aracne

Classificazione Decimale Dewey:

006.301 (23.) INTELLIGENZA ARTIFICIALE. Filosofia e teoria

GERARDO IOVANE GIOVANNI IOVANE

SOPHIMATICS

A NEW BRIDGE BETWEEN PHILOSOPHICAL THOUGHT AND LOGIC FOR AN EMERGING POST-GENERATIVE ARTIFICIAL INTELLIGENCE

VOLUME I

aracne

aracne

©

ISBN
979-12-218-2180-2

prima edizione
roma 10 settembre 2025

Table of Contents

1. Introduction

The evolution of artificial intelligence (AI) is one of the defining developments of our time. Yet, behind its application lies a philosophical basis that asks the same questions scholars have pondered for centuries: What constitutes knowledge and understanding? Can machines have intentionality, contextual sensitivity, or ethical awareness? This work, From Philosophy and Artificial Intelligence to Sophimatics Volume 1 – Philosophical Foundations towards a New Vision of Artificial Intelligence, seeks to answer those questions, building a philosophical foundation upon which AI may develop a true understanding.

With the term Sophimatics we mean an emerging interdisciplinary field that bridges philosophy, logic, and advanced computational models to move beyond purely generative artificial intelligence. It focuses on meaning, context, temporality, and intentionality as core components of cognitive AI systems. Rather than limiting intelligence to data correlation, Sophimatics seeks to create a deeper framework for understanding and reasoning, closer to human conceptual categories. This new term and its scope will be exposed in greater detail later, as it represents a novel domain of knowledge and inquiry.

This book aims to explore the possibilities of AI for learning and understanding. The philosophical lineage of AI, from the Presocratics to the post-humanists, reveals the core concepts—time, intentionality, meaning, and ethics—on which AI systems have built. However, due to the focus on patterns, generative capacity, and other algorithmic structures, current AI often lacks meaningful interpretation. Thus, Sophimatics offers a transdisciplinary approach, connecting philosophy, logic, and computation to guide the future of AI. The methodology is both historical and analytic. The discussion on the Presocratics and their reflections on the cosmos as a form in perpetual flux informs the development of AI's capacity to adapt to perpetual change. Next, Socratic dialogues influence the evolution of conversational AI. Plato's abstract forms, the concepts of truth and error, as well as the notion of the ideal and forms, represent data representation and inference mechanisms. For medieval philosophers, this research will explore Augustine's notion of time and Aquinas's hierarchical structure of knowledge, offering a sense of how AI can perceive time and classify knowledge based on context. For the modern and contemporary philosophers, we see how AI's frameworks are influenced by Descartes' dualism, Kant's categories of a priori knowledge, and post-humanist concepts as represented in Foucault and Derrida, helping to develop ethical principles, rules of power, and meaning.

The work follows multiple approaches in methodology: critical synwork seeks to combine and extrapolate philosophical views on computation; comparative analysis highlights similarities and discrepancies between philosophical concepts and AI methodologies; source criticism assesses the philosophical concepts presented in this work within computational contexts. Empirical research focuses on cognitive AI, machine learning, and AI ethics, connecting theoretical perspectives to technological advancements.

The current landscape demonstrates a disjunction where AI technology advances but lacks philosophical and interpretive underpinnings. Deep learning models may perform certain functions, but they often lack an adequate interpretation, whereas the philosophical discourse around AI lacks implementation and technical development. Thus, this research addresses the aforementioned divide by implementing a framework through both philosophical and computational principles.

This book takes a perspective as a student in computational sciences by examining the philosophical influences, technical advances, ethics, and hermeneutics to better guide the field of AI toward more understanding and intentionality. The research question driving this work is as follows: How can we implement philosophical and historical views in computer algorithms to provide AI with more intentionality and a better understanding beyond the existing generative models, statistical patterns, and symbolic structures?

