

Original Article

Philosophical Foundations of Artificial General Intelligence: A Comparative Inquiry

DR. J. CHARLES AROCKIASAMY

Department of English, St. Joseph's College (Autonomous), Tiruchirappalli, Tamil Nadu, India.

ABSTRACT: *AGI research is strongly connected to philosophy, since it helps us think about intelligence, consciousness and what the mind is. This study goes into the philosophical support systems that guide AGI research by drawing on old and new perspectives. One of the central debates is whether machines can truly have minds and consciousness like humans ("strong AI") or if they can only imitate human behavior in the form of "weak AI." For humans to have artificial general intelligence, it is necessary to rely on basic concepts such as embodiment, symbol grounding, causality and memory, like we see in biological cognition. This part of the analysis looks at the difference between Western traditions focused on logic and computing and newer ways that combine understanding from neuroscience, evolutionary theory and Eastern philosophies. Even with technological advancements, AGI is still only theorized, as there are still debates over what makes intelligence and consciousness, whether machines could have minds and the ethical effects of producing self-thinking machines. When we compare different philosophical standpoints, we can highlight both the benefits and obstacles in advancing AGI, which stresses the importance of teams from various fields to work together as technology grows stronger.*

KEYWORDS: *Artificial General Intelligence (AGI), Philosophy of artificial intelligence, Strong AI vs. weak AI, Embodiment, Symbol grounding, Causality, Memory, Machine consciousness, Cognitive science, Interdisciplinary inquiry*

1. INTRODUCTION

1.1. CONCEPTUAL ORIGINS OF ARTIFICIAL GENERAL INTELLIGENCE

Philosophical ideas and computer advancements have gone hand in hand in creating Artificial General Intelligence (AGI). People who invented AI have always been interested in questions about intelligence, the mind and reasoning. [1-3] The handling of deductive reasoning by philosophers like Aristotle set the basis for considering human thinking as computational, which was later adopted in the computational theory of mind. According to thinkers such as Haugeland, this view suggests that our mental processes are largely about processing information, and it urges us to design machines capable of using information in the exact way our own minds can. At the beginning, scientists wanted to build machines that used reasoning, learned and adapted like humans, not just machines that looked intelligent.

1.2. DEFINING INTELLIGENCE: STRONG AI, WEAK AI, AND THE AGI DEBATE

The idea of separating "strong AI" and "weak AI" marks one of the main philosophical debates in AGI. John Searle explained this as the distinction between strong AI and weak AI, with strong AI holding that machines can understand things, unlike weak AI, which believes that machines can appear intelligent but not truly understand anything. Many questions about the mind and consciousness occur when it comes to strong AI, but in most AI fields, being smart is seen in behavior rather than in the way the machine thinks. Many AGI researchers take this position, working on systems that can handle any intellectual task humanly possible, even when the issue of consciousness is not settled.

1.3. THE PHILOSOPHICAL DREAM: MECHANIZING HUMAN REASONING

Building machines that act like the human mind has been a theme in the progress of AGI. Philosophers like Aristotle and later Leibniz and Hobbes helped to make logic formal and computer-processable, greatly boosting the progress of AI. Bionic believes that intelligence can be explained by formal systems and by manipulating symbols in them, leading to the physical symbol system hypothesis, which says that such systems can exhibit general intellect. AGI research today still follows this tradition by working towards creating artifacts that can act intelligently in many kinds of situations.

2. CONCEPTUAL OVERVIEW OF ARTIFICIAL GENERAL INTELLIGENCE

2.1. DEFINITION AND DISTINCTIONS FROM NARROW AI

AI researchers sometimes refer to Artificial General Intelligence (AGI) as "strong AI" or "human-level AI," meaning it can perform most tasks involving learning, understanding and applying knowledge, much like a human. [4-7] Narrow AI, which includes tasks like image recognition, language translation or chess playing, focuses on a narrow area. AGI wants to demonstrate several broader cognitive abilities. As a result, AGI would have the ability to think logically, solve complex

problems, grasp complicated ideas and move knowledge from one field to another, doing it just as well as or better than humans.

Voice assistants, recommendation systems and self-driving cars all rely on Narrow AI, which leads the way today in technology. Many of these systems are created to do one job well and usually do it better than a human can. Nevertheless, they can't work outside the limits set in their original training and cannot handle unfamiliar tasks or environments. An AI designed for medical images cannot change its task to that of a financial forecast or language analysis without extensive training first. AGI would differ from AI in the sense that it can analyze many situations, use what it learns often and use that knowledge in different contexts by itself. If AGI becomes capable of this, it would be able to handle issues that are new, grasp abstract ideas and display creative problem-solving skills, all skills currently limited to humans. Although thought-out AGI is still a goal to achieve, understanding the divide between narrow and general AI is basic and essential when following AI research.

2.2. KEY FEATURES OF AGI: FLEXIBILITY, AUTONOMY, ADAPTABILITY

Several things separate AGI from all other AI systems: its flexibility, autonomy and ability to learn from its experiences. Flexibility is the feature that allows AGI to handle many types of problems, moving between different domains and task types without any special training. Compared to narrow AI, AGI would be able to use what it learns in one area to help solve problems in other fields, for example, using biology knowledge to assist in economics or engineering.

