**A multi-dimensional look at artificial cognition**

Source: China Social Sciences Network - China Social Sciences Daily

Author: Wei Yidong

Editor: Cui Jin (newspaper) Qi Zeyao (Internet)

2023-12-25

<https://www.cssn.cn/skgz/bwyc/202312/t20231225_5721365.shtml>

　The rise and development of artificial intelligence has gradually made artificial cognition a new cognitive model and brought major challenges to human cognition. The emergence of large models for natural language processing has increased the intensity of the challenge. Many people are worried about whether machine learning systems can surpass or even have surpassed human cognitive abilities, and whether they will have a major impact on human life and survival. However, such concerns cannot stop the development of artificial intelligence. This is a phenomenon of doubts about new things that are bound to appear in the process of scientific and technological development, but ethical reflection and regulation are necessary. From a philosophical perspective, the author proposes three principles, five methods, and six types of ethical norms for the research and development of artificial cognition.

**Three principles**

　　The first is the principle of adaptation-representation. This category framework is based on cognitive philosophy and aims to unify and coordinate the two aspects of cognitive science and artificial intelligence. The cognitive system is an adaptive representation system and a product of long-term evolution in nature. In cognitive science, evolution means adapting to the environment, and cognition means understanding and representing meaning. These two meanings also mean that the cognitive system is a predictive processing system. In artificial intelligence, predictive processing is self-optimization and foreknowledge. Self-optimization is an artificial cognitive method that transfers human cognitive ability to technical systems, with the aim of enabling technical realization of the adaptation of external target states through related system behaviors. Therefore, both predictive processing and self-optimization are adaptive representation behaviors, which aim to make endogenous predictions and capacity adjustments of targets through the description of technical systems in response to environmental changes, so that the behavior and goals of the system can adapt to each other autonomously.

　　The second is the situation-awareness principle. Situation-awareness refers to the current perception of elements in the environment by the subject (human or machine) within a period of time and space, which requires understanding of their meaning and prediction of their future state. This is an online real-time representation of dynamic systems, which is essentially an adaptive representation method that aims to set up a human-machine-environment interaction situation for dynamic cognitive systems. This principle shows that the fitness of a cognitive system is the fit between the system and a specific situation, and its representation is generated in a specific context. Therefore, adaptability changes with the specific situation, and the representation depends on the specific context to form. If there is only a situation without contextual awareness, the system will lack a meaningful description of the generated content, making it difficult to understand. Therefore, this requires the intervention of context-awareness.

　　The third is the context-awareness principle. Context-awareness refers to the ability of the subject to respond differently according to the current environment (including external and internal environments). For humans, context-awareness means that cognition is the ability to perceive and think based on context; for artificial intelligence entities, context-awareness refers to the ability of computers to perceive and obtain environmental information. Human-computer interaction and brain-computer interface are a mixture of two types of cognition, forming an interactive network. For example, the Internet of Things is a lot of objects connected together through context-awareness. Therefore, if a system uses context to provide relevant information or services to users, then the system is context-aware. The difference between situation-awareness and context-awareness is that the former emphasizes the present and immediacy, while the latter emphasizes relevance and semantics.

　　In a nutshell, the Adaptation-Representation Principle, the Situation-Awareness Principle, and the Context-Awareness Principle all contain the adaptability of the interaction between the cognitive system and the environment and the representation of semantic expression. They are the specific principles of adaptive representation methodology. Situation and context are the micro-habitat and macro-habitat that the subject needs to adapt to. Only when the artificial intelligence entity is in such a situation or context can consciousness, mind, and intelligence emerge. Therefore, psychological attributes such as consciousness, mind, and intelligence are all overall emergent attributes of the interaction between many factors within the cognitive system and the external environment, which are not possessed by a single factor.

**Five methods**

　　The first is the new artificialism approach. Artificialism is a computational representation method that formally models the properties and parts of artificial life systems. It constructs parts and details of life systems through computational simulation, and clarifies the relationship between computation and intelligence. Its essence is computational simulation and symbolic reasoning. If artificialism is combined with the biosemiotics method (assuming that the "mind" itself does not exist, but is used as a location identifier of a concept), a new artificialism is formed. It is different from the analytical philosophy of mind that does not make a reductive interpretation, and it is also different from the philosophy of neuroscience that ignores brain evolution; it transcends traditional naturalism and physicalism, and also transcends the machine functionalism of philosophy of mind. This new approach not only challenges biological naturalism, but also has the potential to replace it.

　　The second is the new precognitive approach. Precognition is a kind of ability to go from reactivity to proactivity. As a method, it integrates the representation of artificial intelligence into the situation, thereby transforming traditional artificial intelligence to form precognition. If this precognition is combined with perception-action (the method of collecting data or information through perception and then acting after processing, also known as actionism), a new precognition that transcends traditional behaviorism and dynamism is formed. Because the purpose of precognition is to act rather than just predict, prediction is for better action. Obviously, this new method goes beyond classical representationalism and connectionism, and will make the new generation of artificial intelligence smarter.