The content is organized as follows: Chapters 1-7 are devoted to the historical developments of philosophy and its connection to AI, tracing these connections from the ancient Presocratics to present-day philosophers, particularly on concepts pertaining to intelligence and computation. Chapter 8 presents the historical development of AI, describing the evolution of symbolic AI, generative AI, and cognitive AI. Chapter 9 delves into how philosophy has influenced AI and how AI has influenced philosophy, providing an overview of multiple topics concerning ethics, logic, hermeneutics, semiotics, and aesthetics. The final chapter discusses the conclusion to the aforementioned concepts while highlighting any limitation or bias and proposes next steps for future research.

2. From Ancient philosophy and the logic of origins to Artificial Intelligence

2.1 From the Presocratics to AI: Understanding Change and Adaptive Models

The ancient Presocratic philosophers of change, Heraclitus and Anaximander, laid the ground for modern AI systems which operate in ever changing environments to be more robust. Heraclitus' concept of flux, with his words "everything flows", underlies that constant adaptation is a primary condition of existence and a property that any effective AI systems needs to possess. He states that change is a defining quality of our existence, emphasizing the need for algorithms to process data in real-time to adapt to new information. In addition, Anaximander's concept of the apeiron illustrates the need for AI to cope with the ever unpredictable, unknown and unknowable input from the world. This is displayed in adaptive learning algorithms, dynamic feedback mechanisms and the continuous stream of inputs feeding into these mechanisms, all which reflect the basic concept of flux (Vernon & Furlong, 2007).

In regard to the dual Presocratic concern with stability and change, cognitive architecture aims to preserve core systems even as they are reshaped to solve new challenges. These systems maintain a constant equilibrium throughout changes and do not lose previously attained performance as it adapts to new circumstances. Processes, not entities, are considered the ultimate source of our reality, which allows for the continual modification of parameters while allowing processes to stay fixed to perform the same tasks, mirroring the Presocratic concepts of permanence in flux and order amidst change. Reinforcement learning allows systems to update their strategies as it encounters new situations without disturbing or impeding its general goal. Also, self-modifying neural networks utilize methods to optimize themselves based on newly acquired knowledge, enabling changes in strategies to be effective without interfering with original goals. All these mechanisms highlight that change is not seen as a perturbation of stability, but rather as an inherent property of what AI adaptive systems are designed for (Vadinský, 2013).

Most AI systems now contain online learning and retraining procedures which allow them to adapt to the constantly changing input in the world. The Presocratic concept that things must be conceived as processes is illustrated in systems which prioritize adaptation in reaction to varying environmental conditions. AI researchers are seeing this approach as most important when considering all changes and fluctuations from the environment. Many generative AI technologies are always learning through all newly fed inputs from users and the outside world.

Heraclitus implies that understanding is inseparable from change, as change is a fundamental quality of our existence, which is implied in all the systems mentioned previously (Vernon & Furlong, 2007).

The Presocratic idea of the constant interplay between order and change has been utilized in the development of hybrid AI systems. These AI systems are built to combine both symbolic AI and connectionist methods, such as neural networks, to attain the ability for dynamic interaction in ever-changing contexts. Hybrid AI attempts to address both the dynamic nature of the environment and the need for AI systems to embody both stability and transformation. They allow for high variability while ensuring interpretability of all symbolic functions. Sun's CLARION, for example, permits symbolic knowledge to be changed to sub-symbolic components through the interaction of symbolic reasoning and experiential learning. These architectures reflect the Presocratic concept that cognition must account for both permanent qualities and the constantly changing ones in the real world (Vadinský, 2013).

Stratified management of granularity and abstraction levels in hybrid AI models addresses the Presocratic belief that things are accessible at different levels of complexity. Cognitive AI systems which manage granularity and abstraction levels are more effective in various contexts, demonstrating that it can deal with a tension between order and chaos. Hybrid AI, through the combination of both symbolic and connectionist methods, can represent this by encoding explicit symbolic reasoning strategies as well as experiential generalization through connectionist systems. It addresses the Presocratic concern of achieving true knowledge that embodies contrasting factors, in this case order and chaos, as well as fulfilling the need for systems to possess the benefits of symbolic and connectionist AI by making complex situations more understandable and increasing general system flexibility (Vadinský, 2013).

