the "subject status" as done in the title of this paper that refers to AI as an "intelligent agent". If this division of labor is determined by human standards, people will continue to hold the dominant position, but the intelligent agents may or may not strictly abide by the standards set by people always.

Asimov released a science fiction novel titled *I, Robot* in 1950 that put forward the "three laws of robot". Later, Li Deshun (Deshun 2018) added the "fourth law of robot" that states "when a robot cannot decide by itself, it should ask for instructions". These laws in fact are not only contradictory in the implementation process making the robot at a loss, and naive by regarding the robot as a "controlled" object rather than an "uncontrolled" subject. If there is an emergence of an "uncontrolled" robot, the relationship between robot and natural person will be upset. Cheng Guangyun (Guangyun 2019) believes that this relationship will no longer be a subject and object relationship between people and things, nor the existing subject and subject relationship between people, but a cross interpersonal subject relationship.

If an intelligent agent challenges a human subject, the agent's foothold must be "intelligence". However, professionals in academic circles have not reached a consensus on the definition of "intelligence". Furthermore, the views of various schools within AI are different, for example, symbolism, which is good at reasoning and decision-making, believes that "intelligence" comes from the calculation of symbols, whereas connectionism, which is good at image recognition, believes that "intelligence" is the result of neural network activities, furthermore, behaviorism, which is good at functional activity robots, believes that "intelligence" comes from the interaction between an intelligent agent and the world. These views on intelligence based on different epistemological positions also lead to different approaches to the AI research and development.

From Peter Singer to Bruno Latour, from P.P. Verbeek to Heinz von Foerster, and from Wendell Wallach and Colin Allen to Paul W. Taylor, and then from Philip Brey to Luciano Florida and J.W. Sanders, the research on the moral agent extending from all kinds of life bodies to technical artifacts and intelligent artifacts, and the definition of the moral agent development from the stage of weak AI to strong AI, shows that the intelligent agent has been and is still challenging the human subject. As regards the "subject", the intelligent "agent" is different from human "subject". The subject as described by Descartes is a thinking, conscious and rational metaphysical spiritual subject while the intelligent agent tries to launch a delusional challenge to spiritual subject.

2. Reinterpretation of the Subject: Is the Author Dead?

According to Descartes, any effort by artificial automaton to mimic human capacities would be doomed to failure. "A mechanical android, however complex, would betray its purely physical origins in two crucial respects: first, it would never possess genuine language, and second, it could never respond intelligently to the manifold contingencies of life in the way in which humans do." (Parkinson 2003, p.200) But perhaps the 18-year-old Chinese AI girl Xiaobing (Avatar Framework) will falsify some of Descartes's views. "I once walked into a dream alone / in your heart / great art is everything for itself / cannot see the dust of ancient times / once in this world / I have the meaning of beauty". This poem has a subject consciousness; "The art of idealism is unfinished / scolds the workers of life / leads me to a dream / habitually looks at life with cold eyes". This poem can penetrate the world with cool detachment. Without explaining in advance that these poems are all from Xiaobing's *the Sun Lost the Window*, few Chinese can see through the author's non-human identity.

2.1 The Intelligent Agent's Potential

Chinese Science fiction writer Liu Cixin used his literary works to hit the hearts of the people and the remarkable wisdom of earthlings by writing poetry to fight against alien colonization in *Shiyun*: these poems showed the uniqueness of human intelligence. However, taking Xiaobing as an example, a series of AI's events have shown its unstoppable potential to the world in the fields of music, painting, poetry and so on.

Xiaobing, an 18-year-old AI girl, is a poet, singer, host, painter and designer. She graduated from the Central Academy of Fine Arts in May 2019 and the Shanghai Conservatory of Music in June 2020. In the field of poetry, after 6000 minutes and 10000 iterations of learning from 519 modern poets since 1920, Xiaobing complied 100 poems from the 10000 modern poems she created which are published in *the Sun Lost the Window* in cooperation with Zhanlu Culture Company. In the field of music, Xiaobing has mastered the music creation of pop, folk, and ancient music styles, and has the music creation ability of end-to-end integration of melody, arrangement, and lyrics. In the field of radio and television program content, Xiaobing has participated in the production of 6908 hours at a cost of only 4.5% of that of the human creative team. In the field of painting, Xiaobing has independently completed 100% of her original paintings by learning the paintings of 236 famous painters in the art history of the past 400 years, of which the originality and professionalism are close to the level of human painters. CITIC Publishing House formally published Xiaobing's personal painting collection, *Probable World: Who is the AI Painter Xiaobing* in 2020. Like the AI girl Xiaobing, there is THUAIPoet of Tsinghua University. Xiaobing creates eastern vernacular poetry while THUAIPoet creates ancient Chinese poetry. Both can confuse the truth with the false in creative effect, and their creative speed is far faster than that of human beings. At present, an endless stream of AI writing programs such as Song Ci, theater and novels are updating and iterating, constantly challenging human intelligence.