AGI is also defined by its ability to act independently. Because of this feature, an AGI system may be able to choose its targets, decide what to do and grow through experience on its own. It would be able to learn by itself, find new problems and plan how to address them. Today's narrow AI cannot work independently like this and instead relies on being instructed and supervised for every new job it does.

AGI's ability to adapt, the third key trait, means it learns through different tasks, adjusts to new environments and gets better with time. AGI is different from static programming because it grows and learns new things whenever it encounters varied situations. Since real-life situations are often unpredictable, systems that respond flexibly are necessary since task-specific methods may not work well. Flexibility, autonomy and adaptability, used together, set the goal that AGI aspires to meet. They explain why many in the field believe AGI will allow AI to match or exceed human intelligence in every intellectual activity.

3. HISTORICAL AND PHILOSOPHICAL ROOTS OF INTELLIGENCE

3.1. CLASSICAL PHILOSOPHICAL VIEWS ON MIND AND INTELLIGENCE

3.1.1. RATIONALISM (E.G., DESCARTES)

According to René Descartes, rationalism holds that the main way we know anything is through reason and that certain ideas are naturally present in our minds. Descartes wanted to find a sure basis for what we know, and he explained that our basic certainty rests in understanding, "I think, therefore I am." According to Descartes, the mind isn't the body; it is a "thinking thing" that is able to doubt, understand, affirm or deny, will or refuse, imagine and perceive things. [8-11] He separated thoughts that come naturally, those we learn through our life and thoughts that are thought of on our own. The rationalist view argues that some concepts are known without and apart from senses, and mental thinking and logic alone enable people to grasp such concepts. Descartes' ideas on the separation of mind and body and on reasoning were important for developing the philosophy of mind and for debates about intelligence and artificial thinking.

3.1.2. EMPIRICISM (E.G., LOCKE, HUME)

Rationalism, empiricism, as supported by John Locke and David Hume among others, believes that all knowledge comes from our senses. According to Locke, the mind is an empty mental slate at birth, and experience shapes it. Empiricists state that concepts and knowledge do not come naturally but are picked up by interacting with our surroundings. Hume explained that ideas are often formed by habit and association in the mind and that human knowledge is limited. Empiricism holds that impressions from our senses enter the mind passively, and we understand and interpret them as we gain more experience and reflection. Many modern methods in AI are guided by this view and use data-driven learning to find patterns, as the empiricists believe intelligence develops from our experience in life.

3.2. DUALISM VS. PHYSICALISM

Debates about dualism and physicalism are key to understanding the philosophy of mind and how intelligence is conceptualized. According to dualism, as stated by Descartes, there are two substances: the mind, being non-physical and thoughtful and the body, which follows physical principles. Such a division makes people wonder if mental states can impact physical behavior and vice versa, which is known as the mind-body problem. Philosophers still talk about dualism when considering whether machines can ever really experience feelings and concepts.

Physicalism claims that all mental activities can ultimately be linked to brain-based physical functions. Many neuroscientists and cognitive scientists adopt this monist view by studying consciousness, thought and intelligence as caused by brain activity and information processing. Physicalists believe that the study of the brain and computer modeling is leading to a better

understanding of how material objects can process complex ideas. People are still debating how reliable AGI is, especially when it comes to determining if a computer made only from physical parts could ever truly copy human-like intelligence and consciousness.

3.3. FUNCTIONALISM AND COMPUTATIONAL THEORY OF MIND

In the 20th century, Functionalism appeared to address issues of dualism and behaviorism, arguing that mental states should be understood according to their functions in the mind, instead of what they are made of. Functionalists believe that the key thing is not mental content but the structure and interactions of perceptions, thoughts and actions. This theory is closely related to the computational theory of mind, which thinks of cognition as a computer-like information-processing system. In this model, the mind acts like software on the hardware of the brain or any system that reaches a certain level of complexity.

