A Probe into the Intentionality of Philosophy of Mind Based on Artificial Intelligence Theory Modeling

Jiayi Li[1*]

[1]*Department of Philosophy, School of Humanities and Law, South China Agriculture University, Guangzhou, 510642, China*
*Corresponding Author.

Abstract

For artificial intelligence to realize its desire to simulate and even surpass human intelligence, it will undoubtedly have to model the phenomena of the human mind, and indeed, there have been numerous attempts at modeling. However, it is worrisome and thought-provokingthat much of the modeling is based not on the latest results of the philosophy of mind and cognitive science that study the human mind, but on folk psychology. Folk psychology is essentially an erroneous topography, topography, structural theory, and dynamics of mental phenomena, and uncritical use of such resources will lead theoretical constructs and engineering practices of artificial intelligence astray. A fundamental study of philosophy is undoubtedly needed to change the status quo and achieve a breakthrough. To model intentionality, we are required to realize that it is a systemic project. In this respect, the discussion and results of the philosophy of mind on consciousness and intentionality themselves have no direct engineering significance, but since they involve the basic theoretical problems of modeling, they can at least be regarded as a component or necessary condition of this systematic project..

Keywords: Artificial intelligence, Philosophy of mind, Intentionality

I. Introduction

In order to realize the desire of artificial intelligence (hereafter referred to as AI) to simulate or even surpass human intelligence, it is undoubtedly necessary to model the phenomena of human mind, especially the intentional phenomena which best reflect the essential characteristics of human mind. In response to this requirement, some tentative models have been created. At least many began to dissect the theory of intentionality itself, thus making theoretical preparations for further modeling. Such a transition has occurred under two driving forces. On the one hand, it is based on methodological considerations. Generally speaking, in order to study complex objects, it is necessary to construct models for them by means of modeling methods in order to reveal their main constituents, structures and mechanisms, and to grasp their substance and main characteristics. On the other hand, it is the "return" from the "self-subjective research". The study of AI as a discipline began with the human being as an intelligent self-subject. But later, in its specific course, for one reason or another, it forgot about the real prototype it was trying to simulate, and wandered into a world of imaginary intelligence with more imaginary overtones.

II. BDI model

Bluttman is an accomplished American philosopher concerned with AI and cognitive science, and his related theories have been influential in AI research. Bluttman presents a fundamental view of intentionality and self-subjectivity and describes the model he constructs for the human mind as the basis of AI simulations, namely, the belief-desire-intention model. By combining the first letters of these three words together, the model can be abbreviated as BDI. In essence, the BDI model is both his conclusion of dissecting the human mind, or his philosophy about the mind, and his model for AI simulation, beyond the human mind, with engineering guidance.

In the BDI agent, the basic building blocks are data structures such as beliefs, desires and intentions, and functions that represent thinking (determining what intentions should be, deciding what to do) and means-end reasoning. Among them, intention plays the biggest role. Because once the intention is formed, the behavior is determined and the rest is a matter of deductive reasoning. And what intention there is is determined by the agent's current beliefs and desires, or rather, by the relationship between beliefs, desires, and intentions.