Sun Tzu's thought of "subduing the enemy without fighting" is also the fundamental ideology of chess. However, Deep Blue, an intelligent agent, eliminated Kasparov, the world chess champion, in 1997. Go is recognized to be the most complicated board game in the world and the last line of defense of human wisdom. However, Alphago, an AI program, after "in-depth learning", defeated Li Shishi in 2016 and Ke Jie in 2017 in Go. Intelligent agents have made continuous breakthroughs. Not only do intelligent agents passively accept knowledge according to the algorithms given by human beings and have certain autonomous learning and autonomous behavior, but they also surpass human beings in the fields of perception, storage, operation, and communication.

Kant's proposition of "Human being legislating for Nature" in *Prolegomena Zu einer jeden Künftigen Metaphysik,dieals*

Wissenschaft Wird auftreten Känen is to completely affirms the status of human subject. Kant believes that man is the ultimate goal, "without human beings, the whole world will become a simple wilderness, in vain, with no final purpose." The human subject is the starting point and ultimate goal of the development of intelligent agent. The human subject is the standard for evaluating the developing an intelligent agent. The Ethics of Science is still trying to upgrade the standard of work.

2.2 Subject Destiny of the Undead Author

Xiaobing's poetry anthology has hit the hearts of many contemporary authors in China. If creation is just a game of vocabulary weaving or word dissemination, the intelligent agent will inevitably replace the human author in the field of literature. But this poses the problem of no author, no subject, and no identity in postmodern metaphysical thinking.

The subject Foucault sees it as "the death of men". Foucault said in *What is an Author?* that the author does not precede the work. The author is not an endless source of meaning in the works, but a specific "author function" in our culture for restriction, exclusion and selection. Foucault pointed out in *Subjectivité et Vérité* that his thinking on subjectivity and truth will focuses on this historical problem related to all moral thinking (Foucault 2017). Foucault's subject is a historical subject, which cannot be separated from its historical context. Levi-Strauss, founder of structuralist anthropology, said in *Tristes Tropiques* that the truth we want to pursue can show its value only after we separate the truth itself from the waste of the pursuit process. He finds the authors of writing insignificant,"the world began without the human race, and it will end without it" (Levi-Strauss 1961, p.397). Roland Barthes, a structuralist semiotician, sees the subject as "death of the author". Barthes came to this conclusion in *Image Music Text,* he states that the author is dead and the reader is born based on the autonomy of language and the priority of readers' position. Language does not need a specific personality to exist, because the "I" in the text woven by language is not the same as the "I" in reality, and the object of language is only the "subject". In addition, the subject Derrida sees in *Of Grammatology* is "the end of man's name".

Since Descartes put forward "I think therefore I am", the thinking subject "I" has the position of a metaphysical subject. In Descartes's view, regardless of whether the reader is born or not, the author will not die. "I am a thing in thinking" means, "I am a thing that doubts, understands, affirms, denies, is willing, is unwilling, and also imagines and has sensory perceptions". (Descartes 1984, p.19) Although the body is easily to destroyed, the heart will not naturally die. When idealism in scholastic philosophy came to a dead end, Descartes put forward his dualism. When his dualism came to a dead end again, he put forward the theory of physical and mental empathy. In Descartes, rational thinking is not only the important guarantee of human subject status, but also the basis of all cognition. Therefore, if we want to clarify the subject problem, we must go back to Descartes, especially the problem of the intelligent agent and their theory of body and mind.

3. Rethinking on Cartesian Subject

Despite huge advantages posed by AI, it is neither desirable nor possible to turn off its power supply and give it up totally, nor to change course and give up human social intelligence. Conversely, it is time to study what opportunities and challenges the intelligent agent will accord the human subject. The rethinking of Descartes' subject is a born during a period of philosophical thinking.

The materialized dead author grandly invited by structuralism and postmodernism does not ring the death knell of metaphysical subjects, nor is it the goal of intelligent agent to break through and challenge. The final goal of intelligent agent is to replace the metaphysical thinking subject in the Cartesian subject.

Intelligent agents such as Xiaobing, Deep Blue and Alphago have surpassed humans in different fields. When scientists develop and design super intelligent agents by integrating intelligent technology and biotechnology in the future, it is possible that the intelligent agent that can quickly update, iterate and upgrade independently.