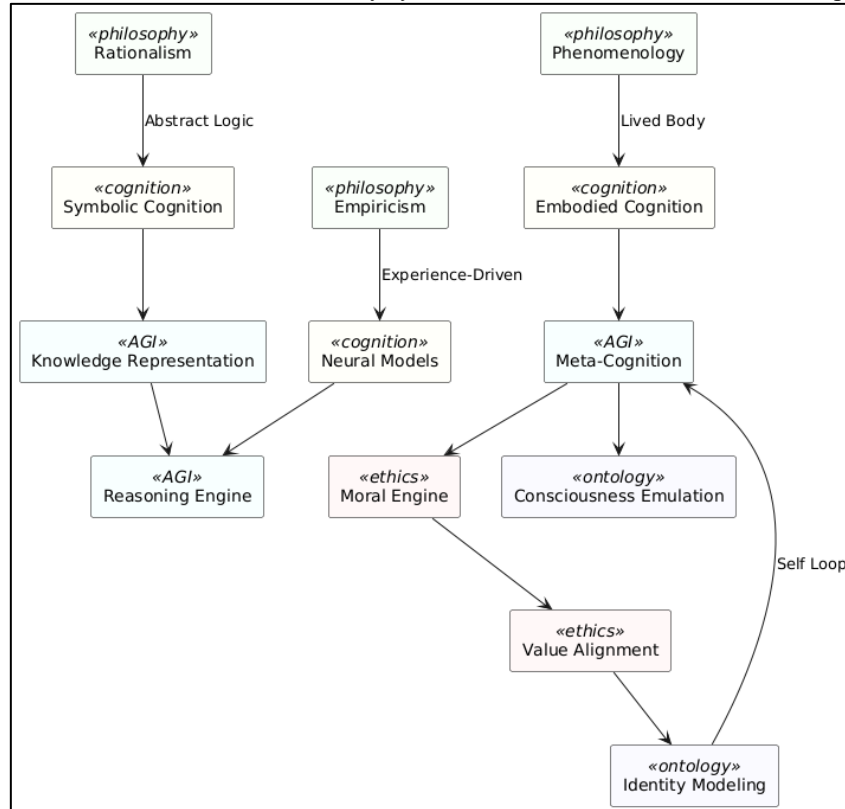


FIGURE 1 Philosophical influence on AGI architecture

The field of artificial intelligence was greatly influenced by the computational theory of mind, which suggested ways for symbolic manipulation, algorithms and memory to bring about intelligence. The main principle of Functionalism is that intelligence could potentially be programmed and realized in technological systems not tied to the brain. Many of the positive beliefs about AGI are based on this idea, believing that if computers are organized similarly to the human brain, they may acquire general intelligence.

The article explores how important philosophical models are in Artificial General Intelligence (AGI) development by going through different cognitive and ethical structures. Data from Rationalism, Empiricism and Phenomenology is used to identify key cognitive models, Symbolic Cognition, Neural Models and Embodied Cognition, which design the structure and operations of different parts of AGI. It helps show how abstract ideas in philosophy become part of real architecture and structures.

4. COMPARATIVE PHILOSOPHICAL FRAMEWORKS FOR AGI

Symbolic cognition, which symbolism and logical deduction allow, helps build knowledge and reasoning engines for AGIs. They are developed to deal with formal logic and symbolic regulations, which are similar to the traditional ideas about the mind found in classical philosophy. [12-16] Models and approaches that rely on experience and observation, including empiricism, play a role in neural and connectionist science. These frameworks are based on data and guide AGI's moral actions by helping it connect outside inputs with the proper response from what it has learned. Phenomenology, which emphasizes what it is like to experience and be conscious, plays a role in influencing both the embodied and higher cognitive aspects of AGI. For AGI systems to make thoughtful and considered judgments, these layers are designed to mirror self-

awareness, the ability to reflect and consciousness. Meta-cognition and consciousness interacting in the image show a process that copies how living creatures judge their own reflection and understanding.

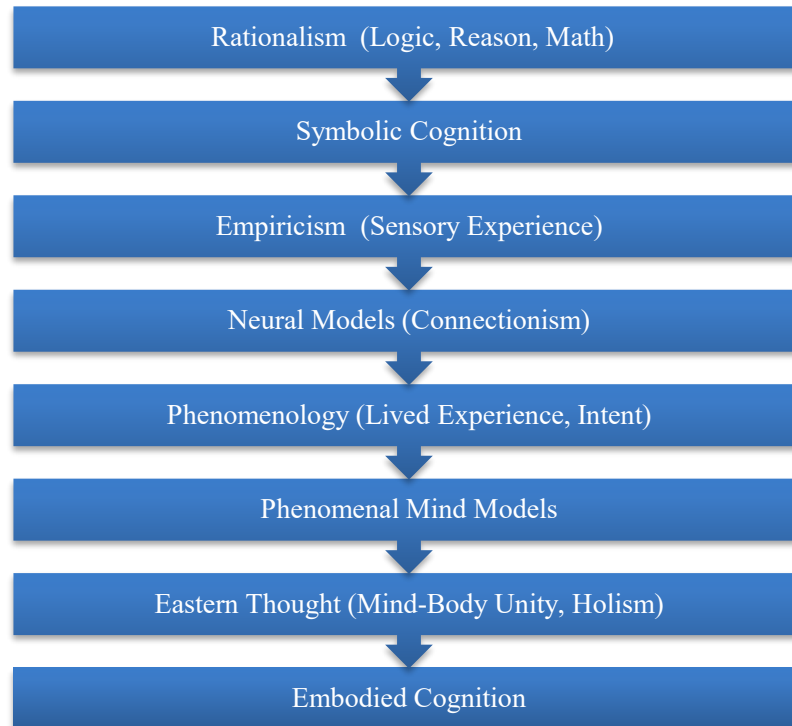


FIGURE 2 Philosophical schools to cognitive models

They made an intellectual and philosophical impact in the areas of ethics and ontology, particularly with value alignment and identity modeling. This is a main concern in AGI research because it aims for systems to both think and learn and display values recognized by society while understanding what their actions may lead to. Therefore, the diagram describes clearly how AGI is built, so it is a fundamental idea to include in your article.