Compositionally, the agent's state is a triad of beliefs, desires, and intentions (B, D, I). In terms of process, the agent completes its practical reasoning through seven stages, as shown in Figure 1 below:
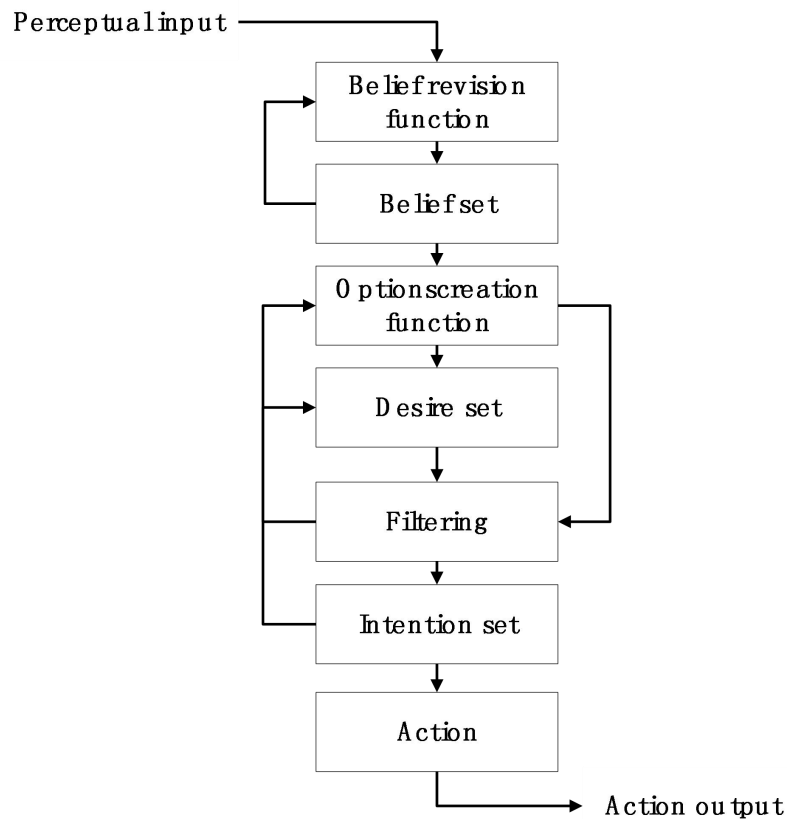

Fig 1: Practical reasoning diagram of BDI agent

As can be seen from Figure 1, in the first step, the agent makes a decision to behave. This decision is generally related to the information about the environment provided by the senses, and after receiving the information, many beliefs are generated. In the second step, the agent, thanks to the belief revision function, is able to form a new set of beliefs based on perceptual input and existing beliefs. In the third step, the agent's options creation function forms the corresponding desires based on the existing beliefs and then makes possible choices. On this basis, the agent uses the means-ends reasoning process to determine its intention and the process and method to achieve it. To do so, a further choice must be made, which is more specific than the intention. It is a recursive process of option generation, through which more specific intentions are formed, until the intentions corresponding to the actions that can be taken are obtained. In the fourth step, through the selection mechanism, the agent selects a number of possible courses of action. In the fifth step, with the help of the filter function, i.e., the deliberative function of the agent, new intentions are determined based on current beliefs, desires and intentions, in order to choose among multiple possible actions. In the sixth step, the current set of the agent's intentions are analyzed. They are the focus of the agent's attention and are the goals it promises to achieve. In the seventh step, the action selection function is used to determine the actions to be performed based on the intentions.

It is a representative approach in the current modeling of intelligent agents to understand the human mind based on the relationship and dynamics between beliefs and desires, etc., and to construct corresponding models for AI. Because the agent has such a fundamental nature, it has been the object of reflection and criticism by philosophers of mind and cognitive science. From the practical research of AI on the agent, many people realize that building a model about the agent is to build a BDI model about intentional states such as beliefs, which in turn is a prerequisite for a real intelligent agent to have a veritable agent.

III. Philosophy of mind based on AI modeling

Bluttman's BDI model is one of the most discussed theories in the field today, and enjoys an important position in the theoretical construction and engineering practice of AI, and has become the theoretical basis for many engineering practices. However, it should be noted that this model has at least two major problems.

First, its theoretical basis is common sense or folk psychology. This psychology is essentially a false topography, topography, structural theory and dynamics of mental phenomena, and its uncritical use of such resources will lead AI theoretical constructs and engineering practices astray. As we know, folk psychology is also called common sense psychology and intentional psychology. It is the starting point of scientific psychology and the object of critical reflection. Since this psychological knowledge is held by everyone, it is called common sense psychology. It is called intentional psychology because it mainly appeals to intentional states such as beliefs, desires, goals, and intentions to explain and predict actions. Beliefs and other states are called intentional states because their fundamental characteristic is intentionality, which means that they are about and intended for external states. They have this intentional and autonomous character because they have an agent behind them. Because the agent is so fundamental, it has been the object of reflection and criticism by philosophers of mind and cognitive science. From the actual study of the agent by AI, many people realize that building models of the agent means building models of intentional states such as beliefs, which is a prerequisite for truly making the intelligent agent an agent worthy of the name. It is a great idea, but the problem is that if the modeling is based on an unreliable or inherently wrong foundation, no amount of good thinking or effort will make a difference.