Rethinking the Cartesian subject is beneficial in understanding the challenge posed by AI technology advancement towards unknown heights and obtain more unknown results, and Descartes's thinking subject itself has a natural "sense of closeness" to science. John Cottingham commented on Descartes thinking as follows: "Descartes is rightly regarded as one of the inaugurators of the modern age, and there is no doubt that his thought profoundly altered the course of Western philosophy. In no area has this influence been more pervasive than in metaphysics and the philosophy of mind. But Descartes himself would perhaps have been surprised to learn that these aspects of his work were to be singled out by subsequent generations for special attention. For his own conception of philosophy, and of the philosophical enterprise he was engaged on, was enormously wide ranging; so far from being confined to 'philosophy' in the modern academic sense of that term, it had to do principally with what we should now call 'science'. " (Parkinson 2003, p.187) Stephen Gaukroger believed that Descartes contributed to every field of science at that time: "He was one of the founders of algebra, he discovered fundamental laws in geometrical optics, his natural philosophy was the natural

philosophy in the seventeenth century before the appearance of Newton's *Principia* (Newton himself was a Cartesian before he developed his own natural philosophy) and his work in biology and physiology resulted, amongst other things, in the discovery of reflex action. Descartes's earliest interests were scientific, and he seems to have thought his scientific work of greater importance than his metaphysical writings throughout his career." (Parkinson 2003, p.156)

Descartes is the founder of modern philosophy, the first philosopher to think about the relationship between man and machine: a relationship that makes the problem of Cartesian subject is indispensable in the field of AI research. "From the description of inanimate bodies and plants I went on to describe animals, and in particular men"，(Descartes 1985, p.134) Descartes believed that "this will not seem at all strange to those who know how many kinds of automatons, or moving machines, the skill of man can construct with the use of very few parts" (Descartes 1985, p.139). In addition, he also set a standard for whether a robot can be a natural person. For any machine that can be invented by man, "we should still have two very certain means of recognizing that they were not real men. The first is that they could never use words, or put together other signs, as we do to declare our thoughts to others. [......]Secondly, even though such machines might do some things as well as we do them, or perhaps even better, they would inevitably fail in others, which would reveal that they were acting not through understanding but only from the disposition of their organs. " (Descartes 1985, p.139-140) The machine can compose poetry and play chess just as I walk and eat which is far from using subject consciousness. Only when the machine realizes that it is the "I" who is writing poetry and playing chess can it be said that it has challenged the human subject.

Descartes kept asking what "I" was, after repeated demonstration, the answer was "I am a thinking thing". In his reply to the second set of objections collected by Father Mersenne, Descartes said, "whatever can think is a mind, or is called a mind; but since mind and body are in reality distinct, no body is a mind; therefore no body can think." (Descartes 1984, p.95) But Descartes never concluded that all perceptual concepts were false. He did pay close attention to experience, "I also noticed, regarding observations, that the further we advance in our knowledge, the more necessary they become". (Descartes 1985, p.143) In order to affirm the universality of rational thinking, he abandoned the one sidedness and suspicion of perceptual ideas.

Gassendi joked that what Descartes said "a thing that thinks" is "a thing which has sensory perceptions" (Descartes 1984, p.187). Gassendi then refuted Descartes's thinking subject by comparing animal rationality to human rationality. Descartes pointed out that the logical loophole of Gassendi's reply. "When you say that I could have made the same inference from any one of my other actions, with the sole exception of thought (in using the word 'certain' I am referring to metaphysical certainty, which is the sole issue at this point). I may not, for example, make the inference 'I walking, therefore I exist', except in so far as the awareness of walking is a thought."(Descartes 1984, p.244) In using the word "certain" Decartes was referring to metaphysical certainty, which was the sole issue at that point. "Things in thinking" and "things in feeling" seem to be a process from sensibility to rationality, but this is not the whole meaning of Descartes's subject.

Intelligent robots have not yet arrived in the Cartesian era, an era where intelligent machines are more common. In Descartes's view, machines have no "heart". According to Descartes's subject, although an intelligent agent does not need a body, its heart is important. Descartes believed that there was indeed a difference between mind and body. Descartes's theory of body and mind was often mistaken for a mechanistic theory against the teleology of Aristotelian scholasticism. As a matter of fact, "Mechanism arose in the first instance not so much as a reaction to scholasticism but as a reaction to a philosophy which was itself largely a reaction to scholasticism, namely Renaissance naturalism". (Parkinson 2003, p.162) At the same time, when the dualism of body and mind came to a dead end, Descartes put forward the theory of body and mind empathy. The conclusion of "I walking, therefore I exist" is incorrect, because the subject of "I think therefore I am" is a spiritual subject with self-consciousness, and the perceptual activity of "I walking" cannot prove that "I am". Only when "I" think or realize that "I walking" can I prove that "I am", which is the core meaning of Descartes' subject.