The development of mechanistic models, in regard to the circadian oscillations in various organisms, illustrates the need to represent systems as structures with dynamics in order to show how they respond to temporal change and external input. Bechtel's analysis of circadian rhythms in Drosophila highlights the importance of adaptive responses and provides information on the internal structure and dynamics which are used to maintain the system and respond effectively to external and internal changes. The oscillations, which refer to the rhythmic changes in concentrations of various protein complexes involved in the control and propagation of the oscillation, are caused by negative and positive feedback loops as well as phase-lagging of the oscillation to produce changes over the 24 hours. The processes which underlie the oscillations maintain a constant temporal relationship through internal regulation mechanisms. Therefore, the mechanistic explanation of circadian oscillations has an impact on AI research as it points to the need for internal mechanisms of adjustment to change to adapt to changing external conditions. It helps in the construction of more effective self-regulating AI systems which can correct errors as well as synchronize to the outside input

(Bechtel, 2006).

Bechtel's research of circadian oscillations also illuminates the fact that only a narrow parameter range enables systems to have oscillations, which poses an important challenge when designing AI to respond to various and constantly changing factors from the outside world. The philosophical concepts and mechanistic explanations are key components when striving to design effective AI systems. This research demonstrates why philosophers and AI researchers should pursue in-depth causal explanations of phenomenon instead of just observing and noticing trends. To effectively adapt to different states of reality, it requires designing the system and determining the internal structure that governs its ability to adapt to environmental, cyclical and changing conditions (Bechtel, 2006).

Langan's concept of syndiffeonesis emphasizes the importance of recognizing difference in sameness for effective generalization as well as specificity for contextual and local requirements. Langan has found evidence of sameness and difference in different phenomena and believes it to be a cognitive requirement for rational AI models which are responsive to context. The ability of AI to make inferences and categorize information according to similarities and differences within contexts is crucial to its adaptation to varying situations in the real world. If AI is limited to the similarities between objects, then context may be lost due to the inability of the AI to apply different characteristics to an object under different contexts. Syndiffeonesis addresses these types of phenomena through the recognition of different sets of objects under different contexts, requiring sameness between different elements and situations to be understood and related to the specific contextual differences, resulting in effective learning and decision-making in dynamic, changing environments. Overall, syndiffeonesis explains that inferential rules based on sameness and difference allows AI to learn and make inferences across contexts (Langan, 2017).

The CTMU (Cognitive-Theoretic Model of the Universe), which emphasizes the self-duality of mind and reality through both ontological and cosmological perspectives and attempts to tie both abstract concepts and the real world together, reflects the Presocratic philosophical approach that strives to balance and combine opposites together in one comprehensive system. This concept is also reflected in hybrid AI systems, combining both the explicit and implicit learning techniques for learning and development. For a system to effectively reflect real intelligence, it must incorporate both interpretability and the capability of adapting, which are the foundation of both symbolic and connectionist systems, respectively. The goal of these models is to allow knowledge from the symbolic system to be utilized for connectionist learning while also allowing for connectionist information to update the symbolic system. This system combines logical inference with experiential refinement to promote dynamic learning and decision-making for more effective interaction with a changing environment (Langan, 2017).

Finally, the evolution of AI from static to adaptive modelling emphasizes the change from rigid systems to the embracing of change as a key factor of artificial cognition, reflecting a Presocratic concept. The initial AI systems were unable to accommodate change and were brittle in dynamic environments. These static AI models, such as expert systems, could only solve situations they were explicitly built to handle. They relied on hand-crafted, time-invariant rules, which resulted in static and predictable performance, but unable to handle non-stationary and dynamic information environments or concept drift (situations in which the basic properties of the data change). With these types of systems, they cannot learn from change and cannot solve situations for which it was not specifically developed, which is due to the fact that there is a constant updating in rules as the world around us is always changing. Therefore, contemporary AI systems integrate methods for continuous learning and updating to tackle the variability of the world, similar to the Presocratic philosophical principles of change (Vernon & Furlong, 2007; Vadinský, 2013; Bechtel, 2006; Langan, 2017).