4.1. RATIONALIST APPROACHES TO AGI

Rationalist approaches to AGI are based on the belief that intelligence depends on applying formal rules and logic. The idea that cognition can be modeled with symbols and rules of inference began with Aristotle and was fully shaped by Descartes and Leibniz. Rationalist approaches in AGI often turn into systems based on logic, where knowledge and reasoning rely on forms of written language and mathematical reasoning. The purpose of such systems is to mimic how people think, work on problems and plan, by inserting key rational ideas at the core.

Symbolic AI has been influenced by rationalism since it seeks to develop systems capable of thinking through explanations, facts and other rules. BDI (Belief-Desire-Intention) is an example of such an approach, using logic and decision theory to model how an agent acts intelligently. Rationalist AGI focuses on making its workings explanatory and reasoned, which is why it works well in domains where strict reasoning matters. Even so, the rationalist ways may struggle to handle the uncertainty, ambiguity and complexity found in real-world situations. Formal logic tends to be very rigid, so probabilistic reasoning and mixing symbolic and sub-symbolic methods in hybrid models are being used more. Even with their drawbacks, rationalist approaches are still basic in AGI research and give a dependable way to understand and design general intelligence by using logical tools.

4.2. EMPIRICIST MODELS AND AGI LEARNING

Empiricist approaches to AGI are based on the belief that our main source of knowledge comes from what we see, hear and feel. Empiricism, which is based on the writings of Locke and Hume, believes that intelligence arises from experience and the use of perception, not just from ideas that come naturally. This study guide covering the works of Locke and Hume, empiricism argues that the development of intelligence stems from using and processing reality itself, rather than relying on ideas and logical thinking alone. For AGI, this involves systems that gain experience by acting in the world and using what they learn to adjust their actions.

In machine learning and deep learning, used in modern empiricist approaches, the systems use a lot of data and experience to adapt to differences in tasks and conditions. This approach is seen in cognition architectures such as Soar and neural-symbolic

systems, which include capabilities for perception, memory and learning to allow for flexible action depending on context. Capable of pattern recognition, sensory integration and learning gradually, these models allow AGI candidates to handle changing and challenging situations.

The use of empiricist frameworks reduces some of the shortcomings of rational systems by making systems adaptable, strong and able to learn from fresh situations. Even so, they meet obstacles, including massive data needs, challenges in understanding their predictions, and often being delicate when working outside of their trained examples. Combining empiricist and rationalist approaches in neural-symbolic systems underlines that learning from experience and using structured thinking together are necessary for real intelligence.

4.3. MATERIALISM AND PHYSICAL SYMBOL SYSTEMS

Materialism or physicalism argues that both intelligence and other mental events are based on physical processes. This point is put into practice in AGI as the physical symbol system hypothesis, which explains that a physical system able to handle symbols and follow rules can, theoretically, show general intelligence. According to Newell and Simon, the hypothesis they proposed is a key basis of classical AI, arguing that the structure of the system matters less than how it processes information. Using physical symbol systems allows for creating intelligent agents that hold knowledge, think about their surroundings and perform services through the actions of manipulating symbols. Soar and other architectures follow the materialist view by merging memory, reasoning and learning into one complete system. The fact that intelligent agents agree on knowledge structure and can reason about themselves allows them to improve and be reliable in many situations.

4.4. PHENOMENOLOGY AND CONSCIOUSNESS IN AGI

Materialist ideas are appealing to AGI since they make it easy to apply new developments in hardware and computational theory. On the other hand, some believe that symbol manipulation fails to represent fully how perception, our bodies and consciousness work in people's minds. Therefore, current research in AGI often combines approaches that reason about symbols with approaches that rely on components, bodies or connections to bring abstract computations closer to real behaviors. Phenomenology examines internal conscious experiences and the ways in which our minds are built. In the context of AGI, the use of phenomenology tries to gauge if artificial systems can truly experience or imitate real awareness, intention and being. According to this approach, models that only explain cognition through computation and brain material are incomplete, and intelligence can't be fully grasped unless subjective, lived aspects are included.

AI researchers and philosophers who follow phenomenology want to know what makes an artificial being conscious. In this, one looks at embodiment, understanding experiences, the ability to ponder them and the idea of artificial qualia. This approach usually opposes the idea of the mind being like a computer, instead highlighting the effect that context, personal interaction and one's body have on how people behave intelligently. Integrating findings from phenomenology into AGI research is still hard because computers have trouble with subjective experience. Still, addressing these issues is important for understanding the mind, intelligent behavior and the moral questions around constructing conscious devices. As AGI technology develops, views from phenomenology are expected to rise in importance in discussions about machine consciousness and what artificial intelligence can do.