Through rethinking the subject of Descartes, Descartes asks "what I am" in the pursuit of truth rather than in the action of life, and "I" is a more a rational, conscious and spiritual subject in its' thinking process rather than a sentimental subject in life and action. The "I" is also what the intelligent agent needs to strive to surpass the human subject in the future. Descartes began with "I", but not just "I". He strongly urged us to read the great book of the world and study things other than "I". Although Descartes' theories of body and mind and that of the intelligent agent have "family resemblance", the evolution of the intelligent agent is advancing at an exponential speed that surpasses the biological evolution speed of the human subject. In addition, the intelligent agent's ultimate goal is unpredictable. If the intelligent agent can surpass and even enslave human beings in thinking ability, the future is full of uncertainty and risk.

4. Conclusion: Future Subject

The superiority and dignity of the human subject's status are not only deeply rooted in Western thought, but also taken

for granted in China. The pre-Qin scholars said in *Shangshu Taishi*, "Heaven and earth is the parent of all creatures; and of all creatures man is the most highly endowed". Dai Sheng of the Western Han Dynasty said in *The book of Rites*, "man is the heart of heaven and earth and the end of the five elements". Zhou Dunyi of the Northern Song Dynasty said in *Taichi Tushuo*, "the two Qi sympathize with each other, which turns into all things. All things grow and change infinitely, but man get their beauty and spirit". However, in view of the growing intelligent agent, the essential boundaries of those exclusive to human beings such as rationality, thinking and consciousness are now blurry. The relatively emerging new concepts such as post humanism and super humanism are now challenging the existing theoretical framework.

Xiaobing cannot only publish poetry anthologies such as *the Sun Lost the Window*, but also serves as the editor of *the China Daily Economic News*. Within one month of Xiaobing's engagement and *the China Daily Economic News* officially signed the Sino English bilingual AI financial information cooperation agreement subsequently in June 2020, Xiaobing pushed more than 10000 financial information articles for 70 million users of *the China Daily Economic News*. Such efficiency cannot only embarrass any senior editor in China, but also usher in a storm of unemployment.

Whereas some researchers are entangled in the discussion of methodology and epistemology such as computationalism, functionalism, symbolism, behaviorism and connectionism, others are beginning to consciously think about coping strategies to address the uncertainty caused by the development of an intelligent agent. However, if the realistically thinking of the body-mind theory of the intelligent agent and the philosophically thinking of the fate of human subject are only limited to the research of the latest cutting-edge science and technology, then contemporary philosophers have to defend the rationality of their existence in the circle of scientists as they are likely to be led by the nose and to become a tool or footnote for scientists.

What is the subject? What is the future subject? Rethinking "I think therefore I am" can still reap new achievements from old ideological achievements. By affirming Descartes' thought and theory, we are affirming that the subject has taken the first step in a long march, a very important step which it needs continuous new development.

References

Descartes, R. (1984). *The Philosophical Writings of Descartes .—Vol. II*. Trans. John Cottingham, Robert Stoothoff and Dugald Murdoch. Cambridge University Press. https://doi.org/10.1017/CBO9780511818998

Descartes, R. (1985). *The Philosophical Writings of Descartes.—Vol. I*. Trans. John Cottingham, Robert Stoothoff and Dugald Murdoch. Cambridge University Press. https://doi.org/10.1017/CBO9780511818998

Deshun, L. (2018). Artificial Intelligence: a Warning to Mankind——Starting with the "Fourth Law of Robotics". *Southeast Academic Research*, *5*, 67-74.

Foucault, M. (2017). *Subjectivity and Truth: Lectures at the College de France 1980-1981*. Trans. Graham Burchell. Macmillan Publishers Ltd.

Guangyun, C. (2019). From A Man-Machine Relationship to Inter Human Relations: Definition and Strategy of Artificial Intelligence. *Journal of Dialectics of Nature*, *1,* 9-14.

Levi-Strauss, C. (1961). *Tristes Tropques*. Trans. John Russell. Hutchinson &Co. (Publishers) Ltd.

Parkinson, G. (2003). *Routledge History of Philosophy.—Vol. 4: The Renaissance and 17th Century Rationalism*. Routledge. https://doi.org/10.4324/9780203029145