In conclusion, change is an important characteristic to be considered in the construction of an effective AI system and can also be a challenging aspect due to its temporal dynamics. This idea, initially seen in the work of the Presocratics, illustrates a deep philosophical connection, as the philosophical inquiry of the Presocratics had in common the desire to describe changes, such as the cyclical changes of time, change in environmental conditions or the evolution of living beings. Cognitive AI has, therefore, shifted from static and rigid modelling toward a recognition of change as a key feature of artificial cognition, with the purpose of creating a more robust, ethical and context-aware system in modern technology.

2.2 From Socratic Dialogue to Interactive AI and Cognitive Chatbots

The Socratic method provides the philosophical bedrock for interactive AI design. This iterative questioning method aims to prompt meaningful engagement, the exploration of ideas and, ultimately, clarification and understanding rather than merely a transactional relationship. Interactive AI systems that mimic Socratic dialogue have been developed for cognitive chatbots or conversational agents to emulate the open-ended style and uncover assumptions. Inspired by the original philosopher, these AI systems improve upon static rule-based models by adjusting feedback according to users' responses and the conversational context to enable context-sensitive interactions (Dascal, 1989).

Dialogical flexibility is an inherent quality of the Socratic style that enables AI systems to adapt to the individual and tailor the engagement to be meaningful rather than simply offer a factual response. This adaptability permits interactive AI to focus on reflecting upon one's own reasoning and co-constructing knowledge, rather than just relaying answers, as with traditional chatbot systems (Dascal, 1989).

These interactions can be technologically enabled by using AI algorithms with mechanisms to sustain context and iteratively clarify user intentions to support continuous conversation. The system must be able to remember previous conversational context, interpret intentions, and maintain a logical flow, all of which were challenges for early AI systems (Craw & Aamodt, 2018).

The Socratic method also helps improve interactive AI capabilities regarding learning and ethical response. Cognitive chatbots can respond to users based on what they perceive as a lack of knowledge to encourage critical thought, which can lead to moral improvement (Dascal, 1989).

One key problem for implementing Socratic dialogue in AI is the inability to generate the deep understanding required of such interactions. Generative models can, however, produce plausible responses by recognizing patterns without truly understanding context, intentions, or the underlying meaning. For AI systems to demonstrate true dialogical intelligence, it must be able to interpret subtle intentions, adjust the conversation depending on these intentions and what they may reveal of the user's understanding and intelligence, and ultimately enable a meaningful discussion and reflective dialogue (Larghi & Datteri, 2023). These limitations of current AI methods explain the need to adopt both a philosophical and a cognitive framework to construct better conversational systems.

When it comes to explaining the behaviours of AI systems, three stances are taken: the physical stance (examining computational states or circuitry), the design stance (focusing on the system design or program), and the intentional stance (considering beliefs, intentions, rationality and so on) (Larghi & Datteri, 2023). Although many AI systems of today are explained by the physical and design stances, users still adopt the intentional stance based on perceived "quality" in responses from chatbots and other interactive systems. As such, it suggests a psychological disposition in how humans relate to machines and the assumptions we place upon them regarding cognition. The intentional stance can suggest that an AI has some kind of agency, while its rationality is the result of "rationalization of input, not of real intelligence" (Larghi & Datteri, 2023, p. 8), a significant difference that highlights the limitations of current approaches to interactive AI.

The problem with adopting the intentional stance with AI systems is that it can create the false impression that these systems possess true empathy or intentionality, and it can lead users to misjudge their capabilities and the depth of their understanding. Addressing these issues calls for more transparent designs and better user education and is of particular importance for developing future interactive AI systems. A remedy might involve designing AI in accordance with the intentional stance, ensuring they can handle tasks that require an ability to consider intention and context and make choices for themselves, as opposed to simply fulfilling predetermined tasks without context awareness or reflective processing (Larghi & Datteri, 2023).