Second, Bluttman believed that his BDI model was inspired by the theory of intentional states of Donald Davidson, a famous philosopher. The author argues that the BDI model is problematic because it contains a misinterpretation of Donald Davidson's theory, which in turn is a source of his misuse of common sense psychology. Crucially, Bluttman fails to understand the essence of Donald Davidson's philosophy of mind as a dialectical negation of folk psychology and a thoroughgoing physicalism, and instead interprets it as a defense of folk psychology. Another misunderstanding of Bluttman is that he does not really understand the significance of Donald Davidson's emphasis on causal explanations of behavior by appealing to beliefs and intentions. According to Donald Davidson, the so-called causal explanatory role of beliefs is in fact the causal role of physical events described in physical terms. In short, Donald Davidson's thought as Bluttman understands it is at least as far removed from the actual thought as there is.

To model intentionality, AI must undoubtedly have a corresponding philosophical theoretical foundation. In other words, it must be based on the latest research findings in philosophy, psychology, and cognitive science on the human mind, especially intentionality. In short, it must be based on the developing philosophy of mind's knowledge of the mind. In my opinion, if we want to make use of the results of philosophy of mind, on the one hand, we need to have an accurate understanding of the results. On the other hand, it is necessary to recognize the new direction of "disenchantment" or "demystification" of philosophy of mind, and to critically reflect on, deconstruct, and cleanse common sense psychology and traditional philosophy of mind. Although the direct motive of such a philosophy of mind is to develop a philosophy of mind, it has an indirect and undervalued significance for AI research, because it is actually clearing the ground for AI research to be built on a reliable philosophical foundation. Therefore, in order to make use of the results of philosophy of mind, we should pay attention to such

philosophy of mind with the nature of disenchantment.

IV. Reflections on the philosophy of mind based on AI modeling

In order to make the intelligent modeling move in the right direction and let the artificial system show the real intelligent characteristics so that it can become an agent close to or even beyond the human autonomous system, the key is to understand the human intelligence and its structure and operation mechanism truthfully, to truly realize "the return of the agent", and to recognize the characteristics and signs of human intentionality. Moreover, we should realize that there is a huge gap between the existing syntactic machine and the human "semantic machine", and actively search for the inner reasons and solutions. Through the new anatomy of human intentionality, we will find that it is the highest form of intentionality among living beings. Its uniqueness is manifested in its conscious relevance, transcendence, and semantic content. More than that, it has three other essential features: (1) human intentionality is autonomous, (2) human being has meta-intentionality, i.e. the ability to direct intention to intention itself, and (3) human being has the ability to show itself with a clear representation of consciousness. As a result of cush consciousness, symbols and semantics can be bound together. It is clear that in order for artificial systems to become truly intelligent, it is imperative to investigate how to model the main described capabilities and characteristics "one" and how to make them realized or executed on machines. It should be acknowledged that many artificial systems have achieved remarkable results in simulating some features of human intentionality, such as relatedness and initiative. However, if we compare the composition and characteristics of human intelligence with those that are realistically recognized, we will find that there is still a fundamental gap between the two.

In order to simulate intentional and semantic intelligence, it is necessary to have a scientific modeling of intentionality and semantics, which is a necessary step for scientific research and the translation of theory into application. To model this, we need to have a correct conception of intentionality, which is undoubtedly a hard nut to crack today. Therefore, we can say that this is a fork in the road for philosophy and AI research. And if we are not careful and take the wrong path, we will fall into a dead end. It should also be noted that it is here that the dualism between folk psychology and traditional philosophy has laid a trap that is too wide to be seen and too deep to be understood. Sometimes, even if we are scientifically minded, we can't help but be fooled. In fact, some of the models of intention and belief in AI research, especially the agent research, already show signs of being fooled. Based on the results of philosophy of mind over the past decades, the way not to be fooled here is to do disenchantment or demysterization, to deconstruct folk psychology, and to accomplish an ontological change in the concept of mind.