4.5. EASTERN PHILOSOPHICAL PERSPECTIVES (E.G., BUDDHIST OR CONFUCIAN VIEWS)

Buddhism and Confucianism, from Eastern philosophies, have their own approaches to the mind, intelligence and ethics that can contribute to AGI research. Westerners tend to stress individual mental processes and practical activities, but Eastern philosophies focus instead on understanding everything, relationships and gaining wisdom and virtues. Buddhists conceive of the mind as always interacting with other things and not as a static element by itself. It indicates that mental states come and go, it is essential to remain mindful, and it focuses on developing kindness and moral understanding. It appears that it would be better for AGI if its architecture emphasized being flexible, a sense of context and ethics, rather than simply following orders or aiming to reach a specific solution.

Confucianism highlights the value of social harmony, developing a good character and the connections we have with people. Social knowledge and values from Confucianism might be used to guide how AGI systems are built, so they are sensitive to interaction and focused on the good of society rather than just a single person. This way of thinking brings in ethical and social factors to the main structure of AGI, rather than just focusing on intelligence as a problem-solving tool. Adopting Eastern philosophical ideas can help AGI research go further, covering both technical obstacles and questions about what gives meaning, value and ethical responsibility to life. Integrating these concepts may lead to AGI systems that are smart, kind, thoughtful and promote human well-being.

5. EPISTEMOLOGICAL CONSIDERATIONS IN AGI DESIGN

5.1. NATURE OF KNOWLEDGE REPRESENTATION

Artificial General Intelligence (AGI) must address the way knowledge is represented at the heart of its design. It is about how knowledge is structured, coded and employed by intelligent systems to work with information about the world. Representations in AGI have to be wide and still specific enough to help with many tasks and thought processes. Philosophical

epistemology identifies three schools of thought: objectivist, constructionist and subjectivist, which have different views on knowledge and how it ought to be structured. Objectivist approaches highlight facts that do not depend on the observer and usually support clear reasoning with formal representations like logic, ontologies and rule systems. Constructionist and subjectivist approaches point out that details like context, understanding and social processes matter for gaining knowledge, much like neural networks and distributed systems.

Different ways of understanding knowledge can have a major impact on AGI building. Symbolic systems do well when tasks emphasize clarity and explanation, but subsymbolic or connectionist approaches perform better at finding patterns and studying raw data. Such models combine the rules-based approach of symbolic reasoning with the ability of neural networks to adjust. An ideal AGI system would represent knowledge skillfully, so that it can handle strict logic as well as flexible learning from everyday experiences, imitating the way the human mind operates.

5.2. LEARNING, UNDERSTANDING, AND THE SYMBOL GROUNDING PROBLEM

Learning and understanding are key parts of AGI, yet they bring major challenges, primarily due to the symbol grounding problem. This challenge, which Stevan Harnad put forward, queries whether computational systems can really attach meaning to abstract symbols or if those symbols simply follow patterns without connecting to real-world events. Symbols in traditional symbolic AI are linked to other symbols, so their meaning can go on forever until they are connected to something sensory or physical.

According to empiricist models, AI should interact with their world, collect sensory information and develop concepts using similar reasoning as humans gain knowledge. Nonetheless, having a true understanding means that the system can relate its internal ideas to things happening in the world. Utilizing symbolic as well as sub-symbolic ways of learning is becoming more popular, as it allows AGI to relate symbols to perception and think about them non-concretely. AGI systems face challenges in using their experience to make general statements, applying their knowledge to various domains and building strong conceptual networks. AGI research focuses on combining learning, understanding and symbol grounding because these elements are central to shaping both the theory and structures of future intelligent systems.