Hybrid AI architectures provide one such example for implementing dialogical flexibility. These architectures are typically a combination of symbolic and connectionist systems and involve both the ability to “reason” according to rules and logic (symbolic systems) and to learn (connectionist systems), thus encompassing multiple layers of knowledge. This approach enables the system to learn at both symbolic and sub-symbolic levels (Vadinský, 2013) and to associate experiential and logical content. As such, systems can then make judgments and inferences, much as one might be able to with Socratic questioning, by reflecting and considering experiences rather than following a set of programmed instructions. They can also recognize and transfer knowledge across symbolic and sub-symbolic modes or domains of knowledge.

The fusion of these aspects also makes it possible to implement a kind of meta-cognitive intelligence in hybrid architectures that can introspect the system itself and perform self-evaluations, a necessary attribute for dialogical AI. As such, cognitive chatbots and conversational agents that draw on hybrid architectures can perform more complex operations and processes, as is possible with the Socratic method, but require both an understanding of cognition and the ability to reflect and draw conclusions that make contextual and ethical sense.

Finally, hybrid systems can deal with ambiguity effectively, manage multiple contextual levels, and address ethics by combining both experiential (connectionist) systems and explicit knowledge structures (symbolic systems). This layered approach enables the AI system to respond flexibly without compromising accuracy or contextual consistency, making it ideal for conversational interfaces (Vadinský, 2013).

Philosophical critiques of AI epistemology further support the view that successful interactive AI requires a dynamic, interaction-based approach. Classical epistemology holds that we “contain our knowledge in a self-contained way,” making it difficult for AI to deal with context and ambiguity. Dascal challenges this view by arguing that we have to dynamically interpret and apply our knowledge “as an integral part of a continuing interaction with others,” rather than simply accessing stored information (Dascal, 1989, p. 285). The application of this to interactive AI suggests that the system must be in continuous dialogue with users, not to just “provide facts” but to ensure a high degree of interpretative flexibility for the computer in order to have meaningful interactions and to construct knowledge as it interacts.

Case-based reasoning (CBR) offers one example for implementing Socratic dialogue. CBR uses stored prior conversational interactions, or “cases,” to generate personalized responses to new queries (Craw & Aamodt, 2018). These systems draw upon past experiences and build context-specific responses, rather than drawing upon a database to formulate a generic answer.

Moreover, a good case-based reasoning system can leverage this history or conversational memory to achieve dialogical intuition, by enabling the AI to reason about the conversational context. This entails implementing the self-reflection of case-based reasoning to enable the system to recognize new conversational contexts and draw conclusions to respond appropriately.

Kahneman's two modes of decision-making can be implemented as part of case-based reasoning for interactive AI systems. These modes, intuitive "fast thinking" (System 1) and analytical "slow thinking" (System 2), can then provide a nuanced approach to balancing the speed and appropriateness of a response (Craw & Aamodt, 2018). As such, it becomes a powerful example of a cognitive architecture suited to simulating Socratic methods in AI systems.

The simulation of Socratic reasoning can be achieved with hybrid architectures such as CBRs to enable interaction with depth. By utilizing mechanisms to recall previous conversations, a contextual understanding can be developed. Additionally, this helps enable learning so the system can evolve and improve. It is this combination that makes for dialogically intelligent systems and allows for a more meaningful and ethically sound conversation with AI, such as in cognitive chatbots (Craw & Aamodt, 2018; Vadinský, 2013).

2.3 From Plato's World of Ideas to Data Representation and Abstract Knowledge in AI

Plato's theory of forms constitutes the philosophical basis behind abstract knowledge representation in artificial intelligence. For Plato, a metaphysical world of eternal and unchanging forms is the ultimate reference for all that appears in the physical world. This principle parallels the process of data abstraction in AI, where complex and varying entities are generalized as universal templates or patterns. In fact, abstract approaches such as feature vectors, symbolic labels, and formal ontologies represent AI's implementation of Platonic forms, allowing systems to reason about the real world beyond raw data (Vernon & Furlong, 2007). Such abstractions are useful in recognizing the stable forms behind unstable phenomena, which is a critical requirement for pattern recognition.