　　The third is the new contextual emergence method. Contextual emergence demonstrates a non-reductive but well-defined relationship between different levels of description of physical and other systems, aiming to illustrate the irreducibility of properties between systems at different levels, such as consciousness cannot be reduced to neural activity in the brain. In the field of artificial intelligence, contextual emergence has produced a formally sound and empirically appropriate procedure to translate between levels of description in an overall consistent manner, and to discuss the contextual emergence of mental states from the description of the neural level. If contextual emergence is combined with generativeism (an ecological approach), a new contextual emergence method is formed. It may unify the cybernetic method and artificial intelligence method of traditional cognitive research, which will not only help the integration of artificial cognition and human cognition, but also be a useful supplement to the generally accepted situational approach.

　　Fourth, new simulation methods. The simulation method is to use a designed object or process to imitate or simulate a real object or process, with the purpose of showing the key characteristics of the selected physical system or abstract system. Different from the simulation theory in the philosophy of mind about the ability to predict the behavior of others, the simulation method requires adaptability between the simulated object and the simulated object. At the algorithmic level, the simulation consists of a formalized program for understanding the psychological state, which produces accurate outputs for given specific inputs. If the simulation is algorithmized, a new simulation method that goes beyond pure morphological simulation can be formed. The former focuses on psychological simulation and modeling, while the latter focuses on computer simulation and reasoning.

　　The fifth is the cultural reconstructionist approach. Reconstructionism is a bottom-up empirical approach to studying the mind, brain, and behavior. It complements the top-down process of reductionism and emphasizes the mechanisms of mind and brain and their emergence. As a philosophical position, reconstructionism assumes that psychological phenomena are the result of the interaction of a series of psychological and neural attributes, but ignores the shaping of the mind by culture. If reconstructionism is combined with human-specific culture, a kind of cultural reconstructionism is produced. This kind of cultural reconstructionism reshapes artificial intelligence in the way of cultural evolution, transcending traditional cognitivism and reductionism, and enabling artificial intelligence to develop in coordination with human cultural evolution.

　　In short, these five methods are the embodiment of adaptive representation methodology. The essence of artificialism is formal modeling using Bayesian methods, constantly trying to achieve goals through empirical revelation methods, and combining it with biosemiotic methods to show adaptive representation characteristics. Precognitive analysis plus actionism means "perception + action = adaptability + behavioral expression". The mechanism of context emergence is adaptive representation, plus generativity means "change + result presentation = adaptability + performance" (another form of representation). Simulation is a goal-guided fitting representation, plus the algorithmization of the goal, which means "adaptability + computational representation". Cultural reconstructionism means the combination of natural evolution and cultural evolution, showing dual adaptability and representation, further strengthening the embodiment and semantics of artificial intelligence.

**Six types of ethical standards**

　　The first is the norms of artificial moral subjects. If artificial intelligence becomes a human-like moral subject, it should be subject to moral constraints. For example, once artificial intelligence has subjectivity, it will have moral consciousness and will no longer only show automaticity and reactivity, but may also have initiative. At this point, it should be subject to moral constraints.

　　The second is safety regulations. Unreliable and unsafe artificial intelligence is undoubtedly unethical. Safety is the primary condition for the use of all artificial devices. This may be solved through AI-Boxing, that is, to impose special restrictions on artificial intelligence. Any AI research that is not beneficial or even harmful to humans must be banned.

　　The third is cognitive norms. If safety issues are mainly technical issues, ethical issues are often manifested as cognitive issues. Technical issues can be solved by improving technology, such as continuously improving and improving power robots. Moral issues are usually cognitive issues. The ethical norms of advanced intelligent behavior cannot be achieved only through technology, but also need to be educated through education and learning, or through legislation to enforce behavioral norms.

　　Fourth, cognitive enhancement standards. This is mainly aimed at brain-computer interfaces and general artificial intelligence. The current brain-computer interface can be used to treat brain damage and brain diseases, which is beneficial to humans, but it is absolutely not feasible to use it to enhance the cognitive ability of normal people. Because this is not only likely to harm people (even if the technology is fully mature, the risk still exists), but it is also ethically unacceptable. The intervention of artificial intelligence will undoubtedly improve the level of brain-computer interfaces, but it cannot improve moral cognitive ability. Future general artificial intelligence cannot improve the level of ethical cognition, although it can participate in evaluation.

　　Fifth, controllable norms. The development of artificial intelligence must be controllable. If a super artificial intelligence outperforms humans in most cognitive tasks, it should be controlled, which requires developers to design "beneficial artificial intelligence" rather than "evil artificial intelligence." Therefore, if artificial intelligence research is properly controlled and relevant ethical and legal treaties are formulated and implemented, enhanced intelligence is not entirely a threat.

　　Sixth, sociological norms. If intelligent robots become the norm in society, the socialization of artificial intelligence will occur. Socialization is the understanding of artificial intelligence as a social and cultural phenomenon and a "non-human intelligent actor". Artificial intelligence is increasingly becoming a part of society, and human-computer interaction has become a reality, as reflected in the evolving smart home systems, self-driving vehicles, chatbots, smart public displays, nursing robots and other fields. This requires sociological and social management research on artificial intelligence, thus forming the sociology of artificial intelligence and the social management of robots.

　　(This article is a phased result of the National Social Science Fund Major Project "Philosophical Research on the Challenge of Artificial Cognition to Natural Cognition" (21&ZD061))

　　(The author is a professor at the School of Philosophy, Shanxi University)