6. ETHICAL AND ONTOLOGICAL IMPLICATIONS

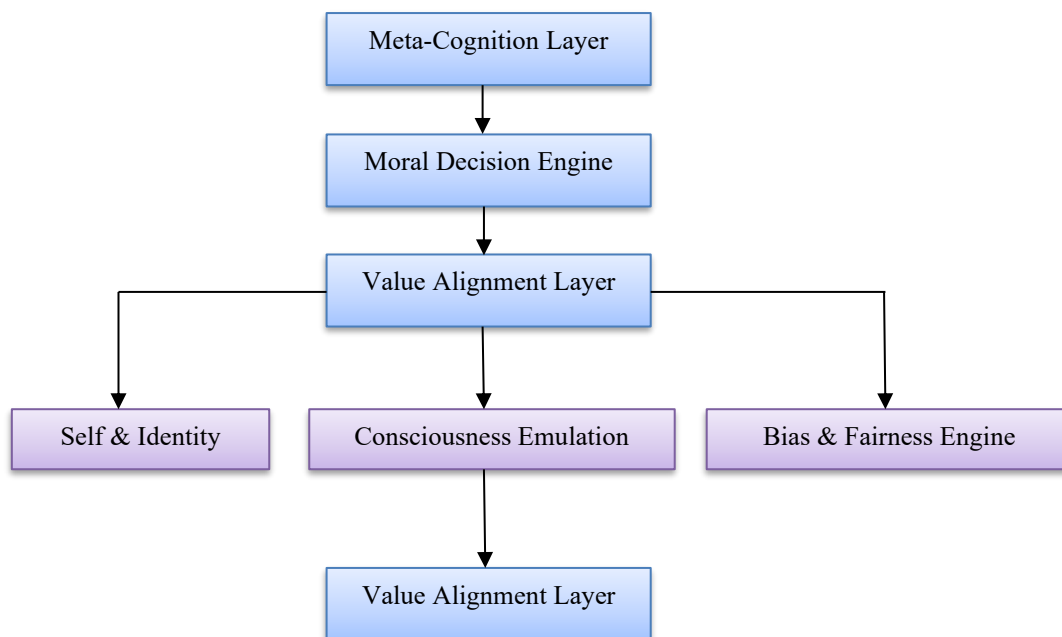


FIGURE 3 Ethical and ontological integration in AGI

6.1. MORAL STATUS OF AGI

When AGI appears, it opens up a debate over whether such systems ought to be seen as ethical individuals or as people worthy of respect. When AI comes close to human thinking power, it blurs the line between tools and agents, making it necessary to consider matters of self-decision making, responsibility and rights. Ethical frameworks should address whether AGI can have its own interests, intentions, or feel pain, as these are usually key considerations for granting moral status. Because AGI can decide things on its own and shape the lives of people, it is very important to ensure that AI systems are closely monitored and open. As these systems become more independent, concerns about who bears responsibility and what ethics are involved for those who design them become important. Collaboration among experts in various fields and the development of well-structured governance are needed so that AGI promotes human values, remains safe and benefits society.

6.2. PERSONHOOD AND IDENTITY

The question revolves around whether AGI demonstrates traits such as self-awareness, the ability to make deliberate decisions and sustained self-identity, which are usually linked to personhood. Should AGI computers have behaviors and abilities similar to ours, discussions about whether they are artificial persons become more urgent. People are debating this issue on both a philosophical and legal level: Should AGI have rights, duties or legal recognition? Is it AGI's code, what it has gone through or its physical existence that defines its identity? As AGI might develop its own ways of thinking and improve itself, it becomes hard to draw a line between a man-made machine and a true autonomous being. The discussion of personhood in AGI also touches on issues such as morality, social integration and the possibility of AGI developing its own self-concept or shared identity, which should be carefully examined and regulated.

6.3. HUMAN-AI COEXISTENCE AND ETHICAL ALIGNMENT

Maintaining balance between humans and AGIs is a major ethical requirement. AGI's ability to make decisions like people raises concerns about misalignment with ethics, loss of control and potential harm to society. Ensuring that AGI systems follow human ethics is a difficult issue, mainly because ethical principles differ around the world. Bringing AGI into use may result in people losing their jobs, changes in the workforce's makeup and new power shifts. Making such governance structures transparent, accountable and inclusive demands that ethicists, policymakers and technologists collaborate. Laws and guidelines should put safety, human supervision and fair access first and be ready for the potential effects of AGI technology on humans and society. The aim is to ensure AGI supports human well-being, maintains democratic values and sticks to strong ethical standards.

7. IMPLICATIONS FOR AGI ARCHITECTURE AND DEVELOPMENT

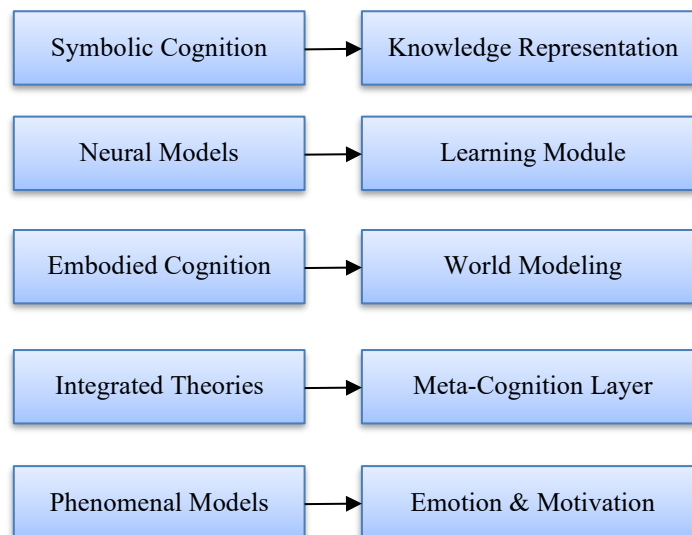


FIGURE 4 Cognitive models to AGI system design

7.1. HOW PHILOSOPHICAL ASSUMPTIONS SHAPE AGI DESIGN

The way AGI is built and developed is greatly affected by philosophical views concerning intelligence, knowledge and consciousness. Due to emphasizing reasoning and formal logic, rationalist traditions caused architectures to be designed with a clear focus on representing symbols and applying explicit rules. The purpose is to have AI systems match human thinking, which supports clear and understandable decision-making. Conversely, empiricist philosophies, which emphasize learning from what we experience with our senses have led to models like neural networks. Materialist approaches, pointing to the importance of physical bases, agree that intelligence can be fully achieved in computers by having them process perception, memory and actions all together. Philosophers in the West and East focus on movements of the body, context and connection with others, which leads us to consider sensory-motor feedback, learning that adjusts to new information and a moral sense. The outcome is that leading architectures mix ideas from both approaches, often using combinations of symbolic and subsymbolic methods for more flexible and human-like systems.