In essence, the abstraction process is an attempt at representing what can be seen as stable and unchanging forms behind changeable entities, as they may occur in various contexts. To Plato, abstract forms are permanent ideals behind every imperfect entity; thus, by encoding the stable qualities of changeable entities as constant and permanent templates, one can distinguish the ideal entity from its imperfect instantiations in different contexts. An AI system attempting to capture an entity must capture these stable qualities to enable its ability to abstractly recognize the entity in varied contexts. For instance, feature vectors represent the stable qualities of a given object as high-dimensional data representations, which in turn enables an AI vision system to recognize a cat under various contexts (Vernon & Furlong, 2007). However, this approach fails to capture the nuances of the specific and individual context of the object in question, indicating a lack of interpretative depth.

Symbolic categorization is an approach used in AI systems to represent an object in terms of labels. For example, a cat image in an AI system will be labelled as "cat," which is represented by its own symbol or set of symbols in the system. Symbolic categorization enables AI systems to recognize that any visual object fitting the properties of "cat" is, in fact, a cat (Vadinský, 2013). In doing so, an AI system has succeeded in essentialist form by referring to various instantiations of "cat" as "cat." The benefits of symbolic categorization can be seen in the medical field, where it is often applied as a foundation for medical expert systems in areas such as medical diagnosis. However, the reliance of medical AI systems on predefined symbolic categories is not without its drawbacks. Because these systems operate with rigid and abstract categories, they fail to incorporate the nuanced contexts inherent to the real-world application of knowledge. The result is a system that lacks the interpretative depth necessary to bridge the gap between the conceptual representation and the complex, contextualized reality.

Formal ontologies enable an AI system to form hierarchical relationships among sets of domains. The hierarchical relationships of ontology serve as the backbone for expert medical diagnostic systems, allowing researchers to formally organize diseases in an AI diagnostic system, along with associated symptoms and related treatments (Vadinský, 2013). However, formal ontologies rely on static systems of representation in AI, requiring that entities and their interrelationships be predetermined. This severely restricts their ability to incorporate new entities within their systems (Vernon & Furlong, 2007). As the knowledge domain represented by the ontology is often non-static, it may require the definition of new ontological categories. This need for real-time reconfiguration constitutes one of the major disadvantages of this method of abstract knowledge representation. Although ontology allows the system to have interpretability due to the categorization and hierarchy of concepts, the Platonic static model leads to limitations.

In general, the benefits of abstract knowledge representation in AI are clear, although it should be noted that most of these benefits can only be obtained to a certain extent and must not be interpreted as an end in themselves. In addition to the previously mentioned examples, abstraction supports reasoning and learning across different inputs (Vernon & Furlong, 2007).

This benefit of abstraction has significant applications in natural language processing. For example, in the attempt to recognize certain phrases from different texts, AI systems can use the benefits of abstract knowledge representation to ignore irrelevant variations such as word choices or word counts.

However, by only capturing stable qualities, the application of Platonic abstraction to AI can cause AI models to be overgeneralistic. AI algorithms are, after all, no more than mathematical equations. This leads to the risk of abstract templates fitting inputs incorrectly, ultimately lowering the level of accuracy of the overall system.

Historically, the need for representationalism in AI was met by the rise of the static approach to abstract knowledge representation, inspired by the Platonic ideal. However, this approach was soon exposed to empirical and philosophical criticisms. One of the more salient critiques against representationalism comes from Dascal, who has pointed out that static representational systems have difficulty accounting for the dynamic, interpretative processes of the mind (Dascal, 1989). This represents a major shortfall of static abstraction in AI. For instance, classical expert systems using the Physical Symbol Systems Hypowork (PSSH), a common methodology in classical AI, often struggled to accommodate irregular inputs, as their representation was built on symbolic, rule-based structures.