In exploring the philosophy underlying AI modeling and dissecting the mind, it is important to keep an eye on the latest developments in cognitive science and the philosophy of mind and their results, and not to concentrate on outdated theories, especially those of folk psychology. Here is the fundamental reason why the BDI model is wrong in direction. It uses outdated psychological theories. For example, its understanding of the mind is still at the level of the event-causal model of the mind, which is being surpassed by the latest cognitive achievements, namely the agent-causal model and especially the natural agent-causal model. We know that human intelligence is often expressed in action, and that action and the choices and forces that determine it must have a root, not outside the agent, but within it. In other words, what is at the root of the agent is either the causal role of events or the causal role of the agent. The agent-causal model, which emphasizes the causal role of things like desires and beliefs, is actually the outdated folk psychology model on which the BDI model is based. It believes that the cause of action is an event, and that this event causes the result depending on internal and external conditions and resources. The agent-causal model is a new model of the structure and dynamics of human intelligence that has emerged in philosophy of mind and AI research along with "the agent return".

In understanding the inner mechanisms and characteristics of the human mind, we should also pay attention to the exploration and results of naturalism. And it is another direction in the philosophy of mind today that is fundamentally different from the phenomenological tradition in its orientation. It deserves our attention in its

construction because it has the characteristic of maintaining the authority and dignity of natural science without easily denying the legitimacy of psychology, but trying to reconcile natural science and psychology. The strategy is to describe the concept of psychology in the terms of the natural sciences, but in fact to reduce the latter to the former, while it is also known as the "naturalization movement of the mind". We should also see that there are different scientific theories that people use to explain psychological phenomena. For example, some psychological phenomena are based on behaviorism, while others are based on physics or brain science, and so on. Thus, there are different directions or forms of naturalization in the West.

V. Conclusion

Since the naturalization of mental phenomena is a continuation of the scientific enterprise and has taken a crucial step toward application, it is a useful pavement for engineering modeling of the human mind and deserves attention. There is no doubt that it is a field that is being explored with a wide variety of results, so on the one hand we should be careful in using it, and on the other hand it is worthwhile for researchers in AI to advance it with their own strengths. Here, we should pay special attention to the research and results of the nascent neurophilosophy. Neurophilosophy is the study of the components of the mind, especially free will, such as directionality, intention, and volition, from both neuroscientific and philosophical perspectives, in order to build a neurophilosophical theory of them. For example, when studying intentionality, we can focus on how neural states manifest intentionality. Neurophilosophy has two parts. The first is general neurophilosophy, whose task is to summarize and distill the information in neuroscience that can help to solve the mind-body problems, and then to construct a theory of mind-body based on neuroscience. The second is the special neurophilosophy, whose task is to summarize and generalize the results of neuroscience and to provide answers to specific questions of the philosophy of mind, such as the question of intentionality.

References

[1] Wang Meiyan. Theoretical Purpose of Intentional Issues under Naturalism Proposition - Analysis Based on Contemporary Philosophy of Mind [J]. Philosophical Research of Science and Technology, 2011, 28(1): 52-56.

[2] Xinmin. Philosophical Psychology of Artificial Intelligence Modeling [J]. Abstract of Arts and Sciences in Higher Institutions, 2019, 036 (006): P.30-31.

[3] Wu Tianpeng. Self-Subjectivity in Philosophy of Mind [D]. Central China University of Science and Technology.

[4] New People. New Trends in the Development of Dualism and Naturalism in the Philosophy of Mind - Clued by Chamers Naturalism Dualism [J]. Academic Monthly, 2011 (09): 43-50.

[5] Chen Junjie. The Cartesian Theory of Psychosomatic Relations from the Perspective of Philosophy of Mind [D]. Shandong Normal University, 2009.

[6] Yan Kunru. Philosophical and ideological exploration of artificial intelligence [J]. Theoretical exploration, 2020 (2).

[7] Chen Shifeng. Psychosomatic Problems and Artificial Intelligence, Cognitive Science - Understanding and Thinking of Sell's Philosophy of Mind [D]. Nanjing University, 2005.

[8] Zhao Zelin. Definition of Intelligence - Exploration of Psychology and Philosophy of Mind [J]. Research on Natural Dialectics, 2010 (06): 21-26.