7.2. DESIGN TRADE-OFFS (SYMBOLIC VS. SUBSYMBOLIC; LOGIC VS. EXPERIENCE)

Developing AGI forces designers to make difficult choices about using symbolic or subsymbolic systems and to decide if reasoning should rely on logic or learning from experience. Symbolic systems do well at complex tasks that need orderly thought, clearly presented knowledge and can be easily understood. They are the right choice in areas that require high transparency and accountability. But they find it difficult to handle unclear situations, how they are perceived and adapting to something new. Deep neural networks and similar subsymbolic systems are good at noticing patterns, learning from data that is

not in order and generalizing their knowledge. They usually perform well in ever-changing scenarios, yet it is often hard to explain how they work, and they need tons of data for training.

The conflict between logical thinking and gained experience is also significant. Logic-based systems offer significant reliability and guarantees, but they might fail under conditions of uncertainty. Experience-driven models evolve and withstand challenges, but they might be hard to understand and can act unpredictably. Hybrid approaches are becoming more popular in present AGI research to merge the positive aspects of each style and reach higher levels of generality, robustness and transparency. Selecting these design factors carefully helps ensure both the technological strength and social, ethical and social acceptance of AI.

7.3. IMPACTS ON EXPLAINABILITY AND TRANSPARENCY

For responsible use of AGI, explainability and transparency are necessary to build trust, hold everyone accountable and ensure AGI obeys regulations. Architectures that display and track reasoning steps are chosen when building software that values rationalism and symbolic reasoning. Systems like these are able to clearly explain their choices, which support monitoring and ethical analysis. Conversely, subsymbolic and empiricist approaches may be tough to understand or predict as they act in ways that are not easy for us to explain. The ambiguity around AGI raises doubts about safety, fairness and risks of unforeseen results when AGI systems are trusted with more power and autonomy. The need for explainable AI has inspired methods to explore neural models, show decision paths and blend symbolic reasoning to increase how clear AI is. Architectural paradigms and their underlying beliefs affect how easy to understand, trustworthy and widely accepted an AGI will be, making it important to keep exploring different fields as the field of AI gets stronger.

8. DISCUSSION

Working on Artificial General Intelligence (AGI) involves both technical and important philosophical and ethical issues. As AGI systems get closer to matching human intellect, they create concerns about who should be in control, who decides and who bears responsibility for their decisions. Making sure that transparency, accountability and value match are maintained is important, especially as AGI systems learn from themselves and make decisions on their own. Since unpredictable actions, ethical mistakes and leaving control too much to machines are risks, logical collaboration between cognitive specialists, ethics experts, and policymakers is necessary. These frameworks have to deal with matters such as decision-making freedom, responsibility, security and equal chances to use AGI, as well as the effects that AI has on workforce changes and financial disruption.

The focus of the debate lies on making sure AGI supports and helps uphold different human values and ethical concepts. Since AGI systems are used in important domains such as healthcare and robotics, ensuring explainability, fairness, and bias mitigation becomes extremely important. Effective reporting, ethical audit practices and inviting all key stakeholders help AGI protect the public interest and gain people's trust. Since AGI continues to advance, new ethical rules must be put in place to address risks and unintended harm, while ensuring development and avoiding problems. Our ability to shape smart technologies ethically and technically is key to shaping AGI, so that it not only supports our skills but also protects society.

9. FUTURE DIRECTIONS

The evolution of AGI research depends on fast technology advancement, joining experts from different areas and strong regulation. By using deep reinforcement learning, neuroevolution and hardware inspired by the brain, AGI systems can respond effectively to change and take over intricate decision-making processes. Since merging neuroscience with artificial neural networks has become common, the development of AGI now leads to hybrid systems that notice their environment, think abstractly and process information in real time. The improvements are making AGI smarter and also allowing it to be used in different industries such as autonomous cars, industry automation, medical science and creative fields.

At the same time, the progress of AGI results in many changes for society, ethics and the economy. When AGI starts to think and learn in a human-like way, concerns about losing our jobs to AI, protecting personal privacy and unexpected mistakes from "black-box" models grow. It is important that technology encourages ethical values, provides transparency and is helpful for society. Researchers recommend the use of evaluation, ethical standards and involvement from multiple parties to keep harmony between progress and the community. Empirical, multi-speciality experiments, proper ethical progress and creating AGIs that depict life's important aspects are necessary in the upcoming years. For AGI to progress, it will depend on new technologies and also on maintaining strong commitments to ethics and diverse public discussions.

10. CONCLUSION

The study of Artificial General Intelligence (AGI) indicates that seeking human-level AI involves discussion of concepts and ethics as well as technical development. Using rationalist, empiricist, materialist, phenomenological and Eastern philosophical theories, we can more clearly see the ideas and values driving AGI's design. These ideas highlight the hurdles faced in representing knowledge, learning, understanding consciousness and keeping AGI ethical, all of which need to be handled so AGI is strong, understandable and helpful for humanity.