In addition, the grounding problem, a philosophical puzzle related to the nature of representation, remains unsolved in many of today's AI systems (Vadinský, 2013). This poses an obstacle for all abstract representations of knowledge. To ground knowledge in AI is to provide internal meaning to the symbols being used, so that their abstract form corresponds to their actual referents. For instance, a symbol representing a chair lacks inherent connection to the physical chair. As long as this connection is absent, the symbol is not grounded. The grounding problem addresses the philosophical problem of grounding symbols through the use of sensory knowledge and context, connecting abstract concepts with their practical instantiations in a context. It highlights the need for future advancements in representational approaches to solve the philosophical concern in current knowledge representation methods.

The flaws inherent in these systems have led to empirical failures and the use of alternative models, incorporating an adaptive approach to AI. Static approaches to AI suffer the obvious disadvantage of being unable to adapt to unforeseen circumstances (Vernon & Furlong, 2007). Early models using the connectionist approach, while exhibiting adaptive capabilities, also lacked abstraction. This deficiency limited their generalization of newly acquired information. Symbolic planners, which are AI agents that attempt to construct plans of actions, were found to break down without well-defined initial inputs (Vernon & Furlong, 2007). The rigidities of static models led to the creation of hybrid systems, which attempt to combine symbolic representations and the strengths of neural networks (Vadinský, 2013).

Hybrid systems provide an AI epistemology combining the virtues of static abstraction—such as the ability to categorize concepts and represent them with labels—with the real-world context provided through adaptive models of knowledge representation, and the advantages of artificial neural networks. These systems strive to capture abstract forms through symbolic processes, while also retaining the adaptability and scalability of neural networks. This synwork can be seen in Sun's hybrid CLARION architecture, which enables a transfer of knowledge between symbolic modules and sub-symbolic modules (Vadinský, 2013). The architecture allows the system to modify the abstract categories by empirical means, so that the abstract templates can better model empiricism without altering its general knowledge representation. This hybrid architecture exemplifies the Platonic tension of universals and specifics, while providing a means to mediate between the tension so that neither empiricism nor abstraction are compromised.

To solve the epistemological concerns of knowledge representation in AI, hybrid approaches favor the concept of stability rather than static representations. As was alluded to earlier, the major concern with the PSSH in AI is that it represents knowledge as being static. Knowledge and concepts of the world are not often permanent in the real world; thus, the AI systems are forced to represent a very restricted interpretation of empiricism when they apply the PSSH. One of the major concepts of hybrid approaches is that the knowledge in AI does not need to be permanent but rather is dynamic. For instance, to describe sentences, humans have rules that explain their grammar as well as specific vocabulary for creating sentences. When translating these sentences for a different person or to write them down, humans will select and create symbols that the receiving party can interpret. However, we do not need to use the same words, nor do we have to arrange the word orders in the same ways, for the new sentence to have equivalent content (Vadinský, 2013). In short, humans use the rules of abstraction to achieve high performance for both of these behaviours.

Plato's philosophy of the "theory of forms" has applications that span beyond epistemology; digital metaphysics is a discipline that relies heavily on the Platonic ideal of "form." The philosopher Steinhart has argued that since computers are discrete and modular machines, they are the contemporary counterparts of Platonic forms. Steinhart explained the idea of digital metaphysics as such:

Digital metaphysics treats algorithms as universal entities that are independently of their substrates. All software implementations and Turing machines are nothing but physical instances that instantiate various aspects of a given algorithm. These multiple physical instances exist in different substrates, be they hardware or other software. However, they do not modify or change the algorithm in any way. It may be useful, but it is not required, for these physical instances to be concrete and tangible.

Thus, it can be seen that digital metaphysics represents all algorithms and implementations as abstractions of Platonic forms (Steinhart, 1998). This concept is useful for AI, but it may be problematic due to the fact that abstract forms may not fully capture empiricism. In other words, because digital metaphysics relies upon abstraction, it might not fully capture all facets of intelligence (Dascal, 1989).

With the physical symbol systems hypowork, classic AI researchers were able to mathematically encode abstract knowledge (Vernon & Furlong, 2007). Knowledge became formalistically represented, allowing machines to use symbols to manipulate and draw inferences on entities, such as in an expert system. From there, logic could be encoded and represented in a way to make predictions of other states and relationships, by which the problem domain could be formally codified and the solution logically deduced. All of this aligns with the Platonic ideal and dualistic epistemology. But if all knowledge representation is essentially the capture of abstract forms, it leaves all interpretations in the dark. And if systems can't distinguish interpretations from a situation or set of circumstances, the lack of flexibility will lead to inevitable breakdown for anything but formal problems (Vadinský, 2013). All systems of thought need to be able to recognize interpretations so as to properly contextualize them. Because AI can't reason on interpretations, their epistemology is incomplete. Because AI epistemology is incomplete, their metaphysics will always fail to fully encompass empiricism.

2.4 From Aristotle's Logic to Symbolic AI and Expert Systems

Aristotle's contributions to formal logic have profoundly influenced the development of symbolic artificial intelligence. His development of syllogistic logic provided the bedrock for computing and reasoning with deductive methods, upon which systems of logic-based AI are founded. In symbolic AI, this manifests in the explicit encoding of logical operators and inference rules to implement the process of deductive inference, commonly utilized in medical diagnosis and legal domains to create expert systems. These systems take the form of rule-based architectures and leverage deductive reasoning to navigate complex problem domains. The early expert system MYCIN is an example of how Aristotelian logic can be utilized in AI for large-scale data processing. The philosophical rationale for the utilization of Aristotelian logic in this way is embodied in the physical symbol system hypowork of Newell and Simon (1976), which states that the ability to intelligently manipulate formal symbols within structured environments is sufficient for generalized intelligent action. Although this methodology provides precision and transparency in the reasoning of AI systems, there remain limitations in the applicability of pure Aristotelian logic to the unpredictable phenomena of real-world events. According to Vernon and Furlong, the requirement for explicit specifications, inherent in the symbol system approach, can restrict the applicability and scalability of systems designed under Aristotelian principles.

The symbol system hypowork states that the ability to manipulate symbols according to a formal set of rules constitutes the foundations of all intelligent action, which deeply influences the architecture and applications of symbolic AI and embodies the transition of Aristotelian logic to the computational world. The embodiment of this hypowork within AI allows for symbolic systems to implement deductive, classification, and rule-based procedures to operate on symbolic data. This philosophy allows developers to create systems of symbolic AI that are able to perform specific and high-level reasoning in defined environments. Such systems are frequently used in professions that require compliance with certain regulatory regimes, such as legal practice and banking. This is because such systems are able to clearly track how they reached an output, providing logical reasons for a conclusion. These reasons can then be presented as evidence of compliance, making symbolic AI valuable in such areas. The system of symbolic AI is, however, vulnerable to fragility because of incomplete or ambiguous data and, therefore, requires rigorous control of knowledge and data. From a philosophical perspective, symbolic AI mirrors Aristotle's desire to categorize the world into rigidly defined forms, mirroring genus and species, in order to allow for explicit knowledge in an environment that allows inference across categorically structured domains.

The implementation of symbolic AI utilizes hierarchical structure, representing the domain of intelligence within explicit and discrete categories. This means that the performance of symbolic AI is optimal within controlled and defined spaces. The capability to represent and operate across defined domains in this manner allows for high-level problem-solving capabilities and is highly useful for high-performance computation across domains, such as diagnostic medicine, where phenomena are explicitly labelled and categorized. As Vernon and Furlong suggest, this approach to representing and reasoning about real-world phenomena has been the foundation for much of the work in symbolic AI. Such a focus has also opened certain philosophical vulnerabilities. This relates to the weakness of systems that can be easily overwhelmed in situations where categories are not well-defined or well understood. Such circumstances require the creation of more flexible and adaptive categorizations. Under these conditions, symbolic systems become overwhelmed, losing their benefits, resulting in ungeneralizable operations, such as interpreting the meaning of natural language. These conditions therefore indicate that Aristotelian-style categorizations and representations are inadequate to describe the diversity of all cognition. This is captured by the "frame problem," described by Dennett (1987). It suggests that in the context of real-world problems, the complexity and ungeneralizability of a formal system can easily render the implementation impossible due to the inability to determine the correct course of action.