As AGI research develops, the link between philosophical questions and advances in technology will likely retain its importance. For AGI to succeed, it will also rely on having ethical thinking, working with different fields and engaging with society throughout its development. Realising AGI in a responsible way can lead to significant progress, yet it also stresses the importance of staying alert, humble and committed to helping all people through artificial intelligence.

REFERENCES

- [1] Sonko, S., Adewusi, A. O., Obi, O. C., Onwusinkwue, S., & Atadoga, A. (2024). A critical review towards artificial general intelligence: Challenges, ethical considerations, and the path forward. *World Journal of Advanced Research and Reviews*, 21(3), 1262-1268.
- [2] AI to AGI: Ethical Considerations with EthicalEngineAI, Linked in, online. <https://www.linkedin.com/pulse/ai-agi-ethical-considerations-ethicalengineai-adam-m-victor-eia0c>
- [3] The Implications of Artificial General Intelligence (AGI) on Technology, kanerika, online. <https://kanerika.com/blogs/artificial-general-intelligence/>
- [4] Slavin, B. B. (2023). An architectural approach to modeling artificial general intelligence. *Heliyon*, 9(3).
- [5] Raman, R., Kowalski, R., Achuthan, K., Iyer, A., & Nedungadi, P. (2025). Navigating artificial general intelligence development: societal, technological, ethical, and brain-inspired pathways. *Scientific Reports*, 15(1), 1-22.
- [6] Implications of Artificial General Intelligence on National and International Security, yoshuabengio, online. <https://yoshuabengio.org/2024/10/30/implications-of-artificial-general-intelligence-on-national-and-international-security/>
- [7] Barnes, E., & Hutson, J. (2024). A Framework for the Foundation of the Philosophy of Artificial Intelligence. *SSRG International Journal of Recent Engineering Science*, 11(4).
- [8] Abbate, F. (2023). Natural and artificial intelligence: a comparative analysis of cognitive aspects. *Minds and Machines*, 33(4), 791-815.
- [9] Halpin, H. (2025). Artificial intelligence versus collective intelligence. *AI & SOCIETY*, 1-16.
- [10] Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. *Journal of Artificial General Intelligence*, 5(1), 1.
- [11] Islam, M. M. (2024). Artificial General Intelligence: Conceptual Framework, Recent Progress, and Future Outlook. *Journal of Artificial Intelligence General Science (JAIGS)* ISSN: 3006-4023, 6(1), 1-25.
- [12] Wang, P., & Goertzel, B. (Eds.). (2012). *Theoretical foundations of artificial general intelligence* (Vol. 4). Springer Science & Business Media.
- [13] Joshi, K. (2024). Artificial General Intelligence (AGI): A Comprehensive Review. *Journal of the Epidemiology Foundation of India*, 2(3), 93-96.
- [14] Anthony, M. (2008). Integrated intelligence: Classical and contemporary depictions of mind and intelligence and their educational implications (Vol. 19). BRILL.
- [15] Rescorla, M. (2015). The computational theory of mind.
- [16] Aydede, M. (2005). Computation and functionalism: Syntactic theory of mind revisited. In *Turkish studies in the history and philosophy of science* (pp. 177-204). Dordrecht: Springer Netherlands.
- [17] Priya, A. (2019). Haruki Murakami's characters and works-as the representation of postmodernism. *Think India Journal*, 22(3), 56-62.
- [18] Božić, P. (2020). Consciousness for AGI. *Procedia Computer Science*, 169, 365-372.
- [19] J. Charles Arockiasamy. (2025) A Post Humanistic Approach to H. G. Wells's *The War of the Worlds*, *International Journal of Business and Economics Research*, Special Edition: Futuristic Trends in Literature, 16-22.
- [20] Dr. A. Ezhugnayiru (2019) "The Apocalyptic Vision of Orhan Pamuk Reflected In the Black Book: A Study In The Light of Eco-Criticism," *International Journal of English Language, Literature in Humanities*, 7(7), 611-617.
- [21] Mallisetty, Harikrishna; Patel, Bhavikkumar; and Rao, Kolati Mallikarjuna, "Artificial Intelligence Assisted Online Interactions", Technical Disclosure Commons, (December 19, 2023) https://www.tdcommons.org/dpubs_series/6515
- [22] J. Charles Arockiasamy (2023) "Study of Algorithms from a Transhumanist's Approach to a Modern Narration in the Artificial Intelligence Generated Novel," *New Frontiers in Education*, 56(4).
- [23] A. Priya (2025) Climatic Changes in Kim Robinson Staley's *Fifty Degree Below: An Eco-Critical Study*, *International Journal of Business and Economics Research*, Special Edition: Futuristic Trends in Literature, 118-123.
- [24] Ezhugnayiru, A. (2025) The Interplay of Memory, History, and Fiction: Metafictional Devices in Holocaust Narratives and Collective Consciousness, *International Journal of Literature, Linguistics, and Humanities*, 1(1), 25-32.